

**JOSÉ PAULO GARCIA DE CASTRO**

**A EPISTEMOLOGIA DA ESCOLHA**

**Sobre a possibilidade da simulação artificial da inteligência humana**

---

**Orientador: Manuel da Costa Leite**

**Universidade Lusófona de Humanidades e Tecnologia**

**Escola de Comunicação, Arquitetura, Artes e Tecnologias da Informação**

**Lisboa**

**2014**

**JOSÉ PAULO GARCIA DE CASTRO**

**A EPISTEMOLOGIA DA ESCOLHA**

**Sobre a possibilidade da simulação artificial da inteligência humana**

---

Tese apresentada para a obtenção do Grau de Doutor em Filosofia na Especialidade de Pensamento Contemporâneo, conferido pela Universidade Lusófona de Humanidades e Tecnologias.

Orientador: Prof. Doutor Manuel da Costa Leite

**Universidade Lusófona de Humanidades e Tecnologia**

**Escola de Comunicação, Arquitetura, Artes e Tecnologias da Informação**

**Lisboa**

**2014**

Ser é escolher.

A Ana Luísa, Madalena, Ricardo e Francisco.

## **AGRADECIMENTOS**

À família, por me acolher.

A meus pais, pelo apoio incondicional.

Ao Professor Manuel da Costa Leite, Scholar e Orientador desta tese, respeitador severíssimo dos meus timings extensos e aleatórios.

Aos Mestres que me marcaram ao longo do meu percurso académico, Pofessores Jorge Crespo, José Anes, Carlos de Jesus, Augusto Mesquitela Lima, João Andrade e Silva, João Resina Rodrigues, Luís Moniz Pereira, João Bau, José Luís Simões da Fonseca, José Ramalho Croca e Rui Nobre Moreira.

Aos meus amigos, pela generosa paciência.

## Resumo

A simulação do pensamento humano é um dos objectivos principais da Inteligência Artificial. No entanto, a sua exequibilidade é um problema que permanece em aberto. O presente trabalho tem como objectivo contribuir para a discussão filosófica da exequibilidade desse programa, aduzindo um argumento especulativo que se designou por «Argumento da Variabilidade Nomotética». O argumento conduz a uma contradição que sugere a impossibilidade computacional daquela simulação e a existência de livre arbítrio na natureza.; e deixando em aberto a possibilidade de uma teoria sobre a natureza que inclua a volição como uma das suas propriedades. Apresenta-se ainda, num plano meramente conjectural, especulativo, uma estrutura de análise possível para a acção volitiva humana, como contributo para o que designámos uma Epistemologia da Escolha.

*Palavras chave: Inteligência Artificial; Livre arbítrio, Volição, Argumento da Variabilidade Nomotética; Epistemologia da Escolha.*

**Abstract**

The simulation of the human thought is one of the main objectives of Artificial Intelligence. Its feasibility, however, is a problem that remains open. The present work aims at contributing to the philosophical inquiry into the feasibility of this program, introducing a philosophically exploratory “argumento da variabilidade nomotética” [nomothetic variability argument]. This argument leads to a contradiction that suggests the computational impossibility of that simulation and the existence of free will in nature. Also, the argument gives way to the possibility of a theory of nature which includes *volition* as one of its properties. A framework for the analysis of human volition is also introduced, contributing to what is anticipated as an Epistemology of Choice.

*Keywords:* Artificial Intelligence; Human Thought; Volition; Free Will, Nomothetic Variability; Epistemology of Choice.

# A Epistemologia da Escolha

Sobre a possibilidade da simulação artificial da inteligência humana

---

## ÍNDICE GERAL

<b>INTRODUÇÃO.....</b>	<b>11</b>
<b>A SIMULAÇÃO ARTIFICIAL DA INTELIGÊNCIA HUMANA.....</b>	<b>11</b>
<b>1º CAPÍTULO</b>	
<b>CRÍTICAS À SIMULAÇÃO ARTIFICIAL DA INTELIGÊNCIA HUMANA.....</b>	<b>14</b>
1.1. AS OBJECÇÕES DE HUBERT DREYFUS PARA A COMPUTABILIDADE DO SIGNIFICADO.....	14
1.2. O ARGUMENTO GÖDELIANO DE JOHN LUCAS.....	17
1.3. A CRÍTICA MORAL DE JOSEPH WEIZENBAUM.....	20
1.4. JOHN SEARLE E O ARGUMENTO DA SALA CHINESA.....	23
1.5. HILARY PUTNAM E O MODELO FUNCIONALISTA.....	26
1.6. A TESE FISCALISTA DE RODNEY BROOKS.....	30
1.7. O MODELO ORCH OR DE PENROSE-HAMEROFF.....	32
<b>2º CAPÍTULO</b>	
<b>BREVE HISTÓRIA DA INTELIGÊNCIA ARTIFICIAL.....</b>	<b>44</b>
2.1. OS PARADIGMAS SIMBÓLICO E CONEXIONISTA DA IA.....	44
2.2. O PROBLEMA DO CONHECIMENTO DE SENSO COMUM.....	46
2.3. AS TRÊS FASES HISTÓRICAS DA INTELIGÊNCIA ARTIFICIAL.....	48
2.3. OS CONTRIBUTOS FORMALISTAS.....	51
2.4. OS CONTRIBUTOS FISCALISTAS.....	52
<b>3º CAPÍTULO</b>	
<b>FÍSICA E COMPUTAÇÃO.....</b>	<b>55</b>
3.1. TERMODINÂMICA DA COMPUTAÇÃO: IRREVERSIBILIDADE LÓGICA E O PRINCÍPIO DE LANDAUER.....	56
3.2. COMPUTAÇÃO NA NATUREZA E O PRINCÍPIO DA EQUIVALÊNCIA COMPUTACIONAL.....	60
3.3. COMPUTAÇÃO QUÂNTICA E O PRINCÍPIO DE DEUTSCH-TURING.....	63
3.4. IDENTIDADE ONTOLÓGICA ENTRE REALIDADE E COMPUTAÇÃO: TEORIAS ALGORÍTMICAS DE TUDO.....	67
3.5. INFORMAÇÃO E LIMITES ENTRÓPICOS NA NATUREZA.....	70
3.6. O PRINCÍPIO HOLOGRÁFICO E A CONSERVAÇÃO DE INFORMAÇÃO NA NATUREZA.....	74
3.7. TEORIAS DA UNIFICAÇÃO, INFORMAÇÃO E QUANTIFICAÇÃO DO ESPAÇO.....	88
<b>4º CAPÍTULO</b>	
<b>SISTEMAS FORMAIS E COMPORTAMENTOS NA NATUREZA.....</b>	<b>94</b>
4.1. PROPRIEDADES DA INTELIGÊNCIA HUMANA.....	94
4.2. SISTEMAS FORMAIS.....	95
4.3. PROCESSOS DE PROVA NUM SISTEMA FORMAL E COMPORTAMENTOS NA NATUREZA.....	103
4.4. APLICAÇÃO DO PROCEDIMENTO NOMOTÉTICO.....	117
4.5. ARGUMENTO DA VARIABILIDADE NOMOTÉTICA PARA A EXISTÊNCIA DE VOLIÇÃO NA NATUREZA.....	127
<b>5º CAPÍTULO</b>	
<b>PARA UMA EPISTEMOLOGIA DA ESCOLHA.....</b>	<b>131</b>
5.1. UMA ESTRUTURA POSSÍVEL PARA A ESCOLHA.....	131
5.2. CASOS ESTRUTURAIS DA ESCOLHA.....	139
5.3. APLICAÇÕES DO ESQUEMA DA ESCOLHA.....	145
5.4. UMA ONTOLOGIA DA ESCOLHA.....	152
<b>CONCLUSÃO.....</b>	<b>161</b>
<b>BIBLIOGRAFIA.....</b>	<b>162</b>
<b>ÍNDICE REMISSIVO.....</b>	<b>175</b>

<b>APÊNDICE 1 .....</b>	<b>180</b>
<b>UMA VERSÃO FISCALISTA DO ARGUMENTO DA VARIABILIDADE NOMOTÉTICA .....</b>	<b>180</b>
<b>1. PRINCÍPIO DA RESOLUÇÃO HOLOGRÁFICA .....</b>	<b>181</b>
<b>1.1. Entropia termodinâmica e informação binária.....</b>	<b>181</b>
<b>1.2. Limite entrópico, entropia termodinâmica mínima de um sistema físico e área mínima associável à observação de um sistema físico. ....</b>	<b>197</b>
<b>1.3. O paradoxo holográfico.....</b>	<b>212</b>
<b>1.4. A condição de não contradição na natureza e a relativização do observador.....</b>	<b>217</b>
<b>1.5. O critério lógico de realismo, o Princípio da Resolução Holográfica e a solução do paradoxo holográfico.....</b>	<b>220</b>
<b>1.6. Informação e o Princípio da Resolução Holográfica. ....</b>	<b>224</b>
<b>2. ARGUMENTO DA VARIABILIDADE NOMOTÉTICA (VERSÃO FISCALISTA). ....</b>	<b>232</b>
<b>2.1. Coordenação holográfica de um evento numa região de espaço.....</b>	<b>232</b>
<b>2.2. Propriedades de uma teoria sobre a natureza. ....</b>	<b>239</b>
<b>2.3. A lista de todos os valores de abcissa numa coordenação holográfica de eventos..</b>	<b>241</b>
<b>2.4. Proposições universais e contingentes numa teoria sobre a natureza.....</b>	<b>245</b>
<b>2.5. Dedução do Argumento da Variabilidade Nomotética.....</b>	<b>250</b>

## Índice de Figuras

FIGURA 5.1.....	131
FIGURA 5.2.....	133
FIGURA 5.3.....	140
FIGURA 5.4.....	143
FIGURA 5.5.....	144
FIGURA 5.6.....	148
FIGURA 5.7.....	149
FIGURA 5.8.....	151
FIGURA A2.1.....	238
FIGURA A2.2.....	244
FIGURA A2.3.....	249
FIGURA A2.4.....	250

## Introdução.

### A simulação artificial da inteligência humana.

A ideia de que é possível construir um robô com capacidades intelectuais e emocionais humanas impregna a imaginação popular nas culturas ocidentalizadas e está subentendida em grande parte da literatura científica sobre inteligência artificial. Reemerge frequentemente nos filmes e romances de ficção científica e representa um feito tecnológico que se crê estar no horizonte próximo da ciência. Este optimismo idealista não é novo. Acompanha a disciplina desde o momento da sua inepção.

Crevier (1994) descreve como na primeira conferência sobre Inteligência Artificial, em 1956, no Dartmouth College (Hanover, EUA), parecia seguro que a solução do problema estava próximo. Ali, alguns dos pioneiros da IA definiram as linhas de um programa de investigação, no mínimo, ambicioso. John McCarthy [1927 - 2011], Claude Shannon [1916 - 2001], Marvin Minsky [1927 - ] e Nathaniel Rochester [1919 - 2001] sugeriram levar a cabo uma investigação de *dois meses* com o objectivo de encontrar soluções computacionais para algumas das propriedades da inteligência humana com uma equipa de dez cientistas:

“We propose that a 2 month, 10 man study of artificial intelligence be carried out during the summer of 1956 at Dartmouth College in Hanover, New Hampshire. The study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it. An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves. We think that a significant advance can be made in one or more of these problems if a carefully selected group of scientists work on it together for a summer”

(McCarthy *et al*, 1956)

O desafio não seria resgatado nas décadas seguintes. Ao entusiasmo inicial de Dartmouth seguir-se-ia um discurso progressivamente mais conservador sobre as condições que, hipoteticamente, permitem efectivar a simulação artificial do pensamento humano. Numa página da Internet, alojada na Universidade de Stanford, escreveria McCarthy, cinquenta anos depois:

“A few people think that human-level intelligence can be achieved by writing large numbers of programs of the kind people are now writing and assembling vast knowledge bases of facts in the languages now used for expressing knowledge. However, most AI researchers believe that new fundamental ideas are required, and therefore it cannot be predicted when human-level intelligence will be achieved.(...)”

The philosopher John Searle [1932 -] says that the idea of a non-biological machine being intelligent is incoherent. He proposes the Chinese room argument. The philosopher Hubert Dreyfus [1929 -] says that AI is impossible. The computer scientist Joseph Weizenbaum [1923 - 2008] says the idea is obscene, anti-human and immoral. Various people have said that since artificial intelligence hasn't reached human level by now, it must be impossible. Still other people are disappointed that companies they invested in went bankrupt.“

(McCarthy, 2008)

Ao fim de meio século, a simulação artificial do pensamento humano resiste de forma formidável às abordagens computacional e robótica. Poderá suceder que este objectivo programático seja tarefa obviada por escolhos epistemológicos? Poderá suceder que a realidade seja demasiado complexa para que possa ser descrita por sistemas informáticos, implementáveis num robô, ou que a mimetização electrónica de um cérebro se revista de uma complexidade estrutural, tecnologicamente inalcançável? Existirá um limite que impeça a simulação artificial da inteligência humana? Estará qualquer linguagem de computação com esse objectivo sujeita a uma impossibilidade fundamental, por mais sofisticada que seja a sua performance lógica ou a forma como descreve a realidade?

O presente trabalho tem como objectivo contribuir para a discussão filosófica sobre a exequibilidade deste programa. Apresentando de forma especulativa um argumento que sugere a existência de volição na realidade, como propriedade não capturável de ponto de vista computacional. O argumento parte da hipótese fundadora que os comportamentos dos objectos na natureza são regidos logicamente por um sistema formal, uma estrutura ontológica abstracta, que se aceita actuar na natureza de forma mecânica e determinista, sem interpretações associadas.

A estrutura especulativa do argumento envolve os seguintes passos:

i) Descreve-se, *arguendo*, uma estrutura de sistema formal (baseado referencialmente em Kleene), considerando-se, por um lado, que parte dos procedimentos conducentes a uma prova no sistema correspondem a comportamentos na natureza e, por outro, que o sistema pretende formalizar a generalidade das teorias matemáticas informais

disponíveis. Considera-se ainda que sobre esse sistema formal pode sempre conceber-se uma metateoria que analise os processos conducentes a uma prova no sistema.

ii) Define-se um procedimento metamatemático de validação das fórmulas e das inferências no sistema chamado procedimento nomotético e uma função sobre os números naturais que utiliza o resultado dessa validação.

iii) Constrói-se um procedimento de análise metamatemática, agenciado por um sistema físico que se aceita, *arguendo*, exibir comportamento “inteligente”. O procedimento total envolve, entre outros, o procedimento nomotético.

iv) Analisam-se as consequências lógicas dos procedimentos metamatemáticos ao sistema. Essa análise revela, porém, a ocorrência de uma contradição.

v) Ocorrência de contradição: discute-se a origem da contradição e sugere-se que a causa se pode atribuir à aplicação do procedimento nomotético.

vi) Uma vez que a aplicação do procedimento nomotético, por parte do agente, corresponde, como vimos presumidamente, a um comportamento na natureza, sugere-se que a sua concatenação com outros procedimentos metamatemáticos sobre o sistema pode ocorrer, sem que haja uma influência sobre a integridade lógica do sistema, produzindo contradição.

vii) No seguimento do que foi proposto, sugere-se que podem existir concatenações de comportamentos físicos, que, isoladamente, correspondem a processos de prova no sistema, mas cuja concatenação global não corresponde a um processo de prova, de acordo com as regras de transformação do sistema.

viii) Atribui-se a origem dessa concatenação de processos físicos, independente do sistema formal, à existência de uma propriedade que se sugere poder ser identificada com *volição*.

ix) Desenvolve-se um esquema representando, presumivelmente, uma estrutura possível para o acto de escolha humano, como contribuição provisória para uma Epistemologia da Escolha.

x) Conclui-se a dissertação, discutindo as consequências da presente conjectura para a exequibilidade da simulação artificial da inteligência humana.

## **1º Capítulo**

### **Críticas à simulação artificial da inteligência humana.**

Desde o início da segunda metade do século XX que têm sido apresentadas críticas ao oásis programático da inteligência artificial, questionando sobretudo a ideia de que o fenómeno inteligente possa ser produzido através de um cálculo simbólico. Algumas dessas vozes dissonantes, pela ordem cronológica das suas intervenções, foram as de Hubert Dreyfus [1929 - ]; John Lucas [1929 - ]; Joseph Weizenbaum [1923 - 2008]; Hilary Putnam [1926 - ]; John Searle [1932 - ]; Rodney Brooks [1950 - ], Roger Penrose [1931 - ] e Stewart Hammerof [1947 - ].

#### **1.1. As objecções de Hubert Dreyfus para a computabilidade do significado.**

Hubert Dreyfus [1929 - ], o primeiro pensador a criticar o programa racionalista da IA, é admirador e estudioso diligente da obra do filósofo alemão Martin Heidegger [1889 - 1976]. Ensina a fenomenologia existencialista heidegueriana e também a do filósofo francês Maurice Merleau-Ponty [1908 - 1961] na Universidade da Califórnia, em Berkeley (UC Berkeley).

A sua crítica ao programa da AI remonta a 1965, quando era professor no M.I.T. A pedido de Paul Armer [1924 - ], responsável pelo departamento de análise numérica da RAND Corporation - Research and Development. Dreyfus avaliou, nessa altura, o estado da simulação computacional da inteligência. Dessa análise resultou «Alchemy and AI» (Dreyfus, 1965), um texto irónico que comparava a utilização de códigos intrincados para simular a inteligência humana às tentativas dos alquimistas para transformar chumbo em ouro. Receando causar um estado de sítio entre a comunidade da IA, Armer adiou a publicação do relatório, vindo a fazê-lo, apenas mais tarde, sob a forma de um memorando da RAND.

Dreyfus recenseou quatro teses sobre a simulação artificial da inteligência humana que ainda se mantêm actuais:

## a) A tese biológica:

O cérebro processa informação através de procedimentos efectivos, discretamente, passo a passo, utilizando sistemas biológicos que são análogos a interruptores binários («on/off switches»).

## b) A tese psicológica:

O funcionamento da mente é totalmente equivalente ao funcionamento de um dispositivo que processa bits, de acordo com um conjunto de regras.

## c) A tese epistemológica:

Todo o conhecimento sobre o mundo pode ser formalizável num discurso lógico que representa o que se passa no mundo.

## d) A tese ontológica:

O mundo é um conjunto de factos independentes representável por um conjunto de símbolos independentes.

(Dreyfus, 1992)

Será a validade das três últimas teses que Dreyfus recusará liminarmente, em face do pensamento holístico de Heidegger (Dreyfus, 1992). O idealismo cartesiano mantém que o universo da *res extensa* pode ser descrito por um discurso lógico que articula símbolos associados a factos. Símbolos que são independentes entre si e que não necessitam de qualquer informação contextual sobre a natureza específica da realidade para que, sobre ela, integrem um discurso causal, consistente e verdadeiro. Fortemente influenciado pela tradição científica da física e, em especial, pelo êxito da mecânica, o racionalismo cartesiano sustenta que o discurso sobre a natureza, sendo uma mescla elegante entre linguagem natural e matemática, se poderá aproximar de um cálculo simbólico, dispensando os escolhos subjectivos da interpretação humana. Pensar é computar, apenas outra forma de fazer matemática: mais complexa e sofisticada, criando a beleza ilusória da consciência. Mas pode o pensamento humano prestar-se a este tipo de descrição? À medida que a IA progredia, já depois de «Alchemy and AI», Dreyfus apontava as dificuldades que sistemas

automatizados pareciam encontrar quando se deparavam com os factos simples da vida, aquilo a que o autor chamou «the commonsense knowledge». A mais inocente das situações poderia revelar-se problemática. Uma frase inocente como «Mary saw a dog in the window. She wanted it» poderia facilmente induzir o computador em erro. Terá a máquina compreendido que é o animal o objecto do desejo de Mary e não a janela? (Dreyfus, 1992, p. xix) Para a ortodoxia da IA a solução deste tipo de problemas corresponderia apenas à codificação adequada da informação, quer em termos de quantidade, quer em termos da complexidade das relações entre os dados. Programme-se em extensão suficiente e o computador finalmente “compreenderá”. Para Dreyfus, pelo contrário, a situação não era assim tão simples. A existência de um contexto significativo corresponde a uma dimensão hermenêutica e holística da realidade que nenhuma linguagem de computação consegue capturar. Seguindo Heidegger, a acção humana sobre o mundo é, por definição, uma construção teleológica de significados, em que cada ser humano age em direcção a um conjunto de objectivos. Essa demanda existencial implica capacidade de interpretação e de atribuição de significados à realidade. Um objecto, por exemplo, que não é decorativo pode ter uma forma inadequada para essa função, mas tornar-se-á num, assim que um sujeito o interprete como tal. Há sempre oportunidade para que se forme um novo significado projectado sobre a realidade, rebelde e indomável em relação a qualquer tentativa prévia de o capturar, programando. A actividade holística de interpretação do mundo que Dreyfus atribui ao pensamento humano implica que, por muitos factos que o idealismo cartesiano aconselhe a programar, haverá sempre um elemento de imponderabilidade no processo de relação entre a máquina e o meio ambiente. Na perspectiva de Dreyfus, um ser humano escolhe o modo como pretende existir, interpretando-se a si e à realidade envolvente, à medida que prossegue. A primitividade da escolha e a invisibilidade dessa interpretação em relação aos procedimentos mecânicos de uma linguagem lógica, fazem destes dois requisitos de humanidade propriedades aparentemente não computáveis. Propriedades cuja descrição está para além das possibilidades computacionais de qualquer sistema. Esta conclusão parece substanciada pelo desenvolvimento da IA nas últimas décadas.

No prefácio de «What Computers *Still* Can't Do» (Dreyfus, 1992), escreve o autor:

“This edition of What Computers Can't Do marks (...) a change of status. The book now offers not a controversial position in an ongoing debate but a view of a bygone period of history. For now that the twentieth century is drawing to a close, it is becoming clear that one of the great dreams of the century is ending too. Almost half a century ago computer pioneer Alan Turing suggested that a high-speed digital computer, programmed with rules and facts, might exhibit intelligent

behavior. Thus was born the field later called artificial intelligence (AI). After fifty years of effort, however, it is now clear to all but a few diehards that this attempt to produce general intelligence has failed. This failure does not mean that this sort of AI is impossible; no one has been able to come up with such a negative proof. Rather, it has turned out that, for the time being at least, the research program based on the assumption that human beings produce intelligence using facts and rules has reached a dead end, and there is no reason to think it could ever succeed. Indeed, what John Haugeland has called Good Old-Fashioned AI (GOFAI) is a paradigm case of what philosophers of science call a degenerating research program”.

(Dreyfus, 1992, p. ix)

Concordando-se com a posição de Hubert Dreyfus, será aqui proposto que a adoção de um certo conjunto de asserções, compatíveis com uma descrição logicista da realidade, conduz a um argumento que sugere fortemente que aquela simulação é procedimentalmente impossível. Embora a inteligência, enquanto fenómeno, possa emergir associada a um conjunto de fenómenos físicos, será proposto que essa emergência não pode ser reproduzida recorrendo a uma estratégia racionalista de descrição da realidade. Noutros termos, será sugerido que a inteligência humana é um fenómeno, em essência, não computável, nos termos em que qualquer discurso lógico se desenvolve. A razão fundamental apontada é a de que no seio da actividade inteligente existe, manifestando-se de forma irrevogável, uma propriedade a que se chamará *volição*, cuja expressão corresponde à efectivação de escolhas no mundo. Será ainda proposto que a volição é uma propriedade metalógica no universo, independente de qualquer imperativo lógico-formal apriorístico sobre comportamentos na natureza. A sua actividade sendo, portanto, independente das engrenagens de qualquer fenomenologia mecanicista, cuja descrição possa ser feita de forma lógica. Nestes termos, entende-se que a atribuição de significado à realidade, invocada por Dreyfus, é uma instanciação de um acto de volitivo, como tal não computável.

## 1.2. O argumento gödeliano de John Lucas.

John Randolph Lucas [1929 - ] foi professor de filosofia na universidade inglesa de Oxford até 1996, altura em que se jubilou. Lucas relacionou o primeiro teorema da incompletude de Kurt Gödel [1906 - 1978] (Gödel, 1931) com o conceito de livre arbítrio,

defendendo que existe uma diferença, em termos procedimentais, entre um computador e um ser humano.

Gödel no seu artigo «On Formally Undecidable Propositions» restringiu-se ao sistema formal definido nos Principia Mathematica, publicados em 1925 por Bertrand Russell [1872 - 1970] e Alfred North Whitehead [1861 - 1947]. Porém, os teoremas de incompletude com o seu nome aplicam-se a sistemas formais, contendo axiomas e regras mecânicas de prova, que podem ser modelos abstractos de teorias sobre números naturais. Gödel mostrou como um predicado teórico-numérico pode ser associado a uma noção metamatemática sobre o sistema. Essa noção é pensada numa metateoria sobre as capacidades de prova do sistema formal, da qual este é a teoria objecto. O predicado  $p$  que dá origem ao primeiro teorema da incompletude é tal que, interpretado intuitivamente na metateoria do sistema formal, afirma sobre si mesmo que o sistema não consegue provar o predicado  $p$ . Esse predicado, por outro lado, codificada na teoria informal numérica, de que o sistema é também um modelo abstracto, corresponde a uma afirmação aritmética que é vista ser verdadeira por mera inspecção. O primeiro dos teoremas de Gödel, resultante destes factos, estabelece que «qualquer teoria gerada, utilizando procedimentos finitistas, capaz de expressar a aritmética elementar não pode ser simultaneamente consistente e completa. Em particular, para qualquer teoria formal consistente e gerada utilizando aqueles procedimentos, provando certa verdade da aritmética básica, existe uma afirmação aritmética que é verdadeira, mas não demonstrável na teoria» (Kleene 1967, p 250).” O valor lógico do predicado  $p$  não advém de uma prova a partir dos axiomas e regras de inferência válidos no sistema. Pelo contrário, a verdade daquele predicado é atribuída de forma intuitiva pelo sujeito humano. De forma metalógica em relação ao sistema. Isto é; independentemente das possibilidades dedutivas daquele, por mera inspecção do conteúdo do predicado. Em certo sentido, até de forma oposta àquelas possibilidades, porquanto, para o sistema, o reconhecimento da verdade sobre o que quer que seja é equivalente à sua prova. Ora, no caso do primeiro teorema de Gödel (1931), de ponto de vista metateórico, o sujeito humano é capaz de reconhecer que o sistema afirma sobre si mesmo ser incapaz de provar a verdade de uma certa proposição e, portanto, de a reconhecer, à sua medida, mecanicamente. O predicado verdadeiro  $p$ , que afirma que o sistema não consegue provar  $p$ , é exactamente a afirmação que fica por provar no sistema. Ao matemático, ao ser humano sentiente, torna-se inescapável a conclusão de que, tendo sido bem formulado o predicado e correspondendo aquele a uma verdade algébrica intuitivamente constatável, o sistema não pode ser simultaneamente consistente e ser tal que todas as fórmulas ali escritas resultem de uma prova. Porquanto, afirmar que o sistema é completo conduz à

negação metateórica de  $p$  e, conseqüentemente, à negação da afirmação algébrica verdadeira que lhe corresponde, produzindo uma inconsistência no sistema.

Esta capacidade da inteligência humana para discernir a verdade de forma metatórica ou, como diremos, metalógica, distingue-nos de uma máquina que manipula símbolos, seguindo regras mecânicas, de forma completamente indiferente em relação aos significados que a combinação daqueles símbolos importa. Não pode assim um ser humano ser uma máquina cibernética como afirma John Lucas em «Minds, Machines and Gödel» onde gizou o seu argumento pela primeira vez:

“Gödel's theorem must apply to cybernetical machines, because it is of the essence of being a machine, that it should be a concrete instantiation of a formal system. It follows that given any machine which is consistent and capable of doing simple arithmetic, there is a formula which it is incapable of producing as being true - i.e., the formula is unprovable-in-the-system-but which we can see to be true. It follows that no machine can be a complete or adequate model of the mind, that minds are essentially different from machines”.

(Lucas, 1959)

A comparação entre agente humano e máquina sugere a presença de uma forma de indeterminismo lógico na acção do agente, sugerindo a existência de livre arbítrio na acção humana, temática posteriormente desenvolvida pelo autor em «The Freedom of the Will», em 1970. O argumento de Lucas inaugura as demonstrações ditas godelianas da existência de uma diferença substancial entre seres humanos e máquinas na crítica ao mecanicismo computacional da IA, de que outro exemplo se pode encontrar em Roger Penrose [1931 - ], descrito mais à frente.

O argumento apresentado nesta dissertação pretende ir mais além do que um argumento do tipo godeliano. Não pretendemos comparar as capacidades entre máquinas e seres humanos na atribuição de um valor de verdade a uma certa afirmação. Considerando a mesma acepção racionalista que vê o pensamento humano como uma manipulação de objectos, de acordo com regras mecânicas, assumiremos, de forma geral, que todos os comportamentos na natureza podem ser vistos como manipulações de objectos, sujeitas a regras. Essas regras serão intuitivamente conotadas com as leis da natureza, sendo a acção humana de pensar um daqueles comportamentos, em particular. Esperaremos, desse modo, dar força ao argumento de Lucas e fundamentar a possível existência de livre arbítrio no comportamento humano.

### 1.3. A crítica moral de Joseph Weizenbaum.

Joseph Weizenbaum [1923 - 2008] é um dos fundadores da inteligência artificial como disciplina científica. Nascido em Berlim, emigrou para os Estados Unidos em 1936 onde viria a estudar matemática e, mais tarde, a participar no desenvolvimento do computador. Em 1964 tornou-se professor de computação no M.I.T. («Massachusetts Institute of Technology»), onde, dois anos depois, escreveria o código de «Eliza» um dos programas mais emblemáticos da história da IA. As suas reflexões sobre a relação entre humanos e computadores, a que chamou artefactos ilusoriamente inteligentes, foram o início de uma linha de pensamento sobre a natureza ética dessa relação. Nomeadamente, sobre os problemas éticos que uma visão excessivamente racionalista do mundo pode colocar. A antropomorfização dos robôs e, em consequência, a empatia que os seres humanos poderão estabelecer de forma natural com máquinas, produzem a ilusão de que há uma interacção social genuína com um ser cognoscente, quando se trata afinal de um mecanismo sofisticado de relojoaria. Quando muito, de um espelho que reflecte humanidade; por vezes de forma perversa. Depois de uma catarse pessoal, face aos efeitos perversos de Eliza, Weizenbaum escreve, em 1977, «Computer power and human reason», onde descreve a forma como o público reagiu a Eliza:

“My own shock was administered (...) by some people who insisted on misinterpreting a piece of work I had done. I write this without bitterness and certainly not in a defensive mood. Indeed, the interpretations I have in mind tended, if anything, to overrate what little I had accomplished and certainly its importance (...)

The work was done in the period 1964-1966 (...) I composed a computer program with which one could "converse" in English. The human conversationalist partner would type his portion of the conversation on a typewriter connected to a computer, and the computer, under control of my program, would analyze the message that had so been transmitted to it, compose a response to it in English, and cause the response to be typed on the computer's typewriter (...) the program was constructed in a two-tier arrangement, the first tier consisting of the language analyzer. and the second of a script. The script is a set of rules rather like those that might be given to an actor who is to use them to improvise around a certain theme. (...)

For my first experiment, I gave ELIZA a script designed to permit it to play (I should really say parody) the role of a Rogerian psychotherapist engaged in an initial interview with a patient. The Rogerian psychotherapist is relatively easy to imitate because much of his technique consists of drawing his patient out by reflecting the patient's statements back to him."

(Weizenbaum, 1977, pp. 2-3)

"The shocks I experienced as DOCTOR became widely known and "played" were due principally to three distinct events.

1. A number of practicing psychiatrists seriously believed the DOCTOR computer program could grow into a nearly completely-automatic form of psychotherapy. (...)

2. I was startled to see how quickly and how very deeply people conversing with DOCTOR became emotionally involved with the computer and how unequivocally they anthropomorphized it. Once my secretary, who had watched me work on the program for many months and therefore surely knew it to be merely a computer program, started conversing with it. After only a few interchanges with it, she asked me to leave the room. Another time, I suggested I might rig the system so that I could examine all conversations anyone had had with it, say, overnight. I was promptly bombarded with accusations that what I proposed amounted to spying on people's most intimate thoughts; clear evidence that people were conversing with the computer as if it were a person who could be appropriately and usefully addressed in intimate terms. (...)

3. Another widespread, and to me surprising, reaction to the ELIZA program was the spread of a belief that it demonstrated a general solution to the problem of computer understanding of natural language. (...) I had tried to say that no general solution to that problem was possible, i.e., that language is understood only in contextual frameworks, that even these can be shared by people to only a limited extent, and that consequently even people are not embodiments of any such general solution. But these conclusions were often ignored".

(Weizenbaum, 1977, pp. 4; 6-7)

A respeito da diferença entre decisão e escolha, entendida como expressão de volição, escreve ainda Weizenbaum:

"[people] make decisions all day long, every day. But they appear not to make choices. They are, as they themselves testify, like Winograd's robot. One asks it

‘Why did you do that?’ and it answers ‘Because this or that decision branch in my program happened to come out that way.’ And one asks ‘Why did you get to that branch?’ and it again answers in the same way. But its final answer is ‘Because you told me to.’ Perhaps every human act involves a chain of calculations at what a systems engineer would call decision nodes. But the difference between a mechanical act and an authentically human one is that the latter terminates at a node whose decisive parameter is not ‘Because you told me to’ but ‘Because I chose to’.

(Weizenbaum, 1977, pp. 259-260)

A crítica fundamental de Weizenbaum à IA incide, não sobre a sua possibilidade procedimental, mas sobre as respectivas consequências éticas da existência de sistemas artificiais inteligentes para a experiência emocional humana.

“There are (...) two kinds of computer applications that either ought not be undertaken at all, or, if they are contemplated, should be approached with utmost caution.

The first kind I would call simply obscene. These are ones whose very contemplation ought to give rise to feelings of disgust in every civilized person. The proposal I have mentioned, that an animal's visual system and brain be coupled to computers, is an example. It represents an attack on life itself. One must wonder what must have happened to the proposers' perception of life, hence to their perceptions of themselves as part of the continuum of life, that they can even think of such a thing, let alone advocate it. On a much lesser level, one must wonder what conceivable need of man could be fulfilled by such a ‘device’ at all let alone advocate it.

I would put all projects that propose to substitute a computer system for a human function that involves interpersonal respect, understanding, and love in the same category. I therefore reject (...) [the] proposal that computers be installed as psychotherapists, not on the grounds that such a project might be technically infeasible, but on the grounds that it is immoral”.

(Weizenbaum, 1977, pp. 268-269)

A crítica à possibilidade da simulação artificial da inteligência pode assim ser de natureza moral. Embora não seja esse o trajecto argumentativo escolhido neste trabalho, não podemos deixar de o referir como linha de reflexão interessante, podendo vir a revelar uma relação entre lógica e ética.

#### 1.4. John Searle e o argumento da sala chinesa.

John Searle [1932 - ], professor de filosofia na Universidade da Califórnia, em Berkeley (UC Berkeley), ocupa-se de questões ligadas à filosofia da mente e da linguagem. Desferiu um ataque crítico à possibilidade da simulação artificial da inteligência humana quando, em 1984, foi orador convidado para leccionar as Palestras Reith, organizadas anualmente pela BBC (British Broadcasting Corporation) <sup>1</sup>. Preparou então um conjunto de seis conferências a que chamou «Minds, Brains and Science» (Searle, 1984), transmitidas radiofonicamente e depois compiladas num livro com o mesmo nome.

A crítica de Searle incide sobre o funcionalismo simbólico ou o computacionalismo, doutrina que propõe que a actividade inteligente humana pode, em princípio, ser reproduzida por qualquer artefacto capaz de manipular símbolos, de acordo com um conjunto de regras sintáticas. Searle começa por reconhecer a validade epistemológica dos conceitos a que chama «fiscalismo naïve» e «mentalismo naïve»:

“(...) on my view, the mind and the body interact, but they are not two different things, since mental phenomena just are features of the brain. (...) Suppose we define 'naive physicalism' to be the view that all that exists in the world are physical particles with their properties and relations. (...) And let us define 'naive mentalism' to be the view that mental phenomena really exist. (...) Naive mentalism and naive physicalism are perfectly consistent with each other. Indeed, as far as we know anything about how the world works, they are not only consistent, they are both true”.

(Searle, 1984 pp. 26 - 27)

Embora Searle admita que estados mentais são aspectos do funcionamento do cérebro, adoptará uma posição crítica em relação à possibilidade de simular esses estados através de uma manipulação mecânica de símbolos. Searle desenvolverá a sua crítica, usando a expressão «strong AI» (inteligência artificial forte):

---

<sup>1</sup> Esta série de palestras começou, em 1948, com o filósofo e matemático inglês Bertrand Russell [1872 – 1970].

“The prevailing view in philosophy, psychology, and artificial intelligence is one which emphasises the analogies between the functioning of the human brain and the functioning of digital computers. According to the most extreme version of this view, the brain is just a digital computer and the mind is just a computer program. One could summarise this view – I call it 'strong artificial intelligence', or 'strong AI' – by saying that the mind is to the brain, as the program is to the computer hardware”.

(Searle, 1984 p. 28)

Esta asserção será denominada, neste trabalho, «tese forte da inteligência artificial». Searle defende que a manipulação mecânica de símbolos, respeitando um conjunto de regras sintáticas, não é no entanto suficiente para simular inteligência. Tal como sucedeu com Eliza, o programa de Weizenbaum, essa manipulação pode criar a ilusão da presença de uma inteligência, mas não a gera. Dirá Searle que aos estados internos da máquina, regulados pelo software, faltará o conteúdo semântico, no sentido de que o sistema será incapaz de reconhecer o significado das suas próprias acções.

Procurando ilustrar essa ausência o autor propõe a célebre «experiência da sala chinesa»:

“Imagine that a bunch of computer programmers have written a program that will enable a computer to simulate the understanding of Chinese. So, for example, if the computer is given a question in Chinese, it will match the question against its memory, or data base, and produce appropriate answers to the questions in Chinese. Suppose for the sake of argument that the computer's answers are as good as those of a native Chinese speaker. Now then, does the computer, on the basis of this, understand Chinese, does it literally understand Chinese, in the way that Chinese speakers understand Chinese? Well, imagine that you are locked in a room, and in this room are several baskets full of Chinese symbols. Imagine that you (like me) do not understand a word of Chinese, but that you are given a rule book in English for manipulating these Chinese symbols. The rules specify the manipulations of the symbols purely formally, in terms of their syntax, not their semantics. (...)

Now the point of the story is simply this: by virtue of implementing a formal computer program from the point of view of an outside observer, you behave exactly as if you understood Chinese, but all the same you don't understand a word of Chinese”.

(Searle, 1984 pp. 32 - 33)

Sobre a possibilidade da simulação artificial da inteligência humana, conclui Searle:

“The question we wanted to ask is this: 'Can a digital computer, as defined, think?' That is to say: 'Is instantiating or implementing the right computer program with the right inputs and outputs, sufficient for, or constitutive of, thinking?' And to this question, unlike its predecessors, the answer is clearly 'no'. And it is 'no' for the reason that we have spelled out, namely, the computer program is defined purely syntactically. But thinking is more than just a matter of manipulating meaningless symbols, it involves meaningful semantic contents. These semantic contents are what we mean by 'meaning'.”

(Searle, 1984 p. 36)

A argumentação de Searle exhibe, do nosso ponto de vista, uma fragilidade fundamental. Não é absolutamente claro que o reconhecimento por parte de um agente inteligente sobre o conteúdo semântico dos seus pensamentos, não corresponda, afinal, a um estágio mais complexo de processamento computacional. Por um lado, Searle não produziu uma prova negativa que impeça a emergência de estados mentais a partir de manipulações simbólicas. Por outro, uma vez que afirma que cérebros produzem mentes («Brains cause minds») (Searle, 1984, p. 39), de acordo com um fisicalismo naïve, um cérebro é um sistema físico sujeito às leis da natureza. E o conjunto de predicados e regras numa teoria informal empírica podem corresponder a um sistema formal, sem qualquer interpretações ou significados associados. Aqui, em especial, a argumentação de Searle não é convincente. Porquanto, quando o cérebro produz fenómenos mentais, pode-se considerar que o mesmo procede como se instanciasse fisicamente um procedimento lógico associado a um sistema formal hipotético, cujas regras podem intuitivamente ser conotadas com as leis da natureza. De forma metafórica, a emergência de uma mente num cérebro pode ser entendida como uma computação executada pelas leis da natureza. Estas correspondem tão somente a um programa de computador mais complexo que, para todos os efeitos, actua sobre objectos no mundo, como se de símbolos se tratassem. Pode-se dizer que qualquer crítica à IA, que se confine a comparar o cérebro com as propriedades de um computador, depara-se sempre com a mesma objecção. A de que a natureza, utilizando um conjunto de regras que, aparentemente dão origem a padrões e regularidades, terá de alguma maneira produzido criaturas inteligentes. Esta constatação chama a si a ideia de que usando de alguma combinação apropriada de certos materiais será, enfim, possível construir um sistema físico inteligente, tal como um cérebro. A afirmação de que este objectivo pode ser alcançado, recorrendo a uma teoria adequada sobre a natureza será denominada neste trabalho «tese fisicalista da inteligência artificial». Nesta dissertação,

tendo-se adoptado a abordagem descrita na introdução, acreditamos ter-se procedido a uma análise suficientemente abrangente de forma a colocar em questão as teses forte e fisicalista da inteligência artificial.

### 1.5. Hilary Putnam e o modelo funcionalista.

Hilary Whitehall Putnam [1926 - ], filósofo norte-americano, é professor emérito do departamento de filosofia da universidade de Harvard nos Estados Unidos. Estudou matemática e filosofia, doutorando-se nesta disciplina em 1951 com a tese «The Meaning of the Concept of Probability in Application to Finite Sequences», orientada pelo positivista alemão Hans Reichenbach (1891 – 1953). Desenvolveu o seu pensamento no âmbito da filosofia da matemática, filosofia da mente, filosofia da linguagem, meta-filosofia e epistemologia. No final dos anos sessenta, propôs uma teoria para o funcionamento da mente a que chamou «funcionalismo». Mais tarde, refutou-a na obra publicada em 1988, «Representation and Reality».

“The desire that (...) once gripped me was the desire to make belief-desire psychology ‘scientific’ by simply identifying it outright with computational psychology. When I proposed this program (under the name ‘functionalism’), I thought that the way to effect it was simple : we simply postulate that desires and beliefs are ‘functional states’ of the brain (i.e., features defined in terms of computational parameters plus relations to biologically characterized inputs and outputs). For example, one might postulate that *believing there is milk in the supermarket* is displaying one of the formulas in the *lingua mentis* whose translation is ‘there is milk in the supermarket’ in a special ‘belief register.’ Displaying another formula in a ‘desire register’ could be *desiring milk for tomorrow’s breakfast*. And going from these two computational states to the resultant, i.e., the action of going to the supermarket and buying milk, might be the result of a certain algorithmic procedure on these displayed formulas (as well as on others). In such a picture, ordinary-language mentalistic psychology, folk psychology, is a rough approximation to an ideal computational model of what goes on in the brain. An *ideal* belief-desire psychology would be isomorphic to (a part of) the computational description of what goes on in the brain. Make that assumption and you have *mentalism* in its most recent form.

Mentalism is just the latest form taken by a more general tendency in the history of thought, the tendency to think of concepts as scientifically describable

(‘psychologically real’) entities in the mind or brain . And it is this entire tendency that, I shall argue, is misguided”.

(Putnam, 1988b, p. 7)

A crítica posterior de Putnam à sua própria teoria incide sobre a inadequação do funcionalismo uma vez mais para explicar a produção de significado. Falando sobre o trabalho do seu antigo aluno de Doutorado, o norte-americano Jerry Fodor [1935 - ], Putnam especifica que tipo de teoria funcionalista pretende criticar:

“The way of thinking I am going to discuss first of all is expounded by Jerry Fodor in *The Language of Thought*. Fodor acknowledges that he owes a great deal of his inspiration to the work of Noam Chomsky. However, Chomsky has never committed himself to the possibility of finding ‘psychologically real’ entities which have enough of the properties we preanalytically assign to ‘meanings’ to warrant an identification . The ‘representations’ and ‘innate ideas’ of which Chomsky writes are deep syntactic structures and syntactic universals. Fodor’s program is thus not identical with Chomsky’s but rather a daring extension of it. (...)

At any rate, given that the key ideas of Chomsky’s theorizing are (1) the idea of Linguistic Universals, (2) the Innateness Hypothesis, and (3) the recent idea of modularity, the form that one can expect a Chomskian theory of the semantic level to take is relatively clear (and Fodor’s theory does take the expected form), even if the details can take various shapes. A Chomskian theory of the semantic level will say that there are ‘semantic representations’ in the mind/brain; that these are innate and universal; and that all our concepts are decomposable into such semantic representations. This is the theory I hope to destroy”.

(Putnam, 1988b, pp. 4 - 5)

Uma teoria funcionalista do significado reconheceria, no espaço externo ao sistema onde se instancia o pensamento, um referente absoluto a partir do qual as sentenças mentalmente formuladas podem ser aferidas. Identificando o significado dessas sentenças com a enunciação das condições que permitem a validação das mesmas.

“The doctrine called ‘meaning holism’ arose as a reaction to logical positivism; it offered arguments refuting positivist attempts to show that every term we can understand can be defined in terms of a limited group of terms (the ‘observation terms’) (...)

Holism is thus, in the first instance, opposed to positivism. The positivist view of language insists that all meaningful descriptive words in our language must have

definitions in terms of words in a 'basic' vocabulary, a vocabulary which consists of words which stand for notions which are *epistemologically more primitive* than, say, the theoretical terms of science. The favorite candidate of positivists was a vocabulary which consists of sensation terms, or, at any rate, terms for what is supposed to be 'observable'. If we formulate positivism as a thesis about the truth conditions for sentences rather than as a thesis about the definability of terms, we may say that, as a historical fact, positivists originally insisted that the meaning of a sentence should be given by (or be capable of being given by) a rule which determines in exactly which experiential situations the sentence is assertable".

(Putnam, 1988b, p. 8)

A utilização, permanente e inesperada, por parte dos agentes humanos, de ligações a informação não explícita, para além da realidade imediatamente observada, torna difícil implementar numa linguagem de programação todas as relações significativas relevantes, recorrendo estritamente a procedimentos lógicos. Nas palavras do autor:

"[Daily life language] sentences meet the test of experience "as a corporate body' and not one by one. (Hence the term 'holism').

(...) If someone tells you, for example, that *the thief entered through that window, and there is muddy ground outside the window*, you will 'deduce' that *there are footprints in the mud*. But this is not a logical consequence of the facts stated, for you obviously have made use of an unstated auxiliary hypothesis to the effect that if the thief entered through that window, he walked on the ground to get to the window, and other items of general information as well. If your informant says 'No, he was wearing stilts' then instead of expecting to find shoe prints in the mud you will now expect to find holes of a different shape. What has experiential import is the corporate body of statements, and this import is not the simple sum of the experiential imports of the individual statements".

(Putnam, 1988b, pp. 8 - 9)

Seres humanos são os agentes de uma dinâmica hermenêutica no mundo que torna difícil ou eventualmente impossível algoritmizar esse processo. Uma sugestão que já tínhamos encontrado nas reflexões de Hubert Dreyfus. Trata-se ainda do problema da representação do conhecimento de senso comum («the commonsense knowledge problem») e da falta de critérios algorítmicos para produzir as ligações semânticas relevantes entre a informação. Putnam identifica ainda outro aspecto da produção de significado que parece escapar às capacidades computacionais de um sistema lógico. Ligado

agora à alteração do entendimento de um conceito, resultante da variação dos sistemas de crença («believes») dos agentes.

“(...) a theory of synonymy would be a theory that decided questions of interpretation . Consider, however, just how subtle questions of interpretation can be, even when we deal with texts that aren't particularly 'literary' (...) [An] example is our knowledge of the fact that when Bohr used the word 'electron' (*Elektron*) in 1934, he was talking about the same particles he called 'electrons' in 1900. We do not determine this by comparing the *theories and descriptions* of the electron that Bohr gave at these two different times and seeing that they were much the *same*, because they weren't. The 1900 theory said that electrons go around the nucleus just as planets go around the sun, i .e., electrons have trajectories, whereas the 1934 theory (which is, in essence, the present quantum theory) says that an electron never has a trajectory - in fact, it never has a position and a momentum at the same time. Yet a physicist might well describe the development of the later theory from the earlier in this way: in the nineteenth century we discovered that electrons have a certain mass-charge ratio by deflecting electron beams in a magnetic field ; later we discovered by another experiment what the electron charge is (and hence what the value of the electron mass must be); we discovered that electric current is a stream of electrons; we discovered that every hydrogen atom consists of one electron and one proton ; we thought for a time that electrons had trajectories, but then we discovered the Principle of Complementarity; and so on. In short, he would tell the story as a story of successive changes of belief about the same objects, not as a story of successive 'changes of meaning' And [a] kind of 'general intelligence' is involved in his decision to treat all of these occurrences of 'electron' as synonymous as is involved in his decision to treat later research programs in the story as extensions of the earlier ones; a kind of decision that plays a central role in theory evaluation”.

(Putnam, 1988b, pp. 12 - 13)

O pensamento parece ser, por conseguinte, uma actividade significativa holística, dificilmente implementável num sistema lógico, quer por questões de dificuldade na identificação da relevância das relações entre factos, quer pela variação inesperada dessas relações. Parece existir uma transiência significativa, dir-se-ia de origem metalógica.

O que se afirmou parece então sugerir que a tese forte da IA pode não ser verdadeira e que só a tese fisicalista pode legitimar epistemologicamente a simulação artificial do pensamento humano.

## 1.6. A tese fisicalista de Rodney Brooks.

Rodney Brooks [1950 - ], de naturalidade australiana, é professor de Robótica no MIT (Massachusetts Institute of Technology), onde já deteve o prestigiado cargo de director do Laboratório de Inteligência Artificial. Fundou e dirige as companhias americanas Heartland Robotics Inc., empresa especializada na construção de robôs para fins industriais, e iRobot Corp.<sup>2</sup>, produzindo robôs para o mercado de consumo. Na página inicial da Heartland Robotics pode ler-se:

“They will have intelligence and awareness. They will be teachable, safe and affordable. They will make us productive in ways we never imagined”.<sup>3</sup>

Durante a década de oitenta, Rodney Brooks desenvolveu um novo tipo de arquitectura para os seus robôs, denominada «subsumption architecture» (Brooks, 1986). O conceito fundamental é o de que o funcionamento de cada uma das partes («modules») do robô dispõe de autonomia. Espera-se assim que da interacção entre módulos se produza a «emergência» de um comportamento complexo, que não se obteria a partir da operação isolada de cada uma das partes. Cada módulo é constituído por camadas («layers»), organizadas por níveis de abstracção, em que cada layer tem como função fazer com que o robô atinja determinado objectivo. Na estrutura hierárquica lógica do sistema a camada que ordena ao robô que explore o mundo está acima daquela que o faz procurar uma fonte de energia. Esta, por sua vez, está acima da camada que faz com que o robô evite obstáculos. Dessa forma, se surgir um obstáculo, a operação da camada logo acima será interrompida, sendo esta inibida de enviar sinais aos actuadores do robô. O objectivo mais geral será retomado, assim que o robô transponha o obstáculo.

A disposição das camadas é feita segundo um design denominado arquitectura «bottom-up». Cada camada tem acesso à informação dos sensores e gera informação para os actuadores do robô. As camadas mais baixas operam como mecanismos de adaptação rápida, enquanto que as camadas mais elevadas actuam para que os objectivos mais gerais possam ser cumpridos. A abordagem de Brooks tem a virtude de produzir mecanismos com capacidades adaptativas, em relação ao meio ambiente, que podem ser observados actuar, pragmatizando o que se poderia chamar uma perspectiva evolucionista da tese fisicalista da inteligência artificial.

---

<sup>2</sup> «iRobot Corporation: Home Page», 2011. « [www.irobot.com](http://www.irobot.com) »

<sup>3</sup> «Heartland Robotics - Designing and Building Next-Generation Industrial Robots», 2011. « [www.heartlandrobotics.com](http://www.heartlandrobotics.com) »

Esta abordagem foi apresentada num artigo, escrito dez anos depois do início do desenvolvimento da «subsumption architecture», com o título sugestivo: «Elephants Don't Play Chess» (Brooks, 1990). Brooks criticou a tese forte da inteligência artificial, enunciada na forma da hipótese dos sistemas físicos simbólicos («symbol system hypothesis»), introduzida originalmente, em 1955, por Allen Newell (1927 - 1992) e Herbert Simon (1916 - 2001), na origem da tese forte da IA.

“There is an alternative route to Artificial Intelligence that diverges from the directions pursued under that banner for the last thirty some years. The traditional approach has emphasized the abstract manipulation of symbols, whose grounding, in physical reality has rarely been achieved. (...)

In this paper we argue that the *symbol system hypothesis* upon which *classical AI* is base is fundamentally flawed, and as such imposes severe limitations on the fitness of its progeny. Further, we argue that the dogma of the symbol system hypothesis implicitly includes a number of largely unfounded great leaps of faith when called upon to provide a plausible path to the digital equivalent of human level intelligence. It is the chasms to be crossed by these leaps which now impede classical AI research.

But there is an alternative view, or dogma, variously called *nouvelle AI*, *fundamentalist AI*, or in a weaker form *situated activity*. It is based on the *physical grounding hypothesis*. It provides a different methodology for building intelligent systems than that pursued for the last thirty years.”

(Brooks, 1990 pp. 3)

E continua Brooks, referindo-se à «physical grounding hypothesis», enunciada em consistência com a sua arquitectura:

“Nouvelle AI is based on the physical grounding hypothesis. This hypothesis states that to build a system that is intelligent it is necessary to have its representations grounded in the physical world. Our experience with this approach is that once this commitment is made, the need for traditional symbolic representations soon fades entirely. The key observation is that the world is its own best model. It is always exactly up to date. It always contains every detail there is to be known. The trick is to sense it appropriately and often enough.

To build a system based on the physical grounding hypothesis it is necessary to connect it to the world via a set of sensors and actuators. Typed input and output are no longer of interest. They are not physically grounded.

Accepting the physical grounding hypothesis as a basis for research entails building systems in a bottom up manner”.

(Brooks, 1990 p. 5)

Brooks parece pois admitir que, por via de uma ligação sensorial entre um mecanismo e o seu meio ambiente, a inteligência emerge de forma natural, tal como sucedeu ao longo da evolução, com os sistemas biológicos na natureza.

“Like the advocates of the symbol system hypothesis, we believe that in principle we have uncovered the fundamental foundation of intelligence”.

(Brooks, 1990 p. 12)

A hipótese de Brooks é uma instanciação da hipótese fisicalista. A ideia de que alguma combinação adequada de material físico produzirá inevitavelmente um sistema inteligente. Se, no entanto, aceitarmos que os comportamentos na natureza são regidos por alguma estrutura lógica hipostasiável, poderá suceder que a tese fisicalista, em qualquer versão, se preste também a críticas formais, a exemplo do que sucede com a tese forte da inteligência artificial. Apesar da inteligência humana ter emergido na natureza, poderá suceder que a sua reprodutibilidade se encontre para além de qualquer estratégia formal. Noutros termos, poderá tratar-se de um fenómeno absolutamente não computável, cuja ocorrência não é previsível, nem mesmo no contexto de uma teoria informal empírica. Um fenómeno que, conseqüentemente, não pode ser tecnologicamente controlado. Será justamente essa a visão favorecida nesta tese.

### **1.7. O modelo Orch OR de Penrose-Hameroff.**

Roger Penrose [1931 - ], matemático e físico-matemático inglês, a ele se devem contribuições nos domínios da matemática e da relatividade generalizada. É professor emérito da cadeira de matemática de Rouse Ball, no Instituto de Matemática da Universidade de Oxford, em Inglaterra. Stuart Hameroff [1947 - ] é médico anestesista e investigador no Centro Médico da Universidade do Arizona, nos Estados Unidos.

Formularão o modelo da Redução Objectiva Concertada («Orch OR - Orchestrated Objective Reduction») para a produção de consciência no cérebro.

Penrose publica, em 1994, «Shadows of the Mind: A Search for the Missing Science of Consciousness», onde apresenta um argumento do tipo godeliano para a não computabilidade do pensamento humano. Afirma que a sua versão do primeiro teorema de

Gödel, a que chama o argumento de Gödel-Turing, mostra que a consciência humana é de natureza não-algorítmica e que aquela não pode ser computacionalmente simulada numa máquina de Turing. Sugere, além disso, que fenómenos quânticos enquadrados no que chama uma teoria quântica correcta da gravidade» («CQG – Correct Quantum Gravity») devem desempenhar um papel na emergência de consciência no cérebro. Nomeadamente, produzindo-se ali fenómenos físicos não-computáveis no decurso do colapso das funções de onda associadas a estruturas cerebrais.

O matemático inglês Alan Turing [1912 – 1954] propôs, em 1936, que um número real seria computável, se o seu desenvolvimento decimal fosse calculável numa máquina conceptual, hoje conhecida por «máquina de Turing» (Turing, 1936). O conceito actual de computabilidade refere-se a um problema que pode ser representado e resolvido de forma mecânica, numa sequência finita de passos, através do que se chama um procedimento efectivo. Dado um problema, se não existir um procedimento efectivo, mediante o qual esse problema possa ser implementado e resolvido numa máquina de Turing, esse problema dir-se-á «não computável» no sentido de Turing. Pode-se entender um algoritmo como um conjunto de instruções que leva um computador a executar um conjunto de acções fisicamente possíveis. Utilizando o mesmo conceito de computabilidade de Turing, Penrose construiu um argumento através do qual afirmou ter identificado um elemento de não computabilidade no pensamento humano (Penrose, 1994, pp. 74 - 75). Entendendo por algoritmo que actua sobre um número natural, um conjunto de procedimentos que efectua um cálculo a partir desse número, representou por  $C_m(n)$  o  $m$ -ésimo algoritmo actuando sobre o natural  $n$  e considerou a lista formada por todas as computações de todos os algoritmos sobre todos os naturais. Para cada um destes números  $n$  existe uma sequência de computações  $C_1(n); C_2(n); C_3(n) \dots C_m(n) \dots$  que podem computá-lo. De modo que podemos construir a lista de todas as computações possíveis que actuam sobre cada um dos naturais:

$$C_1(1); C_2(1); C_3(1) \dots C_m(1) \dots C_1(2); C_2(2); C_3(2) \dots C_m(2) \dots C_1(n); C_2(n); C_3(n) \dots C_m(n) \dots$$

A lista pode ser computacionalmente obtida e é denumerável. A notação « $C_m(n)$ » representa a  $m$ -ésima computação actuando sobre o  $n$ -ésimo número natural ou, utilizando uma representação informal:

$$C_m(n) \equiv C_{m\text{-ésima computação}} \text{ (sobre o } n\text{-ésimo natural)} \text{»}$$

Hipostasia-se agora a existência de um algoritmo  $A(m, n)$  capaz de actuar sobre a  $m$ -ésima computação, que por sua vez actua sobre o  $n$ -ésimo natural, isto é; capaz de

actuar sobre  $C_m(n)$ . O algoritmo  $A(m,n)$  tem como objectivo avaliar se a computação  $C_m(n)$  parará ou não, providenciando um resultado final. Turing demonstrou que esse algoritmo não existe (Turing, 1936). Não é possível construir um procedimento efectivo geral que, sendo implementável numa máquina de Turing, permita decidir se um algoritmo sob análise produzirá ou não um resultado. A célebre resposta ao problema da decisão («Entscheidungsproblem») colocado, em 1928, pelo matemático alemão David Hilbert [1862 – 1943] (Copeland, 2004) que consiste em saber se toda a proposição bem formulada num sistema formal é ali decidível. Se um algoritmo representar o processo lógico pelo qual determinada proposição é provada, esse algoritmo pertencerá à classe de todos os algoritmos sob inspecção no esquema de Turing. Sucede, no entanto, que se não existe um procedimento que permita distinguir os algoritmos que produzem provas daqueles que as não produzem, não haverá também maneira de distinguir entre proposições decidíveis e indecidíveis. O argumento de Penrose mostrará que existe uma situação em que é humanamente possível determinar o valor lógico de uma proposição, ainda que se demonstre a inexistência de um algoritmo capaz de produzir o mesmo resultado. A conclusão é então a de que agentes cognitivos humanos são capazes de determinar o valor lógico de uma dada proposição sem utilizarem um algoritmo implementável numa máquina de Turing e que, conseqüentemente, o pensamento humano não é computável, invalidando a tese forte da inteligência artificial. O argumento é o seguinte. Admita-se que existe um algoritmo *são* («sound»)  $A(m,n)$  com a seguinte propriedade, dada uma computação  $C_m(n)$  sob a inspecção de  $A(m,n)$ , sucede que:

- a) Se a computação  $A(m,n)$  pára, então a computação  $C_m(n)$  não pára.
- b) Se a computação  $A(m,n)$  não pára, então a computação  $C_m(n)$  pára.

A primeira condição especifica que se o decisor  $A(m,n)$  parar, produzindo uma prova, isso implicará que a computação  $C_m(n)$  nunca chegará a produzir uma prova. A segunda parte da definição especifica que se, pelo contrário, a computação  $A(m,n)$  não parar, não chegar a produzir uma prova, então a computação  $C_m(n)$  parará, produzindo uma prova. É importante referir que a paragem da máquina significa que um dado resultado pode ser computacionalmente obtido. Pelo contrário, a circunstância em que uma computação não pára determina que um resultado não pode ser obtido computacionalmente. Penrose agora assume que para todo e qualquer computação  $C_m(n)$

existirá um número inteiro positivo  $n$  que é igual ao número  $m$  associado a esse algoritmo;  $n = m$ , doutro modo a lista das computações considerada não estaria completa. A instanciação das duas condições anteriores a) e b), para  $n = m$ , é a de que dada uma computação  $C_n(n)$  sob inspecção de  $A(n, n)$ , sucede que;

a') Se a computação  $A(n, n)$  pára, então a computação  $C_n(n)$  não pára.

b') Se a computação  $A(n, n)$  não pára, então a computação  $C_n(n)$  pára.

A computação  $A(n, n)$ , que actua sobre o número natural  $n$ , encontra-se necessariamente na lista de todas as computações que actuam sobre todos os naturais. Se  $A(n, n)$  corresponder à computação  $C_k(k)$ , isto é, fazendo  $A(n, n) = C_k(k)$ , ter-se-á para a lista de todas as computações sobre todos os números naturais:  $C_1(1); C_2(1); \dots C_k(1) \dots C_m(1) \dots C_1(2); C_2(2) \dots C_k(2) \dots C_m(2) \dots C_1(k); C_2(k) \dots C_k(k) \dots C_m(k) \dots$

Resultando:

$C_1(1); C_2(1); \dots A(k, 1) \dots C_m(1) \dots C_1(2); C_2(2) \dots A(k, 2) \dots C_m(2) \dots C_1(k); C_2(k) \dots A(k, k) \dots C_m(k) \dots$

A nova instanciação das duas proposições a') e b') para  $A(m, n)$  é agora a de que dada a computação  $A(k, k)$ , actuando sobre o natural  $k$ , sob a inspecção de  $A(k, k)$ , sucede que;

a'') Se a computação  $A(k, k)$  pára, então a computação  $A(k, k)$  não pára.

b'') Se a computação  $A(k, k)$  não pára, então a computação  $A(k, k)$  pára.

A consideração conjunta destas duas afirmações leva a que se afirme, de forma metateórica, que a computação  $A(k, k)$  simultaneamente pára e não pára, obtendo-se uma contradição no sistema formal de uma máquina de Turing onde se tente implementar esse algoritmo. Isso significa o seguinte, qualquer processo computacional, do tipo  $A(m, n)$ , que se pretenda implementar numa máquina de Turing, será eventualmente bem sucedido num subconjunto de computações possíveis, falhando apenas na situação identificada de auto-referenciação. Porém, se assim é, para esse caso, de acordo com a) e b), devemos concluir que  $A(m, n)$  não irá parar, porque falha a produção de uma prova. Chegámos pois, de forma intuitiva, ao valor lógico de uma proposição, a afirmação de que para, pelo menos, um

número natural  $k$ , a computação  $A(k,k)$  não pára, afirmação cuja prova, recorrendo ao próprio algoritmo  $A(m,n)$  é manifestamente impossível. Uma prova obtida de forma metalógica, para além das capacidades de uma máquina de Turing, sugerindo que agentes cognitivos humanos não utilizam um algoritmo *são* (sound) que possa ser especificável para chegarem à verdade matemática. Sobre o mesmo tema, Penrose apresenta as visões de Gödel e Turing.

“It is instructive to consider the viewpoints of the two leading mathematical figures who are essentially responsible for the very argument that has led us to our conclusion  $\mathcal{G}$ . [Human mathematicians are not using a knowably sound algorithm in order to ascertain mathematical truth]. What, indeed, did Gödel think? What did Turing think? It is remarkable how, when presented with the same mathematical evidence, they came to basically opposite conclusions. It should be made clear, however, that both these outstanding thinkers expressed viewpoints in accordance with the conclusion  $\mathcal{G}$ . Gödel's viewpoint seems to have been that the mind is indeed not constrained by having to be a computational entity, and is not even constrained by the finiteness of the brain. (...) while accepting both of Turing's implicit contentions that 'the brain functions basically like a digital computer' and that 'the physical laws, in their observable consequences, have a finite limit of precision', Gödel rejected Turing's other contention that 'there is no mind separate from matter', referring to it as 'a prejudice of our time'. Thus, Gödel appears to have taken it as evident that the *physical* brain must itself behave computationally, but that the mind is something beyond the brain, so that the mind's action is not constrained to behave according to the computational laws that he believed must control the physical brain's behaviour. (...) On the other hand, Turing appears to have rejected such a mystical standpoint, believing (as Gödel did) that the physical brain, like any other physical object, must act in a computational way (...) Thus, he had to find some other route around the case provided, in effect, by  $\mathcal{G}$ . Turing made a great point of the fact that human mathematicians are very capable of making mistakes; he argued that for a computer to be able to be genuinely intelligent, it, also, would have to be allowed to make mistakes:

In other words then, if a machine is expected to be infallible, it cannot also be intelligent. There are several theorems which say almost exactly that. But these theorems say nothing about how much intelligence may be displayed if a machine makes no pretence at infallibility.

The 'theorems' he had in mind are undoubtedly Gödel's theorem and others related to it, such as his own 'computational' version of Gödel's theorem. Thus he appears to have regarded the inaccuracy of human mathematical thinking as essential, allowing the mind's (supposed) inaccurate algorithmic action to provide a greater power than that which would be achievable by means of any completely sound algorithmic procedure. Accordingly, he suggested a way of evading the conclusions of the Gödel argument: the mathematician's algorithm would be technically unsound, and it certainly would not be 'knowably sound'. (...) it seems probable that he would have agreed with viewpoint [that "All thinking is computation; in particular, feelings of conscious awareness are evoked merely by the carrying out of appropriate computations"].

(Penrose, 1994, pp. 128 - 129)

Divergindo da hipótese metalógica de Gödel e aproximando-se do computacionalismo fisicalista de Turing, Penrose argumentará que é na própria natureza e, em particular, na física do cérebro que devemos procurar a origem da não computabilidade do pensamento humano. Uma instanciação da tese fisicalista da inteligência artificial.

"Appropriate physical action of the brain evokes awareness, but this physical action cannot even be properly simulated computationally".

(Penrose, 1994, p. 12)

"Whatever it is that controls or describes the mind must indeed be an integral part of the same grand scheme which governs, also, all the *material* attributes of our universe".

(Penrose, 1994, p. 213)

Juntamente com Stewart Hameroff, Penrose publicará um artigo, em 1996, onde se descreverá o modelo da redução objectiva concertada («Orch OR - Orchestrated Objective Reduction»), pretendendo relacionar o comportamento físico do cérebro com a emergência de consciência, e a ocorrência de processos mentais não computáveis.

In this model, quantum-superposed states develop in microtubule subunit proteins ("tubulins"), remain coherent and recruit more superposed tubulins until a mass-time-energy threshold (related to quantum gravity) is reached. At that point, self-collapse, or objective reduction (**OR**) abruptly occurs. We equate the pre-reduction, coherent superposition ("quantum computing") phase with pre-conscious processes, and each instantaneous (and non-computable) **OR**, or self-collapse, with a discrete conscious event. Sequences of **OR** events give rise to a "stream" of

consciousness. Microtubule-associated-proteins can "tune" the quantum oscillations of the coherent superposed states; the **OR** is thus self-organized, or "orchestrated" ("**Orch OR**"). Each Orch OR event selects (non-computably) microtubule subunit states which regulate synaptic/neural functions using classical signaling.

The quantum gravity threshold for self-collapse is relevant to consciousness, according to our arguments, because macroscopic superposed quantum states each have their own space-time geometries (Penrose, 1994; 1996). These geometries are also superposed, and in some way "separated," but when sufficiently separated, the superposition of space-time geometries becomes significantly unstable, and reduce to a single universe state. Quantum gravity determines the limits of the instability; we contend that the actual choice of state made by Nature is non-computable. Thus each Orch OR event is a self-selection of space-time geometry, coupled to the brain through microtubules and other biomolecules. basis of mind!"

(Hameroff, Penrose, 1996)

Hameroff descreverá num artigo posterior, em 2012, de que forma o modelo Orch OR pode ajudar a explicar a existência de livre arbítrio como consequência de processos físicos quânticos no cérebro. Começa por evidenciar três argumentos criticando a existência de livre arbítrio no comportamento humano.

#### “THREE PROBLEMS WITH FREEWILL

We have the sense of conscious control of our voluntary behaviors, of free will, of our mental processes exerting causal actions in the physical world. But such control is difficult to scientifically explain for three reasons:

#### CONSCIOUSNESS AND CAUSAL AGENCY

What is meant, exactly, by “we” (or “I”) exerting conscious control? The scientific basis for consciousness, and “self,” are unknown, and so a mechanism by which conscious agency may act in the brain to exert causal effects in the world is also unknown.

#### DOES CONSCIOUSNESS COME TOO LATE?

Brain electrical activity correlating with conscious perception of a stimulus apparently can occur after we respond to that stimulus, seemingly consciously. Accordingly, science and philosophy generally conclude that we act non-

consciously, and have subsequent false memories of conscious action, and thus cast consciousness as epiphenomenal and illusory (e.g., Dennett, 1991; Wegner, 2002).

#### DETERMINISM

Even if consciousness and a mechanism by which it exerts real-time causal action came to be understood, those specific actions could be construed as entirely algorithmic and inevitably pre-ordained by our deterministic surroundings, genetics and previous experience.

We do know that causal behavioral action and other cognitive functions derive from brain neurons, and networks of brain neurons, which integrate inputs to thresholds for outputs as axonal firings, which then collectively control behavior. Such actions may be either (seemingly, at least) conscious/voluntary, or non-conscious (i.e., reflexive, involuntary, or “auto-pilot”). The distinction between conscious and non-conscious activity [the “neural correlate of consciousness (NCC)”] is unknown, but often viewed as higher order emergence in computational networks of integrate-and-fire neurons in cortex and other brain regions (Scott, 1995). Cortical-cortical, cortical-thalamic, brain-stem and limbic networks of neurons connected by chemical synapses are generally seen as neurocomputational frameworks for conscious activity, (e.g., Baars, 1988; Crick and Koch, 1990; Edelman and Tononi, 2000; Dehaene and Naccache, 2001), with pre-frontal and pre-motor cortex considered to host executive functions, planning and decision making”.

(Hameroff, 2012, p.1)

Contrapondo da seguinte maneira:

“But even if specific networks, neurons, membrane, and synaptic activities involved in consciousness were completely known, questions would remain. Aside from seemingly occurring too late for conscious control, neurocomputational activity fails to: (1) distinguish between conscious and non-conscious (“auto-pilot”) cognition, (2) account for long-range gamma synchrony electro-encephalography (“EEG”), the best measurable NCC (Singer and Gray, 1995), for which gap junction electrical synapses are required, (3) account for “binding” of disparate activities into unified percepts, (4) consider scale-invariant (“fractal-like,” “1/f”) brain dynamics and structure, and (5) explain the “hard problem” of subjective experience (e.g., Chalmers, 1998). A modified type of neuronal network can resolve some of these issues, but to fully address consciousness and free will, something else is needed. Here I propose the missing ingredient is finer scale, deeper order, molecular-level quantum effects in cytoskeletal microtubules inside brain neurons.

In particular, the Penrose–Hameroff “Orch OR” model suggests that quantum computations in microtubules inside brain neurons process information and regulate membrane and synaptic activities. Microtubules are lattice polymers of subunit proteins called “tubulin.” Orch OR proposes tubulin states in microtubules act as interactive information “bits,” and also as quantum superpositions of multiple possible tubulin states (e.g., quantum bits or qubits). During integration phases, tubulin qubits interact by entanglement, evolve and compute by the Schrödinger equation, and then reduce, or collapse to definite states, e.g., after 25 ms in gamma synchrony. The quantum state reduction is due to an objective threshold [“objective reduction (OR)”] proposed by Penrose, accompanied by a moment of conscious awareness. Synaptic inputs and other factors “orchestrate” the microtubule quantum computations, hence “orchestrated objective reduction (Orch OR).”

Orch OR directly addresses conscious causal agency. Each reduction/conscious moment selects particular microtubule states which regulate neuronal firings, and thus control conscious behavior. Regarding consciousness occurring “too late,” quantum state reductions seem to involve temporal non-locality, able to refer quantum information both forward and backward in what we perceive as time, enabling real-time conscious causal action. Quantum brain biology and Orch OR can thus rescue free will”.

(Hameroff, 2012, p.2)

Parece-nos que a explicação de Hameroff é insuficiente, levantando mais problemas do que aqueles que resolve, em especial, quando se diz que:

“The quantum state reduction is due to an objective threshold [“objective reduction (OR)”] proposed by Penrose, accompanied by a moment of conscious awareness”.

É a consciência um resultado causal do processo objectivo de redução OR das funções de onda nos microtúbulos? Precede-o ou surge com ele? É-lhe equivalente, de que maneira? E o processo de redução objectiva, corresponde a algum algoritmo computacionalmente superior às capacidades de uma máquina de Turing ou é algo absolutamente não computável? E se o é, o que quererá isto dizer? Pode a consciência no modelo de Penrose-Hameroff ser redutível às leis da natureza, parecendo no entanto sugerir que a consciência actua independentemente dos processos físicos, no momento do colapso das funções de onda? A insistência em reduzir tudo o que se passa na realidade a instanciações das leis da natureza pode, afinal, encontrar dificuldades até nas propostas sofisticadas de Penrose e de Hameroff. Poderá suceder que no processo computacional

proposto para o cérebro em que se afirma «Each reduction/conscious moment selects particular microtubule states which regulate neuronal firings, and thus control conscious behavior.» (Hameroff, 2012, p.2) uma nova propriedade se manifesta na natureza? Poderá suceder que na natureza existe volição, como uma propriedade geral, de que o livre arbítrio no pensamento humano é apenas um caso possível? Essa propriedade manifestar-se-ia de forma metalógica, independentemente de qualquer esquema determinista, expressando-se através dos mecanismos físicos do cérebro e tornando possível a experiência fenomenológica de controlo das nossas acções.

A existência de volição poderia também enfrentar as objecções à existência de livre arbítrio, recenseadas por Hameroff. A primeira «CONSCIOUSNESS AND CAUSAL AGENCY», referindo uma inadequação da definição do Eu subjectivo e o desconhecimento do mecanismo de agencição da consciência sobre o cérebro, poderia ser abordada da seguinte maneira. Se considerássemos que as leis da natureza são o correlato na realidade de regras num sistema formal, entenderíamos que a volição actuaria de forma independente, em relação a esse sistema. Isso significaria que o Eu subjectivo resultaria da expressão de volição através dos sistemas físicos apropriados e não o contrário. O Eu subjectivo que produz um exercício de livre arbítrio teria uma origem metalógica, sem necessidade de uma definição operativa. Deixaria também de fazer sentido descrever o mecanismo pelo qual a intervenção da volição se verificaria, porquanto a sua existência e expressão são ontologicamente anteriores às determinações de qualquer sistema físico, sendo a expressão da volição a causa de todos os padrões e regularidades observadas na natureza e não o contrário.

A existência de volição na natureza pode também remover a segunda objecção «DOES CONSCIOUSNESS COME TOO LATE?». No início dos processos eléctrico-fisiológicos no cérebro que dão origem a uma acção motora volicional, surgem picos de actividade eléctrica chamados potenciais evocados («readiness potentials»). Em 1985, Benjamin Libet [1916 – 2007] determinou experimentalmente que os potenciais evocados ocorrem num intervalo de tempo entre 350 e 400 milisegundos, prévio ao instante em que o agente afirma ter decidido levar a cabo uma acção motora (Libet, 1985). Libet propôs que a consciência seria uma instância de controlo sobre os ímpetos volicionais que emergiriam a nível subconsciente, promovendo ou inibindo a prossecução da acção motora desencandeável em sua consequência.

“The role of conscious will would be not to initiate a specific voluntary act but rather to select and control volitional outcome. It is proposed that conscious will can function in a permissive fashion, either to permit or to prevent the motor implementation of the intention to act that arises unconsciously. Alternatively, there

may be the need for a conscious activation or triggering, without which the final motor output would not follow the unconscious cerebral initiating and preparatory processes”.

(Libet, 1985, p. 529)

“If the brain can initiate a voluntary act before the appearance of conscious intention, that is, if the initiation of the specific performance of the act is by unconscious processes, is there any role for the conscious function? It is of course possible to believe that active conscious intervention to affect or control a cerebral outcome does not exist and that the subjective experience of conscious control is an illusion (e.g., Harnad 1982). (...)

I propose that conscious control can be exerted before the final motor outflow to select or control volitional outcome. The volitional process, initiated unconsciously, can either be consciously permitted to proceed to consummation in the motor act or be consciously "vetoed." In a veto, the later phase of cerebral motor processing would be blocked, so that actual activation of the motoneurons to the muscles would not occur”.

(Libet, 1985 p.537)

As experiências de Libet foram recentemente repetidas pelo neurocientista britânico Patrick Haggard (Haggard, 2008). O autor promove hipótese de Libet, afirmando que o processo volitivo pode ter como princípio computacional a existência de um ruído basal de potenciais evocados, em que apenas alguns deles chegam a produzir acções motoras.

“Voluntary action as random behavioural noise. (...) intrinsic neural noise ensures an element of randomness in actions, which could explain why one action is chosen over another. Several experimental and computational studies suggest that there is competition between codes for alternative possible actions, such as movements to targets or to distractors. Inhibitory links between action codes mean that small, random variations in the activation level of one code can overcome the activation of another code by a ‘winner takes all effect’, which produces one result rather than another. For example, Cisek found that random variations in the firing of populations of neurons in the premotor cortex before a ‘go’ signal could predict which of two possible responses an animal would make”.

(Haggard, 2008 p. 940)

A situação experimental de Libet e a sua hipótese parecem despromover a existência de volição. No entanto, outra interpretação é possível. Se a volição for uma

propriedade não computável, na acepção metalógica aqui defendida, a sua manifestação física não tem que corresponder estritamente aos potenciais evocados que surgem inicialmente no cérebro. Pode suceder que essa manifestação corresponda à totalidade do processo físico que decorre entre a emergência desses potenciais e o momento em que é efectivada a acção motora. Isso significaria que a sensação consciente de controlo por parte do sujeito faz parte, enquanto experiência fenomenológica, de um processo mais vasto que depende de uma propriedade mais fundamental, manifestando-se através dele. Essa propriedade não depende do estado do físico do cérebro no momento em que os potenciais evocados ocorrem. Sendo que o início e prossecução de todo o processo volitivo que corresponde à sua expressão não tem que ter uma causa em nenhum processo físico. De facto, é mais simples pensar que o conjunto dos processos físicos descritos é o efeito da manifestação de volição, que lhes é ontologicamente anterior.

Estas observações conduzem-nos à última das objecções à existência de livre arbítrio a que Hameroff chama «DETERMINISM» e onde se afirma que livre arbítrio são o resultado de processos complexos de base física. Esta será uma hipótese a que nos oporemos. Apresentaremos um argumento que sugere a existência de volição como uma propriedade na natureza, de que o livre arbítrio constitui uma expressão possível. Essa propriedade não é capturável por nenhum discurso lógico formal. Este facto, por sua vez, implica que se o livre arbítrio for aceite como característica da inteligência humana, então a sua simulação por meios deterministas não será viável.

## 2º Capítulo

### Breve história da Inteligência Artificial.

#### 2.1. Os paradigmas simbólico e conexionista da IA.

No que se segue, apresentamos uma história sumária da inteligência artificial, respeitando Hubert Dreyfus no artigo «Making a Mind Versus Modeling a Brain: Artificial Intelligence Back at a Branchpoint» (Dreyfus & Dreyfus, 1988).

O texto fundador de Dartmouth, subscrito em 1956 por McCarthy, Shannon, Minsky e Rochester, pressupunha que a inteligência humana poderia ser simulada por procedimentos automáticos implementados numa máquina. Esta hipótese vê num ser humano um sistema físico, cujo comportamento exhibe regularidades e cujo funcionamento está sujeito às mesmas regras fixas a que estão sujeitos todos os objectos na natureza.

Esta convicção, de que o mundo é um relógio sem relojoeiro, remonta ao ideário setecentista, mantendo que o comportamento de todas as coisas no mundo se caracteriza por relações de causalidade, de acordo com regras pré-estabelecidas. E dado que pedras e paus, quando caem, não têm sentimentos de qualquer espécie, torna-se expectável que acções inteligentes, só aparentemente dotadas de subjectividade, possam ser mimetizáveis por um sistema físico, configurado para o efeito. Hipótese tanto mais clara, quanto mais sistemática for a abordagem do problema prático no mundo, para o qual se perfila o raciocínio matemático, como paradigma do que é pensar correctamente. Tudo isto sugere assim a possibilidade de formulação de um discurso lógico, de natureza simbólica, marcadamente matemática, que faz do acto de pensar uma computação.

Ser-se humano é assim pensar de forma lógica, resolvendo problemas práticos no mundo, como idealiza parte do discurso da IA. Chegamos assim ao conceito de algoritmo ou programa, onde são enunciadas as regras gerais de produção de um certo discurso lógico. E também ao conceito de computador, máquina capaz de produzir as manipulações simbólicas numa linguagem onde deve estar representada, pelo menos, parte da realidade.

A metodologia parece, pois, clara: utiliza-se uma linguagem de computação suficientemente rica em informação sobre a realidade, implementada na parte de um sistema físico, que é um computador. Com base nos inputs do meio ambiente que chegam aos sensores da máquina, processam-se esses inputs no quadro da informação disponível, e em seguida actuam-se outras partes do sistema que constituem os seus apêndices robóticos. A implementação mecânica de um sistema de manipulação simbólica supõe-se assim ser equivalente à simulação artificial da inteligência humana.

Este conjunto de ideias está de alguma forma expresso na «hipótese dos sistemas físicos simbólicos» («the physical symbol system hypothesis»), proposta em 1955 pelos cientistas da computação Allen Newell [1927 – 1992] e Herbert Simon [1916 – 2001]. A hipótese propõe que um sistema físico simbólico exhibe as condições necessárias e suficientes para agir de forma inteligente. Significando, portanto, que qualquer sistema físico exibindo um comportamento inteligente, deverá corresponder a um sistema físico simbólico. E, reciprocamente, que qualquer sistema físico simbólico, dispondo de recursos adequados (velocidade de processamento, memória e mecanismos de interacção com o mundo), deverá corresponder a um sistema físico com comportamento inteligente. Um sistema físico simbólico substancia-se num computador, um mecanismo universalmente equivalente a qualquer sistema que exhiba comportamento inteligente. Uma outra forma de expressar a tese forte da inteligência artificial.

Outra linha programática de investigação desenvolvera-se no âmbito da neurociência; a partir dos esforços conjuntos do neurofisiologista americano Warren McCulloch [1899 – 1969] e do lógico americano Walter Pitts [1923 – 1969]. Introduzem, em 1943, o primeiro modelo matemático de uma rede neuronal, cuja unidade fundamental é desde então conhecida pelo «neurónio de McCulloch-Pitts». Pretendia-se simular, através de um sistema electrónico o comportamento do próprio cérebro. Uma abordagem fiscalista do problema. Em 1949, o psicólogo canadiano Donald Olding Hebb [1904 - 1985] sugere que o fortalecimento das ligações entre neurónios é proporcional à frequência da sua estimulação pelos neurónios adjacentes. Permite dessa forma a operacionalização da aprendizagem nas redes neuronais. Afirmará em 1962 o cientista da computação Frank Rosenblatt [1928 - 1969] que o objectivo da inteligência artificial consiste, justamente, em automatizar os procedimentos através dos quais as redes neuronais no cérebro aprendem a identificar padrões de estímulos.

Enquanto o programa racionalista procurava desenhar uma teoria geral sobre a manipulação de símbolos, capaz de representar a realidade exterior e de simular o pensamento humano sobre ela, o programa conexionista tentava construir um dispositivo físico que exhibisse comportamento inteligente, através de circuitos, mimetizando a neurobiologia do cérebro. Acrescente-se, de forma vagamente aproximada. Promissoras, as duas abordagens revelaram-se de início profícuas.

Em 1956, Newell e Simon implementam um programa informático [General Problem Solver (GPS)] capaz de resolver teoremas do cálculo proposicional. De forma talvez optimista, afirmariam que as propriedades da cognição, como a intuição, a criatividade e a aprendizagem, deixavam de constituir processos exclusivamente pertencentes à natureza humana, passando a ser simuláveis num computador. No extremo oposto da pesquisa,

Rosenblatt concebia um dispositivo a que chamou «perceptrão» («perceptron»), imitando o comportamento das redes neuronais (Rosenblatt 1957). Ainda em 1957, o sistema reconheceria padrões específicos de estímulos, de entre um conjunto maior de padrões, reproduzindo assim o acto inteligente de identificar. Também Rosenblatt acreditava na possibilidade da simulação artificial da inteligência humana a curto prazo, recorrendo às configurações físicas adequadas. Rosenblatt foi o promotor inicial da tese fisicalista da inteligência artificial.

Porém, à medida que os problemas colocados aos dois tipos de abordagens apresentavam crescente complexidade, a capacidade de computação requerida aos dois tipos de sistemas aumentava desproporcionalmente, ao ponto de tornar qualquer novo desafio uma tarefa intratável. Em 1969, Marvin Minsky e o matemático Seymour Papert [1928 - ] notavam, justamente, que o perceptrão de Rosenblatt parecia exibir uma deterioração computacional perante tarefas progressivamente mais difíceis (Minsky & Papert, 1969). Na década de setenta, a estratégia das redes neuronais sofria assim um retrocesso, face às capacidades heurísticas dos programas informáticos de Newell e Simon, comparativamente mais eficientes. A linha de pensamento racionalista da inteligência artificial ganhava força. Mas face ao problema da explosão computacional perante problemas práticos, quotidianamente simples, uma teoria geral do pensamento era cada vez mais necessária, mesmo que utilizasse estratégias heurísticas para a construção de relações entre objectos no mundo.

## **2.2. O problema do conhecimento de senso comum**

Platão [428 a.C. - 347 a.C.] afirmara sobre a matemática e sobre a ética que o sujeito cognoscente quando reflecte sobre a realidade, aplica regras explícitas, independentes do contexto físico ou social. Regras que têm origem num universo metafísico, contendo todo o conhecimento apriorístico sobre a realidade. Seria esta forma de pensar o mundo que se estenderia através de Descartes [1596 - 1650], Leibniz [1646 - 1716] e Kant [1724 - 1804], ao longo do que Dreyfus chama a «Tradição Filosófica» (Dreyfus & Dreyfus, 1988, p. 26), até à IA convencional. A abordagem simbólica para a inteligência artificial procuraria, precisamente, cumprir o programa platónico, tentando identificar no pensamento humano as representações mentais de elementos primitivos na realidade. Elementos que deverão ser independentes de todo o contexto observável ou como refere Dreyfus «context-free elements» e só posteriormente relacionados.

Uma solução formal para um sistema que identificaria elementos significativos no meio, seria proposta pelo linguista norte americano Marvin Minsky [1927 - ], semelhante ao esquema filosófico de Edmund Husserl [1859 - 1938] sobre o significado do acto de conhecer. No âmbito do programa fenomenológico, Husserl argumentou que um acto de consciência, de *noesis*, a consciencialização do objecto é feita através de uma estrutura conceptual de significados agregados ao objecto, denominada *noema*. Este, por sua vez, está associado ao acto de intencionalidade do sujeito cognoscente sobre o objecto consciencializado. O noema referencia o objecto através de um processo apriorístico de predicacão que utiliza as propriedades atómicas do objecto a que se refere. Os predicados em questão são então combinados, formando descrições complexas do objecto em causa. Para Husserl o noema contém, desse modo, uma hierarquia de regras que dirigem a actividade inteligente, volitivamente despoletada pelo ser cognoscente sobre o objecto consciencializado, providenciando o horizonte estrutural que permite organizar as informações dos sentidos. Esse horizonte fornece o contexto necessário para que o objecto possa ser conhecido, enumerando as possibilidades de relação entre esse e outros objectos.

Com base na proposta de Husserl, Minsky propôs em 1973 a implementação computacional de uma estrutura para a representação do conhecimento da experiência quotidiana ou «conhecimento de senso comum» («the commonsense knowledge») (Dreyfus & Dreyfus, 1988, p. 28). Minsky sugeriu representar informação empírica num esquema («frame») constituído por nós, ligando proposições. Os níveis superiores do esquema seriam fixos, correspondendo ao conjunto das afirmações que são sempre verdadeiras para um dado tipo de situações. Nos níveis inferiores existiriam extremidades a ser preenchidas com informação específica, relativa a cada situação em concreto. A contextualização necessária é, desse modo, obtida através do conhecimento específico para lidar com a situação em causa. Os esquemas de Minsky correspondem, conseqüentemente, a uma implementação de regras que, em princípio, permitem ao sistema identificar características expectáveis em situações mundanas, uma propriedade desejável do comportamento humano. De um modelo passivo, em que só existe processamento lógico, manipulando asserções sobre o mundo, transitava-se para outro, em que um agente, interagindo com a realidade, procede ao seu reconhecimento.

Porém, qualquer teoria sobre o pensamento, em que os elementos primitivos na realidade não fossem encontrados, revelar-se-ia infrutífera. E os esquemas de Minsky arriscavam-se a ser demasiado limitados, perante aquele facto. Dreyfus descreve como a descrição husserliana do acto cognoscente humano foi criticada por Heidegger [1889 – 1976]:

Heidegger, before Wittgenstein, carried out, in response to Husserl, a phenomenological description of the everyday world and everyday objects like chairs and hammers. Like Wittgenstein, he found that the everyday world could not be represented by a set of context-free elements. It was Heidegger who forced Husserl to face precisely this problem by pointing out that there are other ways of "encountering" things than relating to them as objects defined by a set of predicates. When we use a piece of equipment like a hammer, Heidegger said, we actualize a skill (which need not be represented in the mind) in the context of a socially organized nexus of equipment, purposes, and human roles (which need not be represented as a set of facts). This context, or world, and our everyday ways of skillful coping in it, which Heidegger called "circumspection," are not something we think but part of our socialization, which forms the way we *are*. Heidegger concluded:

The context ... can be taken formally in the sense of a system of relations. But ... [t]he phenomenal content of these 'relations' and 'relata' ... is such that they resist any sort of mathematical functionalization; nor are they merely something thought, first posited in an 'act of thinking'. They are rather relationships in which concerned circumspection as such already dwells.

(Dreyfus & Dreyfus, 1988, pp. 28 - 29)

Poderá então existir uma teoria geral capaz de descrever as experiências mundanas, vividas no domínio do senso comum? Existindo situações e objectos cujo conhecimento não parece ser directamente formalizável, de acordo com Heidegger, mesmo recorrendo a uma estrutura de contextualização apriorística, os esquemas de Minsky também não pareciam poder dar a resposta adequada. A sua efectividade no processo parecia ser demasiado mecânica e limitada.

### **2.3. As três fases históricas da Inteligência Artificial.**

A partir da linha programática iniciada por Newell e Simon, a história da IA conheceu três períodos importantes. Entre 1955 e 1965, a representação simbólica da realidade e a metodologia heurística dominaram os esforços de investigação, incidindo sobre o que se chamou «simulação cognitiva» (Dreyfus & Dreyfus, 1988, p. 31). Neste período, Newell e Simon mostraram como um computador poderia ser utilizado para lidar com uma determinada classe de problemas, utilizando um princípio heurístico de procura da informação denominado «análise de optimização dos meios para um fim» («means-end

analysis»). O método utilizava uma métrica abstracta, para avaliar a distância entre a descrição da situação inicial e a descrição de um dado objectivo final, associado à solução do problema. A adopção do mesmo princípio de actuação heurística daria origem em 1957 ao já referido «General Problem Solver» («GPS»). O GPS tinha como objectivo resolver qualquer problema simbolicamente representável: um teorema, um problema geométrico ou o desfecho de uma sequência de jogadas numa partida de xadrez. Operava, diferenciando entre o reconhecimento de determinada dificuldade e a geração de um conjunto optimizável de estratégias para a ultrapassar. Com essa metodologia resolveu-se o célebre «problema das Torres de Hanoi», um dos marcos na história da inteligência artificial. Pragmatizando a hipótese dos sistemas físicos simbólicos, o GPS evoluiria, em 1983, para uma versão melhorada denominada «State, Operator And Result» («SOAR»). Uma arquitectura simbólica desenvolvida pelo cientistas americanos da computação John Laird [1954 - ], Paul Rosenbloom e Allen Newell [1927 – 1992], que ainda é objecto de investigação na Universidade do Michigan nos Estados Unidos <sup>4</sup>.

A fase puramente simbólica da IA dará lugar a um novo período, entre 1965 e 1975, liderado por Minsky e Papert no MIT (Massachusetts Institute of Technology). Procurando-se, então, identificar que factos e regras deveriam ser representados simbolicamente numa estrutura de contextualização («frame») aplicada aos dados da experiência. Concretamente, pretendia-se a implementação de procedimentos destinados a lidar, de forma sistemática, com domínios isolados na realidade denominados «micro-mundos» («microworlds»), (Dreyfus & Dreyfus, 1988, p. 31). Nesta linha de investigação contam-se ainda os «guiões» («scripts») criados pelo cientista da computação americano Roger Schank [1946 - ], utilizando a metodologia «Case-Based Reasoning» («CBR»), em que a resolução de cada novo problema recorre, de forma heurística, soluções encontradas em situações anteriores. E também o programa «SHRDLU», escrito pelo cientista da computação Terry Winograd [1946 - ] no M.I.T., entre 1968 e 1970. O SHRDLU, cujo nome provém do romance de ficção científica «Etaoin Shrdlu» <sup>5</sup> era capaz de reconhecer instruções escritas em inglês e de actuar sobre um «universo elementar» («a toy universe»), composto por sólidos geométricos [blocos] amovíveis. Dominados por este tipo de sistemas, acreditava-se que micro-mundos e universos elementares, progressivamente mais realistas, poderiam fazer evoluir os programas, bem sucedidos a nível elementar, a uma inteligibilidade da realidade, comparável à humana.

---

<sup>4</sup> «Soar Home», 2011. « [sitemaker.umich.edu/soar/home](http://sitemaker.umich.edu/soar/home) »

<sup>5</sup> Winograd, Terry. «How SHRDLU got its name», 2011. « [hci.stanford.edu/~winograd/shrdlu/name.html](http://hci.stanford.edu/~winograd/shrdlu/name.html) »

Esta aspiração veio a revelar-se infrutífera. Tanto os esquemas de Marvin Minsky como os guiões de Roger Schank vacilaram perante situações reais mais complexas (Dreyfus & Dreyfus, 1988, p. 32). A quantidade avassaladora de informação que as tipificam (física, psicológica, social), a combinatoria possível das relações entre esses tipos de informação e a eventual impossibilidade de a representar simbolicamente, condicionaram o sucesso daquelas abordagens.

As dificuldades de representação adequada da realidade num sistema simbólico, motivaram justamente o início da terceira fase de desenvolvimento da Inteligência Artificial. A partir da década de oitenta adoptou-se um novo paradigma devido aos psicólogos americanos David Rumelhart [1942 - ] e James McClelland [1948 - ]. Divergindo da linha anterior de pensamento, a nova estratégia considerava possível actuar inteligentemente sobre a realidade, sem que isso implicasse a formulação de uma representação simbólica da realidade. Renascia em 1986 o modelo conexionista, voltando à hipótese de que o comportamento das redes neuronais no cérebro poderia ser mimetizado por redes neuronais artificiais. Introduzindo-se, além disso, o conceito de «processamento distribuído paralelo» («parallel distributed processing») (Dreyfus & Dreyfus, 1988, p. 34). A uma rede neuronal corresponde um cálculo matricial que a representa, com matrizes de inputs e de outputs, mediadas por matrizes de transformação, que representam as ligações entre neurónios. Os outputs podem estar ligados a actuadores físicos que exibem o comportamento apropriado, mediante o estímulo apresentado à matriz de inputs, Cada rede neuronal deveria ser “treinada” para reconhecer determinado padrão, isto é, afinados ao parâmetros das matrizes de transformação, de modo a que perante um estímulo, o sistema responda com a resposta certa. Uma rede neuronal é assim um programa de computador que utiliza uma estratégia diferente da usada no paradigma racionalista. A codificação empregue está longe de ser explícita, no sentido de que a cada grupo de neurónios corresponda, de modo inequívoco, um conjunto de padrões. O cientista britânico da computação Geoffrey Hinton [1947 - ] apresentou exemplos em que os mesmos neurónios, cuja activação dependia de padrões de estímulos especificamente associados a certo tipo de conceitos, eram também activados por padrões associados a outros conceitos (Dreyfus & Dreyfus, 1988, p. 35). O processo de diferenciação entre padrões parece depender de toda a rede, tornando difícil o seu reconhecimento.

A partir do modelo conexionista de Rumelhart e McClelland reacende-se a disputa intelectual entre os dois paradigmas principais propostos para a simulação artificial da inteligência. A discussão que opõe representação lógica, atomística do mundo, independente do contexto onde os objectos se situam, à simulação neuronal da actividade inteligente, de natureza experimentalista, privilegiando o comportamento físico do cérebro.

Nenhuma das estratégias terá concludentemente imperado sobre a outra. Com efeito, também as redes neuronais encontraram dificuldades, como Geoffrey Everest Hinton viu [], na forma pouco discriminada como actuam. Também as redes neuronais parecem, pois, vulneráveis ao problema da representação do conhecimento do senso comum, mas num sentido inverso. Parece agora que reside no programador a dificuldade em encontrar a estrutura neuronal específica para uma dada tarefa. Embora esta metodologia, de cariz experimentalista, esteja a salvo das críticas feitas ao idealismo racionalista, porquanto tenta pragmatizar a tese fisicalista da IA. O sucesso da arte parece pois prometido ao connexionismo, imitador do cérebro humano, produto da evolução biológica que terá produzido a solução ideal para a expressão da inteligência na natureza. O paradigma connexionista não se encontra comprometido com nenhuma posição filosófica idealista, apenas com o facto de que a natureza produziu sistemas físicos inteligentes.

### 2.3. Os contributos formalistas.

No final do século XIX, assistiu-se a uma crise dos fundamentos da matemática, com a identificação de antinomias na teoria de conjuntos. Esses paradoxos, de que o mais célebre, é o de Russell, motivariam o matemático alemão David Hilbert [1862 – 1943] em 1899 a considerar possível a formalização da generalidade das teorias matemáticas intuitivas, de que se falará na secção 2.2. deste trabalho.

Hilbert considerava que o fulcro do pensamento matemático consistia na manipulação de símbolos, de acordo com regras fixas, sem interpretações associadas sobre a natureza dos objectos que representavam nas teorias matemáticas. Associado à formalização da matemática, invocou Hilbert a necessidade um estudo, que denominou metamatemático, sobre as capacidades lógicas do sistema, visto como um todo. Esse estudo deveria permitir construir uma teoria da prova, determinando se o sistema formal e, por conseguinte, as teorias matemáticas que formalizava, seriam consistentes, completas e decidíveis. Respectivamente, livres de contradições, provando todas as sentenças formuláveis e capazes de dar uma solução a todos os problemas matemáticos. O matemático ~~alemão~~ austríaco Kurt Gödel [1906 – 1978], utilizando procedimentos metamatemáticos, determinaria em 1931 a incompletude de teorias matemáticas capazes de operações algébricas elementares e a incapacidade dessas teorias para provar a sua própria consistência. Alan Mathison Turing [1912 – 1954], matemático inglês, determinou a impossibilidade de determinar a decidibilidade de um sistema, quando abordando o célebre problema da decisão de Hilbert («*Entscheidungsproblem*»), provou em 1936 a existência de

números reais não computáveis e, com isso, que não existia nenhum procedimento matemático geral para determinar a decidibilidade de um problema matemático. O modelo matemático da máquina de Turing seria, mais tarde, desenvolvido por John von Neumann, matemático húngaro naturalizado americano que lhe acrescentou um mecanismo de sincronização, dando origem ao ENIAC (Electronic Numerical Integrator Analyzer and Computer). Começava a época histórica dos computadores.

Este conjunto de pensadores formalistas, entre outros, dariam origem às teorias dos autómatos finitos e da computação, que sobretudo se caracterizam pela determinação de processos computacionais, recorrendo a procedimentos finitistas. E onde reflexões epistemológicas sobre o que é o pensamento se confundem com a identificação de limites computacionais, determinando que manipulações simbólicas são ou não possíveis num sistema de processamento electrónico.

#### **2.4. Os contributos fisicalistas.**

Talvez o primeiro contributo fisicalista para a inteligência artificial tenha sido o do matemático americano Norbert Wiener [1894 - 1964], quando em 1948 publica «Cybernetics or the Control and Communication in the Animal and the Machine». Wiener pretendia que todo o sistema físico fosse um objecto capaz de processar informação. De a receber, transformar e emitir, onde intervêm mecanismos de controlo. Essa ideia estava em contraponto com o conceito de entropia, devido a Boltzmann, Maxwell e Gibbs, na termodinâmica estatística. Quantidade de informação, afirmava Wiener num sentido construtivo, é o valor simétrico, negativo, da quantidade de entropia. Criaturas inteligentes são, portanto, capazes de produzir ordem por entre a desordem prevalecente no Universo, de acordo com a segunda lei da termodinâmica. Esta metáfora, colocando nos processos físicos o conceito de informação, estava associada à definição matemática de transmissão de informação, através de suportes físicos, proposta pelo matemático americano Claude Shannon [1916 - 2001], também em 1948 (Shannon, 1948).

A metáfora seria estendida ao funcionamento do próprio universo, quando a informação deixa de ser uma propriedade secundária, reconhecida no funcionamento de alguns sistemas físicos, para ser o próprio fundamento dos processos físicos. Num processo computacional que segue um programa de computador, as leis da física, para processar um output, o universo. É célebre a frase do físico americano John Wheeler «*it from bit*».

It from bit. Otherwise put, every "it" - every particle, every field of force, even the space-time continuum itself - derives its function, its meaning, its very existence entirely - even if in some contexts indirectly - from the apparatus-elicited answers to yes-or-no questions, binary choices, bits. "It from bit" symbolizes the idea that every item of the physical world has at bottom - a very deep bottom, in most instances - an immaterial source and explanation; that which we call reality arises in the last analysis from the posing of yes-or-no questions and the registering of equipment-evoked responses; in short, that all things physical are information-theoretic in origin and that this is a participatory universe.

(Wheeler, 1990, p. 311)

Este seria um dos conceitos fundadores da filosofia digital, um movimento contemporâneo que procura identificar processos físicos com processos de computação, ocorrendo num espaço celular tridimensional. Ideia tributária das propostas do cientista da computação alemão Konrad Zuse [1910 – 1995], que em 1969 escreveu «Calculating Space» (Zuse, 1970), onde teoriza a ideia considerando que sistemas físicos são o equivalente a autómatos celulares. É também devida referência ao físico americano Richard Feynman [1919 – 1988], que em 1982 analisou de que forma podia um computador simular fenómenos físicos, a nível sub-atómico, de acordo com as leis da mecânica quântica (Feynman, 1982). Discussões no contexto da filosofia digital tem sido promovidas pelo matemático americano Gregory Chaitin [1947 - ], pelo cientista da computação americano Edward Fredkin [1934 - ] e pelo físico e matemático britânico Stephen Wolfram [1959 - ].

O mesmo tipo de reflexões formalistas parecem agora incidir sobre o comportamento físico da realidade. E até há, pelo menos, um autor que considera que uma teoria física definitiva corresponde a um programa suficientemente sofisticado. A este respeito cumpre referir o trabalho do cientista da computação alemão Juergen Schmidhuber [1963 - ], que se ocupa das «teorias algorítmicas de tudo» (Schmidhuber, 2000).

A afirmação emblemática de Fredkin – «What cannot be programmed, cannot be physics» (Fredkin, 1992b, p. 6) - resume todas as ideias subjacentes à física e filosofia digitais <sup>6</sup>. A realidade deverá ser composta, a nível fundamental, por elementos finitos e discretos, cujo comportamento pode ser transdúscível num programa de computador que processa bits. A afirmação de que todos os objectos e eventos são descontínuos, na acepção anterior é conhecida em Fredkin pela «hipótese da finitude da natureza» («the finite nature hypothesis»). A quantificação do espaço e do tempo e subsequente relação com o

---

<sup>6</sup> É um referencial importante a página, mantida por Fredkin, «Digital Philosophy», 2011. «[www.digitalphilosophy.org](http://www.digitalphilosophy.org) »

conceito de informação pode, por outro lado, ser encontrada na termodinâmica dos buracos negros, iniciada por Stephen Hawking, em 1970 e culminando nos trabalhos dos físicos holandeses Gerard 't Hooft [1946 - ] e do físico americano Leonard Susskind [1940- ], sobre o chamado Princípio Holográfico, e do físico americano Lee Smolin [1955 - ] sobre a natureza quantificada de espaço e tempo (Smolin, 2001b). Sobre outros tratamentos relacionando computação e processos físicos há ainda a referir o físico israelita David Deutsch [1953 - ], que promove o equivalente quântico da máquina universal de Turing («universal quantum computer») (Deutsch 1985) e físico americano Seth Lloyd [1960 - ], que trabalhou sobre a capacidade computacional a nível fundamental do universo (Lloyd, 1999).

A física parece assim poder ser idealizada como uma teoria da computação, abrindo-se caminho para a possibilidade de identificação na natureza dos processos que estão na base dos comportamentos inteligentes ou até mesmo da consciência. Tornam-se assim hipóteses sob a égide da tese fisicalista, testáveis experimentalmente. É claro que a inteligência humana deve continuar a ser vista como o resultado de uma computação, desta vez, usando entidades fundamentais na natureza.

É quase uma ironia que o primeiro exemplo dessa forma de tratamento se oponha à ideia de que pensamento é um fenómeno algorítmico. A proposta de Roger Penrose e de Stewart Hameroff que, como vimos no capítulo anterior, apresenta um modelo computacional quântico do cérebro, onde o livre arbítrio parece emergir a partir de uma realidade sub-quântica.

Actualmente, os programas de investigação da IA parecem optar pela escola roboticista, mais ortodoxa, iniciada por Rodney Brooks nos anos oitenta.

### 3º Capítulo

#### Física e computação.

No que se segue serão recenseadas algumas das relações entre física e computação. Apresentam-se as contribuições de vários autores, ilustrando a ideia de que a computabilidade é uma propriedade formal que pode ter expressão no mundo físico e que o universo pode ser concebido como um computador executando um programa. Metáfora que será utilizada mais à frente para hoostasiar a existência de um sistema formal abstracto, capaz de reger comportamentos na natureza.

O objectivo deste capítulo consiste em contextualizar a nossa estratégia de abordagem à possibilidade de simulação artificial do pensamento. Se a natureza fôr vista como um computador e as leis da natureza como um programa, cujos processamentos correspondem a comportamentos físicos, então não se torna difícil imaginar que o pensamento é apenas mais um comportamento que resulta desse processamento. Se assim fôr, será em princípio possível implementar os meios computacionais adequados numa máquina, de forma mimetizar o processamento que ocorre na natureza.

Embora, um sistema formal, que é uma teoria matemática, não seja um autómato finito, mais próximo de uma máquina de computar, sugerimos que esse sistema, enquanto entidade abstracta, possa estar ontologicamente mais próximo do que possam ser as leis da natureza. E, portanto, mais próximo do que possa eventualmente ser o programa de computador, cujos processamentos, na metáfora anterior, correspondem a comportamentos físicos. Sugerimos assim, de forma inteiramente especulativa, adoptar a seguinte metáfora:

*Um sistema formal está para as leis da natureza, assim como os seus processos de prova estão para os comportamentos dos sistemas físicos.*

Antes de avançar com a nova metáfora, devemos no entanto dar corpo à anterior. Mostrando de que forma em algumas áreas de investigação na física se identificam processos físicos com processos computacionais, envolvendo informação.

### **3.1. Termodinâmica da computação: irreversibilidade lógica e o Princípio de Landauer.**

A linguagem matemática parece adequar-se à descrição dos fenómenos naturais. Quer pela quantificação dos seus aspectos, quer pela forma como as proposições matemáticas se relacionam entre si, de modo sistemático, obedecendo a regras sintaticamente estáveis, cujos resultados parecem representar de forma fidedigna o que é observável na natureza. Este facto propicia uma capacidade dedutiva admirável, fundamental para a obtenção de previsões sobre comportamentos, permitindo alguma universalidade do pensamento sobre a natureza, independentemente da subjectividade dos sujeitos na produção do conhecimento.

Roger Penrose sugeriu uma situação em que a inteligência humana matemática parece ultrapassar os limites algorítmicos de uma máquina de Turing (Penrose, 1994). Uma certa visão mecanicista da cognição humana pretende que o pensamento humano é um fenómeno que se processa no cérebro que, de acordo com o modelo de Penrose-Hameroff, exhibe propriedades quânticas associadas a configurações espaço-temporais, permitindo a emergência da consciência. Nesse modelo, a rede neuronal no cérebro parece ser capaz de fazer colapsar estados de coerência quântica, existentes a nível sub-neuronal, de acordo com objectivos a nível mais alto, parecendo comportar-se como um sistema dotado de livre arbítrio. A discussão sobre a simulação artificial da inteligência humana desloca-se, assim, para o estudo dos sistemas físicos. O tratamento fisicalista da origem e natureza da inteligência sugere a investigação da relação entre computação, como caso particular da actividade cognitiva, e as condições físicas para a sua efectivação.

Os resultados de Turing (1936) sugerem que essa relação pode ser estabelecida, pelo menos a nível mecânico. Isto é; considerando o comportamento de uma máquina ideal de computar que actua num contínuo imaginário espaço-temporal, uma fita, de extensão infinita, de acordo com um conjunto de regras implementado nesse dispositivo. Esta abordagem, no entanto, exclui a consideração de outras propriedades físicas possivelmente fulcrais numa computação: a quantidade de energia mínima para que aquela se efective; a quantidade mínima de entropia termodinâmica produzida ou a possibilidade de uma computação reversa a partir de um resultado final.

Do mesmo modo que no final do século XVIII se desenvolveu a termodinâmica, incidindo sobre problemas pragmáticos relativos ao funcionamento das máquinas a vapor da revolução industrial, também a partir da década de sessenta se procedeu à aplicação da física ao comportamento dos computadores.

Um passo importante para que esse desenvolvimento começasse pode ser atribuído à teoria matemática da comunicação, publicada por Claude Shannon [1916 - 2001], em 1948.

Shannon considerou os aspectos quantitativos da transmissão de mensagens, através de canais de comunicação físicos, contendo ruído, definindo «quantidade de informação» transmissível e «entropia», com uma medida para a sua deterioração (Shannon, 1948).

O físico norte-americano Edwin Jaynes [1922 - 1998] notou, em 1957, que o princípio da maximização da entropia termodinâmica poderia ser visto como um caso particular da aplicação da teoria da informação (Jaynes, 1957a, 1957b). Quantidade de informação e entropia termodinâmica, associável à energia de um sistema e à quantidade de calor libertado, podiam deste modo ser relacionadas. O estudo de um sistema físico do ponto da vista da quantidade de informação transferida pode assim ser feito no âmbito da termodinâmica.

A segunda lei da termodinâmica estabelece que a entropia, interpretada como uma medida de desordem do sistema, se conservará ou aumentará num sistema fechado, cuja fronteira impede a troca de energia ou matéria com o exterior. Num sistema fechado não pode, portanto, existir nenhum mecanismo natural que no decurso do tempo diminua o estado de desordem do sistema. Poderão existir fases em que o sistema se torna temporariamente mais organizado, como parece suceder nos sistemas biológicos mediante processos homeostáticos. Porém, qualquer sistema passado o tempo suficiente acabará por assumir um estado de desordem máxima, atingindo o equilíbrio.

Ludwig Boltzmann [1844 - 1906] sugeriu que a temperatura de um sistema físico resulta do estado médio de movimento das suas partículas. E que os comportamentos dos elementos microscópicos no sistema são, em geral, relevantes para a observação dos estados macroscópicos do sistema como o volume, a pressão ou a temperatura.

Dado que cada partícula assume um movimento distinto cujo conhecimento é na prática impossível, tendo em vista o elevado número de partículas, existirão várias possibilidades de configuração microscópica do sistema, originando os mesmos estados macroscópicos. Facto que deverá ser tratado estatisticamente.

Boltzmann propôs que num sistema físico onde o comportamento de cada um dos componentes microscópicos é independente dos restantes, a quantidade de entropia termodinâmica do sistema fosse proporcional ao logaritmo natural do número de configurações microscópicas. O significado da segunda lei da termodinâmica é então o de que existe um número dessas configurações, produzindo estados macroscópicos desordenados, incomparavelmente superior ao número de configurações que originam

estados macroscópicos ordenados e que, conseqüentemente, todos os sistemas físicos tenderão a adoptar um estado associado a uma das configurações mais abundantes.

Poderia esta tendência ser contrariada?

Em 1867 Maxwell [1831 – 1879] concebeu uma experiência mental em que um demónio de proporções microscópicas actuaria sobre um reservatório fechado, com dois compartimentos interiores. A criatura mediria a velocidade de cada molécula de gás ali presente, dispondo-a num ou no outro compartimento, de acordo com o resultado da medição. Seria assim possível construir uma determinada configuração microscópica do sistema físico, em que o gás num dos compartimentos apresentaria temperatura superior à da quantidade de gás no outro, revertendo o processo termodinâmico em direcção a um estado macroscópico ordenado. Tal facto significaria que um sistema físico poderia evoluir para um tipo de configurações microscópicas menos abundantes, diminuindo a sua entropia.

Em face deste resultado Maxwell concluiu que a segunda lei da termodinâmica não era uma lei determinista, mas simplesmente a enunciação de um comportamento possível, eventualmente mais comum. A análise veio a ser conhecida como o «paradoxo do demónio de Maxwell» (Leff, Rex ,2002).

A intervenção do demónio de Maxwell num sistema físico permitiria assim que a partir da diferença de temperatura se extraísse energia desse sistema.

Analisando o mesmo problema, o físico húngaro Leó Szilárd [1898 - 1964] e, posteriormente, o físico franco-americano Léon Brillouin [1889 - 1969] (Leff, Rex ,2002) demonstraram que a acção de medição do demónio de Maxwell consome uma determinada quantidade de energia, não recuperável, superior à quantidade de energia recuperável a partir do reservatório, findo o processo. Demónio e gás interagem. O consumo de energia na medição das velocidades moleculares produzirá no demónio um aumento da sua entropia, superior à diminuição da entropia do reservatório.

O paradoxo do demónio de Maxwell é assim solucionado. A apreensão de informação sobre a velocidade de uma molécula e respectiva colocação num ou no outro compartimento do reservatório são operações que produzem entropia termodinâmica. Este facto sugere que existe uma relação entre aquisição da informação, respectivo processamento e entropia termodinâmica gerada.

O físico alemão Rolf Landauer [1927 - 1999], investigando os limites físicos da computação, sugeriu que determinados processos envolvendo a aquisição de informação não implicariam um aumento de entropia, desde que fossem termodinamicamente reversíveis. Esta situação corresponderia à produção de computações em que a informação tratada não seria apagada.

A proposta, conhecida pelo «Princípio de Landauer», afirma que por cada *bit* de informação apagada no universo é produzida uma quantidade mínima de entropia  $S = k \ln 2 \text{ J.K}^{-1}$ , onde  $k$  é a constante de Boltzmann;  $k = 1,38065 \times 10^{-23} \text{ J.K}^{-1}$  (joule por kelvin).

A esta entropia encontra-se associada uma dissipação mínima de energia não recuperável  $E = kT \ln 2 \text{ J}$  por cada *bit* de informação processada, em que  $T$  é a temperatura do sistema físico.

Em 1966, John von Neumann [1903 - 1957] sugeriu justamente que este seria o valor mínimo de energia consumido por cada *bit* de informação processado na natureza, motivo pelo qual esse valor é conhecido pelo «limite de von Neumann-Landauer» (Leff e Rex ,2002).

A computação reversível tem-se revelado uma área importante de investigação.

O cientista da computação norte-americano Charles Bennett [1943 - ] investigou a relação entre operações lógicas irreversíveis e as propriedades termodinâmicas dos sistemas físicos onde aquelas são implementadas. Numa operação lógica irreversível é impossível recuperar as condições iniciais de um problema a partir da solução final.

Para compreender o conceito, consideremos a multiplicação de dois inteiros. O produto de 2 por 100, por exemplo. O resultado, 200, também se pode obter pelo produto de 1 por 200. Apague-se agora na computação a informação de que os factores iniciais foram 2 e 100. A partir do resultado final, 200, sem mais informação, é impossível conhecer os factores iniciais que deram origem àquele resultado. Existe mais de uma possibilidade para os números que podem ser utilizados inicialmente. Trata-se de uma «computação irreversível». A obliteração de informação no universo fez com que o conjunto das condições iniciais possíveis aumentasse.

Num sistema físico onde se implemente a multiplicação dada, torna-se claro que o registo dos números iniciais deve corresponder a um determinado estado do sistema. Dessa maneira, o aumento do número de hipóteses possíveis originando o mesmo resultado para o cálculo dado, faz com que o conjunto dos estados do sistema que as representam aumente correspondendo a um aumento da sua entropia termodinâmica.

A existência de uma relação entre a obliteração da informação, processamento lógico irreversível e entropia termodinâmica é um facto notável. Essa relação sugere que uma computação, sendo uma abstracção do pensamento humano, pode ser fisicamente caracterizada no mundo.

Ainda no âmbito da computação reversível, assinala-se a «computação por choques elásticos» («billiard ball computation») desenvolvida pelos cientistas da computação Edward Fredkin [1934 - ] e Tommaso Toffoli [1943 - ] (Adamatzky 2002).

Um conjunto de objectos esféricos é preparado de modo a que se produzam colisões elásticas numa superfície plana sem atrito, de acordo com a mecânica de Newton, assumindo-se conservação da energia, sem ocorrência de dissipações irreversíveis e, por conseguinte, sem aumento da entropia termodinâmica no sistema. A computação reversível permite então estudar a optimização energética dos mecanismos que a processam e, desse modo, conceber estratégias para o aumento da sua eficiência. Em 1973, o cientista da computação Charles Bennett [1943 - ] demonstrou que é possível construir uma máquina de Turing computacionalmente reversível (Bennet, 1973).

Fredkin e Toffoli, por outro lado, introduziram, de forma independente, interruptores lógicos adequado à computação reversível («Fredkin gate»; «Toffoli gate»), (Adamatzky 2002). Os dois interruptores são universais permitindo a construção de qualquer circuito lógico efectuando uma computação reversível.

### **3.2. Computação na natureza e o Princípio da Equivalência Computacional.**

Um autómato celular corresponde a um estado numa quadrícula ou célula, pertencente a uma grelha quadrangular plana, visível num monitor ou numa folha de papel, em que as regras que regem o estado de cada célula envolvem os estados das células que lhe são adjacentes. O sistema evolui a partir de uma distribuição inicial dos estados na grelha, os quais são sucessivamente actualizados em função da distribuição imediatamente anterior.

Verifica-se então que condições locais, universalmente válidas para todos os autómatos, originam ao longo da evolução dos estados locais na grelha, padrões globais emergentes. Os autómatos celulares, foram definidos, em 1940, por John von Neumann [1903 - 1957] (von Neumann e Burks, 1966).

Uma forma de computação, recorrendo a autómatos celulares, foi desenvolvida pelo matemático inglês John Conway [1937 - ], que concebeu as regras locais do que ficou conhecido pelo «jogo da vida de Conway» («Conway's life game»<sup>7</sup>), (Gardner, 1970).

Contribuições significativas na mesma área são também devidas a Tommaso Toffoli e Norman Margolus (Toffoli e Margolus, 1987) e ainda ao cientista da computação britânico Stephen Wolfram [1959 - ], que desenvolveu extensamente o estudo dos autómatos celulares (Wolfram, 2002).

Wolfram sugeriu que qualquer computação, implementável numa máquina de Turing, pode ser executada num sistema de autómatos celulares. Os dois modelos são

---

<sup>7</sup> Veja-se o software de autómatos celulares em Wojtowicz, Mirek. «Mirek's Celebration - 1-D and 2-D Cellular Automata viewer, explorer», 2011. « <http://psoup.math.wisc.edu/mcell/> ».

equivalentes, resolvendo o mesmo tipo de problemas. Reflectindo sobre essa equivalência, Wolfram enunciou um princípio de maior amplitude, chamado «Princípio da Equivalência Computacional», que estabelece que todos os sistemas físicos na natureza só podem efectuar computações com uma complexidade inferior ou igual a um limite máximo. Assumindo que o poder computacional de todos os sistemas físicos é equivalente no sentido de que podem resolver o mesmo tipo de problemas. Dessa maneira, a implementação de uma computação num sistema físico, destinada a resolver um problema, reduz-se à representação apropriada desse problema no sistema, existindo sempre forma de a implementar. O princípio parece consistente com a equivalência computacional estabelecida pelo neurocientista e cientista da computação britânico Michael Arbib [1940 - ] entre uma máquina universal de Turing e uma rede neuronal (Arbib, 1964). Encontramos também reflexos do mesmo princípio na «computação molecular», também chamada computação por ADN (ácido desoxirribonucleico), tirando partido do processamento paralelo distribuído.

Recorrendo a fenómenos bioquímicos envolvendo a molécula de ADN, é possível resolver problemas, cujo tempo de processamento é drasticamente reduzido em consequência do elevado número de moléculas actuando simultaneamente para o efeito.

Este tipo de computação foi introduzido pelo cientista da computação norte-americano Leonard Adleman [1945 - ], que em 1994 demonstrou que a computação molecular poderia ser utilizada para resolver o «problema do grafo de sete vértices de Hamilton» (Adleman, 1994).

Outra forma de computação, inspirada na biologia evolutiva, diz respeito aos «algoritmos genéticos». Um algoritmo genético consiste numa técnica de procura iterada de um conjunto de resultados que aproximam o conjunto das soluções de um problema dado.

O processo de procura envolve a transmissão hereditária da informação nos resultados sucessivamente produzidos, assim como a mutação, selecção e recombinação («crossover») das partes constituintes desses resultados. Os algoritmos genéticos são implementados num computador fazendo com que a população dos resultados evolua em direcção ao conjunto das soluções do problema. Os resultados correspondem a indivíduos na população, representados sob a forma de sequências de zeros e uns («binary strings»).

A evolução começa por uma população formada por sequências geradas aleatoriamente. Em cada um dos ciclos do processo, a geração de resultados é testada em termos da sua capacidade de resolução do problema dado («fitness»). No decurso do teste anterior, alguns dos indivíduos são seleccionados e modificados, quer através da mutação aleatória das suas partes constituintes; quer através da recombinação dessas partes com os elementos constituintes de outros indivíduos seleccionados («crossover»). A população obtida é então utilizada no próximo ciclo do algoritmo. O processo termina quando um

número máximo pré-estabelecido de gerações foi produzido ou quando se tiver atingido uma população de resultados que aproximam satisfatoriamente o conjunto de soluções do problema em questão.

Os algoritmos genéticos são aplicados na biogenética, na matemática, na economia e na inteligência artificial, associados à aprendizagem simulada em redes neuronais.

Entre as contribuições mais significativas no domínio da chamada «computação evolutiva» (evolutionary computation) contam-se os trabalhos do norte-americano John Henry Holland [1929 - ] (Holland, 2000) e dos cientistas da computação alemães Ingo Rechenberg [1934 - ] e Hans-Paul Schwefel [1940 - ] (Michalewicz, 1996).

Na linha da tese fisicalista da inteligência artificial, a robótica assume-se como uma área de investigação importante. Com algumas ligações à cibernética, fundada por Norbert Wiener [1894 - 1964], (Wiener, 1961).

A construção de robôs pretende mimetizar o comportamento de sistemas vivos, através da implementação de dispositivos mecânicos e electrónicos (e.g. sensores). Procura-se determinar os meios que promovem a emergência de comportamentos inteligentes a partir da interacção de mecanismos simples com o meio ambiente.

Nesta linha de investigação, destacam-se os trabalhos do neurocientista austro-italiano Valentino Braitenberg [1926 - ] (Braitenberg, 1984), do especialista em robótica australiano Rodney Brooks [1954 - ] (Brooks, 2003) e do cientista austríaco Hans Moravec [1948 - ] (Moravec, 1990, 1998).

Moravec notou que processos cognitivos de alto nível (e.g. partidas de xadrez, resolução de problemas matemáticos) exigem menor quantidade de recursos computacionais, comparativamente àqueles necessários para realizar tarefas simples, como pegar num lápis ou contornar um obstáculo (Moravec, 1990). Esta observação torna particularmente útil a abordagem robótica, que se inspira em muitos dos mecanismos biológicos para dotar robôs das mesmas capacidades de interacção com o meio.

É também interessante o contributo do roboticista britânico Mark Tilden [1961 - ], pela estratégia que adoptou, denominada «BEAM robotics – Biology, Electronics, Aesthetics and Machines (Hrynkiw, Tilden, 2002 ). Utilizando circuitos analógicos simples, em vez de microprocessadores, Tilden produziu robôs de design simples, colocando ênfase na flexibilidade.

A mimetização por meios artificiais do comportamento biológico, utilizando quer programas informáticos, quer robôs, deu origem ao desenvolvimento da área denominada «vida artificial». Um exemplo importante dessa área de investigação, ligado à nanotecnologia, consiste no laboratório no M.I.T. «The Center for Bits and Atoms», liderado

pelo físico norte-americano Neil Gershenfeld (2011). Outros modelos computacionais importantes dizem respeito aos sistemas de «computação distribuída» - inspirados no comportamento colectivo de insectos, com níveis elevados de

A abordagem física da computação persegue, assim, dois objectivos fundamentais de estudo. Por um lado, define modelos físicos onde seja possível a representação de problemas matemáticos. Por outro lado, estabelece os limites computacionais desses modelos.

Esta abordagem pode entretanto ser generalizada, considerando que o universo é um dispositivo computacional de vastas proporções e que os fenómenos físicos são o resultado das computações efectivadas nesse meio.

A primeira proposta concreta nesse sentido foi avançada pelo cientista alemão Konrad Zuse [1910 - 1995].

Estabelecendo uma relação entre a representação matemática dos fenómenos físicos e a teoria dos autómatos celulares, Zuse propôs que o comportamento de alguns sistemas físicos poderia ser simulado, utilizando aproximações numéricas das equações diferenciais associadas ao seu movimento (Zuse, 1970).

Referenciamos em decurso o desenvolvimento actual destas ideias.

### **3.3. Computação quântica e o Princípio de Deutsch-Turing.**

Se os comportamentos na natureza forem considerados o resultado de uma computação, torna-se legítimo identificar se existe e qual é o sistema físico elementar que as pode perfazer, assim como determinar o modelo teórico que descreve o seu comportamento. Os sistemas físicos ao nível fundamental da matéria são descritos pela mecânica quântica. Consequentemente, a construção de um modelo computacional no âmbito dessa teoria parece constituir uma escolha natural para o sistema físico procurado.

Em 1982, Richard Feynman demonstrou que qualquer máquina de Turing que simulasse um fenómeno quântico, demoraria um período de tempo exponencial em relação à duração do fenómeno simulado (Feynman, 1982). A simulação apropriada desse tipo de eventos só se poderia obter recorrendo a um dispositivo com propriedades quânticas. Feynman determinou ainda que existe uma classe de sistemas quânticos que podem simular o comportamento de qualquer objecto físico finito.

Em 1985, o físico israelita David Deutsch [1953 - ] propôs o modelo do «computador quântico universal» (Deutsch, 1985). Um computador quântico universal consiste num modelo matemático, formulado no âmbito da física quântica, que descreve o

comportamento de um sistema físico, com propriedades quânticas. capaz de simular o comportamento de qualquer sistema físico de dimensões finitas.

Um sistema físico com propriedades quânticas é, por seu lado, uma entidade à qual se encontra associada uma onda, denominada «vector de onda» ou «onda  $\Psi$ » («psi»), cuja intensidade é representada, recorrendo ao plano complexo, em espaços de Hilbert (Young, 1988).

O quadrado do módulo da intensidade do vector de onda representa a densidade da probabilidade - numa dada posição e num dado instante - de que um evento atómico ocorra na vizinhança dessa posição, nesse instante. A onda  $\Psi$  está construída com base na ontologia de onda de Forier e é suposto estar presente em todos os locais do espaço. O vector de onda cumpre o «Princípio da Sobreposição de Ondas». A intensidade resultante nos pontos em que duas ondas se cruzam é, portanto, igual à soma das intensidades das ondas consideradas separadamente. As ondas sobrepõem-se inteferindo de maneira construtiva ou destrutiva. O vector de onda associado a um sistema físico codifica a probabilidade de ocorrência de cada um dos estados desse sistema.

A aplicação do princípio da sobreposição a um vector de onda sugere que o mesmo resulta da sobreposição de vários vectores de onda, tantos quantos os estados assumidos pelo sistema físico. Nos sistemas materiais descritos pela mecânica de Newton ou pelo electromagnetismo de Maxwell, a informação é codificada estabelecendo-se uma relação de um para um entre o conjunto dos símbolos representáveis no sistema e o conjunto das observações dos estados físicos utilizados para os representar.

A extracção de 1 *bit* de informação de um sistema físico clássico significa que no instante em que é feita uma observação sobre o sistema, apenas um dos estados pode ser observado. No modelo da computação quântica a situação é radicalmente diferente. O vector de onda contém todos os estados possíveis do sistema e, como tal, no instante em que se processa uma observação é possível obter qualquer um daqueles estados. A observação de um estado corresponde ao que é conhecido na ortodoxia da mecânica quântica como «colapso do vector de onda».

Uma medição sobre o sistema levará a que este assuma, com uma dada probabilidade de ocorrência, apenas um dos seus estados possíveis.

Todas as propriedades que caracterizam o comportamento do sistema como uma onda, cessarão no instante de tempo em que a observação se processa, fazendo com que o sistema exiba todas as propriedades corpusculares de uma entidade material localizada numa região finita de espaço. As duas descrições de um sistema atómico, ondulatória, antes da observação, e corpuscular, feita a medição, são consideradas complementares e

mutuamente exclusivas, de ponto de vista lógico, facto conhecido pela dualidade corpúsculo-onda, enunciada no princípio da complementaridade de Bohr.

É importante enfatizar que a observação directa de um vector de onda, como entidade física, nunca se verifica. O mesmo é considerado uma entidade abstracta que rege o comportamento físico das entidades corpusculares observáveis a que está associado.

A sua possível existência e respectivo comportamento são, portanto, inferidos a partir de uma quantidade, estatisticamente representativa de observações, sobre um ou mais sistemas corpusculares (e.g. fótons, electrões, neutrões).

Um computador quântico universal permite codificar toda a informação representada pela totalidade dos estados no vector de onda e executar todas as computações possíveis envolvendo essa informação. A codificação é feita associando cada par de estados mutuamente exclusivos aos símbolos 0 e 1, de uma codificação binária.

Embora o colapso de onda ocorra aleatoriamente, a evolução do vector da onda procede de modo determinista, de acordo a equação diferencial de Schrödinger» proposta em 1926 por Erwin Schrödinger [1887 - 1961], (Wick, 1998).

O determinismo da evolução ondulatória do sistema permite, mediante a preparação do mesmo, construir situações físicas onde se implementa a computação de problemas físicos. Dois estados do sistema, em correspondência biunívoca, com uma string de dois bits, evoluirão simultaneamente, dando origem a um ou outro desses estados, de acordo com as interacções impostas ao sistema.

Num computador quântico a informação codifica-se na forma de um *qubit* (pronunciado «kiubit»). Um *qubit* é uma entidade abstracta que contém, simultaneamente, os dois estados possíveis associados a um *bit* («binary digit») clássico.

O estado que representa o símbolo «0» e o estado que representa o símbolo «1».

A cada um desses estados, simultaneamente presentes num *qubit*, encontra-se associada a respectiva probabilidade de que, no decurso de uma observação, o sistema tenha como output «0» ou «1». No registo de um computador quântico podemos representar vários *qubit* que, no decurso da evolução do vector de onda, se alterarão em função das suas transformações no algoritmo.

Uma técnica importante na computação quântica consiste em idealizar uma situação física através da qual seja possível aumentar a probabilidade dos estados finais do sistema que correspondem a soluções do problema. O modo como os vectores de onda são alterados, em consonância com o que se pretende de um algoritmo quântico, é matematicamente descrito através de «transformações unitárias», preservando a evolução determinista do vector de onda.

O facto de que uma só acção física pode alterar simultaneamente todos os estados representados no vector de onda faz com que uma computação quântica exiba o que é conhecido por processamento paralelo massivo. A situação clássica correspondente obrigaria a alocar uma máquina de Turing a cada um dos estados presentes, atingindo-se rapidamente o limite dos recursos finitos disponíveis numa computação clássica.

É por este motivo que um computador quântico é mais eficiente do que uma máquina de Turing. Enquanto que o tempo de processamento de uma máquina de Turing depende exponencialmente do número de transformações unitárias utilizadas, o tempo de processamento de um computador quântico depende só polinomialmente do mesmo número de transformações (Nielsen, Chuang, 2000).

Uma propriedade quântica de grande importância para este tipo de computação consiste no que se chama «estado de correlação» ou de «coerência quântica» («quantic entanglement») entre dois ou mais sistemas quânticos. Quando dois sistemas quânticos interagem, os respectivos vectores de onda formarão um novo vector de onda, cujo comportamento global influenciará o estado local de cada um dos sistemas, para qualquer observação posterior à interacção, independentemente da distância a que os sistemas se encontram. A previsão deste fenómeno surgiu em resultado da experiência mental de EPR proposta por Albert Einstein [1879 - 1955], Boris Podolsky [1896 - 1966] e Nathan Rosen [1909 - 1995], em 1935, tendo sido observado por Alain Aspect [1947 - ], em 1982 (Wick, 1998).

Na computação quântica recorre-se à coerência quântica para a observação do estado de um *qubit* de controlo. Este encontra-se em coerência com um *qubit* utilizado na computação. Em vista da natureza aleatória da medição do estado de um sistema quântico, muitos dos algoritmos utilizados incluem passos onde se utilizam métodos de computação clássicos, verificando, nomeadamente, se algum dos estados observados no final do processo representa uma solução do problema.

Entre as contribuições mais relevantes para a computação quântica destacamos o algoritmo do lançamento da moeda de David Deutsch, o algoritmo de Shor, proposto pelo matemático americano Peter Shor [1959 - ], factorizando um número natural e o algoritmo de pesquisa numa base de dados, da autoria do cientista americano Lov Kumar Grover [1961 - ] (Siegfried, 2001).

O conceito de computador quântico universal promove a investigação da relação entre realidade física e computação a nível fundamental. David Deutsch refere a tese de Church-Turing como a afirmação de que toda a função intuitivamente considerada computável é computável por uma máquina de Turing. Tendo em vista a relação entre realidade física e computação, Deutsch generalizou o princípio, para a sua versão física,

nomeadamente; que o comportamento de todo o sistema físico finito é fidedignamente simulado por uma máquina quântica de computar universal, operando com meios finitos (Deutsch, 1985).

Esta afirmação, denominada «Princípio de Deutsch-Turing», estabelece uma relação de equivalência entre a implementação física de modelo computacionalmente universal e qualquer fenómeno físico, como resultado possível dessa computação.

Física e computação convergem assim para um só corpo de conhecimento em que a descrição dos fenómenos pode ser feita recorrendo a teorias de cariz computacional.

### **3.4. Identidade ontológica entre realidade e computação: teorias algorítmicas de tudo.**

Em face do que foi afirmado é razoável supor-se a existência de uma identidade ontológica entre realidade física e computação. Esta hipótese encontra-se substanciada na célebre expressão «*it from bit*» do físico americano John Wheeler [1911 – 2008] (Wheeler, 1990) e ainda no trabalho do físico americano Seth Lloyd [1960 - ], sobre a capacidade computacional do universo (Lloyd, 1999).

Seguindo a mesma linha de pensamento, mas assumindo uma descrição determinística da realidade, destacam-se as ideias do cientista da computação alemão Jürgen Schmidhuber [1963 - ] sobre as teorias algorítmicas de tudo.

Reflectindo sobre a assumpção de Konrad Zuse (1970), de que todo o fenómeno físico pode ser simulado recorrendo a autómatos celulares, Schmidhuber propôs, em 1997, que a história do universo é computável no sentido de Turing, utilizando algoritmos que não incluam instruções de paragem (Schmidhuber, 2000).

Para o efeito, Schmidhuber introduziu dois modelos computacionais. O primeiro, denominado «máquina generalizada de Turing» («generalized Turing machine»), apresenta as mesmas propriedades de uma máquina de Turing, sendo além disso capaz de reescrever o resultado da computação imediatamente anterior, sem interrupções.

O segundo modelo, chamado «gerador de sequências numeráveis» («enumerable output machine»), com as mesmas propriedades do anterior, é tal que as sequências que produz não podem decrescer lexicograficamente. Se o universo pode ser considerado o resultado de uma computação, deve-se então procurar um algoritmo capaz de computar todos os universos possíveis.

Schmidhuber observou que uma versão modificada do algoritmo de pesquisa de dados do cientista da computação russo Leonid Levin [1948 - ] poderia computar todos os

universos possíveis, utilizando recursos denumeráveis no espaço e no tempo.

O programa que processa todos os universos possíveis utiliza a lista de todos os algoritmos concebíveis numa máquina de Turing e é tal que toda a sequência produzida será concluída no mesmo período de tempo usado pelo mais rápido dos algoritmos na lista. Desse modo, o algoritmo de Schmidhuber processa todos os universos computáveis no menor tempo possível.

Do ponto de vista da previsibilidade dos eventos observáveis, a computação em causa levanta um problema. Se o algoritmo gera a evolução de todos os universos possíveis, como identificar aquele em que vivemos?

Uma possibilidade consiste em assumir um critério de escolha baseado na simplicidade de descrição de um universo. Esse critério constitui um refinamento da «teoria da inferência indutiva» do matemático americano Ray Solomonoff [1926 - 2009], usando o conceito de probabilidade algorítmica (Solomonoff, 1964a, 1964b), relacionada com a complexidade de Kolmogorov, inventada pelo matemático russo Andrey Nikolaevich Kolmogorov [1903 - 1987]. Dada uma sequência de símbolos cujo algoritmo se desconhece, a teoria de Solomonoff fornece uma resposta em termos probabilistas para a próxima ocorrência na sequência. Para uma sequência de símbolos, a complexidade de Kolmogorov – também relacionada com o conceito de «irreduzibilidade algorítmica», introduzido por Gregory Chaitin [1947 - ], - define que o algoritmo mais simples será sempre aquele cuja extensão é menor em comparação com a extensão dos algoritmos restantes (Chaitin, 1987).

Schmidhuber observou, no entanto, que o tempo de computação de um algoritmo, identificado utilizando o critério anterior, poderia ser demasiado extenso. Sugeriu assim que o requisito que permite identificar o universo procurado devesse também incluir a otimização temporal da computação.

O critério seguido para definir a probabilidade de ocorrência de um universo deverá, portanto, considerar não só a extensão do algoritmo que o produz, mas também o tempo em que esse universo é computado. Na ontologia de Schmidhuber cada universo é o resultado de um processo determinista correspondente à evolução de vectores de onda, computado sob restrições temporais e sujeito a ruído proveniente de subrotinas pseudo-aleatórias que produzem o colapso de onda. Este ruído justifica a presença aparente de acaso. A complexidade de computação de um universo encontra-se assim associada ao tempo da sua produção, em que os universos mais prováveis de serem produzidos, são aqueles cuja computação demora menos tempo.

Dado um conjunto de sequências finitas de eventos observáveis, Schmidhuber desenvolveu uma forma de atribuir a cada uma das sequências uma probabilidade de processamento. Essa probabilidade, denominada «velocidade apriorística de um universo»

(«speed prior») (Schmidhuber, 2002), depende do número de recursos no espaço e no tempo necessários para computar a sequência a que se refere.

Utilizando em seguida a regra de Bayes, para a probabilidade condicional, o mesmo autor definiu, num conjunto de sequências de eventos, a probabilidade de que uma delas seja a continuação de uma sequência dada. A nova probabilidade, dependendo da velocidade apriorística de um universo, confirma a asserção inicial de que universos computáveis no menor tempo possível são aqueles com maior probabilidade de ocorrerem.

Esta estratégia permite concluir o seguinte. Sequências aparentemente aleatórias de eventos, por serem mais longas, com menor compressibilidade algorítmica, corresponderão sempre a computações mais lentas, existindo a alternativa de as mesmas serem computadas por processos semi-aleatórios, mais rápidos, utilizando algoritmos deterministas. Esta afirmação opõe-se à visão ortodoxa da mecânica quântica, que pressupõe a existência de uma forma de aleatoriedade absoluta, para além das capacidades de um algoritmo pseudo-aleatório para a simular.

Outra linha de investigação relacionando física e computação é conhecida por «filosofia digital» (Fredkin, 2011), em que a abordagem computacional da física pressupõe que todas as quantidades na natureza são discretas. Esta perspectiva implica que na natureza não existem infinidades, quantidades infinitesimais, ou fenómenos aleatórios. Todos os fenómenos físicos são o resultado de processos discretos ocorrendo na natureza, cujo estudo dever ser feito no âmbito de uma «mecânica digital», representando processos quânticos no sentido do computador quântico universal de Deutsch.

A abordagem física da computação conduz, deste modo, à hipótese extrema de que a própria realidade possa corresponder à manifestação de processos de manipulação simbólica, envolvendo transferência de informação. Assumida a diversidade algorítmica na produção de universos, existe uma tendência para definir critérios que associam o universo observável a um programa universal que o gera, processado num modelo universal de computação. De certa forma entendendo-se as regras nos algoritmos como leis físicas que regem fenómenos.

A tese fisicalista da inteligência artificial resulta assim reforçada a partir de uma abordagem computacional absoluta da realidade, quer esta se desenvolva através do computador quântico universal, quer através do modelo computacional determinista de Jürgen Schmidhuber, envolvendo algoritmos pseudo-aleatórios.

Deverá, por outro lado, existir um conjunto fundamental de proposições que permanecem inalteráveis em relação a qualquer observador no espaço e no tempo. Sejam os axiomas da mecânica quântica, seja o critério de optimização temporal para a probabilidade de ocorrência de um algoritmo, sugerido por Schmidhuber (2000).

Na metáfora que será proposta nesta dissertação essas proposições sugere-se poderem ser análogas às que descrevem a estrutura de um sistema formalontológico, universalmente válido, de onde derivaria a inteligibilidade das descrições sobre fenómenos na natureza.

### **3.5. Informação e limites entrópicos na natureza.**

De acordo com o que se viu, na secção anterior, uma computação pode ser entendida como um fenómeno físico, ocorrendo no espaço e no tempo, sujeito às leis da natureza. Supondo-se essa identidade ontológica entre um comportamento natural e uma manipulação simbólica, aos limites computacionais do universo, devem ser aduzidas as restrições impostas à transferência de informação no espaço e no tempo.

O estudo da relação entre o volume ocupado por um sistema físico e a quantidade de informação que o caracteriza teve origem nos trabalhos de Jacob Bekenstein [1947 - ] e Stephen Hawking [1942 - ] que, durante os anos setenta, contribuíram para o desenvolvimento da termodinâmica dos buracos negros.

Em 1972, Bekenstein sugeriu que se um objecto fosse absorvido por um buraco negro, a entropia do universo diminuiria necessariamente, violando o Segundo Princípio da Termodinâmica (Bekenstein, 1972). Sucede que um buraco negro pode ser completamente caracterizado por três grandezas físicas; massa, carga eléctrica e momento angular, independentemente da natureza da matéria absorvida, enunciado conhecido pela expressão « no-hair theorem» (Susskind e Lindesay, 2005).

Uma vez que a informação estrutural dos sistemas físicos que caíram num buraco negro é substituída por uma descrição envolvendo apenas massa, carga eléctrica e momento angular, obtemos um caso em que um conjunto mais vasto de estados físicos se parece reduzir a um conjunto menor de estados na mesma descrição. Isto é; deparamo-nos com uma situação em que a entropia termodinâmica do universo diminuiu, no momento em que um sistema físico desapareceu no buraco negro.

Bekenstein sugeriu que a um buraco negro deve, por isso, estar associada uma certa quantidade de entropia termodinâmica, capaz de compensar a quantidade de estados físicos em falta. No ano seguinte, Stephen Hawking co-autorou com os físicos americano James Bardeen [1939 - ] e australiano Brandon Carter [1942 - ] um artigo onde se estabelecia a existência de uma analogia entre o comportamento de buracos negros e o comportamento de sistemas termodinâmicos. Demonstrou-se, em especial, que a área do horizonte de um buraco negro nunca decresce e também que a área total dos horizontes de

dois buracos negros quando se fundem é sempre maior ou igual à soma das áreas originais (Bardeen *et al.*, 1973). Algo que fazia lembrar o comportamento entrópico dos sistemas, justamente previsto pela segunda lei da termodinâmica.

Porém, reconhecendo mais do que uma simples analogia entre a área do horizonte de um buraco negro e a entropia termodinâmica, Bekenstein propôs, em 1974, que a primeira era efectivamente uma medida da segunda. Tornando-se então compreensível que o aumento da área do horizonte de um buraco negro corresponde a um aumento da entropia termodinâmica, compensando a diminuição anómala de entropia no universo, quando um sistema físico cai num buraco negro. Estavam reunidas as condições para que Bekenstein enunciasse o «Segundo Princípio Generalizado da Termodinâmica», estabelecendo que a soma entre as entropias de sistemas, gravitacionalmente estáveis, e as entropias dos buracos negros é sempre maior ou igual aos valor total da entropia no universo (Bekenstein, 1974).

No entanto, um sistema físico exibindo entropia deve também apresentar, uma temperatura termodinamicamensurável. A medição desta grandeza num buraco negro - um sistema físico do qual radiação não parece poder escapar - constituía um contra-senso, motivo pelo qual Hawking, de início, recusou a proposta de Bekenstein.

Analisando o problema, Hawking descobriu que, em consequência de fenómenos quânticos, no horizonte de um buraco negro, este emite uma forma de radiação térmica, mais tarde conhecida pela «radiação de Hawking» (Hawking, 1974).

No seguimento deste novo resultado e refinada a intuição inicial de Bekenstein, Hawking mostrou que a entropia termodinâmica de um buraco negro  $S_{bn}$  era dada pela expressão:

$$(1) \quad S_{bn} = \frac{k \cdot A}{4 \cdot l_p^2}$$

Onde  $k = 1,38065 \times 10^{-23} \text{ J.K}^{-1}$  é a constante de Boltzmann,  $l_p = 1,61624 \times 10^{-35} \text{ m}$  é o comprimento de Planck e  $A$  é a área do horizonte do buraco negro.

A entropia termodinâmica de um sistema físico é uma medida da quantidade máxima do número de estados independentes que o sistema pode assumir. Pode também assumir-se representar uma medida da informação sobre o comportamento de um sistema físico, no sentido que invoca o número de estados distintos necessários para o descrever.

Estas observações levaram, por sua vez, Bekenstein à formulação de uma relação de natureza universal entre a informação atribuível a um sistema físico e a quantidade de área envolvente associável ao mesmo (Bekenstein, 1981).

De acordo com o Segundo Princípio Generalizado da Termodinâmica, a quantidade de entropia mensurável através do incremento da área do horizonte do buraco negro deve corresponder literalmente à quantidade de informação necessária para descrever o sistema físico absorvido. Bekenstein sugeriu que o princípio permite estabelecer uma relação quantitativa entre a entropia termodinâmica e a área envolvente de um sistema físico, independentemente da sua estrutura inicial, antes de ser absorvido por um buraco negro.

Bekenstein representou a situação dizendo que se  $S_m$  for a entropia de um sistema físico de energia  $E_m$ , num espaço-tempo euclidiano, confinado à esfera de menor raio  $r$  capaz de o conter, então verifica-se a seguinte expressão:

$$(2) \quad S_m \leq 2 \pi E_m r \frac{k}{\hbar \cdot c}$$

A aplicação desta expressão a um buraco negro de Schwarzschild, esféricamente simétrico, sem carga eléctrica e com momento angular nulo, permite derivar a seguinte expressão, escrita em unidades de Planck (Susskind e Lindesay, 2005).

$$(3) \quad S_{bn} \leq \frac{A_{bns}}{4}$$

A igualdade em (3) mostra que um buraco negro de Schwarzschild satura o limite superior entrópico universal de Bekenstein.

Posteriormente, o físico americano Leonard Susskind [1940 -] propôs uma experiência conceptual denominada «processo de Susskind», através da qual mostrou que o Segundo Princípio Generalizado da Termodinâmica implica a existência de um limite entrópico na natureza (Bousso, 2002). Raciocinou da seguinte maneira:

Seja um sistema físico esférico, de massa  $E$ , num espaço-tempo assintoticamente plano, permitindo a formação de buracos negros.

Seja  $A$  a área da esfera de menor raio  $r$  que envolve o sistema.

Chame-se  $M$  à massa do buraco negro formado pelo sistema coberto por uma superfície esférica fechada, de área aproximadamente igual a  $A$  e de massa  $M-E$ .

Finalmente, seja  $S_m$  a entropia do sistema físico imediatamente antes do colapso gravitacional.

O Segundo Princípio Generalizado da Termodinâmica estabelece que:

$$(4) \quad S_{bn}(depois) + S_m(depois) - [S_{bn}(antes) + S_m(antes)] \geq 0$$

O valor de entropia termodinâmica de um sistema físico absorvido por um buraco negro é, por definição,  $S_m(depois) = 0$ , visto que o sistema deixa de poder ser observado.

Por outro lado, de acordo com o Segundo Princípio Generalizado da Termodinâmica, a entropia do buraco negro  $\Delta S_{bn} = S_{bn}(depois) - S_{bn}(antes)$  deve aumentar.

Resultando, a partir de (4):

$$(5) \quad S_m \leq \Delta S_{bn}$$

Dado que o buraco negro imediatamente antes da sua formação não existe, a sua entropia inicial deverá ser nula, isto é;  $S_{bn}(antes) = 0$ . Por outro lado, a sua entropia, imediatamente a seguir ao colapso, de acordo com a expressão de Hawking (3), é  $S_{bn} = A/4$ .

Tendo-se, finalmente:

$$(6) \quad S_m \leq \frac{A}{4}$$

Expressão cuja formulação depende da validade do Segundo Princípio Generalizado da Termodinâmica. A expressão (6) é conhecida pelo «limite entrópico esférico» (Susskind, 1995). Este limite entrópico pode ser considerado um caso particular do «limite entrópico espacial» (Bousso, 2002), cuja definição é a seguinte: a entropia contida em qualquer região tridimensional de espaço não deve exceder um quarto da área, medida em unidades de área de Planck, da menor região bidimensional envolvendo a região tridimensional dada.

Os limites entrópicos «superior universal de Bekenstein», «esférico» e ainda «espacial» são falsificados por um conjunto de contra-exemplos em que se considera o espaço-tempo curvo e em que a distinção entre a zona interior e zona exterior da região bidimensional que envolve a região tridimensional dada é difícil de definir.

Este facto levou o físico israelita Raphael Bousso [1972 - ] a propor uma versão mais geral do limite entrópico, através da aplicação de superfícies bidimensionais formadas por raios de luz (Bousso, 1999). Bousso chamou-lhe o «limite entrópico covariante», definindo-o da seguinte maneira.

Seja  $B$  uma superfície bidimensional qualquer, num espaço-tempo quadridimensional, e seja  $L(B)$  uma superfície bidimensional chamada folha de luz, formada por todos os raios de luz que originam perpendicularmente a  $B$ , no espaço-tempo considerado e que se cruzam num vértice chamado ponto focal.

A entropia  $S_m$  de todos os sistemas físicos observáveis no interior de  $L(B)$  será sempre tal que:

$$(7) \quad S_m[L(B)] \leq \frac{A(B)}{4}$$

### 3.6 O Princípio Holográfico e a conservação de informação na natureza.

Assumindo-se uma equivalência entre entropia e informação, um limite entrópico estabelece que a quantidade de informação suficiente para descrever um sistema físico é proporcional à área da superfície que o envolve e não ao volume ocupado pelo sistema.

Este resultado é contra-intuitivo. Seria expectável que o volume de um sistema físico fosse relevante para a determinação da quantidade de informação necessária para descrever o sistema. A assumpção anterior é, portanto, desconcertante e não seria insensato considerá-la associada apenas à formação de buracos negros, tal como a contabilidade da entropia termodinâmica feita por Bekenstein parecia sugerir.

Porém, a existência de uma inconsistência no contexto da física, na segunda metade do século XX, conduziu as circunstâncias da sua resolução à generalização do limite entrópico, aplicável a qualquer situação física.

A necessidade dessa conclusão encontra-se associada à enunciação do Princípio Holográfico, cuja génese intelectual começa com o «paradoxo da informação» («information paradox»), formulado por Steven Hawking [1972 - ].

A conservação da informação física está associada ao conceito de reversibilidade.

Um sistema físico pode sofrer várias transformações ao longo do tempo, mas com a quantidade ideal de informação, será sempre possível, em princípio, reconstruir o que sucedeu ao longo do tempo a esse sistema. O princípio da causalidade mantém-se e com isso a conservação da informação física.

Seria, portanto, desastroso se as próprias teorias que descrevem a natureza permitissem a destruição de informação física e, com isso, a violação do princípio da causalidade. Porém, foi justamente isso que Hawking propôs.

Depois de ter trabalhado em 1973 na termodinâmica dos buracos negros e previsto no ano a seguir a radiação com o seu nome, Hawking anunciou uma situação física em que parecia existir destruição de informação (Hawking, 1976). O seu ponto de partida foi precisamente a natureza da radiação emitida por um buraco negro.

A existência de um buraco negro produz uma deformação extrema do espaço-tempo. No entanto, para um buraco negro de grandes dimensões, essa deformação tende a ser menos proeminente junto ao horizonte de eventos, onde se pode considerar válida a aplicação de uma teoria quântica de campo. A situação torna-se então análoga àquela que ocorre num espaço de Rindler, isto é; num espaço-tempo plano («flat spacetime»); (Susskind, Lindesay, 2005).

Desse modo, em virtude das flutuações energéticas do vácuo, produzem-se nessas regiões pares partícula/anti-partícula, junto ao horizonte de eventos.

Sucedem que, por vezes, uma das partículas, de energia positiva, é irradiada como parte da radiação de Hawking, enquanto que a outra, de energia negativa, é absorvida pelo buraco negro.

Cada um dos pares de partículas produzidas nunca chega a interagir com qualquer outro par (Mathur *in* Papantonopoulos, 2009). Em consequência deste facto cada partícula está em correlação com a respectiva anti-partícula e a observação conjunta do estado de cada um daqueles objectos define o que em mecânica quântica se chamaria um «estado puro».

A entropia de von Neumann, que mede o grau de correlação entre os objectos componentes de um sistema («entanglement entropy»), para o caso em que uma das partículas do par colapsa no buraco negro é, nesta situação, não nula. Uma vez que a descrição do sistema exclui agora metade do mesmo, ocultando informação. Por este motivo se diz que o sistema se apresenta um estado misturado («mixed state»).

A informação oculta é aquela que combinada com a referente à outra metade do sistema observado, permite reconstruir todas as propriedades do sistema formado pelas duas partículas e obter o seu conhecimento pleno, fazendo com que a entropia de von Neumann seja nula.

A possibilidade de observação de apenas uma das partículas no par corresponde fisicamente à existência de uma radiação desordenada, emitida pelo buraco negro, que pode ser interpretada como a entropia termodinâmica do sistema.

De acordo com Susskind, é possível definir a quantidade de informação, apreendida por um sujeito quando observa um sistema físico, como a diferença entre os valores da entropia de von Neumann calculados, respectivamente, para a situação em que a observação se faz sobre um estado puro (entropia nula) e para a situação em que apenas se pode fazer observações sobre um estado misto (Susskind, Lindesay, 2005).

Consequentemente, a informação de qualquer sistema físico pode ser definida como a sua entropia termodinâmica, no sentido acima descrito.

Na ortodoxia da mecânica quântica, a evolução de um vector de onda, representando um estado puro, em que existe uma correlação entre dois objectos, enquanto não for colapsado por uma observação, é um processo reversível. O sistema evolui de maneira determinista. O que no jargão técnico da mecânica quântica se chama uma evolução unitária.

Consideremos, pois, os pares partícula/anti-partícula na origem da radiação de Hawking. Neste caso, a observação síncrona dos estados de cada objecto componente é em princípio possível mas para observadores que jamais contactarão, uma vez que um desses objectos, um por cada par produzido, se encontra no interior do horizonte de eventos do buraco negro. Porém, embora os resultados não possam ser mais tarde comparados, a capacidade previsiva da mecânica quântica – na forma de uma teoria de campo, válida na orla do horizonte de eventos, garante a reversibilidade do processo de produção da radiação. Isto porque se assume naquela teoria que as anti-partículas absorvidas pelo buraco negro continuam a existir, ainda que veladas à observação de um observador externo ao horizonte.

A radiação de Hawking é formada pela nuvem daquelas partículas de energia positiva, separadas do seu anti-par, cuja entropia de von Neumann é precisamente igual ao valor da entropia termodinâmica do buraco negro calculada por Hawking (Bardeen *et al*, 1973).

A consideração desta entropia e da outra que – de natureza idêntica – diz respeito às anti-partículas ainda existentes no interior do buraco negro (Mathur *in* Papantonopoulos, 2009) permitiria, em princípio, reconstituir todas as características do sistema e da sua evolução unitária. Assumindo-se que parte alguma da informação sobre o mesmo poderia cessar de existir, preservando-se o princípio da causalidade.

Porém, no contexto da relatividade generalizada Hawking descobriu que um buraco negro, passada a quantidade necessária de tempo, perderia a quantidade de

radiação termodinâmica suficiente para que desaparecesse, como afirmou, se evaporasse.

A emissão de partículas de energia positiva ou, equivalentemente, a absorção de anti-partículas de energia negativa corresponde a uma diminuição da energia do buraco negro e, portanto, à diminuição da sua massa.

Esta diminui até que o buraco negro irradia o último quanta de energia e desaparece da realidade física. A combinação das duas teorias, mecânica quântica e relatividade generalizada, deveria possibilitar a formulação de um estado final, a partir do qual, fosse possível traçar a evolução de cada um dos pares de partículas, desde o início do colapso de uma das partículas, até à situação final de evaporação do buraco negro.

Essa descrição só poderia ser obtida se se tivesse acesso às duas formas de radiação entrópica, no interior e fora do buraco negro. Uma vez que já não há interior, sucede que um estado de evolução determinista previsto na mecânica quântica, passou de puro a misto sem que houvesse uma observação sobre o sistema. Esta situação contraria a condição de que uma evolução de um estado puro deve permanecer assim indefinidamente. Hawking interpretou este facto como uma destruição de informação física, conduzindo à violação do princípio da causalidade na natureza. Este é o paradoxo da informação de Hawking.

Defendendo uma posição epistemológica de imprevisibilidade absoluta sobre a realidade, escreve Hawking no resumo do seu artigo:

“A singularity is a place where the classical concepts of space and time break down as do all the known laws of physics because they are all formulated on a classical space-time background. In this paper it is claimed that this breakdown is not merely a result of our ignorance of the correct theory but that it represents a fundamental limitation to our ability to predict the future, a limitation that is analogous but additional to the limitation imposed by the normal quantum-mechanical uncertainty principle. The new limitation arises because general relativity allows the causal structure of space-time to be very different from that of Minkowski space. The interaction region can be bounded not only by an initial surface on which data are given and a final surface on which measurements are made but also a "hidden surface" about which the observer has only limited information such as the mass, angular momentum, and charge. Concerning this hidden surface one has a "principle of ignorance": The surface emits with equal probability all configurations of particles compatible with the observers limited knowledge. It is shown that the ignorance principle holds for the quantum-mechanical evaporation of black holes: The black hole creates particles in pairs, with one particle always falling into the hole and the other possibly escaping to infinity. Because part of the information about the state of the system is lost down

the hole, the final situation is represented by a density matrix rather than a pure quantum state. This means there is no S matrix for the process of black-hole formation and evaporation. Instead one has to introduce a new operator, called the superscattering operator, which maps density matrices describing the initial situation to density matrices describing the final situation”.

(Hawking, 1975)

A posição definitiva de Hawking era a de que informação física era necessariamente destruída.

Anos mais tarde, Leonard Susskind [1940 - ] Gerard 't Hooft [1946 - ] ouviam perplexos estas ideias, enquanto Stephen Hawking, com o auxílio de um intérprete (o físico Martin Rocek), paulatinamente as expunha em 1981, num seminário organizado em São Francisco pelo filantropo Werner Erhard.

Susskind descreve o sucedido no seu livro «The Black Hole War – My battle with Stephen Hawking to make the world safe for quantum mechanics»:

“Stephen (through Martin) said something that I thought was wrong. I turned to Martin and asked him for clarification of the physics. He looked at me like a deer caught in the headlights. Later he told me what had happened. It seems that translating Stephen's speech required such intense concentration that he was usually unable to keep track of the conversation. He barely knew what we were talking about (...)

During the EST conference, I found talking to Stephen very difficult. It took a long time for him to answer, and his answers were usually very brief. These short, sometimes single-word answers, his smile, and his almost disembodied intellect were unnerving. It was like communicating with the Oracle at Delphi. When someone submitted a question to Stephen, the initial response was absolute silence,

and the eventual output was often incomprehensible. But the knowing smile said, ‘You may not understand what I'm saying, but I do, and I am right.’

(...) his boldest move ever was the bomb he dropped in Werner's attic.

I can't remember how his lecture worked at EST. Today a physics seminar given by Stephen has him sitting quietly in his chair while a disembodied computer voice lectures from a previous recording. That computerized voice has become Stephen's trademark; as flat as it is, it is full of personality and humor. But back then, maybe he talked and Martin translated. However it happened, the bomb fell with full force on Gerard and me.

Stephen claimed that ‘information is lost in black hole evaporation,’ and, worse, he seemed to prove it. If that was true, Gerard and I realized, the foundations of our

subject were destroyed. How did the rest of the people in Werner's attic receive the news? Like the coyote in the roadrunner cartoon who overruns the edge of the cliff: the ground had disappeared beneath their feet, but they didn't know it yet.

It is said of cosmologists that they are often in error but never in doubt. If so, Stephen is only half a cosmologist: never in doubt but hardly ever wrong. In this case, he was. But Stephen's 'mistake' was one of the most seminal in the history of physics and could ultimately lead to a profound paradigm shift about the nature of space, time, and matter.

Stephen's lecture was the last that day. For about an hour afterward, Gerard stood glaring at the diagram on Werner's blackboard. Everyone else had left. I can still see the intense frown on Gerard's face and the amused smile on Stephen's. Almost nothing was said. It was an electric moment.

On the blackboard was a Penrose diagram, a type of diagram representing a black hole. The horizon (the edge of the black hole) was drawn as a dashed line, and the singularity at the center of the black hole was an ominous-looking jagged line. Lines pointing inward through the horizon represented bits of information falling past the horizon into the singularity. There were no lines coming back out. According to Stephen, those bits were irretrievably lost. To make matters worse, Stephen had proved that black holes eventually evaporate and disappear, leaving no trace of what has fallen in.

Stephen's theory went even further. He postulated that the vacuum - empty space was full of 'virtual' black holes that flashed into and out of existence so rapidly that we didn't notice them. The effect of these virtual black holes, he claimed, was to erase information, even if there was no "real" black hole in the vicinity.

(...) this was an unmitigated disaster: 't Hooft and I knew it, but the response from everyone else who heard about it that day was 'Ho hum, information is lost in black holes.' Stephen himself was sanguine. For me the toughest part of dealing with Stephen has always been the irritation I feel at his complacency. Information loss was something that just could not be right, but Stephen couldn't see it. The conference broke up, and we all went home. For Stephen and Gerard, that meant back to Cambridge University and the University of Utrecht, respectively; for me a forty-minute drive south on Route 101 back to Palo Alto and Stanford University. It was hard to concentrate on the traffic. It was a cold day in January, and every time I stopped or slowed down, I would draw the figure from Werner's blackboard on my frosty windshield".

(Susskind, 2008, p. 20-23)

Passaria algum tempo até que Susskind, analisando a situação se apercebesse de uma discrepância em relação às descrições sobre os mesmos fenômenos por parte de dois observadores, um externo e outro, interno ao horizonte de eventos de um buraco negro.

Um buraco negro, porque emite radiação, deve ter uma temperatura junto ao horizonte de eventos. Essa temperatura é registada por todos os observadores que orbitam em redor do horizonte, acelerando de forma contrária à atracção gravitacional do buraco negro. Trata-se da temperatura prevista por Bill Unruh associada ao efeito com o seu nome (Unruh, 1976). Esta situação contrasta vivamente com a descrição feita por um observador inercial que cai através do horizonte de eventos. O Princípio de Equivalência formulado por Einstein, no contexto da relatividade generalizada, estabelece o seguinte. Um observador que a partir dos movimentos dos objectos junto a si, verifique a existência de uma aceleração comum ao movimento de todos esses objectos, não consegue determinar se a aceleração é devida a um campo gravitacional externo ou a uma causa associada ao local onde se encontra, a aceleração devida a um foguete que o faz ascender, por exemplo.

Nem tão pouco, na ausência de semelhante aceleração, consegue esse observador distinguir se numa situação de queda livre, se aproxima de um corpo que o atrai gravitacionalmente ou se se encontra em movimento inercial, isto é; livre de acelerações, a uma velocidade constante no espaço.

Um referencial acelerado por uma pseudo-força é, portanto, equivalente a um referencial num campo gravitacional assim como um referencial em queda livre, num campo gravitacional, é equivalente a um referencial inercial.

Desse modo, um sujeito caíndo em direcção a um buraco negro, atravessa o horizonte de eventos na qualidade de observador inercial, sem portanto registar qualquer indício da temperatura de Unruh ou da radiação de Hawking.

Como explica Susskind, chamando Fidos («fiducial observers») aos observadores externos e Frefos («free falling observers») aos internos ao horizonte de eventos, estamos em presença de outra contradição ainda:

“Thus, according to the Fidos, the fluctuation lasts for an infinite time and is therefore not virtual at all. Real particles are seen being injected into the Rindler space from the horizon, and eventually fall back to the horizon. To state it differently, the horizon behaves like a hot membrane radiating and reabsorbing thermal energy.

A natural question to ask is whether the thermal effects are ‘real’. For example, we may ask whether any such thermal effects are seen by freely falling observers carrying their thermometers with them as they pass through the horizon. Obviously the answer is no. A thermometer at rest in an inertial frame in the Minkowski vacuum will record zero temperature. It is tempting to declare that the thermal effects seen by the Fidos are fictitious and that the reality is best described in the frame of the Frefos. However, by yielding to this temptation we risk prejudicing

ourselves too much toward the viewpoint of the Frefos. In particular, we are going to encounter questions of the utmost subtlety concerning the proper relation between events as seen by observers who fall through the horizon of a black hole and those seen by observers who view the formation and evaporation process from a distance. Thus for the moment, it is best to avoid the metaphysical question of whose description is closer to reality. Instead we simply observe that the phenomena are described differently in two different coordinate systems and that different physical effects are experienced by Frefos and Fidos. In particular a Fido equipped with a standard thermometer, particle detector, or other apparatus, will discover all the physical phenomena associated with a local proper temperature (...) By contrast, a Frefo carrying similar apparatuses will see only the zero temperature vacuum state. Later we will discuss the very interesting question of how contradiction is avoided if a Frefo attempts to communicate to the stationary Fidos the information that no thermal effects are present”.

(Susskind, Lindesay, 2005, p. 41-42)

Reformulando e estendendo o trabalho anterior de t' Hooft (1985, 1990), Susskind, juntamente com os físicos LÁrus Thorlacius [1977 – ] e John Uglum, sugeriu a existência de duas descrições complementares verdadeiras para a mesma situação (Susskind, *et al*, 1993). Uma ideia que não era estranha ao princípio da complementaridade onda-partícula em mecânica quântica e que, agora, se aplicava agora à nova situação como o «Princípio da Complementaridade dos buracos negros» («Black Hole Complementarity»).

As duas descrições aduziriam um elemento de não localidade às teorias físicas, porquanto uma diferença de ponto de vista, implicaria uma descrição radicalmente diferente do mesmo fenómeno. A sua possibilidade conjunta não acarretaria, por outro lado, uma inconsistência, se fosse impossível aos dois observadores a comparação *in loco* dos seus registos, facto impedido pelas restrições de natureza gravitacional impostas pelo próprio buraco negro (Susskind, Lindesay, 2005).

Os três postulados na base da formulação desta complementaridade, conforme apresentados no artigo de Susskind, eram os seguintes:

“Postulate 1: The process of formation and evaporation of a black hole, as viewed by a distant observer, can be described entirely within the context of standard quantum theory.

In particular, there exists a unitary S–matrix which describes the evolution from infalling matter to outgoing Hawking-like radiation.

Postulate 2: Outside the stretched horizon <sup>8</sup> of a massive black hole, physics can be described to good approximation by a set of semi-classical field equations.

Postulate 3: To a distant observer, a black hole appears to be a quantum system with discrete energy levels. The dimension of the subspace of states describing a black hole of mass  $M$  is the exponential of the Bekenstein entropy  $S(M)$ .”

(Suskind *et al*, 1993, p. 2-3)

O terceiro postulado, em particular, procurava eliminar a anomalia que surge na aplicação de teorias quânticas de campo para determinar a entropia de um buraco negro, em que o respectivo valor diverge:

“Quantum field theory has too many degrees of freedom to consistently describe a gravitational theory. The main indication that we have seen of this overabundance of degrees of freedom is the fact that the horizon entropy-density is infinite in quantum field theory. The divergence arises from modes very close to the horizon. One might think that this is just an indication that a more or less conventional ultraviolet regulator is needed to render the theory consistent. But the divergent horizon entropy is not an ordinary ultraviolet phenomenon. The modes which account for the divergence are very close to the horizon and would appear to be ultra short distance modes. But they also carry very small Rindler energy and therefore correspond to very long times to the external observer. This is an example of the weirdness of the Ultraviolet/Infrared connection.

A quantitative measure of the overabundance of degrees of freedom in QFT is provided by the Holographic Principle. This principle says that there are vastly fewer degrees of freedom in quantum gravity than in any QFT even if the QFT is regulated as, for example, it would be in lattice field theories.

The Holographic Principle is about the counting of quantum states of a system. We begin by considering a large region of space. For simplicity we take the region to be a sphere. Now consider the space of states that describe arbitrary systems that can fit into such that the region outside is empty space. Our goal is to determine the dimensionality of that state-space”.

(Suskind, Lindesay, 2005, p. 101)

A complementaridade dos buracos negros é uma conjectura, portanto, conducente

---

<sup>8</sup> O «horizonte alargado» («the stretched horizon»), constitui neste contexto uma membrana envolvendo a superfície do buraco negro, separada do mesmo de uma distância da ordem de um comprimento de Planck. Essa membrana tem temperatura, absorvendo e irradiando energia térmica. Pode-se considerar ainda uma espécie de atmosfera térmica do buraco negro.

à ideia de que o comportamento de um buraco negro pode ser integralmente descrito por uma teoria quântica de campo, válida junto ao horizonte de eventos. Mais ainda, que essa teoria necessita apenas de registrar os eventos próximos do horizonte, ainda observáveis por um sujeito externo. E, finalmente, que a quantidade de entropia que contabiliza é, justamente, o valor previsto por Hawking.

As condições de enunciação do Princípio Holográfico começam pois a surgir.

Os três postulados indicados sugerem-no mas procuram apenas a resolução do paradoxo imposto por duas descrições contraditórias da realidade, junto ao horizonte de eventos do buraco negro:

“The (...) [the stretched event horizon] is very real to an outside observer. For example, if such an observer is suspended just above (...) [it], he or she will observe an intense flux of energetic radiation apparently emanating from the membrane. If provided with an electrical multimeter, our observer will discover that the membrane has a surface resistivity of 377 ohms. If disturbed, the stretched horizon will respond like a viscous fluid, albeit with negative bulk viscosity. And finally, the observed entropy of the massive black hole is proportional to the area of the stretched horizon. If, on the other hand, the observer attempts to determine if the membrane is real by letting go of the suspension mechanism and falling freely past the stretched horizon, the membrane will disappear. However, there is no way to report the membrane’s lack of substance to the outside world. In this sense, there is *complementarity* between observations made by infalling observers who cross the event horizon and those made by distant observers.

We believe that Postulates 1-3 are most naturally implemented by assuming that the coarse grained thermodynamic description of an appropriately defined stretched horizon has an underlying microphysical basis. In other words, from the point of view of an outside observer, *the stretched horizon is a boundary surface equipped with microphysical degrees of freedom that appear in the quantum Hamiltonian used to describe the observable world*. These degrees of freedom must be of sufficient complexity that they behave ergodically and lead to a coarse-grained, dissipative description of the membrane.

Much of this paper is concerned with the illustration of the concept of the stretched horizon in the context of two-dimensional dilaton gravity, for which a semi-classical description has been formulated”.

(Susskind *et al*, 1993, p. 6)

Mas que dizer da matéria no interior do buraco negro e das objecções levantadas pelo paradoxo da informação de Hawking. Não ficará ainda parte da realidade por descrever? Poderia, de outro modo, suceder que, no contexto da complementaridade

apresentada, a informação estendida e termalizada do ponto de vista do observador externo constituísse uma forma de codificação sobre o que se encontra no interior do horizonte?

A hipótese da complementaridade será, finalmente, estendida até à máxima consequência e o primeiro a fazê-lo é Gerard t'Hooft, em 1993, num ensaio dedicado ao físico paquistanês e prémio Nobel Abdul Salam [1926 – 1996]:

“The requirement that physical phenomena associated with gravitational collapse should be duly reconciled with the postulates of quantum mechanics implies that at a Planckian scale our world is not 3+1 dimensional. Rather, the observable degrees of freedom can best be described as if they were Boolean variables defined on a two-dimensional lattice, evolving with time. This observation, deduced from not much more than unitarity, entropy and counting arguments, implies severe restrictions on possible models of quantum gravity. Using cellular automata as an example it is argued that this dimensional reduction implies more constraints than the freedom we have in constructing models. This is the main reason why so-far no completely consistent mathematical models of quantum black holes have been found”.

(t' Hooft, 1993, p. 1)

Pouco depois, em 1994, Susskind no artigo «The World as a Hologram» dava substância às propostas de t' Hooft no contexto da teoria das cordas:

According to 't Hooft the combination of quantum mechanics and gravity requires the three dimensional world to be an image of data that can be stored on a two dimensional projection much like a holographic image.

The two dimensional description only requires one discrete degree of freedom per Planck area and yet it is rich enough to describe all three dimensional phenomena. After outlining 't Hooft's proposal I give a preliminary informal description of how it may be implemented.

One finds a basic requirement that particles must grow in size as their momenta are increased far above the Planck scale. The consequences for high energy particle collisions are described.

The phenomena of particle growth with momentum was previously discussed in the context of string theory and was related to information spreading near black hole horizons. The considerations of this paper indicate that the effect is much more rapid at all but the earliest times. In fact the rate of spreading is found to saturate the bound from causality. Finally we consider string theory as a possible realization of 't Hooft's idea”.

(Susskind, 1994, p. 1)

Os fenómenos, aparentemente inacessíveis, no interior do buraco negro reduzir-se-iam-se agora a uma forma de física dimensional, de natureza quântica e cuja evolução é unitária fora do horizonte de eventos. Não havendo, portanto, destruição de informação, uma vez que a história de todos os objectos que colapsam no buraco negro pode ser teorizada através de processos quânticos que ocorrem na orla do horizonte de eventos.

O Princípio Holográfico estabelece uma ligação causal entre fenómenos no interior do horizonte e imediatamente fora dele, de natureza não local numa situação acentuadamente gravitacional. Sendo ainda consistente com a existência de um limite entrópico:

“Let us consider a region of flat space  $T$ . We have seen that the maximum entropy of all physical systems that can fit in  $T$  is proportional to the area of the boundary  $dT$ , measured in Planck units. Typically, as in the case of a lattice of spins, the maximum entropy is a measure of the number of simple degrees of freedom that it takes to completely describe the region. This is almost always proportional to the volume of  $T$ . The exception is gravitational systems. The entropy bound tells us that the maximum number of non-redundant degrees of freedom is proportional to the area. For a large macroscopic region this is an enormous reduction in the required degrees of freedom. In fact if the linear dimensions of the system is of order  $L$  then the number of degrees of freedom per unit volume scales like  $1/L$  in Planck units. By making  $L$  large enough we can make the degrees of freedom arbitrarily sparse in space. Nevertheless we must be able to describe microscopic processes taking place anywhere in the region. One way to think of this is to imagine the degrees of freedom of  $T$  as living on  $dT$  with an area density of no more than  $\sim 1$  degree of freedom per Planck area. The analogy with a hologram is obvious; three-dimensional space described by a two-dimensional hologram at its boundary! That this is possible is called the Holographic Principle”.

(Susskind, Lindesay, 2005, p. 127-128)

O princípio defende um mecanismo análogo ao que se passa na produção de hologramas, motivo pelo qual se adoptou o termo «holográfico». Afirma que é possível obter informação sobre fenómenos físicos, ocorrendo num espaço tridimensional, efectuando medidas sobre uma região aproximadamente bidimensional.

Num espaço a três dimensões existe assim fundamento para considerar que a descrição física de fenómenos tridimensionais pode ser equivalentemente feita em modelos em que uma das três dimensões pode ser desprezada para os efeitos dessa descrição.

Mediante o estabelecimento de um isomorfismo entre uma teoria, válida para  $D$

dimensões e outra, menos complexa, aplicável a  $D-1$  dimensões, torna-se então possível lidar com problemas de difícil tratamento no primeiro modelo, reduzindo-as a situações factíveis, no segundo.

Apesar das contribuições de Susskind e de t'Hooft, a comunidade científica americana permaneceu relutante em relação ao Princípio Holográfico, pelo menos até à contribuição do físico argentino Juan Maldacena [1968 - ].

Maldacena formulou um enunciado conhecido pela «conjectura da correspondência *AdS/CFT*» (Maldacena, 1998), em que a sigla *AdS* se refere a um espaço anti-de Sitter e o acrónimo *CFT* significa «teoria de campo conformal» («conformal field theory»). Isto é; uma teoria quântica de campo, com simetrias euclidianas e invariância de escala, que preserva ângulos sob aplicação de transformações matemáticas ditas «conformais». Concretamente, uma teoria particularmente adequada para a descrição da interacção entre objectos quânticos, em particular, no contexto das teorias das cordas (Gaberdiel, 1999).

Um espaço anti-de Sitter corresponde a uma solução das equações da relatividade generalizada de Einstein. Um espaço-tempo de curvatura negativa, que não é equivalente ao que caracteriza o nosso universo, mas que permite ilustrar possibilidade de existir uma correspondência entre uma região a três dimensões, mais tempo, com gravidade, e uma região a duas dimensões, mais tempo, sem gravidade.

The correspondence between gravity or its string theoretic generalization in  $AdS(5) \times S(5)$  and Super Yang–Mills (SYM) theory on the boundary (...) states that there is a complete equivalence between superstring theory in the bulk of  $AdS(5) \times S(5)$  and  $N = 4, 3 + 1$ -dimensional,  $SU(N)$ , SYM theory on the boundary of the AdS space. (...) [A SYM theory refers to] (...) a particular version of supersymmetric gauge theory,  $N$  represents the number of supersymmetries, and  $N$  is the dimension of the Yang–Mills gauge theory.

(Susskind, Lindesay, 2005, p. 133)

Ainda que os números de dimensões possam inclusivamente ser outros, a ideia é a de que uma teoria quântica sem gravidade, com menos graus espaciais de liberdade pode descrever uma região onde é válida uma teoria de cordas com maior número de graus de liberdade, contendo gravidade.

There is one question that may be puzzling to the reader. Since  $AdS(5) \times S(5)$  is a 10-dimensional spacetime one might think that its boundary is  $(8 + 1)$  dimensional. But there is an important sense in which it is  $3 + 1$  dimensional. (...) [Rescaling the

metric so there is a 5-sphere plus 3 + 1 dimensions, this] 5-sphere shrinks to zero as the boundary (...) is approached. The boundary of the geometry is therefore 3+1 dimensional”.

(Susskind, Lindesay, 2005, p. 133-134)

A conjectura da correspondência *AdS/CFT* propõe que uma teoria de supercordas, definida num espaço anti-de Sitter  $AdS_5 \times S^5$  (com dez dimensões) é equivalente a uma teoria quântica de campo de Yang-Mills  $SU(N)$  (com quatro dimensões), definida na fronteira do espaço  $AdS_5 \times S^5$ .

O prefixo «super» («SU») refere a utilização de supersimetria nas duas teorias, propriedade que assegura a existência de uma relação, contemplada na teoria, entre um conjunto de partículas chamadas bosões e o conjunto de partículas denominadas fermiões.

A teoria  $AdS_5 \times S^5$  descreve os efeitos da gravidade do espaço-tempo e envolve objectos chamados «d-branes», introduzidas pelo físico norte-americano Joseph Polchinski [1954 - ], onde estão acopladas as cordas:

“D-branes are objects which occur in superstring theory. They are stable ‘impurities’ of various dimensionality that can appear in the vacuum. A Dp-brane is a p-dimensional object. We are especially interested in D3-branes. Such objects fill 3 dimensions of space and also time. Their properties are widely studied in string theory and we will only quote the results that we need. The most important property of D3-branes is that they are embedded in a 10-dimensional space”.

(Susskind, 2005, p. 136)

A correspondência *AdS/CFT* estabelece assim um meio para compreender a gravidade como o resultado de um comportamento quântico, propiciando uma unificação entre as teorias de campo quânticas e relatividade generalizada.

Depois da publicação de Maldacena, o físico americano Edward Witten [1951 - ], aproveitou o resultado, relacionando-o com o Princípio Holográfico (Witten, 1998), como refere Susskind:

I don't know whether Maldacena made the connection between his discovery and the Holographic Principle, but Ed Witten soon did. Just two months after Maldacena's paper, Witten placed his own paper on the Internet, giving it the title ‘Anti De Sitter Space and Holography.’(...)

The moment I saw Witten's paper, I knew that the Black Hole War was finished.

Quantum Field Theory is a special case of Quantum Mechanics, and information in Quantum Mechanics can never be destroyed. Whatever else Maldacena and Witten had done, they had proved beyond any shadow of a doubt that information would never be lost behind a black hole horizon. The string theorists could understand this immediately; the relativists would take longer. But the war was over”.

(Susskind, 2008, p. 418-419)

Seguindo o físico espanhol António Dobado [1967 - ], o Princípio Holográfico afirma que num universo de dimensão  $D$ , a descrição completa de uma região  $R$  de fronteira  $B = \partial R$  com dimensão  $D-1$ , pode ser obtida através da descrição de todos os processos físicos que ocorrem em  $B$  (Dobado, 2004).

Mas de que forma pode o paradoxo da informação de Hawking ser solucionado através da invocação do Princípio Holográfico? A resposta é a de que o Princípio implica que a história de um buraco negro, desde a sua formação até ao momento em que evapora, pode ser feita, recorrendo a uma teoria de campo quântica. No caso, uma teoria que é aplicável numa região fora do horizonte de eventos onde ainda é possível fazer observações e tal que essa teoria descreve processos dinâmicos quânticos cuja evolução é unitária.

Está, portanto, preservada a reversibilidade dos processos descritos e não há razão para supor a ocorrência de destruição de informação ou a violação do princípio da causalidade.

A solução apresentada para o paradoxo de Hawking é de natureza estritamente lógica e baseia-se numa declaração de princípio, matematicamente factível, como provou Maldacena. Não obstante, o Princípio Holográfico é uma proposta filosoficamente sedutora e logicamente credível. Sobretudo se adoptarmos a posição da física digital de que o comportamento do universo é comparável ao de um computador.

### **3.7 Teorias da unificação, informação e quantificação do espaço.**

Nas duas secções anteriores viu-se como um problema de preservação da continuidade histórica dos eventos conduziu a um princípio de informação sobre comportamentos na natureza. A metáfora de que o universo é como um computador parece reforçada. No que segue descreveremos como é que essa interpretação foi utilizada para tentar a unificação entre relatividade generalizada e mecânica quântica, conduzindo à própria discretização do espaço.

A relatividade generalizada descreve de forma determinista o efeito da gravidade no universo e, tal como o electromagnetismo, é uma teoria clássica de campo. em que os parâmetros utilizados na descrição variam de modo contínuo.

As teorias quânticas de campo derivam do tratamento quântico do electromagnetismo, cuja origem remonta a 1926 com os trabalhos de Max Born [1882 - 1970], Werner Heisenberg [1901 - 1976] e Pascual Jordan [1902 - 1980], culminando na «teoria da electrodinâmica quântica» («quantum electrodynamics; QED») desenvolvida em 1940 por Richard Feynman [1918 - 1988], Freeman Dyson [1923 - ], Julian Schwinger [1918 - 1994] e Sin-Itiro Tomonaga [1906 - 1979].

Seria desejável que a teoria da relatividade generalizada pudesse ser também objecto de um tratamento quântico. Porém, existe um problema fundamental. Qualquer teoria de campo quântica envolvendo a gravidade, formulada recorrendo a técnicas análogas às utilizadas no tratamento quântico do electromagnetismo, exhibe uma dificuldade conhecida pelo «problema da normalização». De modo sucinto, o problema diz respeito à impossibilidade de obter um conjunto finito de parâmetros observáveis na descrição matemática das interacções gravitacionais.

Outro problema ainda diz respeito ao que sucede na vizinhança de uma singularidade de um buraco negro, em que as métricas da relatividade generalizada [e.g. «a métrica de Schwarzschild», (Taylor *et al.*, 2000) ] implicam que a matéria se concentra num ponto infinitamente denso, enquanto que a mecânica quântica prevê que a localização de qualquer sistema físico é indeterminada no âmbito de um intervalo de incerteza, em consequência da desigualdade de Heisenberg (Messiah, 2000).

A unificação entre gravidade e mecânica quântica está, portanto, por fazer, embora existam tentativas.

Uma das propostas consiste nas teorias de cordas, cuja origem remonta a 1968 com o trabalho do físico italiano Gabriele Veneziano [1942 - ] sobre a força nuclear forte, posteriormente desenvolvida por Yoichiro Nambu [1921 - ], Murray Gell-Mann [1929 - ], Holger Bech Nielsen [1941 - ], Leonard Susskind [1940 - ], John Schwarz [1941 - ], Joel Scherk [1946 - 1979], Tamiaki Yoneya, Alexander Polyakov [1945 - ], Michael Green [1946 - ], Edward Witten [1951 - ], Joseph Polchinski [1954 - ] e Juan Maldacena [1968 - ].

Qualquer teoria de cordas assume que o mundo resulta dos modos de oscilação de pequenos segmentos vibrantes chamados cordas, cuja extensão é da ordem de grandeza do comprimento de Planck;  $l_p = 1,61624 \times 10^{-35}$  m . A teoria mantém que não é possível fazer medições de comprimento inferior ao mínimo comprimento possível de uma corda, quantificando dessa maneira o espaço. Para além das quatro dimensões habituais, as teorias de cordas utilizam dimensões espaciais adicionais (no total, dez, nas teorias de

cordas até agora formuladas) as quais se assumem compactadas. A compactação de uma dimensão espacial, utilizada nos espaços de Kaluza-Klein (Susskind e Lindesay, 2005) significa que a influência daquela dimensão só se faz sentir a escalas muito pequenas de espaço, de modo a que as evidências empíricas que indiciam a sua existência são difíceis de obter.

Originando a chamada «segunda revolução da teoria das supercordas» («superstring theories second revolution»), Edward Witten sugeriu, 1995, num seminário na USC («University of Southern California») que as teorias das cordas, com dez dimensões, até então formalizadas, poderiam ser unificadas por uma nova teoria com onze dimensões a que chamou «teoria M» (M-theory). Da teoria M apenas se conhecem algumas das propriedades formais que deve cumprir pensando-se que pode constituir uma alternativa para a unificação desejada, nomeadamente, através do princípio holográfico.

Para além das teorias das cordas, existe como alternativa para a unificação entre relatividade generalizada e mecânica quântica, a teoria da gravidade de loops quânticos («quantum loop gravity»). As teorias de cordas são formuladas assumindo uma estrutura geométrica espaço-temporal onde cordas e d-branes existem, produzindo efeitos sobre o estado dessa geometria. Por esse motivo, as teorias em questão dizem-se «dependentes do contexto espaço-temporal» («background dependent»).

Na teoria da gravidade de loops quânticos considera-se que um modelo unificando, satisfatoriamente, gravidade e fenómenos quânticos deverá assegurar, tal como sucede na teoria da relatividade generalizada, que aquela dependência não existe.

A teoria diz-se então «independente do contexto espaço-temporal» («background independent»), sendo capaz de produzir, para além de todos os campos quânticos e das excitações sobre eles, originando as partículas, as próprias geometrias espaço-temporais onde os eventos são coordenáveis.

A teoria da gravidade de loops quânticos foi desenvolvida pelo físico norte-americano Lee Smolin [1955 - ] (Smolin, 2004) e pelo físico italiano Carlo Rovelli [1956 - ] (Rovelli, 2004). Consiste num tratamento quântico do espaço-tempo em que as interações entre eventos se representam num espaço abstracto, recorrendo a relações geométricas entre conjuntos de linhas fechadas, denominadas loops, cuja extensão é sempre um múltiplo inteiro de uma quantidade mínima de referência.

Smolin utilizou estes loops para descrever as geometrias espaço-temporais no conjunto das soluções produzidas por uma teoria com propriedades quânticas, na qual as relações topológicas entre os loops estão associadas à curvatura espaço-temporal e, conseqüentemente, aos efeitos da gravidade (Smolin, 2001b).

Na teoria da gravidade de loops quânticos, a área e o volume são grandezas

físicas quantificadas, para as quais existem valores mínimos mensuráveis, respectivamente; a área de Planck,  $l_p^2$ , e o volume de Planck,  $l_p^3$ , em que  $l_p$  é o comprimento de Planck.

Estão, conseqüentemente, desde logo estabelecidas as condições propícias à introdução quer de um limite entrópico, quer do Princípio Holográfico.

Utilizando uma técnica matemática chamada «integração por caminhos» («path integral») sobre conjuntos de loops, representando, desse modo, histórias espaço-temporais, e recorrendo à representação discretizada de Roger Penrose do espaço-tempo chamada rede de spins («spin network»), Smolin construiu um espaço abstracto, evoluindo dinamicamente, através de operadores quânticos, gerando assim geometrias espaço-temporais. Uma rede de spins é um conjunto de vértices ligados por arestas tal que o número de arestas convergindo num dado vértice corresponde ao número de áreas de Planck, formando a área de uma superfície.

O número de vértices da rede mede, por seu lado, o número de volumes de Planck formando o volume de uma região na geometria espaço-temporal considerada. A evolução dinâmica da rede de spins e os estados sucessivos obtidos representam as transições entre os estados observáveis, associados a eventos numa geometria espaço-temporal possível.

A rede descreve, não um conjunto estático de eventos num espaço-tempo, mas sequências possíveis de processos. Apenas as modificações das relações entre os estados obtidos na rede é fisicamente significativa.

Smolin identificou problemas teóricos para todos os limites entrópicos, tal como foram geometricamente definidos, em consequência da sua formulação numa teoria dependente do contexto espaço-temporal. A assumpção de Smolin de que a realidade deve ser representada por um conjunto de transições entre estados levou-o, por outro lado, a considerar que uma rede de spins deveria ser entendida como uma estrutura onde estão representados processos de transferência de informação física, causalmente relacionados.

Interpretando a proposta de 't Hooft e Susskind como uma versão forte do Princípio Holográfico, Smolin e a física teórica grega Fotini Markopoulou-Kalamara [1971 - ] sugeriram uma versão mais lábil do mesmo (Smolin, 2001a, 2001b).

Chamaram «ecrã» a uma superfície bidimensional, não necessariamente fechada, correspondendo a um conjunto de eventos, na qual é possível efectivar observações no âmbito de uma teoria cosmológica holográfica (Smolin, 2001a).

Convencionou-se também que a área de um ecrã representa é uma medida da capacidade de transferência da informação física através do espaço.

Incluindo estes ecrãs na rede de spins, originalmente utilizada, e definindo um conjunto de regras para a evolução da mesma, preservando a causalidade no decurso da transferência de informação física, obtém-se uma rede de ecrãs cujas áreas determinam a

quantidade de informação que pode ser transferida.

A versão fraca do Princípio Holográfico (Smolin, 2001a, p. 41-43) estabelece então que todos os fenómenos físicos resultam de transferências causais de informação do passado para o futuro de um ecrã, em que:

a) A direcção temporal é definida em relação a um sujeito cognoscente que observa eventos nesse ecrã;

b) As transferências em causa são inteligíveis através da observação de eventos num ecrã, associando esses eventos a grandezas observáveis representadas em espaços de Hilbert, relativos a um e outro lado do ecrã e, finalmente;

c) Existe uma relação de proporcionalidade entre a quantidade de informação física que pode ser transferida no espaço através de um ecrã e a área desse ecrã, sendo válida a seguinte expressão:

$$(8) \quad S = \frac{A}{4}$$

Em que  $S$  é a entropia termodinâmica total associada ao ecrã, representando a quantidade máxima de informação física ali presente, num dado instante, e  $A$  é a área do ecrã, medida em unidades de área de Planck.

O Princípio Holográfico, associado à transferência de informação no espaço-tempo permite relacionar a quantidade de informação sobre um sistema físico com uma área.

A entidade ontológica entre computação e realidade física parece ser uma hipótese plausível. Caracteriza-se quer pelas características computacionais que podem ser associadas aos comportamentos de sistemas físicos, quer pelos limites impostos à quantidade de informação que pode ser transferida ou mantida no espaço, na dependência de um limite superior entrópico, supostamente válido na realidade.

A nossa proposta incidirá sobretudo sobre a natureza das leis que regem os fenómenos físicos, que na metáfora da física digital, correspondem ao programa que simula o universo.

Com base no que acabámos de expôr encontramos fundamento para que esse “programa” ontológico tenha as características de um sistema formal matemático. Uma entidade abstracta, livre de interpretações, procedendo de forma determinista e mecânica, de acordo com regras fixas e em que os respectivos processos de prova correspondem a comportamentos na natureza.

## 4º Capítulo

### Sistemas formais e comportamentos na natureza.

#### 4.1. Propriedades da inteligência humana.

Nesta secção enunciamos sumariamente algumas das propriedades da inteligência humana, tendo em mente que as hipóteses forte ou fisicalista da inteligência artificial podem ser verdadeiras. Estas propriedades são enunciadas do ponto de vista de um observador externo em relação a um sistema físico. O objectivo dessa enunciação consiste em aproximar a actividade inteligente, enquanto fenómeno, do comportamento de um sistema físico, que procede de forma regular, de acordo com as mesmas leis que regem outros objectos na natureza. Partimos assim de um quadro definitório de inteligência humana, conservador e minimalista, de forma a identificar o comportamento inteligente com o de uma máquina complexa.

Propriedades da inteligência humana.

i) A inteligência humana tem um suporte físico biológico: cérebro, sistema nervoso e uma fisiologia controlável.

ii) É capaz de reconhecer factos no meio ambiente e de interagir com outros sistemas físicos.

iii) É dotada da capacidade de estabelecer relações entre factos e de as comunicar.

iv) É capaz de tomar decisões de forma a otimizar critérios ou de forma a atingir objectivos.

v) É dotada de memória, capacidade de aprendizagem e de planificação, construindo, com base em situações anteriores, procedimentos pragmatizáveis no meio ambiente.

vi) É capaz de monitorizar o estado do seu suporte físico e de se auto-regular de modo a cumprir critérios de estabilidade.

Estas propriedades são aquelas que, em princípio, entendemos ser suficientes para simular de forma adequada a inteligência humana. Consideramos que outras propriedades atribuíveis à inteligência humana devem ser entendidas como ilusões subjectivas ou qualias da experiência fenomenológica que, de alguma maneira, são produzidas ou causadas por uma ou mais das propriedades anteriores. Essas ilusões

subjectivas, antecedidas pela nomeação das propriedades que lhes dão origem, são as seguintes:

a) A propriedade iii) produz a ilusão de que existe uma capacidade significativa, praticamente ilimitada, associada a uma pretensa criatividade, e a possibilidade de estabelecer situações de empatia na comunicação com os outros. Como se o agente projectasse sobre a realidade as relações que entende, a seu bel-prazer.

b) As propriedades iv), v) e vi) produzem as ilusões subjectivas de que existem emoções e de que existe livre arbítrio. Entendido como a capacidade de escolha entre opções numa situação, independentemente, de qualquer critério pré-estabelecido, externo ou interno ao agente da escolha. Isto é, de forma a que a intencionalidade do sujeito, dirigida sobre determinado objecto da sua acção, possa ser interrompida a qualquer momento, por nenhum outro motivo que não seja a vontade do agente para o fazer. Acção ilusoriamente atribuível ao exercício de volição.

Estas são as propriedades e ilusões subjectivas associadas que um sistema físico deverá ser capaz de simular, cumprindo os requisitos das hipóteses forte e fisicalista da inteligência artificial.

#### **4.2. Sistemas formais.**

No que se segue, descreveremos a origem e natureza de um sistema formal, seguindo Kleene (1952) na obra «Introduction to Meta-Mathematics». É nossa intenção caracterizar esse tipo de estruturas matemáticas de forma a contextualizar adequadamente o modo como, hipoteticamente, relacionaremos processos de prova num sistema formal com comportamentos na natureza.

No final do século XIX a matemática tinha conhecido avanços significativos em direcção a uma sofisticação do cálculo com infinitudes e ao conhecimento da natureza dos conjuntos infinitos. Nomeadamente, com o desenvolvimento dos métodos de continuidade sobre os reais de Weierstrass (1854), os estudos sobre números reais de Dedekind (1879) e a teoria de conjuntos de Cantor (1884). Conhecendo estas teorias os princípios da sua consolidação, surgiram pouco depois inquietações sobre a natureza das infinitudes também paradoxos ou antinomias que se colocavam a partir da manipulação, aparentemente indevida, da noção de conjunto. O primeiro paradoxo a surgir foi o de Burali-Forti (1897), no contexto da teoria dos ordinais transfinitos, depois o paradoxo de Cantor (1899), em seguida, o paradoxo de Russell (1902), cuja explicitação ilustra o problema comum a todas as antinomias. A definição aparentemente clara de conjunto dada por Cantor é a de que se

trata de qualquer colecção  $M$  de objectos  $m$  definidos e diferenciáveis pela nossa percepção ou pensamento num todo. O que é dito não refere as possibilidades de definição e deixa livre que a mesma se possa fazer mediante a contemplação de um todo. Russell considera o conjunto  $T$  dos conjuntos que não são membros de si mesmos e pergunta se  $T$  é ou não membro de si mesmo. Se  $T$  fôr membro de si mesmo, não poderá estar no conjunto dado pela sua própria definição e, desse modo,  $T$  não pode ser membro de si mesmo. Se, por outro lado,  $T$  não fôr membro de si mesmo, cumpre o disposto na sua própria definição e deverá então, por essa definição, ser membro de si mesmo. O problema surge quando perante uma totalidade, objectos que se candidatam ou não a ser membros dela, são escolhidos com base apenas nas propriedades dessa totalidade, com a qual partilham características comuns. Tornando-se assim objectos de propriedades ditas impredicativas. Justamente, para evitar essas situações, criou Russell, em 1908-1910, a sua teoria ramificada dos tipos, separando as propriedades que podem ser atribuídas a objectos por níveis, numa hierarquia. A propriedade mais elementar, que considera apenas a existência de um objecto encontra-se no tipo mais baixo (tipo 0, os próprios objectos). Propriedades atribuíveis a cada objecto, no tipo imediatamente a seguir, propriedades das propriedades de um objecto, acima e assim por diante. Tipos acima do tipo 0, são ainda subdivididos em ordens. Propriedades definidas sem mencionarem totalidades situam-se na ordem 0, propriedades definidas mencionando a totalidade das propriedades de uma dada ordem pertencem à seguinte. Esta organização das propriedades evita o problema anterior, mas também coloca dificuldades à matemática clássica sempre que esta faz uso, aparentemente adequado, de definições impredicativas, como no caso da definição de número irracional de Dedekind, que utiliza sequências infinitas de racionais. Para resolver o problema Russell introduz o axioma da reducibilidade que estabelece, posto aqui de forma simplista, que para cada propriedade numa dada ordem existe outra propriedade, numa ordem mais baixa, equivalente à primeira propriedade.

Outras antinomia é a do paradoxo de Richard (1905), que lida com a noção de definibilidade finita de um objecto, como se compreende na expressão “the least natural number not nameable in fewer than twenty-two syllables”. Frase, com vinte e uma sílabas, que nomea o menor natural que não pode ser nomeado em menos de vinte e duas sílabas.

A emergência de paradoxos sugeria de alguma forma que a estrutura axiomática ou os processos lógico-dedutivos teriam algum problema fundamental e que a sua correcção eliminaria a ocorrência de antinomias ou outros problemas. As teorias que a isso se prestavam tinham sido desenhadas nas planícies intelectuais da matemática informal, fazendo uso do raciocínio lógico-dedutivo. Os matemáticos supunham lidar com sistemas de objectos, introduzidos de forma genética por definições que pretendiam estabelecer a

natureza completa desses objectos. Na mesma tradição dos Elementos de Euclides (305-275 a.c.). Os teoremas deduzidos eram considerados verdades sobre os sistemas, em vez de proposições hipoteticamente aplicáveis a qualquer sistema de objectos que satisfizesse os mesmos axiomas. A partir das antinomias estabeleceu-se então uma discussão sobre os fundamentos da matemática da qual resultaram três escolas de pensamento.

Os promotores do logicismo pretendiam que a matemática é um ramo da lógica e que noções matemáticas devem ser definidas em termos de noções lógicas. Os teoremas devem ser provados como teoremas da lógica. Frege (1884, 1893, 1903) define noções matemáticas em termos lógicos, Peano (1889, 1894-1908) expressa teoremas matemáticos recorrendo a simbolismo lógico e Peirce trata e estende vários temas, por exemplo, o de número (1881). A matemática como território da lógica é também o pano de fundo dos Principia Mathematica de Whitehead e Russell (1910-13).

A escola intuicionista surge com as críticas de Kronecker a Weierstrass, Dedekind e Cantor, por não fornecerem um meio de determinar se um objecto satisfaz as definições que lhe são apontadas. E depois substancia-se com as ideias de Brouwer que, em 1908, escreve um artigo com o título «The untrustworthiness of the principles of logic», desafiando a ideia original de Aristóteles de que as regras da lógica clássica têm validade absoluta, independentemente do objecto a que se aplicam. As críticas intuicionistas recaem sobretudo sobre os excessos das abstracções feitas a partir de conjuntos finitos, extrapoladas sem justificação para conjuntos infinitos de objectos. O princípio da indução exemplifica a exigência de construção que os intuicionistas defendem, por meios finitos, para chegar a qualquer objecto que se pretende provar. E de facto, é requerimento principal que sobre todo o objecto matemático cuja existência se afirma, seja necessário fornecer os meios finitistas para a sua construção. Mesmo que isso envolva a repetição infinita dos mesmos passos. Na matemática clássica, assim demarcada por ser anterior às críticas intuicionistas, o infinito é tratado como uma entidade actual, completa, extensa e existencial. E mesmo o logicismo ainda trata a existência de números naturais como uma hipótese sobre o mundo. A escola intuicionista, pelo contrário, afirma que não há evidência que suporte a existência de números naturais, até que se introduza um procedimento finitista para os construir. O infinito é visto, assim, como uma entidade potencial, em devir, no âmbito de uma actividade construtiva acessível ao pensamento. Uma arena possível, por desbravar, para a experiência mental da sua construção. Números naturais são entidades mais elementares que as noções de número cardinal e respectivas propriedades usadas pelos logicistas para as caracterizarem. A expressão « $(\forall x)A(x)$ » é para ser interpretada como juízo hipotético, *i.e.*, como uma asserção acerca de cada um dos numerais sob consideração. Este juízo é de facto a articulação de uma lei ou princípio geral que pode efectivamente ser verificado

para cada caso individualmente. (Lourenço, 2004, p. 23). «Uma proposição da forma  $(\exists x)A(x)$  é para ser interpretada como um juízo parcial, isto é, como uma parte incompleta de uma proposição mais rigorosamente determinada e completamente enunciada. Esta determinação pode consistir ou na imediata apresentação de um numeral  $x$  tal que  $A(x)$ , ou na apresentação de um processo que permita a efectiva construção de um numeral  $x$  tal que  $A(x)$ . Requer-se ainda, de harmonia com a exigência de efectividade essencial dos processos a utilizar, que na apresentação de um processo que permita a construção de um  $x$  tal que  $A(x)$ , o número de passos tenha que ser menor ou igual a um dado inteiro  $k$ .» (*ibidem*).

A escola formalista vai fazer uso das exigências finitistas intuicionistas. Protagonizada por Hilbert e depois desenvolvida por ele e pelos seus colaboradores, Bernays, Ackermann e von Neumann.

Para resgatar a matemática clássica das críticas do intuicionismo e simultaneamente analisar a integridade dos seus resultados anteriores, Hilbert propõe, em 1904, um programa de investigação inovador. A matemática clássica será formulada como uma teoria axiomática formal e esta teoria será livre de contradições. Querendo-se dizer que num sistema formal  $A$  e  $\neg A$  não serão teoremas. Para garantir este requisito, Hilbert considerará que um esforço especial deve ser colocado na definição e análise dos processos de prova e que, idealmente, o próprio sistema deverá, por meios puramente inferenciais, produzir uma prova da sua própria consistência. O modo de construção deverá seguir o método formal ou axiomático existencial, distinto daquele usado na matemática clássica, denominado método informal ou axiomático material. O método formal concilia duas descobertas na história da matemática. A possibilidade de utilização de um método axiomático-dedutivo, atribuída a Pitágoras (sec. VI, a.c.) e o tratamento formal da lógica. A possibilidade de descrever o raciocínio dedutivo em frases pela sua forma, atribuída a Aristóteles (384 – 322). Consiste na enunciação de símbolos, indicação clara das suas combinações possíveis no sistema e da forma como essas combinações podem ser derivadas umas das outras, a partir de combinações de símbolos, consideradas axiomas. As combinações possíveis são conhecidas através das regras de formulação do sistema. As derivações, através das regras de derivação.

Os axiomas num sistema informal definem ou descrevem as propriedades de objectos que são considerados evidentes, a partir da constatação inicial das suas características. Pelo contrário, num sistema formal todas as propriedades que correspondem à forma, relações e comportamentos dos objectos ali descritos são definidas a partir de asserções colocadas á priori, independentes do conhecimento de quaisquer

factos externos ao sistema. A Hilbert se deve a ideia de que a formalização de um sistema de objectos envolve a total abstracção de significado. Está proibida qualquer interpretação que aduza propriedades aos objectos que sejam extemporâneas às resultantes dos axiomas e da aplicação das regras do sistema. Ao resultado da aplicação do método formal a um sistema de objectos matemáticos chama-se sistema formal, formalismo, teoria formal ou matemática formal. A preocupação de Hilbert em garantir a integridade das teorias matemáticas formalizadas conduz à meta-análise das propriedades e comportamentos do sistema formal, tomado como um todo. O sistema torna-se assim a teoria-objecto de um estudo matemático a que Hilbert chama metamatemática ou teoria da prova. Do intuicionismo retirarão os formalistas as respectivas consequência que determinam a necessidade de utilização de procedimentos finitistas, quer para dotar o sistema formal de capacidade de prova, quer para dar à metamatemática os utensílios para o analisar. Embora, assumindo uma perspectiva diferente, de forma a ultrapassar as objecções colocadas por Brouwer e outros às partes não finitistas da matemática:

“(...) o uso frequente do raciocínio não-finitista em teorias matemáticas [informais] faz com que Hilbert tenha que, nos sistemas formais que são supostos justificar estas teorias, introduzir regras de derivação que correspondem à parte não-finitista da inferência. Suponhamos agora que um sistema formal  $F$  representa uma teoria  $T$  com inferências não-finitistas, as quais serão por isso representadas em  $F$ . Para Hilbert esta situação não é paradoxal por o sistema  $F$  ele próprio ser construtivamente definido, e por isso ele próprio susceptível de tratamento finitista, visto que  $F$  é um conjunto de sucessões de fórmulas formadas a partir de regras. Nestas condições o programa finitista parece oferecer a possibilidade de legitimar o raciocínio não-finitista.

Para não dar a impressão que o finitismo e o intuicionismo de Brouwer são uma e a mesma coisa, apesar de terem em comum alguns pontos de doutrina (...), Brouwer permite o uso de considerações lógicas gerais, ainda que interpretadas de uma maneira mais restritiva do que na teoria clássica; como permite também o uso dos factos da experiência combinatória, os quais são o paradigma da percepção finitista. No intuicionismo domina a noção de que o objecto matemático é essencialmente uma *experiência mental*, a qual consiste na execução de uma demonstração, enquanto que no finitismo de Hilbert encontramos a noção de que o objecto matemático é produzido por uma experiência levada a efeito com objectos concretos, concebidos como formados por partes discretas e de cuja estrutura se pode ter uma percepção de conjunto. Assim é claro que o intuicionismo inclui o finitismo, uma vez que a imagem de um objecto concreto pode ser usada numa construção mental; mas excede o âmbito do finitismo ao

permitir asserções acerca de todas as construções possíveis, as quais não constituem uma totalidade em sentido finitista”.

(Lourenço, 2009, pp. 26 - 27)

Hilbert distinguirá afirmações com sentido intuitivo, que podem ser abstraídas de situações na realidade, daquelas que não têm sentido intuitivo, como a noção de infinito. De qualquer forma, parecia existir uma metáfora entre o comportamento de um sistema com propriedades lógicas e o comportamento de sistemas físicos.

“Hilbert 1928 says, "It is by no means reasonable to set up in general the requirement that each separate formula should be interpretable taken by itself ..." In theoretical physics "only certain combinations and consequences of the physical laws can be checked experimentally - likewise in my proof theory only the real statements are immediately capable of a verification".

A theory in classical mathematics can be regarded as a simple and elegant systematizing scheme, by which a variety of (presumably) true real statements, previously appearing as heterogeneous and unrelated, and often previously unknown, are comprised as consequences of the ideal theorems in the theory”.

(Cf von Neumann 1947, Einstein 1944 p. 288 *in* Kleene, 1952, p. 57)

Nessa metáfora, um sistema formal estaria para um sistema físico, como a sua metateoria estaria para uma teoria física:

“From the standpoint of the metatheory, the object theory is not properly a theory at all as we formerly understood the term, but a system of meaningless objects like the positions in a game of chess, subject to mechanical manipulations like the moves in chess. The object theory is described and studied as a system of symbols and of objects built up out of symbols. The symbols are regarded simply as various kinds of recognizable objects. To fix our ideas we may think of them concretely as marks on paper; or more accurately as abstracted from our experience with symbols as marks on paper. (Proof theory must be to some extent abstract, since it supposes arbitrarily long sequences of symbols to be constructible, although the quantity of paper and ink in the world is finite.) The other objects of the system are analysed only with regard to the the manner of their composition out of the symbols. By definition, this is all that a formal system system shall be as an object of study for meta-mathematics”.

The metatheory belongs to intuitive and informal mathematics (...) [and] will be expressed in ordinary language, with mathematical symbols, such as metamathematical variables, introduced according to need.

The assertions of the meta-theory must be understood. The deductions must carry conviction. They must proceed by intuitive inferences, and not, as the deductions in the formal theory, by applications of stated rules. Rules have been stated to formalize the object theory, but now we must understand without rules how those rules work. An intuitive mathematics is necessary even to define the formal mathematics.

(Kleene, 1952, p. 62)

As propriedades do sistema serão enunciadas no contexto da sua metateoria e o método da formalização axiomática envolve os seguintes passos, escritos de forma intuitiva nessa teoria informal sobre o sistema:

i) Todos os termos técnicos, incluindo todos os símbolos utilizados no sistema que importam para a dedução dos teoremas, devem ser apresentados, sendo possível tratá-los como palavras sem significado ao longo de uma dedução. Aos termos técnicos indefinidos antes da introdução de outras definições, chamam-se noções indefinidas ou primitivas.

ii) Em seguida, define-se de que modo é que os termos podem ser concatenados numa sequência deles, chamada fórmula. Este passo tem como objectivo explicitar as condições que determinam que sequências de termos são ou não fórmulas no sistema. Que sequências de termos técnicos podem existir no sistema. As especificações que permitem identificar se uma sequência de termos é uma fórmula bem construída no sistema são chamadas regras de formação do sistema.

iii) O terceiro passo na formalização do sistema consiste na especificação das condições que permitem derivar uma fórmula a partir de outra, num processo de dedução ou de prova. Essas definições metamatemáticas dizem-se as regras de dedução ou de transformação no sistema e conferem-lhe a estrutura de uma teoria dedutiva. Para o metamatemático essas regras não são asserções sobre os objectos que regem, pois não têm uma interpretação ou significado específico associada. Mais exactamente, não dizem respeito a factos que possam ser identificados na realidade nem tão pouco a afirmações auto-evidentes. São apenas regras que regem os processos de prova no sistema formal. São escritas na metateoria do sistema formal como postulados, sob a forma de axiomas (fórmulas fixas no sistema), esquemas axiomáticos, cuja substituição por axiomas origina novos axiomas e regras de inferência, esquemas de dedução que indicam quais são as concatenações válidas de fórmulas no sistema, a partir das quais é possível deduzir, como consequência directa das primeiras, chamadas premissas, outra fórmula, chamada teorema.

Diz-se agora que o sistema se encontra formalizado e é tarefa do metamatemático estudar as propriedades do sistema formal como um sistema mecânico em que é possível

justapôr símbolos e sequências destes, de acordo com regras, sem qualquer conteúdo significativo atribuído às fórmulas resultantes. Uma fórmula ou sequência delas (sendo um axioma ou um teorema) pode ser invocado para instanciar um esquema axiomático, sem que isso corresponda à efectivação de uma escolha ou ao cumprimento de qualquer propósito no sistema. Consistindo apenas um caso possível na combinatória de processos de prova autorizados pelas regras do sistema. O sistema formal é uma estrutura lógica cujos resultados se obtêm estritamente por processos deterministas. Todas as fórmulas que podem existir num sistema formal dependem inteiramente dos termos primitivos, regras de formação e regras de derivação.

Existirão símbolos formais, constituindo a primeira das categorias de objectos possíveis. Existirão sequências finitas de ocorrências desses símbolos, isto é; fórmulas ou conjuntos delas, obtidas a partir de esquemas axiomáticos, expressões formais constituindo a segunda categoria de objectos. E, finalmente, existirão sequências finitas de expressões formais, constituindo a última e terceira categoria de objectos. Uma prova formal é, assim, uma sequência finita de ocorrências de fórmulas, em que cada uma na sequência ou é um axioma ou a consequência imediata das fórmulas que a precedem na sequência.

A prova é dita sê-lo da última fórmula na sequência e esta diz-se ter sido formalmente provada ou ser um teorema formal (Kleene, 2009, pag. 83).

De forma sintética enunciemos algumas das características de interesse de um sistema formal:

i) O sistema formal utiliza procedimentos finitistas, envolvendo apenas objectos intuitivamente concebíveis para produzir processos de prova. Assim como a teoria metamatemática que o estuda. No entanto, no contexto desta, é possível utilizar os mesmos processos para investigar entidades matemáticas não intuitivas, como a noção de totalidade infinita, fornecendo procedimentos finitistas que se podem repetir, de forma ideal, indefinidamente.

ii) Nenhuma classe de objectos matemáticos será, por isso, considerada como um todo completo, até que seja fornecido um procedimento finitista, por exemplo, de contagem, que permita justificar essa asserção. Provas de existência serão dadas, fornecendo um método, pelo menos implicitamente, para a construção de um objecto cuja existência se afirma.

iii) A formalização de uma teoria é um método que a reduz a regra e forma. Qualquer interpretação que lhe seja atribuída, é extemporânea ao seu funcionamento lógico e usada apenas para facilitar a análise metamatemática do sistema.

iv) Não existem ambiguidades sobre o que constitui uma fórmula ou uma prova no sistema.

v) Questões sobre os processos lógicos no sistema, em especial, se conduzem a contradições, devem ser investigados na sua metateoria.

As características de um sistema formal parecem viabilizar uma interpretação da realidade, em que aos comportamentos de sistemas físicos na natureza correspondem comportamentos lógicos no sistema formal. Os dois tipos de comportamentos são regidos por regras fixas, de forma determinista. Como também o são os comportamentos de sistemas físicos que mimetizam comportamentos inteligentes, como pretendem as teses forte e fisicalista da inteligência artificial. Na próxima secção daremos corpo a essa interpretação, cuja enunciação intuitiva consiste em dizer que existe um sistema formal que rege comportamentos na natureza.

#### **4.3. Processos de prova num sistema formal e comportamentos na natureza.**

A tese forte da inteligência artificial, formulada por Searle, é a de que o cérebro é apenas um computador e a mente um programa de computador. A ideia de que a manipulação de símbolos, segundo regras fixas, implementadas numa máquina pode conduzir à produção de comportamentos semelhantes aos de um agente humano inteligente. A tese fisicalista da inteligência artificial advoga que uma utilização dos materiais adequados e das suas propriedades físicas viabiliza essa simulação, já que cérebro, sistema nervoso e um corpo fisiologicamente controlável são sistemas físicos. Esta ideia tem outra associada, que parece substanciá-la. A de que os objectos na natureza procedem de forma regular, formando padrões, de acordo com proposições aparentemente fixas, evocáveis a partir da observação e intuitivamente chamadas leis da natureza. No decurso desses comportamentos não deverá existir qualquer intencionalidade ou volição, propósito ou efectivação de escolhas. Neste sentido, torna-se plausível aceitar a metáfora de que o universo é um computador que processa comportamentos de acordo com um programa.

De forma a pragmatizar a visão mecanicista da inteligência humana, assumida nestas teses, o objectivo desta secção é estender essa metáfora, considerando que existe um sistema formal hipotético que rege o comportamento dos objectos na natureza. Esse sistema pede-se que seja considerado uma entidade abstracta ontológica, assumindo-se

que a cada comportamento observável na natureza corresponde um processo de prova no sistema. Assume-se também que símbolos, que representam objectos na natureza, podem ser utilizados para produzir processos de prova no sistema que não correspondem a comportamentos observáveis na natureza, mas que correspondem a processos de prova ideais, abstractos, numa teoria matemática. Essa teoria envolve, por exemplo, o conceito de variável e, embora possam existir comportamentos físicos que manipulam um objecto simbolizando uma variável, esse objecto existe durante o processo físico a que corresponde a manipulação. Ao passo que o significado da variável no sistema, ao longo do mesmo processo, é o de que representa a possibilidade de instanciação de um objecto num conjunto deles, sem que esse objecto tenha sido instanciado. Os dois significados atribuídos ao conceito de variável parecem incompatíveis entre si e incompatíveis com a ideia de que processo matemático ideal e processo físico que o implementa possam ser representados pelo mesmo processo lógico no sistema. O sistema formal, que se coloca como hipótese existir, é suposto formalizar comportamentos físicos, mas também ideias matemáticas abstractas, no pleno sentido do termo.

As fórmulas que compõem processos de prova, que correspondem a comportamentos observáveis, estão sujeitas a restrições no que diz respeito à sua forma. Essas restrições têm, fundamentalmente, como objectivo assegurar que símbolos, que representam objectos observáveis na natureza, permanecem instanciados nas respectivas fórmulas, enquanto esses objectos existirem no decurso dos comportamentos. No mesmo espírito do caso do conceito de variável, não fará também sentido ter uma fórmula sobre comportamentos na natureza onde esteja escrita uma disjunção lógica entre símbolos que representam, visto que o significado intuitivo dessa fórmula seria o de que num comportamento poderiam existir vários objectos, sem que nenhum existisse ainda. No âmbito da mecânica quântica isso poderá parecer viável. Porém sucede que antes e depois do colapso do vector de onda existem sempre e só os mesmos objectos ao longo de todo o comportamento. Nomeadamente, as partículas quânticas que são posteriormente observadas e cujos comportamentos podem ser o de ondas ou de corpúsculos, conforme as condições de observação.

O conjunto de postulados que formam as regras de transformação que regem logicamente os processos de prova, correspondendo a comportamentos observados na natureza, não será explicitado, apenas suposto existir e ser adequado para esse fim. O sistema formal tem uma estrutura lógica determinista, portanto, não aleatória. Suporemos que as suas regras de transformação são tais que permitem a ocorrência de processos de prova similares a situações de aleatoriedade, por meios deterministas, como sucede nos

comportamentos caóticos, ou que essas situações podem ser obtidas a partir da expansão decimal de números irracionais.

As propriedades genéricas do sistema informal hipostasiado são propostas, de forma intuitiva, com significados associados, no contexto da metateoria do sistema. Nessa metateoria autorizam-se referências à realidade observável. As mesmas propriedades são supostas serem pragmatizáveis num sistema operável, por intermédio de um ou mais cálculos lógicos e por um cálculo numérico, embora a explicitação da forma como esses cálculos se processam não se faça aqui. São, portanto, propriedades tão hipotéticas quanto o sistema formal que caracterizam e não esgotam o conjunto total de propriedades que semelhante sistema, caso exista, deverá ter. São propriedades enunciadas de forma a estabelecer um grau mínimo de plausibilidade para a existência desse sistema e de forma a salientar algumas das suas características que serão úteis para a nossa argumentação, mais à frente. O sistema apresentado deverá ser entendido como um esboço possível para um sistema formal que, por hipótese, é capaz de reger formalmente comportamentos na natureza. A caracterização desse sistema não corresponde inteiramente a um sistema formal no sentido habitual (como em Kleene, 1952). Para além dos comportamentos na natureza, o sistema é suposto formalizar a generalidade das teorias matemáticas, estendendo agora as ideias de Hilbert. Por outro lado, quando se fala em análise metamatemática, esta não tem necessariamente de confinar a uma teoria da prova, mas a todas as propriedades do sistema. Um agente metamatemático será entendido como um sistema físico que produz análises metamatemáticas sobre o sistema e que é capaz de deduzir teoremas matemáticos, de modo aparentemente intuitivo e informal, aos olhos de um segundo agente. Esses comportamentos que, de maneira ortodoxa, se entendem ser só ilusoriamente inteligentes (no sentido do que foi disposto na secção 2.1.) correspondem a processos de prova que regem o comportamento físico do agente. Esses processos não têm que corresponder aos processos intelectuais que o agente julga ter. No entanto, definem uma maneira alternativa de provar no sistema o que o agente julga que está a provar, quando este considera os seus argumentos sãos, isto é; livres de objecções matemáticas. Deste modo, o sistema consegue provar alguns dos seus teoremas ideais, isto é, dizendo respeito a entidades abstractas que não existem na natureza, através dos comportamentos de alguns sistemas físicos.

As propriedades do sistema formal considerado são as seguintes:

1. Existe um sistema formal que rege o comportamento de um conjunto finito de objectos diferenciáveis na natureza que podem ser observados e contados.

2. Cada objecto diferenciável na natureza é representado por um só símbolo, atribuído de forma arbitrária, que no formalismo do sistema formal se pode repetir um número infinito de vezes, sem que isso tenha de corresponder a um fenómeno na natureza.

3. Existem regras que definem de que forma é que todos os símbolos no sistema formal podem ser justapostos ou concatenados, dando origem a sequências de símbolos chamadas expressões formais. As regras que definem os modos de concatenação dos símbolos chamam-se regras de formação.

4. Algumas das expressões formais obtidas através da concatenação adequada de símbolos podem ser relacionadas com comportamentos observáveis na natureza, enquanto que outras dizem estritamente respeito a objectos e processos lógicos numa teoria matemática. Às regras de formação que determinam os modos distintivos de concatenação de símbolos que dão origem a expressões formais, relacionáveis com comportamentos observáveis na natureza, chamamos regras de formação ontológicas. Essas regras estão enunciadas em 5.1., quando se fala de símbolos variáveis, e em 5.4., quando se fala de símbolos lógicos.

5. Existem tipos diferentes de símbolos no sistema formal. Esses símbolos estão organizados do seguinte modo:

5.1. Símbolos individuais, que podem ser: i) símbolos que se chamam números e que podem representar objectos diferenciáveis na natureza, ii) símbolos que permitem delimitar conjuntos de símbolos (por exemplo, parentesis), símbolos que são variáveis e que podem ser substituídos por números. Entendendo-se que qualquer expressão formal que contenha um símbolo variável não pode ser relacionada com comportamentos na natureza. Exemplos de símbolos que são números são o "0" (zero) e todos os naturais. Estes símbolos são considerados noções primitivas do sistema.

5.2. Símbolos funcionais, que na metateoria do sistema formal, uma vez concatenados de forma adequada com símbolos que representam objectos diferenciáveis, representam processos no sistema formal, através dos quais, símbolos representando objectos diferenciáveis dão origem a outros símbolos representando objectos diferenciáveis. E que uma vez concatenados com números e símbolos variáveis dão origem a outros números, sempre que o símbolo variável seja substituído por um número. Na teoria matemática formal, formalizada pelo sistema, esses símbolos poderão ser "+" (mais), "." (vezes) e "“ “

(sucessor). Às expressões formais que resultam das concatenações de símbolos que representam objectos diferenciáveis e símbolos funcionais ou que resultam da concatenação de números, símbolos variáveis e símbolos funcionais chamamos expressões funcionais ou termos.

5.3. Símbolos predicativos, que na metateoria do sistema formal, uma vez concatenados de forma adequada com expressões funcionais, com símbolos que representam objectos diferenciáveis e com variáveis, representam os processos referidos em 5.2. e o respectivo resultado. Na teoria matemática formal, formalizada pelo sistema, um desses símbolos poderá ser " $=$ " (igual). Às expressões formais que resultam das concatenações de expressões funcionais e símbolos predicativos chamamos predicções simples.

5.4. Símbolos lógicos ou operadores lógicos que se diferenciam em três tipos:

5.4.1 Símbolos lógicos de contexto, que na metateoria do sistema formal, uma vez concatenados de forma adequada com predicções simples, representam processos no sistema formal, através dos quais, predicções simples dão origem a concatenações dessas expressões, cuja interpretação na metateoria do sistema assume quatro casos diferentes:

5.4.1.1. O de alternativa, em que se assume a ocorrência possível de uma de várias predicções simples possíveis. E para o qual se entende que as predicções simples concatenadas deste modo não podem ser relacionadas com comportamentos observáveis na natureza. Um exemplo na teoria matemática formalizada pelo sistema seria " $\vee$ " (ou).

5.4.1.2. O de sincronidade, em que se assume a ocorrência de todas as predicções simples num conjunto finito delas. E para o qual se entende que as predicções simples concatenadas deste modo podem ser relacionadas com comportamentos observáveis na natureza. Um exemplo na teoria matemática formalizada pelo sistema seria " $\wedge$ " (e).

5.4.1.3. O de causalidade, em que se assume que a ocorrência de predicções simples determina sempre que ocorra outra, embora o recíproco possa não ser válido. E para o qual se entende que predicções simples concatenadas deste modo podem ser relacionadas com comportamentos observáveis na natureza. Um exemplo na teoria matemática formalizada pelo sistema seria " $\supset$ " (implica).

5.4.1.4.) O da negação, em que se assume que uma dada predicção simples não ocorre. E

para o qual se entende que predicacões simples concatenadas deste modo não podem ser relacionadas com comportamentos observáveis na natureza. Um exemplo na teoria matemática formalizada pelo sistema seria " $\neg$ " (não).

5.4.2 Símbolos lógicos de quantificação, que na metateoria do sistema formal, uma vez concatenados de forma adequada com predicacões simples, representam no sistema uma forma económica de agrupar um conjunto finito ou infinito de símbolos representando objectos diferenciáveis ou de isolar um desses símbolos a partir de um daqueles conjuntos, e cuja interpretação na metateoria do sistema assume dois casos:

5.4.2.1. O caso da quantificação universal, em que se assume que todas as predicacões simples num conjunto delas podem ocorrer, através da representação de um símbolo variável na predicacão simples. E para o qual se entende que predicacões simples concatenadas deste modo não podem ser relacionadas com comportamentos observáveis na natureza. Um exemplo na teoria matemática formalizada pelo sistema seria " $\forall$ " (para todos).

5.4.2.2. O caso da quantificação existencial, em que se assume que uma predicacão simples de entre um conjunto delas pode ocorrer, através da representação de um símbolo variável na predicacão simples. E para o qual se entende que as predicacões simples concatenadas deste modo não podem ser relacionadas com comportamentos observáveis na natureza. Um exemplo na teoria matemática formalizada pelo sistema seria " $\exists$ " (existe).

5.4.3. Outros símbolos lógicos necessários para concatenar predicacões simples que podem ser relacionadas com comportamentos na natureza.

7. Às expressões formais obtidas a partir de concatenações de predicacões simples com os símbolos lógicos chamamos predicacões compostas.

8. É permitido representar uma predicacão simples por uma letra maiúscula, chamada símbolo de proposição.

9. Os símbolos lógicos referidos em 5.4. aplicam-se de forma idêntica a predicacões compostas.

10. Às predicacões simples e compostas chamam-se, genericamente, fórmulas.

11. No sistema formal existem regras de derivação ou de transformação que determinam o modo como as fórmulas se podem concatenar numa sequência de fórmulas. Poderão existir regras de derivação que se aplicam estritamente a sequências de fórmulas que correspondem a comportamentos na natureza. As regras de derivação podem ser de diferentes tipos:

11.1. Axiomas, que são fórmulas no sistema formal. Um exemplo na teoria matemática formalizada pelo sistema seria " $\neg a' = 0$ ", em que  $a$  é um símbolo que representa um objecto diferenciável na natureza.

11.2. Esquemas axiomáticos, que são expressões metamatemáticas que fixam certas relações lógicas entre fórmulas. Um exemplo na teoria matemática formalizada pelo sistema seria  $B \supset AVB$ , em que pela substituição de  $A$  e  $B$  por predicções simples quaisquer se obteria outro axioma no sistema. Um esquema metamatemático permite assim obter na teoria matemática formalizada no sistema uma classe infinita de axiomas.

11.3. Regras de inferência, que formalizam as operações de predicções compostas a partir de outras predicções compostas, permitindo a concatenação de fórmulas numa sequência em que as fórmulas consideradas, antes da primeira aplicação de uma regra de inferência, são axiomas. As fórmulas que se seguem aos axiomas chamam-se teoremas. À sequência de fórmulas até ao teorema que imediatamente precede o último chama-se processo de prova do último teorema e diz-se que este foi provado. Um exemplo de uma regra de inferência na teoria matemática formalizada pelo sistema seria  $\frac{A, A \supset B}{B}$ , em que a fórmula  $B$  sob o traço pode ser derivada das fórmulas  $A$  e  $A \supset B$  acima do traço.

11.4. Outros postulados, que são afirmações gerais distintas das anteriores, consideradas válidas, sobre o sistema formal.

12. O sistema formal diz-se consistente simples, se para nenhuma predicção composta fôr possível provar essa fórmula e a fórmula resultante da sua negação por aplicação do operador lógico " $\neg$ " (não). Isto é, se não for possível provar  $A$  e também  $\neg A$ .

12.1. Um sistema diz-se inconsistente simples, se toda a predicção composta  $B$  escrita no sistema puder ser provada, sucedendo que se se fôr possível provar  $A$  e também  $\neg A$ ,

então também será possível provar toda a fórmula  $B$  escrita no sistema.

13. O sistema formaliza a generalidade das teorias matemáticas e, em particular, uma teoria sobre números naturais.

14. Ao conjunto de processos de prova matemáticos a partir dos quais é possível derivar, sucessivamente, todos os números naturais chama-se um conjunto de contagem. O conjunto de contagem é infinito e denumerável. Cada fórmula derivada no conjunto de contagem declara que um número natural é sucessor de outro, dizendo-se que o natural que é sucessor do anterior que foi derivado. Ao conjunto dos números assim obtidos chama-se o conjunto dos números naturais.

14.1. No sistema existe um processo de prova matemático com o qual é possível provar um teorema, cujo significado metamatemático é o de que todo o número natural tem um desenvolvimento de base 2 e esse desenvolvimento é tal que cada um dos seus coeficientes é 0 ou 1. Entende-se que esta afirmação significa que o conjunto dos desenvolvimentos de base 2 de todos os números naturais, não sendo denumerável, é completo.

15. No sistema é possível provar as seguintes fórmulas: " $\neg(0 = 1)$ " e " $\neg(1 = 0)$ ".

16. No sistema estão definidas as seguintes operações com o operador aritmético de subtração "-": " $1-1$ " e " $1-0$ ". É possível provar as seguintes fórmulas: " $1-1=0$ " e " $1-0=1$ ".

17. A um objecto diferenciável na natureza, que é observável, chama-se estado. Estados na natureza podem ser contados.

18. A um conjunto de estados, que podem ser observados de forma a serem agrupados e relacionados entre si, chama-se uma configuração de estados e diz-se ser a configuração de estados de um sistema físico. Configurações de estados podem ser contadas.

19. A um conjunto de sistemas físicos que podem ser observados de forma a serem agrupados e relacionados entre si chama-se um ensemble de sistemas físicos. Ensembles de sistemas físicos podem ser contados.

20. Uma relação entre estados numa configuração de estados é representada no sistema formal por uma predicação simples, de acordo com as regras de formação ontológica.

21. O conjunto das relações entre as configurações de estados de sistemas físicos num ensemble de sistemas físicos é representada no sistema formal por uma predicação composta, de acordo com as regras de formação ontológica.

22. Um comportamento observável de um sistema físico na natureza é uma sequência formada pelas suas configurações de estados, observáveis no espaço e no tempo. Um sistema físico pode ter uma configuração de estados que se mantém a mesma.

23. Um comportamento observável de um ensemble de sistemas físicos é uma sequência formada por conjuntos de relações entre os sistemas físicos que compõem o ensemble, observáveis no espaço e no tempo. Um ensemble de sistemas físicos pode ter um conjunto de relações entre os sistemas físicos que o compõem que se mantém o mesmo.

24. Existem processos de prova que correspondem ou representam comportamentos na natureza, de acordo com 21., 22., 23. e 24. Nesses processos, a fórmula que resulta da concatenação de axiomas, cada fórmula entre aplicações sucessivas das regras de inferência do sistema e a fórmula que corresponde ao teorema provado, correspondem ou a uma configuração de estados de um sistema físico ou a um ensemble de sistemas físicos. A aplicação das regras de inferência do sistema, de que resulta a dedução de um teorema, a partir da fórmula imediatamente anterior, é entendida corresponder a uma mudança na configuração de estados de um sistema físico ou corresponder a uma mudança das relações entre sistemas físicos num ensemble deles.

25. Entende-se de forma intuitiva que as regras de inferência aplicadas a processos de prova que representam comportamentos na natureza correspondem a leis da natureza.

26. Processos de prova que não correspondem a comportamentos observáveis na natureza dizem-se processos de prova matemáticos e distinguem-se daqueles que correspondem a comportamentos, pela aplicação das regras de formação ontológica. Um exemplo de um processo de prova matemático seria aquele que deriva a existência de um número infinito de números primos.

27. Postulado da ligação.

É possível aplicar as regras de derivação do sistema para concatenar processos de prova, que correspondem a comportamentos na natureza, com processos de prova matemáticos. Ao resultado dessa concatenação chama-se processo de prova teórico-matemático.

28. Um procedimento de análise metamatemática sobre o sistema formal consiste num comportamento observável na natureza, em que um sistema físico procede de forma como se analisasse o sistema formal e deduzisse conclusões sobre ele, numa região de espaço finita e num intervalo de tempo finito.

29. Postulado da redução.

A todo o procedimento de análise metamatemática corresponde um processo de prova que representa comportamentos na natureza.

30. Procedimento metamatemático simples.

Existe um procedimento metamatemático que consiste em enunciar uma fórmula que foi provada no sistema. Diz-se um procedimento simples.

31. Redução do procedimento metamatemático simples.

Existe um processo de prova que representa comportamentos na natureza que corresponde à aplicação do procedimento simples, tal que o último teorema corresponde à derivação da fórmula enunciada no procedimento simples. Este processo de prova é distinto do processo de prova matemático pelo qual a prova foi derivada no sistema.

31.1. Para qualquer teorema derivado através de um processo de prova matemático, existe um procedimento metamatemático simples que o enuncia, tal que o processo de prova deste pode ser concatenado com o processo de prova do primeiro, de forma a resultar um processo de prova teórico-matemático, cujo último teorema é aquele enunciado.

32. Cálculo da formulação.

Existe um procedimento metamatemático, chamado cálculo da formulação, que se aplica sobre uma sequência  $\sigma_i$  delimitável de fórmulas concatenadas no sistema, do qual resulta uma operação de cálculo numérico que corresponde ao produto de  $n + 1$  factores, em que  $n$  é o número de símbolos na sequência de fórmulas. Esse produto é representado metamatemáticamente pelo símbolo  $\mathcal{F}_i$  e verifica o seguinte:

32.1. Se o primeiro símbolo da sequência puder ocorrer, de acordo com as regras de formação, o primeiro factor no produto é 1, caso contrário, é 0.

32.2. Se o símbolo que imediatamente se segue ao primeiro símbolo, na sequência, puder ocorrer, de acordo com as regras de formação do sistema, o segundo factor é 1, caso contrário, é 0.

32.3. Dados dois símbolos,  $S_i$  e  $S_{i+1}$  na sequência, em que  $S_i$  se segue imediatamente a  $S_{i+1}$ , se este puder ocorrer, de acordo com as regras de formação do sistema, o  $(i+1)$ -ésimo factor é 1, caso contrário é 0.

32.4. Se o último símbolo  $S_n$  na sequência puder ocorrer, de acordo com as regras de formação do sistema, o  $n$ -ésimo factor é 1, caso contrário é 0.

32.5. Se o número de parentesis à esquerda "(" for igual ao número de parentesis à direita ")", o  $(n+1)$ -ésimo factor é 1, caso contrário é 0.

32.5. Se o produto dos  $n + 1$  factores for igual a 1, diz-se que a sequência de fórmulas foi bem escrita no sistema. Se o produto dos  $n + 1$  factores for igual a zero, diz-se que existe, pelo menos, uma fórmula incorrectamente escrita no sistema.

33. Redução do cálculo da formulação.

Existe um processo de prova que representa comportamentos na natureza que corresponde ao cálculo da formulação, tal que o último teorema corresponde à derivação de dois predicados simples possíveis: " $0 + 1 = 1$ ", em que o símbolo "1" do lado esquerdo do sinal de igual representa o resultado do cálculo de formulação para  $\mathcal{F}_i = 1$  ou " $0 + 0 = 0$ ", em que o segundo símbolo "0" do lado esquerdo do sinal de igual representa o resultado do cálculo de formulação para  $\mathcal{F}_i = 0$ .

34. Procedimento de validação lógica.

Existe um procedimento metamatemático sobre o sistema, chamado procedimento de validação lógica, que se aplica a predicções compostas, quando as predicções simples que as compõem são substituídas por letras de proposição. Esse procedimento de validação

é tal que correlaciona os símbolos lógicos em 5.4.1., incluindo aqueles em 5.4.3., com tabelas de verdade. Nestas tabelas esses símbolos são utilizados para concatenar duas predicacões simples. A cada letra de proposição é atribuído um de dois números 0 ou 1, resultando para a concatenação de duas predicacões simples quatro casos possíveis para esses valores atribuídos: 00, 01, 10 e 11. A cada uma dessas hipóteses corresponde um valor 0 ou 1, de acordo com as tabelas de verdade dos respectivos símbolos lógicos envolvidos. Aos símbolos utilizados como argumentos nas tabelas de verdade só se exige que sejam distintos, podendo ser “0” (zero) e o natural “1”. O cálculo de validação procede sobre uma predicacão composta de forma a simplificá-la por passos sucessivos, obtendo-se o resultado dessa validação. O resultado só ser 0 ou 1. O procedimento de validação é finito.

34.1. Quando o resultado da aplicação do procedimento de validação a uma predicacão composta é tal que o seu resultado é 1, para todas as combinações resultantes da atribuição de 0 ou 1, a cada uma das letras de proposição que a compõem, a predicacão composta diz-se uma fórmula identicamente verdadeira.

34.2. É possível aplicar o procedimento de validação a todas as predicacões compostas, num processo de prova, substituindo todas as predicacões simples por letras de proposição.

34.3. Todas as predicacões compostas num processo de prova, incluindo as predicacões compostas que representam axiomas, são fórmulas identicamente verdadeiras.

34.4 Representa-se por  $T_j$  o resultado da aplicação do procedimento de validação a todas as fórmulas num processo de prova, para uma combinação  $j$ , resultante da atribuição de 0 ou 1 a cada um dos predicados compostos no processo de prova.

35. Chama-se troço de um processo de prova a um conjunto finito de teoremas e de axiomas nesse processo, tal que cada teorema resulta do(s) teorema(s) imediatamente anteriore(s) ou do(s) axioma(s) imediatamente anteriore(s). Troços de um processo de prova podem ser contados.

36. Cálculo de validação lógica.

Está definido um procedimento metamatemático sobre o sistema que utiliza uma aplicação

$\mathcal{L}_i : \mathbb{N} \rightarrow \{0,1\}$ , chamado cálculo de validação lógica, que ao  $i$ -ésimo troço de um processo

de prova faz corresponder o natural 0 ou 1. A aplicação está definida pela expressão

$\mathcal{L}_i = \prod_{j=1}^m T_j$ , em que  $m$  é o número máximo de combinações possíveis, resultantes da

atribuição de 0 ou 1 às letras de proposição nos predicados compostos do  $i$ -ésimo troço do processo de prova.

### 37. Redução do cálculo de validação lógica.

Existe um processo de prova que representa comportamentos na natureza que corresponde ao cálculo de validação lógica, tal que o último teorema corresponde à derivação de dois predicados simples possíveis: " $0 + 1 = 1$ ", em que o segundo símbolo "1" do lado esquerdo do sinal de igual representa o resultado do cálculo de formulação para  $\mathcal{L}_i = 1$  ou " $0 + 0 = 0$ ", em que o segundo símbolo "0" do lado esquerdo do sinal de igual representa o resultado do cálculo de formulação para  $\mathcal{L}_i = 0$ .

### 38. Procedimento nomotético.

Existe um procedimento metamatemático sobre o sistema com as seguintes características:

38.1. Aplica o cálculo da formulação a uma sequência de símbolos, que coincide com o  $i$ -ésimo troço de um processo de prova.

38.2. Aplica o procedimento de validação aos predicados compostos no mesmo troço.

38.3. Efectua o cálculo de validação lógica sobre o procedimento de validação efectuado, calculando o produto entre o resultado do cálculo de formulação e o resultado do cálculo de validação lógica.

Ao procedimento assim obtido chama-se procedimento nomotético, definido pela expressão

$\mathcal{N}_i = \mathcal{F}_i \cdot \mathcal{L}_i$ , chamada função nomotética, e cujo resultado é 0 ou 1.

### 39. Redução do procedimento nomotético.

Existe um processo de prova que representa comportamentos na natureza que corresponde à aplicação do procedimento nomotético, tal que o último teorema corresponde à derivação dois predicados simples possíveis: " $0 + 1 = 1$ ", em que o símbolo "1" do lado esquerdo do

sinal de igual representa o resultado da função nomotética para  $\mathcal{N}_i = 1$  ou " $0 + 0 = 0$ ", em que o segundo símbolo "0" do lado esquerdo do sinal de igual representa o resultado da função nomotética para  $\mathcal{N}_i = 0$ .

39.1. Se ao  $i$ -ésimo troço de um processo de prova, correspondendo a um procedimento metamatemático, tiver sido aplicado o procedimento nomotético de tal modo que ocorre  $\mathcal{N}_i = 1$  para esse troço, o procedimento metamatemático é válido, em condições idênticas de aplicação.

#### 40. Diagonalização.

Existe um procedimento metamatemático sobre o sistema que calcula sucessivamente os coeficientes, 0 ou 1, do desenvolvimento binário de cada número natural e utiliza o  $n$ -ésimo coeficiente, aplicando um procedimento aritmético, para calcular o  $n$ -ésimo coeficiente do desenvolvimento binário de outro número natural, chamado número diagonal. Ao procedimento assim obtido chama-se procedimento de diagonalização e ao número natural, cujo desenvolvimento binário se calcula, chama-se número diagonal. Os números naturais utilizados no procedimento podem estar ou não dispostos pela ordem habitual. Este procedimento é finito.

#### 41. Redução do procedimento de diagonalização.

Existe um processo de prova que representa comportamentos na natureza que corresponde ao procedimento de diagonalização, tal que o último teorema corresponde à derivação do predicado simples " $0 + 1 = 1$ ", se o  $n$ -ésimo coeficiente do número diagonal for 1, ou corresponde ao predicado simples " $0 + 0 = 0$ ", se o  $n$ -ésimo coeficiente do número diagonal for 0 (zero). O  $n$ -ésimo coeficiente do número diagonal é o último a ser calculado no procedimento.

#### 42. Procedimento de pressuposição.

Um procedimento de pressuposição é um procedimento de análise metamatemática, em que um sistema físico é observado proceder, como se considerasse que determinada fórmula pode ser provada. À fórmula considerada chama-se fórmula pressuposta. Este procedimento é finito.

#### 43. Redução do procedimento de pressuposição.

Se a fórmula pressuposta for derivável no sistema, o procedimento de pressuposição corresponde no sistema a um processo de prova que representa comportamentos na natureza, em que a última fórmula corresponde ou à fórmula pressuposta ou a uma fórmula derivável no sistema a partir da fórmula pressuposta.

43.1. Nas condições colocadas em 40., se uma fórmula derivada da fórmula pressuposta corresponder a uma fórmula que corresponde à negação de um teorema provado no sistema, o sistema é inconsistente simples.

44. Um procedimento metamatemático diz-se são, se não existirem objecções de natureza lógica em relação à exequibilidade metamatemática do mesmo. Dois ou mais procedimentos metamatemáticos são podem ser intuitivamente concatenados, desde que o resultado dessa concatenação seja um procedimento metamatemático são. Ao novo procedimento metamatemático chama-se um procedimento metamatemático composto.

#### 45. Redução de um procedimento metamatemático composto.

Um procedimento metamatemático composto corresponde no sistema a um processo de prova que representa comportamentos na natureza, em que o último teorema corresponde à última fórmula do processo de prova, que corresponde ao último procedimento metamatemático no procedimento composto.

#### 46. Postulado da verdade.

Num sistema formal que rege comportamentos na natureza, todos os procedimentos metamatemáticos são devem corresponder a processos de prova no sistema.

### **4.4. Aplicação do procedimento nomotético.**

No que segue apresentaremos um conjunto de acções concatenadas entre si, através das quais é, hipoteticamente, possível analisar um processo de prova no sistema. Aplicam-se, sucessivamente, procedimentos metamatemáticos, considerados são, com o objectivo de determinar as consequências lógicas no sistema resultantes da aplicação desses procedimentos. Entre os mesmos encontra-se o procedimento nomotético (propriedade 38). É suposto que a análise seja feita por um agente metamatemático, que

pode ser humano ou uma máquina “inteligente” e que, idealmente, tem acesso a toda a informação sobre o modo como as fórmulas são escritas e como o processo de prova decorre. Os procedimentos metamatemáticos, sendo comportamentos observáveis na natureza, atribuíveis a esse sistema, são supostos corresponderem a processos de prova no sistema formal. Os procedimentos metamatemáticos sendo sãos, considera-se terem sentido intuitivo lógico e estarem concatenados num procedimento metamatemático composto, construído de forma intuitivamente correcta (propriedade 44). Os processos de prova que lhes correspondem no sistema, é suposto estarem, por sua vez, concatenados, com processos de prova matemáticos (propriedade 27), cujos teoremas são afirmações abstractas ideais, que não dizem respeito a comportamentos na natureza.

A concatenação de todos os processos de prova envolvidos constitui assim um processo de prova teórico-matemático (propriedade 27), sobre o qual um agente metamatemático inteligente executa uma análise metamatemática. Quando se fala de “interpretação intuitiva atribuível ao agente” ou quando se afirma que o agente “assume” determinada pressuposição, não se autoriza ao agente qualquer capacidade para produzir significados sobre uma situação ou facto. As expressões deverão ser entendidas como um artifício de discurso, utilizado para descrever as circunstâncias em que a análise decorre. O agente metamatemático deverá ser entendido como um sistema físico, cujo comportamento é regido de forma determinista por um sistema formal especial, uma entidade ontológica e abstracta que se aceita existir, sem quaisquer propriedades que possam ser identificadas como ilusões subjectivas, conforme o disposto na secção 2.1.

A estrutura da sequência de análise metamatemática é, sumariamente, a seguinte.

i) O agente metamatemático enuncia três teoremas no sistema, que correspondem a proposições matemáticas abstractas, tomadas como premissas para o desenvolvimento da análise.

ii) Calcula os primeiros nove coeficientes do desenvolvimento de base 2 do número natural 3, através de um procedimento de diagonalização aplicado sobre o conjunto dos números naturais.

iii) Aplica o procedimento nomotético ao processo de prova no sistema, que corresponde ao que fez até aí. O procedimento permite determinar se as fórmulas, correspondentes aos procedimentos metamatemáticos aplicados se encontram bem escritas e se as regras de derivação foram correctamente aplicadas. Se assim fôr, todos os procedimentos metamatemáticos aplicados, antes do procedimento nomotético, serão sãos (propriedade 46).

iv) Com base nesta informação e usando o valor da função nomotética (propriedade 38), aplica um segundo procedimento de diagonalização para calcular os nove primeiros coeficientes do desenvolvimento de base 2 de um segundo número natural  $d2$ .

v) Após inspecção do resultado do cálculo anterior, o agente faz quatro pressuposições:

a) Pressupõe que todos os coeficientes do desenvolvimento de base 2 no número diagonal  $d2$  obtido, depois dos primeiros dois, são iguais a 1, para qualquer número de aplicações subsequentes do procedimento de diagonalização.

b) Que o desenvolvimento de base 2 desse número se encontra no conjunto dos desenvolvimentos de base 2, que por um dos teoremas iniciais corresponde a uma lista desses desenvolvimentos que é completa.

c) Que existe, pelo menos, uma reordenação da lista de desenvolvimentos de base 2 dos números naturais que dispõe o número diagonal  $d2$  nos primeiros  $n$  números naturais considerados no procedimento de diagonalização.

d) Que procedimento de diagonalização, uma vez aplicado a esse conjunto, é um procedimento são.

vi) O agente volta a aplicar o procedimento nomotético a todo o processo de prova correspondente às acções executadas até aí.

vii) A partir das pressuposições em v) e do resultado da última aplicação do procedimento nomotético, tenta calcular um dos coeficientes do número diagonal  $d2$  e, enunciando uma das premissas do processo de prova, encontra uma contradição.

A contradição permite concluir que o sistema formal hipostasiado é inconsistente simples (propriedade 12.1). As causas e significado dessa inconsistência serão discutidas em decurso, conduzindo ao argumento da variabilidade nomotética.

Explicitamos, agora, pormenorizadamente cada um dos passos da análise metamatemática sumariamente descrita.

i) O agente metamatemático começa por enunciar três teoremas no sistema que foram derivados através de processos de prova matemáticos. Como tal, esses teoremas não representam nem descrevem objectos na natureza. São afirmações matemáticas abstractas, não obstante deriváveis no sistema, uma vez que este formaliza a generalidade das teorias matemáticas (propriedade 13.). Os actos físicos de enunciação dos três teoremas correspondem a três procedimentos metamatemáticos simples (propriedade 30). Os

processos de prova que lhes correspondem podem, portanto, ser concatenados, respectivamente, com os respectivos teoremas (propriedade 31.1.). As três fórmulas resultantes dessas concatenações serão consideradas premissas, de acordo com as regras de inferência no sistema, no processo de prova teórico-matemático que se começa a desenhar. Os teoremas são os seguintes.

O primeiro teorema T1 afirma a existência do conjunto dos números naturais, conjunto que é infinito e denumerável (propriedade 14.). Consideraremos que na interpretação intuitiva atribuível ao agente metamatemático, essa afirmação significa que é, em princípio, possível repetir um procedimento finito tantas vezes, quantas as que se queira. Isto é, que é possível, de forma ideal, contar indefinidamente.

O segundo teorema T2 afirma que todo o número natural tem um desenvolvimento de base 2 e que esse desenvolvimento é tal que cada um dos seus coeficientes é 0 ou 1. Consideraremos que na interpretação intuitiva atribuível ao agente metamatemático, essa afirmação significa que existe uma lista abstracta, infinita e não denumerável de linhas, onde estão representados os coeficientes do desenvolvimento de base 2 de todos os naturais. A lista dir-se-á, por isso, completa (propriedade 14.1). A forma geral do desenvolvimento de base 2 de um número natural dada por:

$$(1) \quad n = \sum_{i=0}^m a_i 2^i$$

Em que o coeficiente  $a_i$  é 0 ou 1, para algum número  $m$  desses coeficientes.

O terceiro teorema T3 corresponde à fórmula, provada no sistema, " $\neg(0 = 1)$ " (propriedade 15).

ii) Depois da enunciação dos teoremas e da (aparente) assumpção dos respectivos significados, o agente adopta um procedimento de diagonalização  $D1$  sobre o primeiro número na lista assumida (propriedade 40). O procedimento aritmético que utiliza é o seguinte. Para cada número natural  $n$ , toma o  $n$ -ésimo coeficiente do desenvolvimento de base 2 de  $n$ . O resultado dessa operação corresponde ao valor do  $n$ -ésimo coeficiente do desenvolvimento de base 2 do número diagonal  $d1$ . O procedimento aritmético pode ser representado, informalmente, por:

$$(2) \quad a_{d1}(n) = a_n(n)$$

O procedimento de diagonalização pode ser visualizado na tabela seguinte, com os coeficientes usados identificados entre aspas:

	$2^0$	$2^1$	...	$2^j$	...	$2^n$
0	" $a_0(1)$ "	$a_0(2)$	...	$a_0(j)$	...	$a_0(n)$
1	$a_2(1)$	" $a_2(2)$ "	...	$a_2(j)$	...	$a_2(n)$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$		$\vdots$
$j$	$a_j(1)$	$a_j(2)$	...	" $a_j(j)$ "	...	$a_j(n)$
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\ddots$	$\vdots$
$n$	$a_n(1)$	$a_n(2)$	...	$a_n(j)$	...	" $a_n(n)$ "

A forma obtida do número diagonal, para  $n$  passos, dada por:

$$(3) \quad d1 = \sum_{i=1}^n a_i(i)2^{i-1}$$

Substituindo os coeficientes dos primeiros nove números naturais, resulta na tabela:

	$2^0$	$2^1$	$2^2$	$2^3$	$2^4$	$2^5$	$2^6$	$2^7$	$2^8$
1	1	0	0	0	0	0	0	0	0
2	0	1	0	0	0	0	0	0	0
3	1	1	0	0	0	0	0	0	0
4	0	0	1	0	0	0	0	0	0
5	1	0	1	0	0	0	0	0	0
6	0	1	1	0	0	0	0	0	0
7	1	1	1	0	0	0	0	0	0
8	0	0	0	1	0	0	0	0	0
9	1	0	0	1	0	0	0	0	0

O número diagonal, no caso, sendo dado por:

$$(4) \quad d1 = \sum_{i=1}^9 a_i 2^{i-1} = 1 \cdot 2^0 + 1 \cdot 2^1 = 3$$

Diferentes disposições das linhas de coeficientes na lista permitem calcular

diferentes números diagonais que correspondem, todos eles, a desenvolvimentos de base 2 de números naturais. Na disposição adoptada, o número, cujo desenvolvimento se calcula é o natural 3, a que chamaremos  $d1$ , bastando duas aplicações do procedimento de diagonalização para que esse número resulte definido, independentemente do número de passos seguintes no procedimento de diagonalização.

O agente metamatemático calcula os nove primeiros coeficientes do desenvolvimento de base 2 do número 3. Nesta fase, o último teorema no processo de prova (propriedade 41), representado de forma informal, é:

$$(5) \quad "0 + a_{d1}(3) = 0"$$

Uma vez que o terceiro coeficiente do desenvolvimento de base 2 do natural 3 é nulo.

iii) Depois do procedimento de diagonalização anterior, o agente metamatemático aplica o procedimento nomotético (propriedade 38) ao processo de prova correspondente a tudo o que fez até aí. Isto é, determina a correcção das fórmulas e das derivações dos teoremas no processo e calcula o valor da função nomotética  $\mathcal{N}_1$ . Se os procedimentos metamatemáticos aplicados forem sãos, devem corresponder a troços no processo de prova que estão correctamente formulados e concatenados (propriedade 46). Consequentemente, o valor da função nomotética deverá ser  $\mathcal{N}_1 = 1$ . Desse cálculo resulta, como último teorema no processo de prova (propriedade 39), o predicado simples, representado informalmente:

$$(6) \quad "0 + \mathcal{N}_1 = 1"$$

iv) Com base no resultado da função nomotética, o agente calcula os nove primeiros coeficientes do desenvolvimento de base 2 de um segundo número natural  $d2$ , utilizando um procedimento de diagonalização  $D2$ , distinto do primeiro, aplicado sobre o conjunto dos números naturais. O procedimento aritmético que utiliza é o seguinte. Para cada número natural  $n$ , toma o valor do cálculo nomotético anterior e subtrai-lhe o valor do  $n$ -ésimo coeficiente do desenvolvimento de base 2 de  $n$ . Este valor é igual ao  $n$ -ésimo coeficiente do desenvolvimento de base 2 do natural  $d1$ . O resultado da operação

aritmética corresponde ao valor do  $n$ -ésimo coeficiente do desenvolvimento de base 2 do segundo número diagonal  $d_2$ . O procedimento aritmético pode ser representado, informalmente, por:

$$(7) \quad a_{d_2}(n) = \mathcal{N}_1 - a_{d_1}(n)$$

Se as fórmulas analisadas no procedimento anterior, estiverem bem escritas no sistema e se as regras de derivação tiverem sido correctamente aplicadas, resultará  $\mathcal{N}_1 = 1$  e, portanto:

$$(8) \quad a_{d_2}(n) = 1 - a_n(n)$$

A expressão para o desenvolvimento de base 2 do número diagonal  $d_2$ , com um número arbitrário  $n$  de passos é dada por:

$$(9) \quad d_2 = \sum_{i=1}^n a_i 2^{i-1} = 0 \cdot 2^0 + 0 \cdot 2^1 + 1 \cdot 2^2 + \dots + 1 \cdot 2^n$$

Sucedendo que os coeficientes, depois dos dois primeiros, no desenvolvimento de base 2 de  $d_2$ , são todos iguais a 1.

A aplicação de um procedimento aritmético que deriva de informação metamatemática sobre o sistema não coloca problemas. O cálculo envolvendo a função nomotética é um cálculo aritmético simples, que presumivelmente depende directamente das propriedades estruturais do sistema, quando este produz processos de prova. Essas propriedades deverão ser suficientemente abrangentes para que o sistema possa produzir teoremas, que interpretados metateoricamente, correspondem a afirmações sobre o seu próprio modo de funcionamento, alguns dos quais resultam da aplicação do procedimento nomotético. Esta é a ideia fundamental que está na base dos teoremas de Gödel e, também, na base do programa de Hilbert, quando se fala de uma metateoria, entendida como uma teoria da prova. O sistema formal é um sistema de objectos representados por símbolos, cuja actividade se processa de forma indiferente às apreciações subjectivas e ilusórias dos agentes humanos. Supondo-se que o cálculo produzido pela função nomotética é apenas mais um, de vários possíveis.

O nono coeficiente do desenvolvimento de base 2 do diagonal  $d_2$  é dado por:

$$(10) \quad a_{d2}(9) = \mathcal{N}_1 - a_9(9) = 1 - 0 = 1$$

Do processo de prova que corresponde à aplicação do terceiro passo da diagonalização  $D2$  resulta um predicado simples no sistema (propriedade 41), representado de maneira informal:

$$(11) \quad "0 + a_{d2}(9) = 1"$$

v) O agente procede agora ao que, intuitivamente, se pode descrever como uma inspeção dos nove coeficientes do desenvolvimento de base 2 do número diagonal obtido  $d2$ . Verifica que os dois primeiros coeficientes são nulos e que todos os restantes são iguais a 1. A partir desse facto, comportando-se como se tivesse em “mente” o que é afirmado no teorema T1, o agente faz a sua primeira pressuposição, isto é, aplica pela primeira vez o procedimento metamatemático de pressuposição  $P1$  (propriedade 42). Pressupõe que todos os coeficientes, calculáveis depois dos dois primeiros coeficientes de  $d2$ , são iguais a 1, para qualquer número de aplicações subsequentes do procedimento de diagonalização. O resultado do processo de prova que corresponde a  $P1$  é alguma fórmula  $P2$  escrita no sistema, cujo significado metamatemático é o que foi indicado.

Em seguida, comportando-se como se tivesse agora em “mente” a forma geral de  $d2$  e o que é afirmado no teorema T2, aplica pela segunda vez um procedimento de pressuposição  $P2$  e assume que o desenvolvimento de base 2 no número natural  $d2$  se encontra no conjunto dos desenvolvimentos de base 2 de todos os números naturais. O resultado do processo de prova que corresponde a  $P2$  é alguma fórmula  $P2$  escrita no sistema, cujo significado metamatemático é o que foi indicado.

Aplica pela terceira vez um procedimento de pressuposição  $P3$  e assume que existe, pelo menos, um conjunto  $\mathcal{D}$  resultante da reordenação da lista de desenvolvimentos de base 2 dos números naturais, que dispõe o número diagonal  $d2$  entre os primeiros  $n$  números naturais, para além da segunda posição na lista, consideráveis no procedimento de diagonalização. Este conjunto, assim ordenado, é infinito, não denumerável e equipotente ao conjunto dos desenvolvimentos de base 2 dos números naturais, dispostos pela ordem habitual. Sendo completa a lista exibindo todos os

desenvolvimentos em  $\mathcal{D}$ . O resultado do processo de prova que corresponde a  $P3$  é alguma fórmula  $P3$  escrita no sistema, cujo significado metamatemático é o que foi indicado.

Finalmente, o agente aplica ainda, uma quarta e última vez, o procedimento de pressuposição  $P4$ , para assumir que é são um procedimento de diagonalização sobre o conjunto  $\mathcal{D}$ , pressuposto existir em  $P3$ . O resultado do processo de prova que corresponde a  $P4$  é alguma fórmula  $P4$  escrita no sistema, cujo significado metamatemático é o que foi indicado.

vi) Depois da aplicação dos procedimentos de pressuposição  $P1$ ,  $P2$ ,  $P3$  e  $P4$  o agente metamatemático volta a aplicar o procedimento nomotético a todo o processo de prova correspondente às acções executadas até aí, incluindo o troço do processo de prova que corresponde à primeira aplicação do procedimento nomotético e o troço do processo de prova correspondente aos quatro procedimentos metamatemáticos de pressuposição  $P1$ ,  $P2$ ,  $P3$  e  $P4$ .

Considerando-se todos os procedimentos efectuados são e estando os respectivos processos de prova, bem formulados e concatenados no processo de prova total, o resultado da função nomotética no segundo procedimento nomotético é  $\mathcal{N}_2 = 1$ . Desse cálculo resulta, nesta fase, como último teorema no processo (propriedade 39), o predicado simples, representado informalmente:

$$(12) \quad "0 + \mathcal{N}_2 = 1"$$

vii) Já perto do fim da sua análise metamatemática, o agente toma como premissas as fórmulas pressupostas em  $P1$ ,  $P2$ ,  $P3$  e  $P4$  e aplica um terceiro procedimento de diagonalização  $D3$ , idêntico a  $D2$ . Este procedimento utiliza o resultado da função nomotética,  $\mathcal{N}_2 = 1$ , obtido em vi). para calcular um  $n$ -ésimo coeficiente do número diagonal  $d2$ , a seguir ao segundo coeficiente. Pretendendo-se que o faça a partir do  $n$ -ésimo coeficiente do desenvolvimento de base 2 do próprio número. O número encontra-se na  $n$ -ésima posição na lista que corresponde ao conjunto  $\mathcal{D}$  dos desenvolvimentos de base 2 dos números naturais. Da aplicação do procedimento  $D3$  resulta a fórmula, representada informalmente:

$$(13) \quad "0 + a_{d2}(n) = 1"$$

Como no procedimento de diagonalização  $D3$  se verifica agora:

$$(14) \quad a_{d2}(n) = \mathcal{N}_2 - a_{d2}(n)$$

Da fórmula (13) é possível derivar a fórmula, representada informalmente:

$$(14) \quad "0 + \mathcal{N}_2 - a_{d2}(n) = 1"$$

O valor da função nomotética, para a segunda aplicação do procedimento nomotético, é  $\mathcal{N}_2 = 1$ , permitindo derivar da fórmula (14), a fórmula, representada informalmente:

$$(15) \quad "0 + 1 - a_{d2}(n) = 1"$$

Como pelo procedimento de pressuposição  $P1$  se assumiu que todos os coeficientes do desenvolvimento de base 2 de  $d2$  são iguais à unidade, da fórmula (15) deriva-se a fórmula:

$$(16) \quad "0 + 1 - 1 = 1"$$

Da qual se deriva, finalmente, a fórmula:

$$(15) \quad "0 = 1"$$

O agente metamatemático, enuncia então o teorema  $T3$ , que corresponde ao teorema no sistema,  $"\neg(0 = 1)"$  e conclui que o sistema formal é inconsistente simples (propriedade 12.1).

**4.5. Argumento da variabilidade nomotética para a existência de volição na natureza.**

Procedemos agora à discussão do resultado anterior. Antes de o fazermos, expomos esquematicamente a estrutura do processo de prova, indicando por troços de prova, os procedimentos metamatemáticos aplicados ao sistema e as respectivas fórmulas obtidas. Nunca perdendo de vista que se trata de uma prova hipotética numa abordagem especulativa.

Troços da prova	i)	ii)	iii)	iv)
Procedimentos metamatemáticos	Simple sobre T1, T2, T3	D1 sobre os naturais	Nomotético sobre i) e ii)	D2 sobre os naturais
Fórmulas	T1, T2, " $\neg(0 = 1)$ "	" $0 + a_{d1}(3) = 0$ " resultando " $0 + 0 = 0$ "	" $0 + \mathcal{N}_1 = 1$ " resultando " $0 + 1 = 1$ "	$a_{d2}(n) = \mathcal{N}_1 - a_{d1}$ resultando " $0 + a_{d2}(9) = 1$ " resultando " $0 + 1 = 1$ "

Troços da prova	v)	vi)	vii)
Procedimentos metamatemáticos	P1, P2, P3, P4	Nomotético sobre i), ii), iii), iv) e v)	D3 sobre o conjunto $\mathcal{D}$ ; procedimento simples sobre " $\neg(0 = 1)$ "
Fórmulas	P1, P2, P3; P4	" $0 + \mathcal{N}_2 = 1$ " resultando " $0 + 1 = 1$ "	$a_{d2}(n) = \mathcal{N}_2 - a_{d2}(n)$ resultando " $0 + 1 - a_{d2}(n) = 1$ " resultando " $0 = 1$ " e " $\neg(0 = 1)$ " Inconsistência simples

Sintetizando, de forma intuitiva, o significado dos teoremas, procedimentos e pressupostos, temos:

Troço i) Premissas da prova:

T1 - Existe o conjunto dos naturais, que é infinito e denumerável, sendo possível contar indefinidamente.

T2 - Existe uma lista infinita, não denumerável e completa de desenvolvimentos de base 2 de todos os naturais.

T3 - " $\neg(0 = 1)$ "

Troço ii) O terceiro coeficiente do desenvolvimento de base 2 de 3, obtido por diagonalização, é nulo,  $3 = 110000000 (\dots)$  na representação binária.

Troço iii) O processo de prova nos troços i) e ii) está correctamente formulado e os procedimentos aplicados são «sãos».

Troço iv) O nono coeficiente do desenvolvimento de base 2 de  $d2$ , obtido por diagonalização, é igual a 1,  $d2 = 001111111 (\dots)$  na representação binária.

Troço v) Pressuposições na prova:

P1 - Todos os coeficientes de  $d2 = 001111111 (\dots)$ , depois dos dois primeiros são iguais a 1.

P2 - O desenvolvimento de base 2 de  $d2$  está no conjunto dos desenvolvimentos de base 2 de todos os naturais.

P3 - Existe uma lista completa de desenvolvimentos de base 2 de todos os naturais, obtida por reordenação da anterior (teorema T2) onde  $d2$  se encontra entre os primeiros  $n$  naturais, tal que  $n > 2$ .

P4 - O procedimento de diagonalização é são sobre a lista anterior.

Troço vi) O processo de prova nos troços i) e ii) iii) e iv) está correctamente formulado e os procedimentos aplicados são «sãos».

Troço vii) " $0 = 1$ " e " $\neg(0 = 1)$ ", significando que o sistema formal é inconsistente ou há um erro na prova.

Os procedimentos metamatemáticos aplicados parecem ser sãos.

De qualquer forma, o processo de prova colocado é hipotético e, portanto, sujeito a críticas. A primeira é óbvia. Não existe um sistema formal nas condições colocadas. A comportamentos na natureza não podem corresponder processos de prova num sistema formal. Porém, todos os procedimentos metamatemáticos envolvidos parecem poder ser executados por um agente físico, suficientemente complexo para o efeito. De acordo com a metáfora de Hilbert, de que se falou na secção 2.1., parece possível estabelecer um isomorfismo (ou pelo menos uma analogia) entre comportamentos físicos e processos de prova. Utilizar um sistema formal como a origem ontológica das regularidades e padrões observáveis na natureza.

Talvez se coloque que apenas certos comportamentos não possam corresponder a processos de prova no sistema. O único conjunto de comportamentos físicos que parece evidente cumprir esse requisito é o que substancia o procedimento nomotético. Mas porque motivo, então, deixaria o sistema formal de exercer o seu papel regulador sobre certos comportamentos, nomeadamente, sobre aqueles associados ao procedimento nomotético? Não parece existir qualquer razão válida para que a combinação desses comportamentos assuma um carácter excepcional. num sistema que actua de forma mecânica e determinista, indiferente a qualquer significado sobre as consequências lógicas deste ou daquele procedimento.

Talvez a solução passe por introduzir uma proibição no sistema, impedido a aplicação do procedimento nomotético em certas circunstâncias. A saber, em todas aquelas conducentes a uma contradição. Esta exigência parece, uma vez mais, pouco natural num sistema completamente determinista e destituído de capacidades significativas. Tanto mais, que seria necessário que este reconhecesse, previamente, as condições lógicas que conduzem à sua própria inconsistência, o que em si mesmo é uma contradição, porque teria de provar a inconsistência para a poder evitar.

Uma solução recorrendo exclusivamente às capacidades lógicas do sistema não parece pois ser a mais adequada.

Para lidar com o problema assumiremos, pelo contrário, que a contradição encontrada só pode ser evitada se olharmos de fora do sistema. A solução que iremos propôr é, então, a seguinte. Sugeriremos que os comportamentos físicos envolvidos na aplicação do procedimento nomotético ainda correspondem a processos de prova no

sistema, mas que a sua justaposição, com os restantes comportamentos na análise, não depende das regras de transformação do sistema formal. Comportamentos físicos podem ocorrer de forma justaposta, no espaço e no tempo, mas os processos de prova que lhes correspondem não estão logicamente concatenados. Desse modo, não ocorre uma contradição no sistema formal. Essas concatenações físicas são-lhe “logicamente invisíveis”. Esse facto só pode significar uma coisa. Embora possam existir comportamentos que isoladamente correspondem a processos de prova, existem também certas justaposições deles que não correspondem a *concatenações lógicas* de processos de prova no sistema.

Assim sendo, a questão que se coloca é a de saber qual é a origem dessas justaposições. A resposta que nos ocorre é esta. Atribuímos a origem de tais justaposições à existência de *volição* na natureza, entendida como uma propriedade metalógica. Essa propriedade expressa-se, independentemente do sistema, produzindo certas tecituras de comportamentos físicos que não correspondem a processos de prova. A *volição* expressa-se como uma propriedade na natureza, que actua a partir de nível ontológico mais baixo, sobrepondo-se, por vezes, às regras de inferência do sistema.

Para comportamentos na natureza que correspondem a processos de prova, o sistema formal actua de forma determinista, sem que ocorram contradições ou inconsistência. Os comportamentos são regulares e previsíveis, em conformidade com os padrões observados. Pontualmente, verificam-se exceções que se traduzem em combinações particulares de comportamentos físicos, que tomados isoladamente, correspondem isoladamente a processos de prova e são regulares, mas cuja justaposição no comportamento global não se encontra prevista no sistema. À ocorrência desses pontos de justaposição, não previstos logicamente no sistema, chamamos *escolhas*. Estas ocorrem, produzindo interrupções pontuais na estabilidade determinista prevalecente, originando uma forma mínima de variabilidade *nomotética*.

A conclusão parece, então, ser esta: a consistência simples de um sistema formal, onde estão representados comportamentos na natureza, exige que certas combinações desses comportamentos ocorram de forma a não influenciar logicamente o sistema. Ou, de outra forma, a consistência simples de um sistema formal, regendo comportamentos na natureza, parece sugerir a existência de *volição*, na acepção dada.



Preparando o resto da exposição, consideraremos que o procedimento foi agenciado por um ser humano, dispondo de um cérebro e de um corpo, isto é, um soma sensorial e motor. As restantes aplicações esquemáticas tomarão como princípio que o agente de acções inteligentes é sempre um ser humano.

No esquema, existem três níveis de descrição. Começando de baixo para cima, encontramos o nível onde ocorrem os processos de prova no sistema formal. As setas indicam o sentido dos trajectos lógicos de inferência, percorridos na derivação dos processos de prova. As setas são assinaladas pela letra “L”, representando que ocorreu um processo de derivação lógica no sistema. O nível seguinte descreve os estados físicos observados no cérebro e corpo do agente. As setas determinam a sequência temporal com que os estados se sucedem, do passado para o futuro. Estão assinaladas pela letra “F”, querendo-se dizer que houve uma sucessão de eventos físicos observáveis. O terceiro nível descreve a ocorrência de estados mentais ou psicológicos, eventualmente, reportados pelo agente. As setas indicam a sua sucessão temporal e estão assinaladas pela letra “S”, significando que existe uma concatenação simbólica entre os estados mentais. Essa concatenação pode ser de natureza lógica informal, ou não, sendo, no entanto, sancionada pelo agente que a reporta. Os estados mentais assumem-se ser estruturas semânticas com conteúdo, podendo ser representações sobre factos externos, sobre as acções ou sobre estados de intencionalidade do agente. Para enfatizar essa característica, a sua descrição, no interior de cada caixa, está escrita entre aspas. Nos três níveis podemos falar, de forma intuitiva, de causalidade. Nomeadamente, da existência de causalidade lógica no nível ontológico, a qual se aceita poder estar livre de constrangimentos temporais e depender estritamente dos recursos lógicos do sistema. Da existência de uma causalidade física, que corresponde estritamente ao resultado de uma concatenação de eventos físicos, que se sucedem no tempo. E, finalmente, da existência de uma causalidade simbólica ou psicológica, que é determinada pelas asserções que o agente faz sobre a relação entre estados mentais contíguos. Cada elemento, em cada nível, é feito corresponder a um elemento no nível imediatamente superior. Esse é o significado das setas verticais. Entre os níveis mais baixo e intermédio, está representada a correspondência entre processos de prova e comportamentos físicos, a hipótese que foi colocada. Entre o nível intermédio e o mais alto, está agora representada uma nova correspondência, que se considerou útil para os efeitos da esquematização pretendida. Concretamente, a de que a cada estado físico do cérebro e do corpo corresponde um estado mental que equivale a uma estrutura semântica com conteúdo.

No que se fez na secção anterior, assumiram-se duas asserções que agora se



A sua expressão sobre o nível onde actua o sistema formal consiste em efectivar duas acções possíveis. A primeira consiste na escolha de um determinado processo de prova, ao qual corresponde um comportamento físico. Acção representada pela seta que une a primeira instanciação de volição, no esquema, à esquerda, à caixa que representa o Processo de Prova 1. Essa seta está assinalada pela letra “E”, significando a efectivação de uma escolha. A segunda acção consiste ainda na escolha de um processo de prova, mas ao qual se encontra associada uma interrupção na causalidade lógica do sistema formal. Esta acção está representada na parte direita do esquema, em baixo, na segunda instanciação da volição. A interrupção na causalidade lógica está representada por uma seta a pontado, assinalada pela expressão simbólica “ $E=\sim L$ ” que significa que o próximo processo de prova escolhido se encontra inferencialmente dissociado do anterior no sistema formal. Esta dissociação foi interpretada na última secção do capítulo anterior como a ocorrência de variabilidade nomotética. E, no caso, é justamente a presença de uma interrupção inferencial no sistema que previne a ocorrência da contradição que tornaria inconsistente o sistema. Desse modo, a existência de volição é, em primeira análise, a consequência de um imperativo de consistência do sistema formal, se quisermos manter a metáfora de que a cada processo de prova no sistema corresponde um comportamento na natureza.

No esquema estão, finalmente, representadas retroacções entre os estados psicológicos dos agentes e as duas instanciações de volição. Existem dois motivos para que a origem de cada retroacções seja um estado psicológico. O primeiro é o seguinte. Poder-se-ia pensar que o estado físico do cérebro e do soma seriam estruturas candidatas para o controlo retroactivo da volição. Porém, uma das posições filosóficas actuais sobre o cérebro, como a entidade física onde está sedeada a consciência, é a de que se trata de um sistema onde são processadas computações numa estrutura neuronal (Bickle, Mandik, Landreth, 2012, p. 8). Esse facto sugere que um problema análogo ao que foi encontrado no sistema formal possa ocorrer. O segundo motivo é o de que agentes humanos reportam estar em posse de uma intencionalidade no decurso da qual tomam decisões e produzem acções. Desse modo, pelo menos inicialmente, parece adequado considerar os estados psicológicos desses agentes como a origem para uma retroacção sobre a volição. No entanto, o circuito de retroacção assumido ainda coloca um problema de causalidade. Dado que se trata de um circuito de correspondências fechado, o mesmo não revela uma origem particular para a instanciação de uma situação de livre arbítrio. Até aqui, a volição pode ser entendida como um interruptor lógico que só garante a consistência do sistema formal ontológico. É neste ponto que fazemos um cometimento ontológico com uma visão particular sobre o funcionamento da mente.

Recusaremos o que é disposto no materialismo eliminativista («eliminative materialism») que mantém que são falsas e enganadoras as explicações psicológicas sobre os desejos e convicções, nas mentes dos agentes, como causas do comportamento humano. Propondo que essas explicações, em vez de incorporadas ou reduzidas a uma neurociência futura, serão completamente substituídas por ela (Bickle, Mandik, Landreth, 2012, p. 6). Recusaremos, portanto, que do ponto de vista da explicação do comportamento volitivo, existe uma identidade entre estados mentais com uma estrutura semântica de conteúdos e estados físicos simultâneos na estrutura do cérebro e soma, mediante a qual, uns pudessem ser completamente reduzidos aos outros. Aceitamos, no entanto, a existência de uma correspondência entre estados mentais e físicos no cérebro, da mesma forma que se sugeriu existir uma correspondência entre comportamentos físicos e processos de prova num sistema formal. Isto é, assumiremos um fisicalismo moderado que aceita rever conceptualizações na abordagem psicológica da mente com base nos resultados das neurociências.

Sem prejuízo da posição anterior, consideraremos ainda que existem estados mentais de intencionalidade, como explica John Searle:

“Intentionality is that property of many mental states and events by which they are directed at or about or of objects and states of affairs in the world. If, for example, I have a belief, it must be a belief that such and such is the case; if I have a fear, it must be a fear of something or that something will occur; if I have a desire, it must be a desire to do something or that something should happen or be the case; if I have an intention, it must be an intention to do something. And so on through a large number of other cases. I follow a long philosophical tradition in calling this feature of directedness or aboutness "Intentionality", but in many respects the term is misleading and the tradition something of a mess, so at the very beginning I want to make it clear how I intend to use the term and in so doing to dissociate myself from certain features of the tradition.

First, on my account only some, not all, mental states and events have Intentionality. Beliefs, fears, hopes, and desires are Intentional; but there are forms of nervousness, elation, and undirected anxiety that are not Intentional. A clue to this distinction is provided by the constraints on how these states are reported. If I tell you I have a belief or a desire, it always makes sense for you to ask, "What is it exactly that you believe?" or "What is it that you desire? "; and it won't do for me to say, " Oh I just have a belief and a desire without believing anything or desiring anything". My beliefs and desires must always be about something. But my nervousness and undirected anxiety need not in that way be about anything. Such states are characteristically accompanied by beliefs and desires, but undirected

states are not identical with beliefs or desires. On my account if a state *S* is Intentional then there must be an answer to such questions as: What is *S* about? What is *S* of? What is it an *S* that?"

(Searle, 1983, pp. 1 – 2)

Na nossa interpretação da estrutura da escolha, a directividade da mente sobre um objecto ou situação externas corresponde à directividade da volição sobre processos de prova em duas acepções diferentes. Sobre processos de prova que correspondem a estados físicos do cérebro que, por sua vez, correspondem a estados mentais onde estão representados esses objectos ou situações. Sobre processos de prova que permitem a activação motora do soma, por parte do agente, fazendo com que uma acção física se desenvolva, interagindo com o meio. O fundamento desta interpretação não advém directamente da situação que descrevemos anteriormente. Ali, volição é uma propriedade cuja existência se deve apenas a um imperativo lógico de consistência. Esse imperativo fez com que fosse possível a ocorrência de uma interrupção inferencial lógica, que identificámos com a ocorrência de variabilidade nomotética. É, por outro lado, a possibilidade de que essa variabilidade ocorra que sugere que a volição encontrada possa ser importada para um modelo que descreva o funcionamento da mente, permitindo o livre arbítrio. Deste modo, através da volição, o agente actua como se assumisse uma posição externa em relação a um conjunto de processos de prova activáveis. E, conseqüentemente, como se fosse externo ao seu próprio soma físico, tomando decisões sobre que estados físicos no cérebro activar, apesar de ser continuamente informado sobre o estado do seu corpo e do meio ambiente. E apesar do seu comportamento físico se reduzir aos comportamentos físicos do cérebro e do soma.

A volição como propriedade metalógica, que agora deverá ser reconhecida como uma propriedade do comportamento humano, não depende do sistema formal ontológico adoptado, nem dos estados físicos do cérebro. Embora as suas manifestações se resumam a esses estados; permite que o agente, a partir de estados mentais de intencionalidade, produza concatenações de estados físicos no cérebro. Podem existir sequências desses estados que, isoladamente, correspondem a processos de prova no sistema, mas que uma vez concatenados fisicamente, pela intervenção da volição, formarão concatenações que não corresponderão a processos de prova. Uma metáfora possível é a de que o cérebro é como um sistema físico, capaz de processar informação complexa, controlado por um agente externo que, de ponto de vista empírico, coincide com ele. O tratamento do cérebro como se tratasse de um computador não é estranho às neurociências.

“At one level of analysis the basic computational element of a neural network (biological or artificial) is the neuron. This analysis treats neurons as simple computational devices, transforming inputs into output. Both neuronal inputs and outputs reflect biological variables. (...) neuronal inputs are frequencies of action potentials (neuronal “spikes”) in the axons whose terminal branches synapse onto the neuron in question. Neuronal output is the frequency of action potentials in the axon of the neuron in question. A neuron computes its total input (usually treated mathematically as the sum of the products of the signal strength along each input line times the synaptic weight on that line). It then computes a new activation state based on its total input and current activation state, and a new output state based on its new activation value. The neuron's output state is transmitted as a signal strength to whatever neurons its axon synapses on. The output state reflects systematically the neuron's new activation state. Analyzed at this level, both biological and artificial neural networks are interpreted naturally as *vector-to-vector transformers*. The input vector consists of values reflecting activity patterns in axons synapsing on the network's neurons from outside (e.g., from sensory transducers or other neural networks). The output vector consists of values reflecting the activity patterns generated in the network's neurons that project beyond the net (e.g., to motor effectors or other neural networks). Given that neurons' activity depends partly upon their total input, and total input depends partly on synaptic weights (...), the capacity of biological networks to their synaptic weights make them *plastic* vector-to-vector transformers. In principle, a biological network with plastic synapses can come to implement any vector-to-vector transformation that its composition permits (number of input units, output units, processing layers, recurrency, cross-connections, etc.)”.

(Bickle, Mandik, Landreth, 2012, pp. 8 - 9)

A natureza substancial dos estados mentais não será aqui discutida. Ela é, quando muito, identificável parcial ou, em certos casos, totalmente, com padrões de actividade neurológica no cérebro. Pressupondo a existência de volição, diremos antes que o comportamento humano, no modelo proposto, requer que a sua descrição se faça a quatro níveis. O nível em que a volição actua, por retroacção proveniente dos estados mentais de intencionalidade do agente. O nível formal, onde ocorrem os processos de prova, com causalidade inferencial, permitindo a elocução de regras descrevendo padrões e regularidades na natureza. O nível físico de observação empírica e o nível mental onde se pode aplicar ou uma teoria psicológica ou uma teoria psicosemântica sobre estados com conteúdo representacional.

“A psychosemantics seeks to explain what it is for a representational state to be about something: to provide an account of how states and events can have specific representational content. A physicalist psychosemantics seeks to do this using resources of the physical sciences exclusively (...) Neurophilosophers have contributed to two types of physicalist psychosemantics: the Functional Role approach and the Informational approach.

(...)

The core claim of a functional role semantics holds that a representation has its content in virtue of relations it bears to other representations. Its paradigm application is to concepts of truth-functional logic, like the conjunctive ‘and’ or disjunctive ‘or.’ A physical event instantiates the ‘and’ function just in case it maps two true inputs onto a single true output. Thus it is the relations an expression bears to others that give it the semantic content of ‘and.’ Proponents of functional role semantics propose similar analyses for the content of all representations (Block 1986). A physical event represents birds, for example, if it bears the right relations to events representing feathers and others representing beaks. By contrast, informational semantics ascribe content to a state depending upon the causal relations obtaining between the state and the object it represents. A physical state represents birds, for example, just in case an appropriate causal relation obtains between it and birds. (...) A common criticism of informational semantics holds that mere causal covariation is insufficient for representation, since information (in the causal sense) is by definition always veridical while representations can misrepresent. A popular solution to this challenge invokes a teleological analysis of ‘function.’ A brain state represents X by virtue of having the function of carrying information about being caused by X (Dretske 1988). These two approaches do not exhaust the popular options for a psychosemantics, but are the ones to which neurophilosophers have contributed”.

(Bickle, Mandik, Landreth, 2012, p. 18)

A aplicação de teorias psicosemânticas aos estados mentais intencionais parece adequada de forma a estabelecer relações adequadas de correspondência entre a estrutura dos estados intencionais e padrões e regimes de actividade neurológica no cérebro. Sobretudo utilizando uma versão refinada do que se descreveu acima conhecida por espaço de fase semântico («state-space semantics»), em que um conceito não é representado apenas pela localização de um só nó na rede, mas também pela sua proximidade geométrica com outros nós (Bickle, Mandik, Landreth, 2012, p. 19).

Seja como fôr, a existência de uma correspondência entre processos de prova, estados físicos no cérebro e estados mentais de intencionalidade requer que existam dicionários que traduzam, que convertam cada um dos códigos noutra. Desse ponto de

vista, o esquema que aqui se apresenta deve ser visto como um modelo provisório contemplando a possível existência de volição. Uma estrutura possível para a escolha.

## **5.2. Casos estruturais da escolha.**

No que segue apresentaremos casos possíveis de efectivação de escolhas. Uma escolha consiste numa de duas acções fundamentais possíveis. Uma acção, da qual resulta contiguidade lógica no sistema formal e outra acção da qual se obtém uma dissociação lógica entre os processos de prova.

O agente assume um conjunto de estados mentais intencionais ou o mesmo, repetido um número finito de vezes, usando, eventualmente, a memória de um ou mais estados mentais anteriores, activando, por retroacção, a volição. O modo preciso como essa activação ocorre não será aqui discutido. Considera-se que a mesma é possível, em princípio. Essa activação substancia-se numa escolha de um processo de prova, que corresponde ao objecto sobre o qual incide a intencionalidade do sujeito, a nível mais alto na estrutura. Quando a escolha de um processo de prova é tal que este pode ser inferencialmente concatenado com o processo de prova imediatamente anterior, dizemos que o agente volitivo efectuou uma escolha passiva ou que agencia a escolha passivamente. Entre os dois processos de prova obtidos diz-se existir uma contiguidade lógica, de acordo com as regras de transformação do sistema. É o caso ilustrado na figura seguinte.

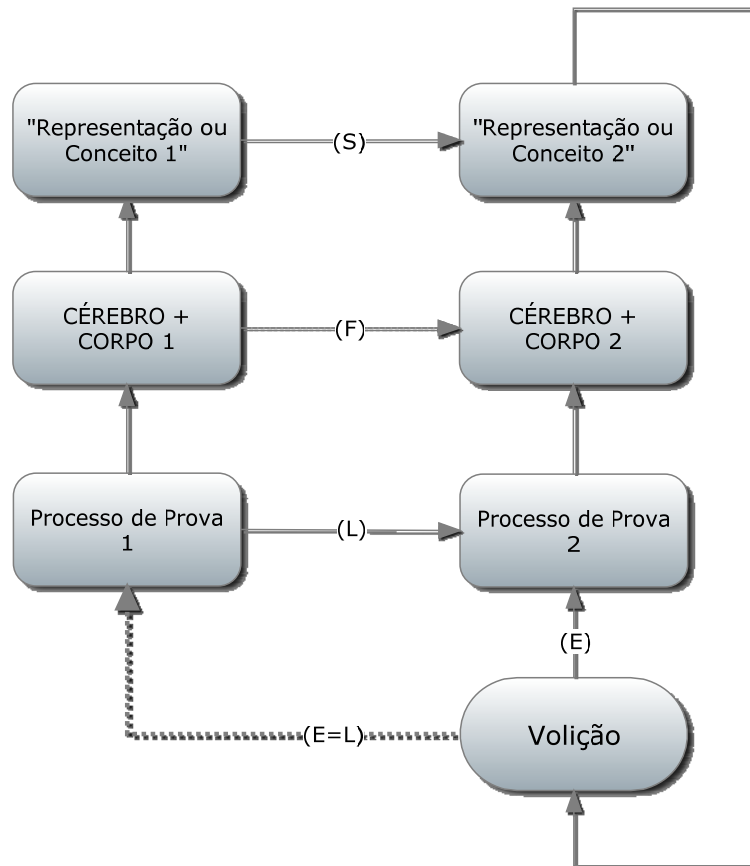


Figura 5.3 - Concatenação de estados mentais com contiguidade lógica no sistema formal.

No nível mais elevado do esquema a referência à representação ou conceito consiste no objecto sobre o qual incide a intencionalidade do sujeito, que se pode traduzir numa focalização da sua atenção. Esse regime mental não deve ser visto como um estado isolado, mas como uma sequência de estados, cada qual exibindo uma estrutura semântica com conteúdo. Ou ser considerado um regime em que o mesmo estado se repete <sup>9</sup>. A sequência de estados ou a repetição de um deles entende-se ser um comportamento mental, ao qual corresponde uma sequência de estados físicos no cérebro e, mais abaixo, o processo de prova evocado. Isso significa que há questões a tratar sobre a natureza do circuito que permite activar a volição. Parece colocar-se, por exemplo, um problema de causalidade, em que o estado mental parece existir antes mesmo de produzir o processo de prova e o respectivo comportamento físico no cérebro, que o substancia fisicamente. Talvez o circuito como foi desenhado dê a falsa impressão de que existe um sinal que o percorre, levando tempo a activar a volição. Mas também pode ser que a volição seja uma

<sup>9</sup> A esse respeito ver as experiências de Ernst Pöppel (1988) sobre a percepção do tempo na consciência, sugerindo a existência de um intervalo de tempo máximo para a atenção sobre um objecto, após o qual, um novo objecto se apresenta à consciência. Interpretamos esse resultado, falando de uma repetição da focalização da atenção, quando um objecto permanece subjectivamente o mesmo na experiência mental.

propriedade fundamental da realidade e que estado mental, estado físico neurológico e processo de prova sejam apenas aspectos síncronos de um só fenómeno. O nosso esquema revelaria então apenas dependências causais abstractas, sem entrar em linha de conta com separações espaço-temporais. Outra solução ainda, mais simples, seria adoptar a explicação de Libet e Haggard que, como vimos no final do primeiro capítulo, consideram o cérebro uma fonte ruído, que o sujeito veta ou promove de forma a ocorrer a activação mecânica do soma. Nesse caso, haveria um processo causal espaço-temporal e poderíamos conceber uma coalecência de um estado mental, inicialmente indefinido, a partir de ruído no cérebro, correspondendo a certos processos de prova, mais elementares. Estes dariam gradualmente lugar a outros, por activações volitivas sucessivas, como se o circuito volitivo fosse uma espécie de mecanismo de controlo cibernético até se chegar a um estado mental intencional familiar. Uma hipótese mais radical pertenceria ao modelo de Penrose-Hameroff, permitindo retroacção temporal a nível subquântico. Estas questões, embora importantes, não têm ainda cabimento num trabalho introdutório. Pertencem a uma fase posterior de pesquisa sobre a natureza exacta do processo volitivo, na continuação do que Libet iniciou. Interessa-nos para já considerar o valor heurístico do esquema colocado e ter presente que o mesmo originou de uma hipótese especulativa que trouxe contradição a um sistema que se considerava suficientemente robusto para a evitar.

Voltando ao esquema apresentado na figura, a contiguidade lógica ou, se se quiser, a causalidade lógica está representada ainda por uma seta a ponteadado, assinalada agora pela expressão simbólica “E=L”. A situação traduz-se por existir estabilidade nomotética e o agente pode entender-se actuar como um sistema físico habitual, cujo comportamento é regido pelo sistema formal ontológico, na base da estrutura. De forma intuitiva diríamos que o agente “escolhe não escolher”.

É suposto o agente ter à sua disposição todos os processos de prova possíveis no sistema formal, evocáveis a qualquer momento.

A segunda possibilidade de acção por parte do agente, utilizando a volição, ocorre quando este assume um estado intencional, nas mesmas condições de há pouco, mas desta vez procedendo a uma escolha tal que o processo de prova evocado não é logicamente concatenável com o anterior, de acordo com as regras de transformação do sistema. Diz-se então que entre os dois processos de prova não há contiguidade lógica e que houve uma interrupção inferencial no sistema formal ontológico.

É suposto que quando isto sucede, o processo de prova escolhido, tenha sido obtido por meios inferenciais no sistema que diferem daqueles utilizados no processo de prova imediatamente anterior. E que a sua evocação seja equivalente a uma derivação imediata. Desse ponto de vista, o conjunto de processos de prova que estão à disposição

da volição pode ser visto como uma estrutura permanente que arboresce a partir de um conjunto de axiomas <sup>10</sup>, de acordo com regras fixas. O sistema formal não tem, em princípio, restrições espaço-temporais, como se esperaria de uma entidade ontológica abstracta, que se comparou às leis da natureza.

Por vezes a acção da volição, por activação do estado mental do agente, conduz necessariamente a um estado de dissociação lógica entre processos de prova ou, como se disse, a uma situação de variabilidade nomotética. Essa situação é justamente aquela que se apresentou para evitar a possibilidade de contradição no sistema formal. É um caso possível adicional para a acção volitiva e parece constituir o limite que se coloca ao livre arbítrio. Não existe um estado mental intencional, mediante o qual um agente produza uma ruptura lógica no sistema formal. Este encontra-se, por assim, dizer protegido desse tipo de interferências. Por outro lado, este imperativo pode ser utilizado para explicar porque motivo podemos mentir ou assumir estados de confusão quando pensamos, mesmo assumindo semelhante controlo sobre parte de um edifício ontológico, por princípio, fundamental. Esses processos que nos fazem actuar de forma inconsistente ou ilógica são porventura acompanhados de uma dissociação lógica inevitável, a nível mais baixo, permitindo que mesmo no meio do caos, conservemos uma certa integridade psicológica. Aquilo a que poderíamos intuitivamente chamar sanidade mental. Mas justamente o que dizer da loucura? Porque ocorre então? Talvez o substrato físico, a estrutura do soma, as contingências do meio ou outras influências determinem o espaço de fase dos processos de escolha, do ponto de vista de um certo cérebro e de um certo soma. Parece assim existir uma certa determinação que o nível intermédio produz sobre os estados mentais que, estes por sua vez, transmitem à volição.

O esquema apresentado revela assim a existência de um equilíbrio ou, mais exactamente, de uma regulação que ocorre através das quatro etapas do circuito volitivo: estado mental, volição, sistema formal ontológico, estrutura física do cérebro e de volta ao estado mental. A consideração deste processo de regulação não constitui novidade. Porém, o elemento adicional que se incorporou, a volição, parece ter trazido uma nova potencialidade heurística de explicação ao processo. A sensação subjectiva de *Self* parece ser relativamente compreensível, em vista do esquema disposto. Uma possibilidade de definição seria a de que consiste no resultado da apreciação permanente e consequente alteração que o agente faz dos seus próprios comportamentos mentais, situação permitida pelo circuito volitivo. O esquema a seguir representa o caso de uma escolha activa, isto é,

---

<sup>10</sup> Supomos finito o número de axiomas que dizem respeito a processos de prova correspondendo a comportamentos na natureza.

onde se verifica uma dissociação lógica entre processos de prova no sistema formal ontológico.

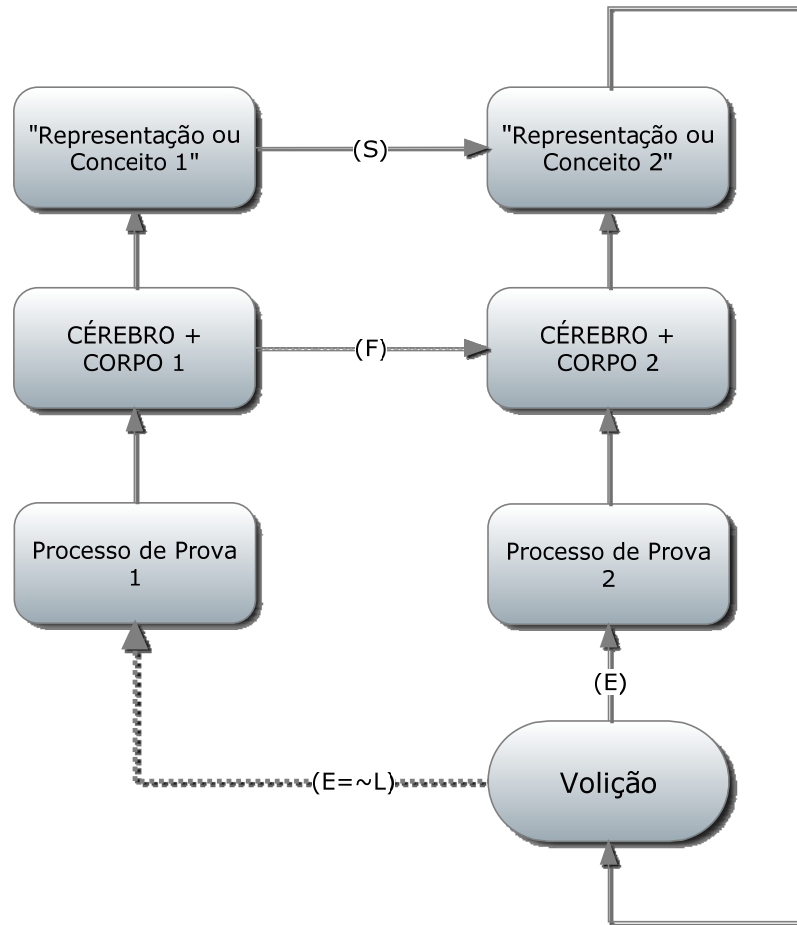


Figura 5.4 - Concatenação de estados mentais com dissociação lógica no sistema formal.

Como no exemplo dado para a aplicação do procedimento nomotético, a seta a pontilhado está agora assinalada pela expressão simbólica “ $E=\sim L$ ”.

A posição do agente face ao meio externo é também representável, como mostra o esquema seguinte. Ilustra o caso em que um agente procede à observação de três comportamentos físicos seguidos no tempo. Esses comportamentos físicos de alguma forma interagem com o soma do agente, produzindo estímulos sensoriais. Estes dão origem a estados mentais intencionais que o agente decide ou não concatenar de ponto de vista simbólico, traduzindo-se essa dissociação simbólica, numa dissociação formal, através da volição. Essa situação sucede na terceira observação.

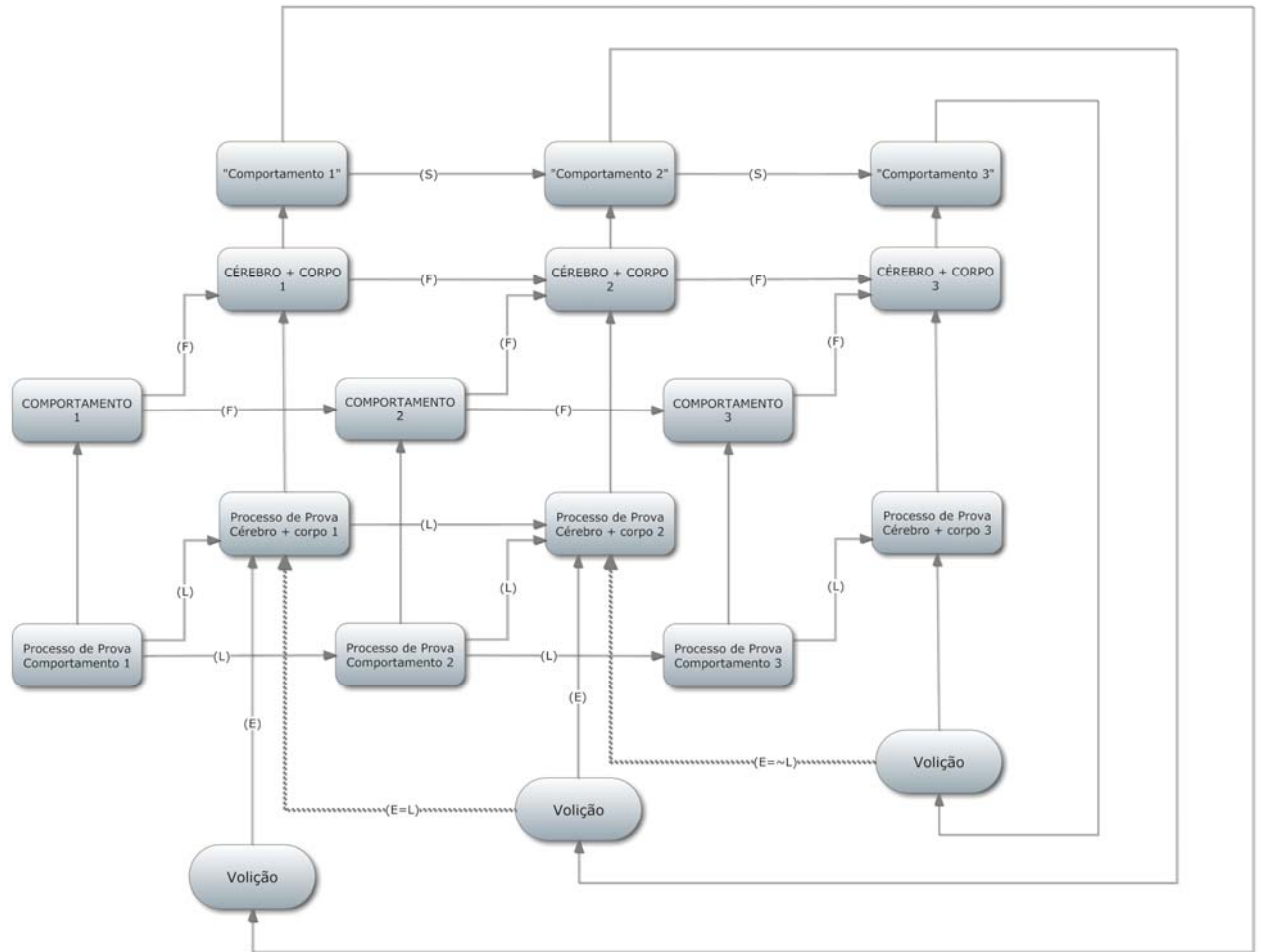


Figura 5.5 - Concatenação de três comportamentos físicos observados (o terceiro não tem contiguidade lógica).

Deve-se enfatizar que a concatenação de estados mentais intencionais, assim como a concatenação de estados físicos é apenas disposta no tempo. Isto é, no primeiro caso não tem de existir uma relação significativa, intuitivamente lógica entre estados mentais. Seres humanos podem pensar de forma imperfeita ou até inconsistente. No segundo caso, pede-se que se aceite que dois regimes neurológicos contíguos no tempo não tenham que ser tais de forma a que o segundo seja o resultado causal inevitável do primeiro. Embora possam existir alguns em que isso possa suceder, correspondendo à situação em que o agente efectua escolhas passivas. Uma situação em que, eventualmente, se vê obrigado a isso no limite das capacidade de plasticidade do cérebro.

No que dispôs, uma sequência de observações pode ser também considerada um processo volitivo.

### 5.3. Aplicações do esquema da escolha.

Nesta secção vamos testar o poder explicativo do esquema apresentado, aplicando-o a três situações. A primeira diz respeito ao problema filosófico dos Futuros Contingentes, descrito pela primeira vez por Aristóteles [384 – 322, a.c.] no segundo livro do *Organon*, *De Interpretatione*.

Sobre o tema respeitaremos o artigo de Peter Øhrstrøm e Per Hasle, "*Future Contingents*", (Øhrstrøm & Hasle, 2011).

O problema coloca-se do seguinte modo:

“On Interpretation, is without doubt the philosophical text which has had the greatest impact on the debate about the relations between time, truth, and possibility. The discussion in this text certainly bears witness to the fact that Ancient philosophy was highly conscious of tense-logical problems. Central to the discussion in this famous Aristotelian text is the question of how to interpret the following two statements:

“Tomorrow there will be a sea-battle”

“Tomorrow there will not be a sea-battle”

Aristotle considered questions like: Should we say that one of these statements is true today and the other false? How can we make a clear distinction between what is going to happen tomorrow and what must happen tomorrow?

(...)

The interpretative problems regarding Aristotle's logical problem about the sea-battle tomorrow are by no means simple. Over the centuries, many philosophers and logicians have formulated their interpretations of the Aristotelian text (...) [such as] an interpretation of the text from the scholastic period and (...) [modern interpretations] (...) based on a three-valued semantics”.

(Øhrstrøm & Hasle, 2011, p. 4)

O esquema da escolha permite resolver uma instância do problema, em que participam agentes humanos, dispostos a travar batalhas navais, no dia seguinte, ou capazes de as evitar. Partiremos da sistematização do problema feita por outro filósofo grego posterior, Diodorus Cronus [340 – 280, a.c.].

“In the generation after Aristotle, Diodorus Cronus (...) analysed similar problems using his so-called Master Argument. This argument was a trilemma. According to Epictetus, Diodorus argued that the following three propositions cannot all be true:

(D1) Every proposition true about the past is necessary.

(D2) An impossible proposition cannot follow from (or after) a possible one.

(D3) There is a proposition which is possible, but which neither is nor will be true.”

(Øhrstrøm & Hasle, 2011, p. 5)

Na interpretação de Øhrstrøm e Hasle, Diodorus parecia empenhado em provar que não existem Futuros Contingentes:

“Diodorus used (...) (D1) and (D2) to argue that (D3) is false. Assuming (D1) and (D2) he went on to define the possible as “that which either is or will be true” and the necessary as “that which, being true, will not be false”. In this way his argument seems to have been designed to demonstrate that there cannot be any future contingency at all. However, little is known about the way in which Diodorus used his premises in order to reach the conclusion. The reconstruction of the Master Argument certainly constitutes a genuine problem within the history of logic”.

(Øhrstrøm & Hasle, 2011, p. 5)

A nossa interpretação do argumento de Diodorus, considerando a do lógico medieval Lavenham [c. 1380] (Øhrstrøm & Hasle, 2011, p.7), é a seguinte:

Considera-se que qualquer proposição verdadeira sobre o passado é necessária. Isto é, que sendo agora verdadeira sobre um evento passado nunca será falsa. Esse poderá ser o significado da sentença (D1):

“(D1) Every proposition true about the past is necessary”.

Por uma proposição “possível” entenda-se uma afirmação que “ou é verdadeira ou será verdadeira”. Considere-se que o valor de verdade de uma proposição verdadeira, evocada sobre outra no passado, se transmite a essa proposição, se essa proposição for possível. De forma que se pode raciocinar da seguinte forma:

i) Tendo sido evocada, no presente, uma proposição  $p1$  sobre o passado, verdadeira porque é necessária, referindo-se a outra  $p2$ , possível sobre o futuro.

ii) Tendo sido evocada no passado a proposição  $p_2$ , que pode ser verdadeira, porque é possível.

iii) De  $p_1$  não é possível derivar a falsidade de  $p_2$ .

Notando que  $p_1$  cumpre também o disposto na definição de “proposição possível” e que, além de necessária, é também possível, a sentença (D2) pode ser entendida como a conclusão em iii).

“(D2) An impossible proposition [ $p_2$ ] cannot follow from (or after) a possible one [ $p_1$ ]”.

As sentenças (D1) e (D2) serão, conseqüentemente, incompatíveis com a sentença (D3):

“(D3) There is a proposition [ $p_2$ ] which is possible, but which neither is nor will be true.”

Diodorus, ao contrário do que faria Aristóteles, considerará falsa a proposição (D3).

O tratamento do problema dos Futuros Contingentes parece assumir que existe uma causalidade lógica ao nível da observação empírica dos factos. Esse pressuposto é retomado nas interpretações modernas, envolvendo lógicas modais temporais em que se introduz um terceiro valor lógico, quantificando a indecidibilidade de uma afirmação sobre o futuro (Øhrstrøm & Hasle, 2011). Através da aplicação do esquema da escolha torna-se claro que não existe causalidade lógica entre eventos físicos, mas tão só uma contiguidade de eventos, que pode ou não ser determinada pelo sistema formal. A causalidade lógica encontra-se mais abaixo, a nível ontológico, e, como vimos, pode ser interrompida. A figura seguinte, tratando a situação em que um dos futuros contingentes é fisicamente actualizado, ilustra a solução que propomos.





A diferença entre este e casos anteriores tratados é a de que não existe um estado mental intencional, correspondendo ao estado físico do cérebro em que surgem os potenciais evocados, cerca de um segundo antes do momento em que o sujeito reporta ter tomado uma indecisão. Esse estado neurológico anterior parece comparável ao estado de um objecto físico, externo ao agente, antes da sua observação. De qualquer forma, o esquema da escolha parece, também neste caso, exibir potencial heurístico suficiente para descrever a situação de forma adequada com os dados experimentais.

Outro caso diz respeito à situação em que o agente produz uma decisão de veto sobre os potenciais evocados, cerca de um segundo antes. Uma vez mais, não existe um estado intencional do agente correspondendo ao estado físico do cérebro, no momento anterior. A descrição de Libet para a decisão de veto é a seguinte:

“The evidence for conscious veto is of two kinds: (a) Subjects in our study of RPs [readiness potentials] and conscious timings reported that during some of the trials a recallable conscious urge to act appeared but was "aborted" or somehow suppressed before any actual movement occurred; in such cases the subject simply waited for another urge to appear, which, when consummated, constituted the actual event whose RP was recorded (...). However, there is presently no technique available for recording and analyzing any RPs that may be associated with such spontaneous, irregularly appearing conscious urges to act that do not lead to an actual motor event, (b) In series of acts to be performed at prearranged times, subjects were instructed in advance to veto the developing intention/preparation to act and to do this about 100 to 200 ms before the prearranged clock time at which they were otherwise supposed to act. In these series a ramplike pre-event potential was still recorded during >1 sec before the preset time (...), even though no actual muscle activation occurred (...). This resembles the RP of selfinitiated acts when preplanning is present (...). The form of the "veto" RP differed (in most but not all cases) from those "preset" RPs that were followed by actual movements; the main negative potential tended to alter in direction (flattening or reversing) at about 150-250 ms before the preset time (...). This difference suggests that the conscious veto interfered with the final development of RP processes leading to action. (...) In any case, the preparatory cerebral processes associated with an RP can and do develop even when intended motor action is vetoed at approximately the time that conscious intention would normally appear before a voluntary act”.

(Libet, 2011, p. 538)

O esquema seguinte ilustra o caso da decisão de veto. Presume-se que através do

circuito volitivo, o agente consegue produzir uma dissociação lógica entre o estado neurológico onde surgem os potenciais evocados e o seguinte. Este seria aquele que daria curso a uma acção motora, por parte do agente, constituindo uma escolha passiva. A intervenção do agente, através do circuito volitivo, permite assim inibir a prossecução dos efeitos dos potenciais evocados no sistema motor do corpo. Constitui assim uma intervenção volitiva activa do sujeito, interrompendo o que de outro modo se poderia considerar um processo físico determinista, conforme se mostra.

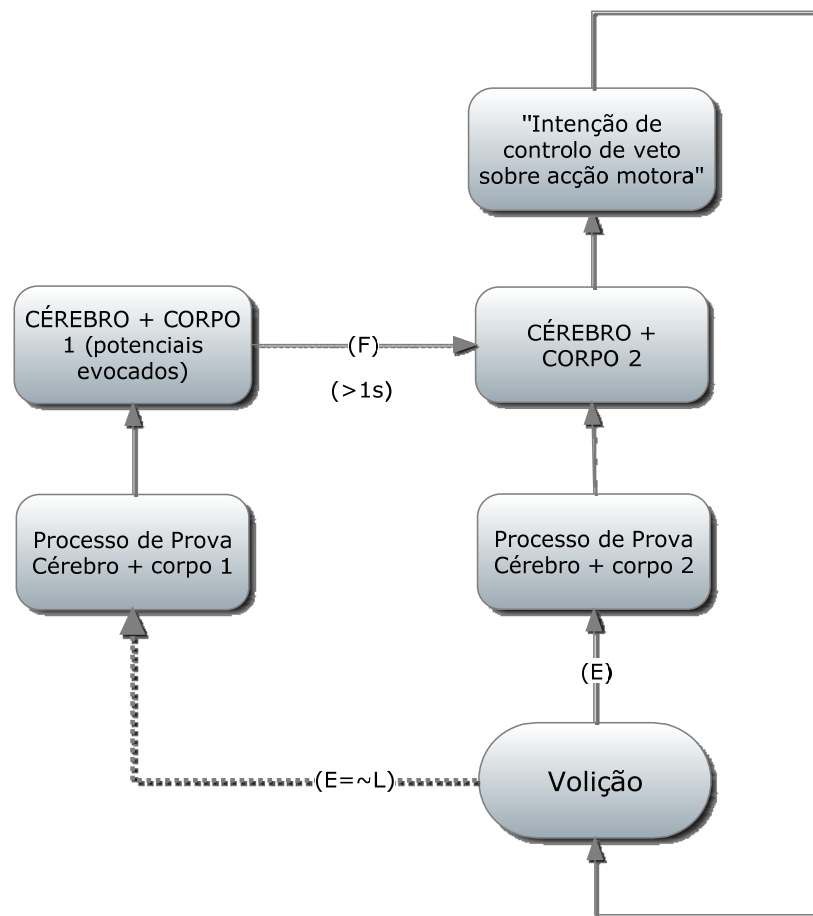


Figura 5.8 – Concatenação de Libet com potenciais evocados, sem contiguidade lógica no sistema formal, inibindo a acção motora.

O esquema da escolha parece ainda adequado à situação observada e teorizada por Libet em que existe uma profusão eventual de potenciais evocados, eventualmente despoletados por estímulos externos ou por algum mecanismo interno no cérebro, que são depois controlados por acção de uma escolha, através do circuito volitivo.

#### 5.4. Uma ontologia da escolha.

Apresentamos, agora, de forma sucinta e sistematizada, um conjunto de afirmações sobre a escolha, de acordo com o que foi dito nas três secções anteriores.

1. Supõe-se a existência de uma entidade ontológica abstracta que é um sistema formal que rege, deterministicamente, comportamentos na natureza. Esse sistema tem as seguintes propriedades:

1.1. Parte dos seus processos de prova correspondem a comportamentos observáveis na natureza.

1.2. O sistema formaliza todas as teorias matemáticas informais concebíveis.

2. Supõe-se a existência de uma propriedade ontológica na natureza chamada volição. Essa propriedade encontra expressão na realidade, através da efectivação de escolhas, agenciadas por agentes humanos, com um cérebro e um corpo, e tem as seguintes propriedades:

2.1. A expressão da volição na natureza não depende do sistema formal, dizendo-se uma propriedade metalógica.

2.2. Não existe nenhuma expressão da volição capaz de produzir uma contradição no sistema formal.

3. Existe uma estrutura da escolha em que comportamento humano pode ser analisado, considerando quatro níveis de descrição:

3.1. O primeiro nível ontológico onde actua a volição, permitindo interferência sobre concatenações de processos de prova no sistema formal.

3.2. O segundo nível onde actua o sistema formal, produzindo concatenações de processos de prova, de acordo com as suas regras de transformação e onde se diz que há causalidade lógica.

3.3. O terceiro nível, onde são observáveis os comportamentos de sistemas físicos

na natureza, concatenados no espaço e no tempo e onde se diz que há causalidade espaço-temporal.

3.4. O quarto nível onde se identificam estados mentais intencionais do agente humano, com estrutura semântica, concatenados no tempo e onde se diz que há causalidade simbólica.

4. Supõe-se existir no esquema da escolha uma correspondência entre cada um dos processos de prova e um comportamento de um sistema físico na natureza. Em particular, a cada processo de prova associado ao comportamento do agente existe um regime neurológico, formado por estados do cérebro e do corpo que se sucedem no tempo.

5. Supõe-se em, pelo menos, certas circunstâncias existir uma correspondência entre estados mentais intencionais do sujeito e regimes neurológicos e físicos no soma. Sempre que existe uma expressão de volição, essa correspondência é válida

6. Supõe-se a existência de um circuito volitivo de retroacção em que um estado intencional do agente pode actuar sobre a volição, de forma a efectivar uma escolha.

7. Uma escolha consiste numa acção da volição sobre o sistema formal, promovendo ou dissociando a contiguidade lógica entre dois processos de prova no sistema.

7.1. Entende-se que quando a concatenação de processos de prova no sistema formal, associados ao comportamento agente, se encontra de acordo com as regras de transformação do sistema, o agente actuou volitivamente, de forma passiva, promovendo o desenvolvimento do processo de prova, comportamentos físicos, do cérebro e soma, e estados mentais intencionais correspondentes. Diz-se que a volição produziu contiguidade lógica entre processos de prova no sistema.

7.2. Entende-se que quando a concatenação de processos de prova no sistema formal é tal que dois processos, aos quais correspondem comportamentos físico seguidos no tempo, não podem ser concatenados de acordo com as regras de transformação do sistema, o agente actuou volitivamente de forma activa. Diz-se que a volição produziu uma dissociação lógica entre processos de prova no sistema.

7.2.1. No caso descrito em 7.2. considera-se que cada processo de prova no sistema formal foi derivado de forma independente, nesse sistema, e que a concatenação dos comportamentos físicos, que correspondem a um e outro processo, não têm influência formal sobre o sistema, de acordo com 2.2.

8. A expressão da volição através de escolhas resume-se, do ponto de vista empírico, à existência de concatenações de comportamentos físicos que não são determinadas pelo sistema formal.

9. Não existem concatenações de estados mentais intencionais no tempo que não correspondam a concatenações de regimes neurológicos e de estados físicos do corpo, embora o inverso possa ser verdadeiro.

10. Existem concatenações de regimes neurológicos e de estados físicos do corpo, no espaço e no tempo, que não correspondem a concatenações de processos de prova no sistema formal. Embora o inverso não seja verdadeiro.

11. Considera-se que o cérebro, pelas suas propriedades plásticas, é um sistema físico especialmente adequado para que nele ocorram processos lógicos ou computacionais que não têm de ter contiguidade lógica, no sentido que se atribuiu em 7.1.

12. Existe um equilíbrio no circuito volitivo, fazendo com que estados mentais intencionais possam exercer controlo sobre o cérebro, mas também fazendo com que as suas propriedades físicas globais determinem, a cada momento, o grau e natureza desse controlo.

Se aceitarmos que a volição é uma propriedade real, capaz de produzir concatenações de comportamentos, de acordo com leis da natureza, embora livres de qualquer esquema determinista, poderá ser que uma teoria sobre a sua acção possa ser edificada. É nesse sentido que as afirmações acima se dispõem. O seu objectivo é providenciar um contributo inicial para uma teoria da escolha.

Há, no entanto, considerações a fazer sobre o significado da existência de volição para a existência humana. O que significa ser, na acepção humana?

Introduzimos este tema com um texto de Karl Popper.

“The doctrine that men are machines, or robots, is a fairly old one. Its first clear and forceful formulation is due, it seems, to the title of a famous book by La Mettrie, *Man a Machine* [1747]; though the first writer to play with the idea of robots was Homer.

Yet machines are clearly not ends in themselves, however complicated they may be. They may be valuable because of their usefulness, or because of their rarity; and a certain specimen may be valuable because of its historical uniqueness. But machines become valueless if they do not have a rarity value: if there are too many of a kind we are prepared to pay to have them removed. On the other hand, we value human lives in spite of the problem of over-population, the gravest of all social problems of our time. We respect even the life of a murderer.

It must be admitted that, after two world wars, and under the threat of the new means for mass destruction, there has been a frightening deterioration of respect for human life in some strata of our society. This makes it particularly urgent to reaffirm in what follows a view from which we have, I think, no reason to deviate: the view that men are ends in themselves and not "just" machines.

We can divide those who uphold the doctrine that men are machines, or a similar doctrine, into two categories: those who deny the existence of mental events, of personal experiences, or of consciousness; or who say perhaps that the question whether such experiences exist is of minor importance and may be safely left open; and those who admit the existence of mental events, but assert that they are "epiphenomena" - that everything can be explained without them, since the material world is causally closed. But whether they belong to the one category or the other, both must neglect, it seems to me, the reality of human suffering, and the significance of the fight against unnecessary suffering.

Thus I regard the doctrine that men are machines not only as mistaken, but as prone to undermine a humanist ethics. However, this very reason makes it all the more necessary to stress that the great defenders of that doctrine - all upholders of humanist ethics. From Democritus and Lucretius to Herbert Feigl and Anthony Quinton, materialist philosophers have usually been humanists and fighters for freedom and enlightenment; and, sad to say, their opponents have sometimes been the opposite. Thus just because I regard materialism as mistaken - just because I do not believe that men are machines or automata - I wish to stress the great and indeed vital role which the materialist philosophy has played in the evolution of human thought, and of humanist ethics”.

(Popper, Eccles 1977 pp. 4 – 5)

A existência de volição no comportamento humano, tal como foi hipostasiado, significa que o mundo não é um mecanismo obedecendo às regras formais de uma relojoaria cósmica, que actua omnipresente. Se fosse esse o caso, aceite o argumento da

variabilidade nomotética, o mecanismo pararia e o divino maquinista ver-se-ia perante uma falha inexorável, uma contradição qfulminando implacavelmente as engrenagens mais subtis do seu relógio.

Existe aparentemente um escolho que interrompe a progressão maquinal do relógio, indiferente a qualquer significado ou razão, e esse escolho é a volição humana, como se o livre arbítrio estivesse incorporado nos processos mais fundamentais da natureza.

Nesta acepção, o comportamento humano não depende de um conjunto de asserções apriorísticas, descrevendo padrões e regularidades, mas de uma perpetuação volitiva das regularidades em causa. A natureza humana posiciona-se na realidade como uma composição ordenada de escolhas sobre o soma físico que as protagoniza. Escolhas condicionadas pelas determinações desse soma, do meio e dos outros. Mas ainda assim escolhas.

“Nunca fomos tão livres como durante a ocupação alemã. Tínhamos perdido todos os nossos direitos, começando pela liberdade de expressão. Todos os dias éramos insultados e suportávamos em silêncio. A troco de qualquer pretexto - trabalhadores, judeus ou prisioneiros políticos - éramos deportados em massa para campos de concentração. Por todo o lado, nos escaparates na rua ou nos jornais encontrávamos a revoltante e insípida imagem de nós mesmos que os nossos opressores queriam que aceitássemos. E por causa de tudo isto éramos livres. Porque o veneno nazi se insidiava nas nossas mentes, cada pensamento acutilante era uma conquista. Porque a polícia poderosa nos silenciava, cada palavra tinha o valor de uma declaração de princípios. Porque éramos perseguidos e eliminados, cada um dos nossos gestos tinha o peso de um compromisso solene.

Exílio, prisão e em especial a morte tornaram-se os motivos habituais de preocupação (...) E cada escolha que cada um de nós fazia (...) era uma verdadeira escolha porque era feita em face da morte”.

(Sartre, 1949, p.2)

O que significa, então, ser na acepção humana? Iremos propôr que:

Ser é escolher.

Chamaremos a esta asserção o Princípio Ontológico da Escolha. Uma condição metalógica da existência que se escapa ao confinamento do discurso formal.

Volição e escolha são assim propriedades ontológicas que não podem ser escamoteadas dos processos de construção do mundo. O termo «ontológico» é utilizado neste contexto como sinonímia do que é fundamental e irrevogavelmente inultrapassável por qualquer tentativa de identificação da sua expressão com uma causa anterior.

A caracterização ontológica permanece ainda assim incompleta. Quando se escolhe é sempre em relação a algo que na estrutura da escolha surge no estamo mental intencional. Há, portanto, uma dimensão semântica que deve ser referida e que estrutura o conceito de propósito. Uma escolha perfaz-se a partir de uma teia de relações entre objectos distintos que são concatenadas com outras, através da volição. Afirmamos, por isso, que ser é também significar, no sentido de que o fluxo mental humano se dispõe como resultado de uma ligação entre estados mentais distintos, em momentos diferentes, de acordo com um propósito simbólico. A mente pode ter sobre si um propósito sancionador, tornando a intervenção volitiva um acto contemplativo e autorizando o curso dos acontecimentos simbólicos. Ou pode ter um propósito transformador, inteferindo nesse fluxo e reorganizando-se, através do circuito volitivo. É requerido, portanto, que se diga também que:

Ser é significar.

Elocução a que chamaremos Princípio Ontológico da Significação.

Mas parece haver mais. As escolhas, embora fundacionalmente livres, estão determinadas pelas condições possíveis da sua efectividade. Pelas imposições do soma, do meio e das escolhas de outros. Essas imposições afectam a estrutura do estado mental do sujeito. Este é obrigado a participar no jogo do possível, coadunando-se consigo e com a realidade externa. Chamaremos ao resultado desse movimento participatório, *concertação*, a capacidade do sujeito se harmonizar ou conflitar com as substâncias que se lhe impõem e diremos, finalmente, que:

Ser é concertar.

Elocução que nomeamos por Princípio Ontológico da Concertação.

A existência humana é, dessa maneira, um estado na realidade, que ordena o caos simbólico, prevalecendo volitivamente sobre o que é formalmente imposto. Aqueles que, existindo, ordenam volitivamente as suas visões do mundo, fazem-no produzindo significado, construindo novas concatenações simbólicas. Que ser é concertar significa que

a existência humana, existe necessariamente *com* e *em relação* a outro existente. Existe perante a condição iniludível de uma alteridade, com a qual deve lidar no decurso da efectivação das suas escolhas.

O mundo é feito de relações entre entidades e esse facto, associado ao que é prescrito nos princípios ontológicos anteriores, implica a efectividade de um estado de concertação do sujeito com todos os sistemas físicos na natureza, animados e inanimados e, em particular, de concertação do sujeito com o seu próprio soma.

O Princípio Ontológico da Concertação é o impeditivo da existência humana solipsista, garantindo que o sujeito encontra a objectividade do mundo. Aquele que escolhe, para que se manifeste como algo que existe; para que exerça a sua marca na realidade, fá-lo necessariamente na presença de uma alteridade. E também, na presença de si, reconhecendo-se no soma, mas também fora dele.

A chegada à existência é assim um acto simultaneamente sincrético e conectivo, por parte de todos os seres humanos na natureza. O universo de escolhas humanas possíveis pode ser mais ou menos extenso, em que o ser humano tem a possibilidade de se dissolver «para-fora-do-mundo»; de simplesmente sair de manifestação ou, pelo contrário, de se destituir – por escolha própria – do exercício da sua volição, escolhendo não escolher e recaindo na forma mais rígida de determinismo.

O estado de concertação, não significa, necessariamente a prevalência do que intuitivamente podemos reconhecer como ordem ou harmonia no mundo. Aquele que, existindo, concerta com alteridade, pode fazê-lo em termos destrutivos, com cortes e rupturas sobre imposições formais. Desse ponto de vista, escolher é destruir para reconstruir.

A última observação serve de mote para que reanalise os conceitos de estabilidade nomotética e de variabilidade nomotética no contexto da epistemologia da escolha. Quando o sujeito escolhe, passivamente, concede estabilidade nomotética ao seu comportamento. Porta-se como uma máquina. Quando, porém, escolhe activamente, releva-se enquanto sujeito humano, produzindo variabilidade nomotética sobre o fluxo do seu pensamento e, assim, sobre o seu comportamento físico. Fá-lo, como se viu, sem que isso implique contradições na máquina do mundo, onde o sujeito se inscreve como corpo.

Nesta dialéctica se inscreve a diferença entre necessidade e contingência, na filosofia aristotélica.

'Aristóteles lançou as bases para a atitude da filosofia para com a Vontade, e ao longo dos séculos a elasticidade delas resistiu aos testes e aos desafios mais cruciais. Segundo Aristóteles, todas as coisas que podem ser ou podem não ser, que aconteceram mas podiam não ter acontecido, são devidas ao acaso (...) - ou,

na tradução latina, acidentais ou contingentes - enquanto distintas do que necessariamente é como é, o que é e não pode não ser. Estas últimas, a que ele chamava o «*hypokeimenon*», jazem sob o que é adicionado pelo acaso, isto é, por tudo o que não pertence à verdadeira essência - tal como a cor é adicionada a objectos cuja essência é independente dessas «qualidades secundárias.» Os atributos que podem ou não ligar-se ao que jaz por baixo deles - o seu *substrato* ou *substância* (as traduções latinas de *hypokeimenon*) - são acidentais. Dificilmente pode existir algo mais contingente do que os actos voluntários, que - na pressuposição da vontade livre - podem todos ser definidos como actos dos quais sei que podia também tê-los deixado por fazer. Uma vontade que não é livre é uma contradição nos termos - a menos que compreendamos a faculdade da volição como um mero órgão auxiliar executivo de seja o que for que a razão ou o desejo propuseram. No quadro destas categorias, tudo o que acontece no reino dos assuntos humanos é acidental ou contingente («*prakton d'esti to endechomenon kai allos echein*», «o que é trazido à existência pela acção é o que poderia também ter sido de outra maneira»): as próprias palavras de Aristóteles indicam o baixo estatuto ontológico desse reino - um estatuto que nunca foi seriamente desafiado até à descoberta por Hegel do Sentido e Necessidade na História.

(...)

O produto humano, esse «composto de matéria e forma» - por exemplo, uma casa feita de madeira segundo uma certa forma pré-existente no espírito (*nous*) do artesão - é claro que não foi produzido a partir do nada, e por isso era compreendida por Aristóteles como pré-existindo «potencialmente» antes de ser actualizado por mãos humanas. Esta noção era derivada do modo de ser peculiar da natureza das coisas vivas, onde tudo o que aparece cresce a partir de algo que contém potencialmente o produto acabado, tal como o carvalho existe potencialmente na bolota e o animal no sêmen.

O ponto de vista que tudo o que é real deve ser precedido por uma potencialidade como uma das suas causas nega implicitamente o futuro como um autêntico tempo verbal: o futuro nada é além de uma consequência do passado, e a diferença entre as coisas naturais e aquelas que são produzidas pelo homem é apenas a diferença entre aquelas cujas potencialidades se tornam necessariamente actuais e aquelas que podem ou não ser actualizadas. Debajo destas circunstâncias, qualquer noção de Vontade como um órgão para o futuro, da mesma maneira que a memória é um órgão para o passado, era inteiramente supérflua; Aristóteles não tinha de estar consciente da existência da Vontade; os gregos «nem sequer têm uma palavra» para o que nós consideramos ser «o móbil principal da acção». (*Thelein* significa «estar pronto, estar preparado para alguma coisa» (...)) os autores muito versados em literatura grega estiveram sempre

conscientes desta lacuna. Assim Gilson considera como um facto bem conhecido «que Aristóteles não fala de liberdade nem de vontade livre ... o próprio termo falta».

(Arendt, 1978, pp. 22 - 23)

É difícil não concordar com Aristóteles, quando fala de formas pré-existentes, que são depois actualizadas. Essas formas corresponderiam, na nossa ontologia, a processos de prova possíveis no sistema formal ontológico. No entanto, isso não implica, como parece resultar a partir do pensamento aristotélico, na interpretação de Hannah Arendt que o universo seja um local onde a existência de vontade é supérflua.

De facto, se considerarmos válido o argumento da variabilidade nomotética aqui apresentado, parece que ocorre justamente o oposto. Que a existência de uma entidade ontológica formal determina a existência de Volição.

Talvez seja possível mostrar o contrário, que a existência de volição determina a existência e, não apenas a forma, de uma entidade formal, um problema que deixamos em aberto.

Fechamos esta secção reafirmando as três sentenças ontológicas propostas para o que é ser, na acepção humana.

O Princípio Ontológico da Escolha.

Ser é escolher.

O Princípio Ontológico da Significação.

Ser é significar.

O Princípio Ontológico da Concertação.

Ser é concertar.

## Conclusão

No início desta dissertação colocou-se o problema de saber qual seria a viabilidade da simulação artificial da inteligência humana, um dos objectivos principais da Inteligência Artificial. No decurso dos trabalhos de investigação construiu-se um argumento especulativo, usando-se como metáfora um sistema formal hipotético, no lugar de leis da natureza, e uma máquina “inteligente” no lugar de um agente humano. A nossa especulação conduziu-nos a uma contradição. E esta contradição conduziu – por sua vez - à hipótese de existência de uma propriedade na natureza, que identificámos com a *volição*.

Uma vez que, metaforicamente, vimos nas leis da natureza um programa de computador, o nosso argumento, caso se aceitem as suas premissas, parece desaproveitar a ideia de que a mente está para um programa de computador, como o cérebro está para um computador, nas palavras críticas do filósofo John Searle.

Se existir volição na natureza, essa propriedade actua paralelamente a qualquer forma de determinismo sobre os comportamentos. E como um programa de computador processado no cérebro seria uma instanciação determinista sobre os comportamentos do agente, essa possibilidade parece inviável, à luz da nossa conjectura.

A tese forte da IA parece assim colocada em causa.

Por outro lado, a tese fisicalista: a ideia de que uma combinação adequada de materiais físicos pode, de alguma maneira, fazer emergir comportamentos inteligentes volitivos, permanece em aberto. Podemos assumir que a combinação material de um cérebro, de um sistema nervoso e de um corpo constituem, justamente, um caso possível do tipo de estruturas procurado. Fica, no entanto, por explicar a natureza da interacção entre mente e corpo, entre estados mentais intencionais e estados neurológicos. Entre consciência e soma. No contexto da estrutura da escolha que sugerimos, fica definitivamente por explicar o modo como é que a volição, caso exista, pode ser actuável por um estado mental intencional, em consonância com um substracto cerebral que materializa a mente. Talvez o tipo de explicação se resuma apenas a afirmar a existência de um circuito volitivo, tal como foi proposto, inferível a partir de dados empíricos consistentes com essa hipótese. Talvez a explicação, de modo inesperado, resida noutro lado. De qualquer maneira, ainda à luz da nossa argumentação, parece viável uma teoria sobre a natureza que inclua a volição como uma das suas propriedades.

## Bibliografia

- Adamatzky, Andrew. (2002). *Collision-based Computing*. Springer-Verlag London Ltd
- Adleman, LM. (1994). «Molecular computation of solutions to combinatorial problems». *Science* 266, n. 5187: 1021 -1024.
- Agar, J. (2001). *Turing and the Universal Machine: The Making of the Modern Computer*. Icon Books Ltd.
- Albeverio, S., Jost, J., Paycha, S. & Scarlatti, S. (1997) *A Mathematical Introduction to String Theory: Variational Problems, Geometric and Probabilistic Methods*. Cambridge University Press.
- Amelino-Camelia, G., Kowalski-Glikman, J. (2005). *Planck Scale Effects in Astrophysics and Cosmology*. Springer.
- Arbib, M.. *Brains*. (1964). *Machines and Mathematics*. NY: McGraw Hill.
- Arendt, H. (1978). *A Vida do Espírito, Volume III – Querer*, Coleção Pensamento e Filosofia, Lisboa, Instituto Piaget.
- Ash, R. (1991). *Information Theory*. New Ed. Dover Publications Inc.
- Audi, R. (1997). *Epistemology: A Contemporary Introduction to the Theory of Knowledge*. Routledge.
- Bardeen, J. M., B. Carter & S. W. Hawking. (1973). «The four laws of black hole mechanics». *Comm. Math. Phys.*
- Barrow, J. (2003). *The Constants of Nature*. New Ed. Vintage.
- Barrow, J. (2005). *Theories of Everything: The Quest for Ultimate Explanation*. New Ed. Vintage,.
- Barrow, J. , Tipler, F. (1988) *The Anthropic Cosmological Principle*. New Ed. Oxford Paperbacks,.
- Bekenstein, J. (1972). «Black holes and the second law». *Nuovo Cimento Letters*.
- Bekenstein, J. (1974). «Generalized second law of thermodynamics in black-hole physics». *Physical Review D* 9, n. 12: 3292.
- Bekenstein, J. (1981). «Universal upper bound on the entropy-to-energy ratio for bounded systems». *Physical Review D* 23, n. 2: 287.

- Bennet, C. (1973). «Logical Reversibility of Computation». *IBM Journal of Research and Development*.
- Bickle, J., Mandik, P. & L., Anthony. (2012). "The Philosophy of Neuroscience", The Stanford Encyclopedia of Philosophy, Edward N. Zalta (ed.).
- Boden, M. (1977). *Artificial Intelligence and Natural Man*. Basic Books.
- Boolos, G., & Jeffrey, R. (1989). *Computability and Logic*. 3rd ed. Cambridge University Press.
- Bousso, R. (1999). «A covariant entropy conjecture». *Journal of High Energy Physics* 1999, n. 7: 004-004.
- Bousso, R. (2002). «The holographic principle». *Reviews of Modern Physics* 74, n. 3: 825-874.
- Braitenberg, V. (1984) *Vehicles: Experiments in Synthetic Psychology*. New Ed. MIT Press.
- Bremer, M. (2005). *Introduction to Paraconsistent Logics, An*. Peter Lang Pub Inc.
- Brooks, R. (1986). «A Robust Layered Control System for a Mobile Robot». *EEE Journal of Robotics and Automation*.
- Brooks, R. (1990). «Elephants Don't Play Chess». *Robotics and Autonomous Systems*.
- Brooks, R. (2003). *Flesh and Machines: How Robots Will Change Us*. Reprint. Vintage Books USA.
- Brooks, R. (2011). «Heartland Robotics - Designing and Building Next-Generation Industrial Robots», <http://www.heartlandrobotics.com>.
- Brooks, R. (2011). «iRobot Corporation», <http://www.irobot.com>.
- Brown, J. (1999). *Minds, Machines, and the Multiverse: The Quest for the Quantum Computer*. Simon & Schuster.
- Calmet, X., Graesser M. & Hsu, S. (2004). «Minimum Length from Quantum Mechanics and Classical General Relativity». *Physical Review Letters* 93, n. 21: 211101.
- Camazine, S. (2001). *Self-organization in Biological Systems*. Princeton University Press.
- Chaitin, G. (1987). *Algorithmic Information Theory*. Cambridge University Press.
- Chaitin, G. (1997). *The Limits of Mathematics: A Course on Information Theory and the Limits of Formal Reasoning*. National University of Singapore, Singapore.
- Chaitin, G. (1999). *The Unknowable*. National University of Singapore, Singapore.

- Chaitin, G. (2000). *Exploring Randomness*. Springer-Verlag.
- Chalmers, D. (1998). *The Conscious Mind: In Search of a Fundamental Theory*. New Ed. Oxford University Press Inc, USA.
- Chiara, M., Giuntini, R. (2001). «Quantum Logic». Cornell University Library, *quant-ph/0101028*. <http://arxiv.org/abs/quant-ph/0101028>.
- Comley, J., Dowe, D. (2005). «Minimum Message Length, MDL and Generalised Bayesian Networks with Asymmetric Languages». Em *Advances in Minimum Description Length: Theory and Applications*. Grünwald, P.; Myung, I. J. & Pitt, M. A. Cambridge: MIT Press.
- Copeland, J. (1993). *Artificial Intelligence: A Philosophical Introduction*. Wiley Blackwell.
- Copeland, J. (2004). *The Essential Turing: Seminal Writings in Computing, Logic, Philosophy, Artificial Intelligence, and Artificial Life plus The Secrets of Enigma*. Oxford University Press, New York.
- Cori, R., Lascar, D. (2001) *Mathematical Logic: A Course with Exercises: Recursion Theory, Gödel's Theorem, Set Theory and Model Theory Pt.2*. Oxford University Press.
- Crevier, D. (1994). *AI: The Tumultuous History of the Search for Artificial Intelligence*. New edition. Basic Books.
- Dancy, J., Sosa, E. (1993) *A Companion to Epistemology*. New Ed. WileyBlackwell.
- Davis, M. (2004). *The Undecidable: Basic Papers on Undecidable Propostions, Unsolvble Problems and Computable Functions*. Dover Ed. Dover Publications Inc.
- Dawson, M. (1998). *Understanding Cognitive Science*. WileyBlackwell.
- Dennett, D. (1993). *Consciousness Explained*. New Ed. Penguin Books Ltd.
- Dennett, D. (2004). *Freedom Evolves*. New Ed. Penguin Books Ltd.
- Deutsch, D. (1985). «Quantum Theory, the Church-Turing Principle and the Universal Quantum Computer». *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences* 400, n. 1818: 97 -117.
- Deutsch, D. (1998). *The Fabric of Reality: Towards a Theory of Everything*. New Ed. Penguin Books Ltd.
- Deutsch, D. (2012). *The Beginning of Infinity: Explanations That Transform the World* (Reprint edition.). London; New York: Penguin Books.

- Dobado, A. (2004). «An elementary introduction to the Holographic Principle». Em *Proceedings of the Fundamental Physics Meeting «Alberto Galindo»*. Madrid2004. <http://arxiv.org/abs/hep-ph/0506027>.
- Dretske, F. (1999) *Knowledge and the Flow of Information*. New Ed. Cambridge University Press.
- Dreyfus, H. (1965). *Alchemy and artificial intelligence*. Rand Corp.
- Dreyfus, H. (1972). *What Computers Can't Do*. New York: MIT Press.
- Dreyfus, H. (1991). *Being in the World: Division 1: Commentary on Heidegger's «Being and Time»*. MIT Press.
- Dreyfus, H. (1992). *What Computers Still Can't Do: A Critique of Artificial Reason*. Revised edition. MIT Press.
- Dreyfus, H., Dreyfus, S. (1988). «Making a Mind Versus Modeling a Brain: Artificial Intelligence Back at a Branchpoint». Em *The Artificial Debate: False Starts, Real Foundations*, 15-43. MIT Press.
- Duff, M., Okun, L. & Veneziano, G. (2002). «Dialogue on the number of fundamental constants». *Journal of High Energy Physics* 2002.
- Edman, I., Schneider, H. (2010). *Arthur Schopenhauer: The World as Will and Idea*. Kessinger Publishing.
- Erickson, G., Fossa, J. (1998). *Dictionary of Paradox*. University Press of America.
- Everett, H. (1957). «Relative State Formulation of Quantum Mechanics». *Reviews of Modern Physics*, n. 29: 454-462.
- Feynman, R. (1982). «Simulating physics with computers». *International Journal of Theoretical Physics* 21, n. 6-7: 467-488.
- Feynman, R. (1990). *QED - The Strange Theory of Light and Matter*. Penguin Books Ltd.,
- Feynman, R. (1992). *The Character of Physical Law*. New Ed. Penguin Books Ltd.
- Feynman, R. (1999). *Six Not-so-easy Pieces: Einstein's Relativity, Symmetry and Space-time*. New Ed. Penguin Books Ltd.
- Feynman, R., Hey, A. (2000). *Feynman Lectures On Computation*. 1st ed. Westview Press.
- Fodor, J. (1975). *The language of thought*. Harvard University Press.
- Franklin, S. (1997). *Artificial Minds (Bradford Book)*. New Ed. MIT Press.

- Fredkin, E. (1992a). «A New Cosmogony». Em *Workshop on Physics and Computation*, 116-121. Dallas, TX.
- Fredkin, E. (1992b) «Finite nature». Em *Proceedings of the XXVIIth Rencontre de Moriond Series: Moriond Workshops*. Les Arcs, Savoie, France: Editions Frontieres.
- Fredkin, E. (2011). «Digital Philosophy». <http://www.digitalphilosophy.org>.
- Gaberdiel, M. (1999). «An Introduction to Conformal Field Theory». arXiv:hep-th/9910156.
- Gambini, R., Pullin, J. (1999). «Nonstandard optics from quantum space-time». *Physical Review D* 59, nº 12.
- Garay, L. (1995). «Quantum gravity and minimum length». *International Journal of Modern Physics A (IJMPA)*. nº 10.
- Gardner, M. (1970). «Mathematical Games - The fantastic combinations of John Conway's new solitaire game "life"». *Scientific American*.
- Gershenfeld, N. (2011). «The Center for Bits and Atoms». <http://cba.mit.edu/>.
- Gödel, K. (1931). *On Formally Undecidable Propositions of «Principia Mathematica» and Related Systems*. New Ed. Dover Publications Inc.
- Gödel, K. (1979). *O Teorema de Gödel e a hipótese do contínuo: antologia*. Fundação Calouste Gulbenkian.
- Gomperz, T. (2012). *Greek Thinkers: A History of Ancient Philosophy, Vol. 1*. Forgotten Books.
- Gram, T., Bornholdt, S., Gro, M., Mitchell M. & Pellizzari, T. (1998). *Non-standard Computation: Molecular Computation-cellular Automata Evolutionary Algorithms-quantum Computers*. Wiley VCH.
- Grand, S. (2001). *Creation: Life and how you make it: Life and How to Make It*. New Ed. Phoenix.
- Gruska, J. (1999). *Quantum Computing*. Osborne/McGraw-Hill, U.S.
- Gurney, K. (1997). *Introduction to Neural Networks*. Routledge.
- Haggard, P. (2008). Human volition: towards a neuroscience of will. *Nature Reviews Neuroscience*, 9(12), 934–946.
- Haken, H. (2000). *Information and Self-organization: A Macroscopic Approach to Complex Systems*. 2nd ed. Springer-Verlag Berlin and Heidelberg GmbH & Co. K.
- Hameroff, S. (1987). *Ultimate Computing*. 1st ed. North Holland.

- Hameroff, S. (2000). *Toward a Science of Consciousness: The Third Tucson Discussions and Debates*. MIT Press.
- Hameroff, S. (2012). «How quantum brain biology can rescue conscious free will». *Frontiers in integrative neuroscience*, eCollection.
- Hameroff, S., Penrose, R. (1996). «Conscious events as orchestrate space-time selections». *Journal of Consciousness Studies*, nº1, Vol. 3.
- Haugeland, J. (1989). *Artificial Intelligence: The Very Idea*. New edition. MIT Press.
- Hawking, S. (1974). «Black hole explosions?» *Nature* 248, n. 5443: 30-31.
- Hawking, S. (1976). «Breakdown of predictability in gravitational collapse». *Physical Review D* 14, n. 10: 2460-2473.
- Hawking, S. (2001). *The Universe in a Nutshell*. Bantam Press.
- Hawking, S., Penrose, R. (1970). «The Singularities of Gravitational Collapse and Cosmology». *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences* 314, n. 1519: 529-548.
- Hawking, S., Penrose, R. (2000). *The Nature of Space and Time*. New Ed. Princeton University Press.
- Heidegger, M. *Being and Time*. (1978). New edition. Wiley-Blackwell.
- Heijenoort, J. (1999). *Frege and Godel: Two Fundamental Texts in Mathematical Logic*. iUniverse.com.
- Hofstadter, D. (2000). *Godel, Escher, Bach: An Eternal Golden Braid*. 20th ed. Penguin Books Ltd.
- Holland, J. (2000). *Emergence: From Chaos to Order*. New Ed. Oxford University Press.
- Houaiss, A. (2001). *Dicionario Houaiss da lingua portuguesa*. 1ª ed. Objetiva.
- Hrynkiw, D., Tilden, M. (2002). *JunkBots, Bugbots, and Bots on Wheels: Building Simple Robots With BEAM Technology*. 1st ed. McGraw-Hill Osborne Media.
- Huang, Kerson. (1988). *Statistical Mechanics*. 2nd ed. John Wiley & Sons.
- Ishiguro, H. (2011). «HIL - Hiroshi Ishiguro Lab». <http://www.geminoid.jp/en/index.html>.
- Islam, J. (2001). *An Introduction to Mathematical Cosmology*. 2nd ed. Cambridge University Press.
- Jaynes, E. (1957a). «Information Theory and Statistical Mechanics». *Physical Review* 106: 620-630.

- Jaynes, E. (1957b) «Information Theory and Statistical Mechanics. II». *Physical Review* 108, n. 2: 171-190.
- Jibu, M. (1995). *Quantum Brain Dynamics and Consciousness: An Introduction*. John Benjamins Publishing Co.
- Kleene, S. (2009). *Introduction to Metamathematics*. (1952). Ishi Press.
- Kolak, D, Thomson G. (2005). *The Longman Standard History of 20th Century Philosophy*. 1st ed. Pearson Education.
- Kosko, B. (1994). *Fuzzy Thinking: The New Science of Fuzzy Logic*. New Ed. Flamingo.
- Kozen, D. (2007). *Automata and Computability*. Springer-Verlag New York Inc.
- Laurent, P., Götz, D., Binétruy, P., Covino, S. & Fernandez-Soto, A. (2011). «Constraints on Lorentz Invariance Violation using integral/IBIS observations of GRB041219A». *Physical Review D* 83, n. 12: 121301.
- Leff, H., Rex, A. (2002). *Maxwell's Demon: Entropy, Information, Computing*. 2nd ed. Institute of Physics Publishing.
- Libet, B. (1985). Unconscious cerebral initiative and the role of conscious will in voluntary action. *Behavioral and Brain Sciences*, 8(04), 529–539.
- Lloyd, S. (1999). «Ultimate physical limits to computation». *Nature*, n. 406: 1047-1054.
- Lloyd, S. (2006). *Programming the Universe: A Quantum Computer Scientist Takes on the Cosmos*. Jonathan Cape.
- Lourenço, M. (2004). *Os Elementos do Programa de Hilbert*. Lisboa. Centro de Filosofia da Universidade de Lisboa.
- Lucas, J. (1963). «Minds, Machines and Gödel (1959)». Em *The Modeling of mind. Computers and intelligence*. Kenneth M. Sayre and Frederick J. Crosson. Notre Dame Press.
- Lucas, J. (1970). *The Freedom of the Will*. 1st ed. OUP Oxford.
- Luger, G. (2001). *Artificial Intelligence: Structures and Strategies for Complex Problem Solving*. 4th ed. Addison Wesley.
- Magueijo, J. (2003). *Faster Than the Speed of Light: The Story of a Scientific Speculation*. William Heinemann Ltd.
- Maldacena, J. (1998). «The Large N Limit of Superconformal Field Theories and Supergravity». *Adv.Theor.Math.Phys.*2:231-252.

- Mavromatos, N. (1996). «Eluding the No-Hair Conjecture for Black Holes». arXiv:gr-qc/9606008.
- McCarthy, J, Minsky, M., Rochester N. & Shannon, C. (2011) «A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence (Dartmouth Conferences, August 1955)». <http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html>.
- McCarthy, J. (2007). «What is Artificial Intelligence?». <http://www-formal.stanford.edu/jmc/whatisai/whatisai.html>.
- McFadden, J. (2000). *Quantum Evolution: Life in the Multiverse*. New Ed. Flamingo.
- Messiah, A. (2000). *Quantum Mechanics 2 Volumes*. Dover Publications Inc.
- Michalewicz, Z. (1996). *Genetic Algorithms + Data Structures = Evolution Programs*. 3rd ed. Springer-Verlag Berlin and Heidelberg GmbH & Co. K.
- Minsky, M. & Papert, S. (1969). *Perceptrons: An Introduction to Computational Geometry*. Cambridge, MA: MIT Press.
- Mitchell, M. (1998). *An Introduction to Genetic Algorithms*. New Ed. MIT Press.
- Mohr, P., Taylor B. & Newell D. (2008). «CODATA recommended values of the fundamental physical constants: 2006». *Reviews of Modern Physics* 80, n. 2: 633.
- Moravec, H. (1990). *Mind Children: The Future of Robot and Human Intelligence*. New edition. Harvard University Press.
- Moravec, H. (1998). *Robot: Mere Machine to Transcendent Mind*. Oxford University Press Inc.
- Moriyasu, K. (1984). *An Elementary Primer for Gauge Theory*. World Scientific Publishing Co Pte Ltd.
- Mumford, S. (2004). *Laws in Nature*. Routledge.
- Nagel, E, Newman, J. (2001). *Godel's Proof*. Rev. Ed. New York University Press.
- Nielsen, M., Chuang, I. (2000). *Quantum Computation and Quantum Information*. Cambridge University Press.
- Nilsson, N. (1986). «Probabilistic Logic». *Artificial Intelligence*. North-Holland, Elsevier Science Publishers B.V. edição.
- Nilsson, N. (1998). *Artificial Intelligence: A New Synthesis*. I.S.ed. Morgan Kaufmann.

- Øhrstrøm, P., Hasle, P. (2011). "Future Contingents", *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.),  
 URL=<http://plato.stanford.edu/archives/sum2011/entries/future-contingents>.
- Osborne, M. (1994). *A Course in Game Theory*. MIT Press.
- Papantonopoulos, E. (2009). *Physics of Black Holes: A Guided Tour*. 1st ed. Springer.
- Penrose, R. (1990). *The Emperor's New Mind: Concerning Computers, Minds and the Laws of Physics*. New Ed. Oxford Paperbacks.
- Penrose, R. (1994). *Shadows of the Mind: A Search for the Missing Science of Consciousness*. Oxford University Press, New York, USA.
- Penrose, R. (2000). *The Large, the Small and the Human Mind (Canto)*. New Ed. Cambridge University Press.
- Penrose, R. (2004). *The Road to Reality: A Complete Guide to the Laws of the Universe*. Jonathan Cape.
- Plunkett, K., McLeod, P. & Rolls, T. (1998). *Introduction to Connectionist Modelling of Cognitive Processes*. Pap/Dsk. Oxford University Press.
- Poppel, E. (1988). *Mindworks: Time and Conscious Experience* (1st edition.). Boston: Harcourt.
- Popper, K., Eccles, J. (1984). *The Self and Its Brain: An Argument for Interactionism* (Reprint edition.). London; New York: Routledge, 2003.
- Putnam, H. (1960). «Minds and Machines». Em *Dimensions Of Mind A Symposium*, 20-33. Sidney Hook. New York University Press.
- Putnam, H. (1988a). «Reflexive Reflections». Em *Rerepresentation: Readings in the Philosophy of Mental Representation (Philosophical Studies Series)*, 211-222. Stuart Silvers (Tilburg University, The Netherlands). London: Kluwer Academic Publishers.
- Putnam, H. (1988b). *Representation and Reality (Representation & Mind)*. MIT Press.
- Rasmusen, (2001). E. *Games and Information: An Introduction to Game Theory*. 3rd ed. WileyBlackwell.
- Reichenbach, H. (1999). *Philosophic Foundations in Quantum Mechanics*. New Ed. Dover Publications Inc.
- Reichenbach, H. (2003). *The Philosophy of Space and Time*. Dover Publications Inc.

- Rosenblatt, F. (1957), *The Perceptron--a perceiving and recognizing automaton*. Report 85-460-1, Cornell Aeronautical Laboratory.
- Rosenblatt, F. (1962). *Principles of Neurodynamics*. Washington, DC:Spartan Books.
- Ross, D. (2003). «Changing Contours of the social science disciplines». Em *The Cambridge History of Science, Volume 7: The Modern Social Sciences*, 205-237. Cambridge University Press.
- Ross, S. (2004). *Aristotle* (6th edition.). London; New York: Routledge.
- Rovelli, C. (2004). *Quantum Gravity*. Cambridge University Press.
- Rovelli, C., Smolin, L. (1995). «Discreteness of area and volume in quantum gravity». *Journal-ref: Nucl.Phys. B*, n. 442: 593-622.
- Rucker, R. (2004). *Infinity and the Mind: The Science and Philosophy of the Infinite*. New Ed. Princeton University Press.
- Russell, B. (1989). *Analysis of Mind*. New Ed. Routledge.
- Russell, B. (1992). *The Analysis of Matter*. New Ed. Routledge.
- Sainsbury, R. (1995). *Paradoxes*. 2nd ed. Cambridge University Press.
- Sakurai, J. (1993). *Modern Quantum Mechanics, reissued 2nd Ed*. 1st ed. Addison Wesley.
- Sartre, J. P. (1949). *Situations III*, Éditions Gallimard.
- Schmidhuber, J. (2000). «Algorithmic Theories of Everything». Cornell University Library, *quant-ph/0011122*. <http://arxiv.org/abs/quant-ph/0011122>.
- Schmidhuber, J. (2002). «The Speed Prior: A New Simplicity Measure Yielding Near-Optimal Computable Predictions». Em *Proceedings of the 15th Annual Conference on Computational Learning Theory (COLT 2002)*, 216-228. Lecture Notes in Artificial Intelligence. Sydney, Australia: Springer. <http://arxiv.org/abs/quant-ph/0011122>.
- Schopenhauer, A. (2003). *On the Fourfold Root of the Principle of Sufficient Reason (1813)*. Open Court Publishing Company.
- Schopenhauer, A. (2010). *The World as Will and Representation Vol 1 & 2 (1819)*. Dover.
- Schutz, B. (1985). *A First Course in General Relativity*. Cambridge University Press.
- Searle, J. (1983). *Intentionality: An Essay in the Philosophy of Mind (Cambridge Paperback Library): An Essay in the Philosophy of Mind*. Cambridge University Press.
- Searle, J. (1992). *Minds, Brains and Science: The 1984 Reith Lectures*. New Ed. Penguin Books Ltd.

- Shannon, C. (1948). «A Mathematical Theory of Communication». *Bell System Technical Journal*,.
- Siegelmann, H. (1998). *Neural Networks and Analog Computation: Beyond the Turing Limit*. Birkhauser.
- Siegfried, T. (2001). *The Bit and the Pendulum: From Quantum Computing to M Theory - The New Physics of Information*. New Ed. John Wiley & Sons.
- Smolin, L. (2001a). «The strong and weak holographic principles». *Journal-ref: Nucl.Phys. B*, n. 601: 209-247.
- Smolin, L. (2001b). *Three Roads to Quantum Gravity: A New Understanding Of Space, Time And The Universe: A New Understanding of Space, Time and the Universe*. New Ed. Phoenix.
- Smolin, L. (2004). «An invitation to loop quantum gravity». *Reviews of Modern Physics*. <http://arxiv.org/abs/hep-th/0408048>.
- Smolin, L. (2008). *The Trouble with Physics: The Rise of String Theory, The Fall of a Science and What Comes Next*. Penguin.
- Solomonoff, R. (1964a). «A Formal Theory of Inductive Inference, Part I». *Information and Control* 7, n. 1: 1-22.
- Solomonoff, R. (1964b) «A Formal Theory of Inductive Inference, Part II». *Information and Control* 7, n. 2: 224-254.
- Susskind, L. (1995). «The World as a Hologram». *J.Math.Phys* 36: 6377-6396.
- Susskind, L. (2008). *The Black Hole War: My Battle with Stephen Hawking to Make the World Safe for Quantum Mechanics*. Little, Brown and Company.
- Susskind, L., Lindesay, J. (2005). *An Introduction to Black Holes, Information and the String Theory Revolution: The Holographic Universe*. World Scientific Publishing Co Pte Ltd.
- Susskind, L., Thorlacius, L. & Uglum, J. (1993). «The stretched horizon and black hole complementarity». *Physical Review D* 48, n. 8: 3743-3761.
- t Hooft, G. (1985). «On the quantum structure of a black hole». *Nuclear Physics B* 256, n. 0: 727-745.
- t Hooft, G. (1990). «The black hole interpretation of string theory». *Nuclear Physics B* 335, n. 1: 138-154.
- t Hooft, G. (1993). «Dimensional Reduction in Quantum Gravity». Cornell University Library, [gr-qc/9310026](http://arxiv.org/abs/gr-qc/9310026).

- Tanaka, K. (1996). *An Introduction to Fuzzy Logic for Practical Applications*. Springer-Verlag New York Inc..
- Taylor, E., Wheeler, J. (2000). *Exploring Black Holes: An Introduction to General Relativity*. Addison Wesley.
- Toffoli, T., Margolus, N. (1987). *Cellular automata machines: a new environment for modeling*. MIT Press
- Tomilin, K. (1999). «Natural Systems of Units. To the Centenary Anniversary of the Planck System.» Em *Proceedings Of The XXII Workshop On High Energy Physics And Field Theory*. Protvino, Russia.
- Turing, A. (1936). «On Computable Numbers, with an Application to the Entscheidungsproblem». *Proceedings of the London Mathematical Society. Second Series*.
- Turing, A. (1950). «Computing Machinery and Intelligence» *MIND*, vol. LIX. N. 236.
- Tymoczko, T. (1998). *New Directions in the Philosophy of Mathematics: An Anthology (Revised and expanded edition): An Anthology*. 2nd ed. Princeton University Press.
- Unruh, W. (1976). «Notes on black-hole evaporation». *Physical Review D* 14, n. 4: 870.  
URL = <<http://plato.stanford.edu/archives/sum2012/entries/neuroscience/>>.
- von Neumann, (1996). *J. Mathematical Foundations of Quantum Mechanics*. New Ed. Princeton University Press.
- von Neumann, J., Burks, A. (2002). *Theory of Self-Reproducing Automata*. UMI Reprint University Illinois 1966 Ed.
- Wald, R. (1999). *Black Holes and Relativistic Stars*. University of Chicago Press.
- Walker, E. (2000). *The Physics of Consciousness: The Quantum Mind and the Meaning of Life*. New Ed. Perseus Books,U.S.
- Wardman, A, Creed, J., Bambrough, R. & Bobzien, S. (2011). *The Philosophy of Aristotle* (Reissue edition.). New York, N.Y: Signet Classics.
- Weizenbaum, J. (1977). *Computer Power and Human Reason: From Judgement to Calculation*. New edition. W.H.Freeman & Co Ltd.
- Wheeler, J. (1990). «Information, physics, quantum: The search for links». Em *Complexity, Entropy, and the Physics of Information*. W. Zurek. Redwood City, CA: Addison-Wesley.

- White, F., (1991). *On Schopenhauer's «Fourfold Root of the Principle of Sufficient Reason»*. Brill.
- Whitehead, A., Russell, B. (1927). *Principia Mathematica 3 Volume Set: v. 1-3 (1910, 1912, 1913)*. 2nd ed. Cambridge University Press.
- Wick, D. (1998). *The Infamous Boundary: Seven Decades of Heresy in Quantum Physics*. 1st ed. 1995. Corr. 2nd printing ed. Springer.
- Wiener, N. (1961). *Cybernetics: Or Control and Communication in the Animal and the Machine*. 2nd ed. MIT Press.
- Wiener, N. (1988). *The Human Use Of Human Beings: Cybernetics And Society*. New York, N.Y: Da Capo Press.
- Witten, E. (1998). «Anti De Sitter Space And Holography». arXiv:hep-th/9802150.
- Wittgenstein, L. (1973). *Philosophical Investigations (1929 - 1949)*. New edition. Wiley-Blackwell.
- Wittgenstein, L., (2001). *Tractatus Logico-Philosophicus (1921)*. 2nd ed. Routledge.
- Woit, P. (2007). *Not Even Wrong: The Failure of String Theory and the Search for Unity in Physical Law*. Basic Books.
- Wolfram, S. (2002). *A New Kind of Science*. Wolfram Media Inc,US.
- Young, N. (1988). *An Introduction to Hilbert Spaces*. Cambridge University Press.
- Zizzi, P., (2003). «Emergent Consciousness: From the Early Universe to Our Mind». *NeuroQuantology* 3: 295-311.
- Zurek, W. (1990). *Complexity, Entropy and the Physics of Information*. Westview Press.
- Zuse, K. (1970). *Calculating Space - MIT Technical Translation*. Massachusetts Institute of Technology. Cambridge.

## Índice Remissivo

- Ackermann, 98  
 Adamatzky, 59, 60, 162  
 adimensionalização das constantes, 191, 228  
 Adleman, 61, 162  
 Agar, 162  
 Albeverio, 162  
 algoritmos genéticos, 60, 61, 62  
 alteridade, 158  
 Amelino-Camelia, 162, 202  
 análise de otimização dos meios para um fim, 48  
 Arbib, 61, 162  
 área de Planck, 73, 91, 92, 185, 199, 206, 207, 208, 214, 218, 222, 224  
 Arendt, 160, 162  
 argumento da variabilidade nomotética, 7, 119, 156, 160, 180, 264  
 Aristóteles, 97, 98, 145, 147, 158, 159, 160  
 Armer, 14  
 Ash, 162  
 Aspect, 66  
 Audi, 162  
 autómatos celulares, 60, 63, 67  
 axiomas, 18, 69, 97, 98, 101, 109, 111, 114, 142  
 background dependent, 90  
 Bardeen, 70, 71, 162, 182  
 Barrow, 162  
 Bekenstein, 70, 71, 72, 74, 162, 185  
 Bennet, 60, 163  
 Bernays, 98  
 Bickle, 131, 134, 135, 137, 138, 163  
 Big Bang, 227  
 Binétruy, 168  
 Bohr, 29, 221  
 Boltzmann, 57, 59, 71, 183, 184, 186, 190, 191, 194, 198, 214  
 Boole, 240  
 Boolos, 163  
 Born, 89  
 Bornholdt, 166  
 bosões, 87  
 bottom-up, 30  
 Bousso, 72, 73, 74, 163, 185  
 Braitenberg, 62, 163  
 Bremer, 163  
 Brillouin, 58  
 Broglie, 200  
 Brooks, 14, 30, 31, 32, 62, 163  
 Brouwer, 97, 99  
 Brown, 163, 172  
 buraco negro de Schwarzschild, 72, 203, 204, 208, 209, 212, 228, 230  
 Burks, 60, 173  
 cálculo da formulação, 112, 113, 115  
 Calmet, 163, 202  
 Camazine, 163  
 Cantor, 95, 97  
 Carter, 70, 162  
 Case-Based Reasoning, 49  
 causalidade física, 132  
 causalidade lógica, 132, 134, 141, 147, 148, 152  
 causalidade simbólica, 132, 153  
 Center for Bits and Atoms, 62, 166  
 Chaitin, 53, 68, 163, 164  
 Chalmers, 164  
 Chiara, 164  
 Chuang, 66, 169  
 Church, 51, 66, 164  
 circuito volitivo, 141, 142, 149, 151, 153, 154, 157, 161  
 classe de observadores, 219, 220, 222, 223, 224, 225, 234, 236, 245, 255  
 colapso, 64, 65, 73, 203  
 Comley, 164, 186  
 compactação, 90  
 computação distribuída, 63  
 computação evolutiva, 62  
 computação irreversível, 59  
 computação molecular, 61  
 computação por choques elásticos, 59  
 computação reversível, 59, 60  
 computador quântico universal, 63, 65, 66, 69  
 concatenação, 13, 106, 107, 109, 111, 112, 114, 117, 118, 131, 132, 144, 153, 154  
 concatenações de comportamentos físicos, 13, 133, 154  
 concatenações de estados mentais, 149, 154  
 concatenações de processos de prova, 152, 154  
 condição de inteligibilidade lógica da natureza, 217, 218  
 condição de não contradição na natureza, 217, 218  
 conexionista, 45, 50, 51  
 configuração, 57, 58, 110, 111  
 conhecimento de senso comum, 16, 47  
 conjectura da correspondência AdS/CFT, 86, 87  
 constante de estrutura fina, 228  
 constante de gravitação universal, 200  
 constante de Napier, 190  
 constante de Planck, 200, 227  
 context-free elements, 46  
 contiguidade lógica, 139, 140, 141, 144, 148, 149, 151, 153, 154  
 contingente, 159, 249, 250, 263  
 coordenação holográfica, 211, 232, 234, 235, 236, 238, 239, 240, 241, 242, 243, 244, 246, 247, 248, 250, 251, 252, 253, 257, 259  
 Copeland, 164  
 Cori, 164  
 correlação, 75, 213  
 Covino, 168  
 Crevier, 11, 164  
 critério lógico de realismo, 220, 221, 222, 223, 224, 225, 234, 239, 241, 242, 245, 262  
 critério lógico-formal para a validade de uma proposição sobre a natureza, 240, 262  
 crossover, 61  
 Dancy, 164  
 Davis, 164  
 Dawson, 164  
 d-branes, 87  
 De Interpretatione, 145  
 Dedekind, 95, 97  
 demónio de Maxwell, 58  
 Dennett, 164  
 denumerável, 33, 110, 120, 124, 128, 180, 239, 251

- Descartes, 46  
 desenvolvimento de base 2, 110, 118, 119, 120, 122, 123, 124, 126, 128  
 desigualdade de Heisenberg, 89, 203, 212  
 Deutsch, 54, 63, 66, 67, 164  
 diagonalização, 116, 118, 119, 120, 121, 122, 124, 125, 126, 128, 180, 232, 257  
 Diodorus Cronus, 145, 146, 148  
 dissociação lógica, 139, 142, 143, 148, 151, 153  
 Dobado, 88, 165  
 Dowe, 164, 186  
 Dretske, 165  
 Dreyfus, 12, 14, 15, 16, 17, 44, 46, 47, 48, 49, 50, 165  
 Duff, 165  
 Dyson, 89  
 ecrã holográfico, 198, 199, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 218, 219, 220, 222, 225, 226, 229, 232, 233, 234, 245, 246  
 Edman, 165  
 efeito de Unruh, 182  
 Einstein, 66, 165, 184  
 Eliza, 20, 24  
 ENIAC, 52  
 enredamento quântico, 66  
 entropia, 52, 57, 58, 59, 60, 70, 71, 72, 73, 74, 75, 76, 82, 83, 92, 181, 182, 185, 186, 187, 190, 191, 193, 194, 195, 196, 197, 198, 208, 209, 210, 211, 213, 214, 218, 227  
 entropia de von Neumann, 75, 76  
 Entscheidungsproblem, 34, 51, 173  
 equação diferencial de Schrödinger, 65  
 Erickson, 165  
 espaço de fase semântico, 138  
 espaço de Hilbert, 64, 182, 183, 184, 191, 192, 194  
 esquema, 47  
 esquema axiomático, 102  
 esquema da escolha, 145, 147, 149, 150, 151, 153  
 estabilidade nomotética, 158  
 estado de correlação, 66  
 estado misturado, 75  
 estado puro, 75, 76  
 estrutura da escolha, 136, 152, 157, 161  
 estrutura semântica, 132, 135, 140, 153  
 Euclides, 97  
 Everett, 165, 229  
 experiência da casa chinesa, 24  
 expressão formal, 106  
 expressões funcionais, 107  
 fermiões, 87  
 Fernandez-Soto, 168  
 Feynman, 53, 63, 89, 165  
 filosofia digital, 53, 69  
 fisicalismo, 23, 25  
 fisicalismo naïve, 23, 25  
 fitness, 31, 61  
 Fodor, 27  
 fórmula, 101, 102, 103, 104, 109, 110, 111, 112, 113, 114, 116, 117, 120, 124, 125, 126, 200  
 fórmulas identicamente verdadeiras, 114  
 Fossa, 165  
 Franklin, 165  
 Fredkin, 53, 59, 60, 69, 166  
 Fredkin gate, 60  
 Frege, 97, 167  
 função nomotética, 115, 116, 119, 122, 123, 125, 126, 247, 248, 249, 250, 252, 255, 261  
 funcionalismo, 23, 26, 27  
 Futuros Contingentes, 145, 146, 147, 148, 149  
 Gambini, 166  
 Garay, 166, 202  
 Gardner, 60, 166  
 Gell-Mann, 89  
 General Problem Solver, 49  
 gerador de sequências numeráveis, 67  
 Gershenfeld, 63, 166  
 Giuntini, 164  
 Gödel, 17, 18, 19, 36, 37, 166, 168, 185  
 Graesser, 163  
 Gram, 166  
 Grand, 166  
 Green, 89  
 Gro, 166  
 Grover, 66  
 Gruska, 166  
 Gurney, 166  
 Guth, 227  
 Haggard, 42, 141, 166  
 Haken, 166  
 Hameroff, 37, 56, 166, 167  
 Hasle, 145, 146, 147, 170  
 Haugeland, 17, 167  
 Hawking, 70, 71, 73, 162, 167, 172, 185, 208  
 Hebb, 45  
 Heidegger, 14, 15, 165, 167  
 Heijenoort, 167  
 Heisenberg, 89, 200, 227  
 Hey, 165  
 Hilbert, 34, 51, 92, 98, 99, 100, 105, 123, 129, 168, 174, 183, 198  
 hipótese da finitude da natureza de Fredkin, 53  
 hipótese dos sistemas físicos simbólicos, 31, 45, 49  
 Hofstadter, 167  
 Holland, 62, 166, 167, 169  
 horizonte do buraco negro, 71, 72, 182, 203, 204, 227, 229  
 Houaiss, 167  
 Hrynkiw, 62, 167  
 Hsu, 163  
 Huang, 167, 186, 190, 191  
 Husserl, 47  
 identidade ontológica entre realidade e computação, 67  
 inconsistente simples, 109, 117, 119, 126  
 informação física, 91, 92, 185, 196, 197, 198, 209, 210, 211, 212, 213, 214, 215, 221, 222, 223, 225, 226, 227, 229, 232, 233, 245, 247  
 inteligência artificial forte, 23  
 intuicionismo, 98, 99  
 irreduzibilidade algorítmica, 68  
 Ishiguro, 167  
 Islam, 167  
 it from bit, 52, 67  
 Jaynes, 57, 167, 168  
 Jeffrey, 163  
 Jibu, 168  
 jogo da vida de Conway, 60  
 Jordan, 89  
 Jost, 162  
 justaposição, 130  
 Kaluza-Klein, 90

- Kant, 46  
 Kleene, 12, 18, 95, 100, 102, 105, 168  
 Kolak, 168  
 Kolmogorov, 68  
 Kosko, 168  
 Kowalski-Glikman, 162, 202  
 Kozen, 168  
 Kronecker, 97  
 Landauer, 58, 196  
 Landreth, 131, 134, 135, 137, 138  
 Lascar, 164  
 Laurent, 168  
 Lavenham, 146  
 Leff, 58, 59, 168  
 Leibniz, 46  
 Levin, 67  
 Libet, 41, 42, 141, 149, 150, 151, 168  
 limite de von Neumann-Landauer, 59, 197  
 limite entrópico covariante, 74  
 limite entrópico esférico, 73  
 limite entrópico espacial, 73  
 Lindesay, 70, 72, 90, 172, 210  
 Lloyd, 54, 67, 168  
 Lourenço, 98, 166  
 Lucas, 14, 17, 19, 168  
 Luger, 168  
 M.I.T., 14, 20, 49, 62  
 Magueijo, 168, 227  
 Maldacena, 86, 89, 168  
 Mandik, 131, 134, 135, 137, 138, 163  
 máquina de Turing, 33, 34, 35, 56, 60, 63, 66, 67, 68  
 máquina generalizada de Turing, 67  
 máquina universal de Turing, 54, 61  
 Margolus, 60, 173  
 Markopoulou-Kalamara, 91  
 Maxwell, 58, 64, 168  
 McCarthy, 11, 12, 44, 169  
 McClelland, 50  
 McCulloch, 45  
 McFadden, 169  
 mecânica digital, 69  
 mentalismo, 23  
 Merleau-Ponty, 14  
 Messiah, 89, 169  
 meta-análise, 99  
 metamatemática, 13, 18, 99, 102, 105, 112, 116, 117, 118, 119, 123, 125  
 metateoria, 13, 18, 100, 101, 103, 105, 106, 107, 108, 123  
 métrica de Schwarzschild, 89, 202, 227  
 Michalewicz, 62, 169  
 microestados, 182, 190, 191, 192, 193, 195  
 micro-mundos, 49  
 Minsky, 11, 44, 46, 47, 49, 50, 169  
 Mitchell, 166, 169  
 Mohr, 169  
 Moravec, 62, 169  
 Moriyasu, 169  
 Mumford, 169  
 Nagel, 169  
 Nambu, 89  
 não computabilidade, 33, 37, 263  
 neurónio de McCulloch-Pitts, 45  
 Newell, 31, 45, 46, 48, 49, 169  
 Newman, 169  
 Newton, 60, 64  
 Nielsen, 66, 89, 169  
 Nilsson, 169  
 noções primitivas, 106  
 noema, 47  
 noesis, 47  
 no-hair theorem, 70  
 Øhrstrøm, 145, 146, 147, 170  
 Okun, 165  
 onda, 33, 64, 65, 66, 68, 200, 201, 202  
 Orch-OR, 32, 37  
 Organon, 145  
 Osborne, 166, 167, 170  
 Palestras Reith, 23  
 Papert, 46, 49  
 paradoxo da informação, 74, 77, 83, 88  
 paradoxo holográfico, 181, 212, 216, 217, 218, 220, 222, 223, 229, 232, 260  
 Peano, 97  
 Peirce, 97  
 Pellizzari, 166  
 Penrose, 14, 19, 32, 33, 34, 37, 38, 44, 46, 48, 51, 52, 56, 91, 100, 101, 136, 137, 138, 145, 146, 150, 167, 170  
 percepção, 46  
 Pitágoras, 98  
 Pitts, 45  
 Platão, 46  
 Plunkett, 170  
 Podolsky, 66  
 Polchinski, 89  
 Polyakov, 89  
 postulados, 81, 83, 101, 104, 109, 218, 239, 241, 262  
 potenciais evocados, 41, 42, 43, 149, 150, 151  
 predicação composta, 109, 111, 114  
 predicação simples, 107, 108, 111  
 premissas, 101, 118, 119, 120, 125, 161  
 Princípio da Complementaridade dos buracos negros, 81  
 Princípio da Equivalência Computacional, 60, 61  
 princípio da indução, 97  
 Princípio da Resolução Holográfica, 181, 220, 222, 223, 224, 225, 226, 231, 232, 233, 234, 239, 241, 242, 245, 253, 257, 258, 260, 262  
 Princípio da Sobreposição de Ondas, 64  
 Princípio de Deutsch-Turing, 63, 67  
 Princípio de Equivalência, 80  
 Princípio de Landauer, 56, 59  
 Princípio do Terceiro Excluído, 239  
 Princípio Holográfico, 74, 85, 88, 91, 92, 181, 197, 207, 211, 232, 233, 234, 241, 245, 262  
 Princípio Ontológico da Concertação, 157, 158, 160  
 Princípio Ontológico da Escolha, 156, 160  
 Princípio Ontológico da Significação, 157, 160  
 problema da normalização, 89  
 problema da representação do conhecimento de senso comum, 28  
 problema das Torres de Hanoi, 49  
 problema do grafo de sete vértices de Hamilton, 61  
 problema do horizonte, 227  
 procedimento de validação lógica, 113  
 procedimento metamatemático composto, 117, 118  
 procedimento metamatemático são, 117  
 procedimento nomotético, 13, 115, 116, 117, 118, 119, 122, 123, 125, 126, 129, 143

- procedimentos finitistas, 18, 52, 99, 102  
 processamento distribuído paralelo, 50  
 processo de prova, 13, 104, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 122, 124, 125, 127, 128, 129, 131, 134, 139, 140, 141, 148, 153, 154  
 processo de Susskind, 72  
 proposição contingente, 232, 245, 246, 248, 249, 250  
 proposição indecidível, 239  
 proposição não computável, 258, 259  
 proposição universal, 232, 245, 246, 247, 248, 249, 250, 251, 252, 253, 255, 256, 257, 258, 263  
 prova formal, 102  
 Pullin, 166  
 Putnam, 14, 26, 27, 28, 29, 170  
 quantidade mínima de entropia, 56, 59, 197, 199, 207, 209, 210, 216  
 qubit, 65, 66, 183  
 radiação de Hawking, 71, 75, 76, 80, 182, 231  
 RAND, 14  
 Rasmusen, 170  
 Rechenberg, 62  
 rede de spins, 91  
 rede neuronal, 45, 50, 56, 61  
 regime neurológico, 153  
 regras de derivação, 98, 99, 102, 109, 112, 118, 123  
 regras de inferência, 18, 101, 111, 120, 130  
 regras de transformação, 13, 104, 130, 139, 141, 148, 152, 153  
 Reichenbach, 26, 170  
 resolução do comprimento determinável, 205, 206, 212  
 resolução holográfica, 220, 222, 223, 224, 225, 226, 227, 228, 229, 230, 232, 233, 234, 235, 236, 238, 239, 242, 245, 246, 247, 248, 252, 255, 257, 260, 261  
 Rex, 58, 59, 168  
 Richard, 53, 63, 89, 96  
 Rochester, 11, 44, 169  
 Rolls, 170  
 Rosen, 66  
 Rosenblatt, 45, 46  
 Rosenbloom, 49  
 Ross, 171  
 Rovelli, 90, 171  
 Rucker, 171  
 Rumelhart, 50  
 Russell, 23, 171, 174  
 Sainsbury, 171  
 Sakurai, 171  
 Sartre, 156, 171  
 Scarlatti, 162  
 Schank, 49, 50  
 Scherk, 89  
 Schmidhuber, 53, 67, 68, 69, 171  
 Schneider, 165  
 Schopenhauer, 165, 171, 174  
 Schrödinger, 65  
 Schutz, 171  
 Schwarz, 89  
 Schwefel, 62  
 Schwinger, 89  
 Searle, 12, 14, 23, 24, 25, 171  
 segunda lei da termodinâmica, 52, 57, 58, 71  
 Segundo Princípio Generalizado da Termodinâmica, 71, 72, 73  
 Self, 142, 163, 166, 170, 173  
 Shannon, 11, 44, 57, 169, 172, 186, 187, 193, 195  
 Shor, 66  
 SHRDLU, 49  
 Siegelmann, 172  
 Siegfried, 66, 172  
 símbolo de proposição, 108  
 Simon, 31, 45, 46, 48, 163  
 sistema de unidades de Planck, 196, 198, 214, 228  
 sistema formal, 12, 13, 18, 25, 34, 35, 41, 51, 55, 93, 95, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 111, 112, 117, 118, 119, 123, 126, 129, 130, 131, 132, 133, 134, 135, 136, 139, 140, 141, 142, 143, 147, 148, 149, 151, 152, 153, 154, 160, 161, 181  
 sistema formal lógico, 217, 231  
 sistema lógico-formal, 18, 29  
 sistemas físicos, 31, 41, 45, 49, 51, 52, 53, 55, 56, 58, 59, 61, 63, 70, 74, 92, 94, 100, 103, 105, 110, 111, 152, 158, 180, 182, 199, 212, 215, 219, 227, 228, 233, 234, 241, 246  
 Smolin, 54, 90, 91, 92, 171, 172, 181, 185, 197, 198, 207, 210, 211, 212, 215  
 SOAR, 49  
 Solomonoff, 68, 172  
 soma, 64, 71, 132, 134, 135, 136, 141, 142, 143, 153, 156, 157, 158, 161  
 Sosa, 164  
 subsumption architecture, 30, 31  
 supersimetria, 87  
 Susskind, 54, 70, 72, 73, 78, 89, 90, 91, 172, 185, 210  
 Szilárd, 58  
 t Hooft, 78, 91, 172, 211  
 Tanaka, 173  
 Taylor, 89, 169, 173  
 teorema, 17, 18, 32, 49, 101, 102, 109, 110, 111, 112, 113, 114, 115, 116, 117, 120, 122, 124, 125, 126, 128, 186  
 teorema da focagem, 210  
 teorema da variabilidade nomotética, 181, 232, 250, 253, 263  
 teoremas da incompletude, 17  
 teoria cosmológica de velocidade da luz variável, 228  
 teoria da electrodinâmica quântica, 89  
 teoria da inferência indutiva, 68  
 teoria da inflação cósmica, 227  
 teoria da redução objectiva concertada, 37  
 teoria de campo conformal, 86  
 teoria final sobre a natureza, 17, 219, 231, 232, 234, 239, 240, 241, 242, 245, 246, 247, 249, 250, 251, 253, 254, 259, 263  
 teoria psicosemântica, 137  
 teoria quântica correcta da gravidade, 33  
 teoria ramificada dos tipos, 96  
 teorias algorítmicas de tudo, 53, 67, 70  
 teorias de cordas, 89, 90  
 termos, 16, 17, 18, 32, 61, 68, 97, 101, 102, 107, 158, 159, 183, 197, 213, 216, 217, 223, 224, 229, 233, 241, 246, 251, 253, 263  
 tese da estabilidade nomotética, 219, 231, 232, 234, 241, 245, 250, 251  
 tese de Church-Turing, 66

- tese fisicalista da inteligência artificial, 25, 30, 46, 62, 69, 231
- tese forte da inteligência artificial, 24, 26, 31, 34, 231
- Thomson, 168
- Tilden, 62, 167
- Tipler, 162
- Toffoli, 59, 60, 173
- Toffoli gate, 60
- Tomilin, 173
- Tomonaga, 89
- transformações unitárias, 65, 66
- Turing, 16, 33, 34, 36, 56, 61, 67, 162, 164, 172, 173, 185
- Tymoczko, 173
- universo elementar, 49
- universos paralelos, 229
- Unruh, 173, 182
- valor mínimo da área de referência, 222, 223, 224, 227
- variabilidade nomotética, 158, 219, 227, 231, 232, 250, 263, 264
- vector de onda, 64, 65, 66, 200, 221
- velocidade apriorística de um universo, 68, 69
- Veneziano, 89, 165
- veto, 42, 150
- vida artificial, 62
- volição, 6, 12, 13, 17, 21, 41, 42, 43, 95, 103, 127, 130, 131, 133, 134, 136, 137, 139, 140, 141, 142, 143, 149, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 180, 181, 264
- von Neumann, 52, 59, 60, 75, 98, 100, 173, 196, 197
- Wald, 173
- Walker, 173
- Webb, 228
- Weierstrass, 95, 97
- Weizenbaum, 12, 14, 20, 21, 22, 24, 173
- Wheeler, 67, 173
- White, 174
- Whitehead, 18, 97, 174
- Wick, 65, 66, 174
- Wiener, 52, 62, 174
- Witten, 87, 89, 90, 174
- Wittgenstein, 174
- Wolfram, 53, 60, 61, 174
- Yoneya, 89
- Young, 64, 174
- Zizzi, 174
- Zurek, 173, 174
- Zuse, 53, 63, 67, 174

## APÊNDICE 1

### Uma versão fisicalista do Argumento da Variabilidade Nomotética

Numa primeira versão da dissertação ambicionámos assertar a existência de volição na natureza, sob a forma de uma propriedade metalógica e ao ponto de encontrar expressão em todos os fenómenos da natureza, atribuindo-se a possibilidade de escolha a todos os sistemas físicos. A versão actual atribui essa capacidade apenas ao agente humano, dispondo de um cérebro e de um corpo. As razões pelas quais se decidiu diminuir a amplitude da hipótese prendem-se com o desenho da argumentação então seguida.

A primeira razão é a seguinte. O argumento da variabilidade nomotética, cujo significado se preservou, obteve-se na primeira versão através de um procedimento de diagonalização sobre um écran holográfico, no sentido de Smolin. Esse écran holográfico, uma superfície imaginária, através da qual se supunha ser transferida informação física, teria de ser infinitamente subdivisível numa infinidade de regiões menores. O conjunto das regiões obtidas teria a mesma potência do denumerável que o respectivo conjunto de coordenadas holográficas. O argumento, desenvolvido deste modo, no entanto, apresenta debilidades. A primeira, desde logo, é a de que não parece existir qualquer evidência empírica desse facto e mesmo que venha a haver, a sua pressuposição será sempre indutiva. Sucede que o procedimento de diagonalização utilizado, a exemplo do que se fez na versão actual, utiliza o conceito de infinidade. E a pressuposição da existência de infinidade, feita por um agente humano, estratégia a que se recorreu na versão actual, é substancialmente diferente da afirmação de que essa infinidade existe na natureza. Atribuir á natureza e ao espaço físico essa propriedade acabou por nos parecer excessivo.

Na primeira versão da dissertação tentou-se justificar a possibilidade da existência de regiões de espaço, infinitamente pequenas, com um raciocínio lógico aplicado sobre uma combinação teórica especulativa entre Relatividade Generalizada e Mecânica Quântica. Essa combinação originou uma contradição lógica a que chamámos Paradoxo Holográfico e do qual extraímos o Princípio da Resolução Holográfica, cujo enunciado permitia a subdivisão pretendida. Embora ainda sustentemos, com honestidade intelectual, a hipótese da correcção do Princípio, a ausência de evidência empírica ou até mesmo o risco da sua impossibilidade, fez-nos optar por reduzir o protagonismo que lhe estava atribuído.

A segunda razão foi a de que o nosso argumento requeria ou recomendava que a teoria sobre a natureza fosse completa, de modo a que a sua capacidade de previsão sobre a existência de todas as regiões infinitesimais de espaço fosse ali previsível. Esta exigência

era e é excessiva; e mostrava-se inclusivamente desnecessária, se adoptássemos outro eixo de argumentação.

A terceira razão colocou-se também no plano conceptual. Embora a metáfora de que *processos de prova correspondem a comportamentos físicos*, utilizada na presente versão, não tivesse sido expressa na primeira, a metáfora de que *o universo está para um computador como um programa está para as leis da natureza*, comparecia. No entanto, esgrimia-se no contexto de uma teoria informal sobre a natureza, embora a imaginássemos, hipoteticamente, completa. Uma teoria capaz de descrever, se não todos, a maior parte dos fenómenos na natureza. Ocorreu-nos então que essa teoria, não sendo um sistema formal, teria associadas as interpretações dos agentes que a formularam. Estaria, por assim dizer, impregnada das convicções intencionais dos agentes que descrevia e, como tal, seria uma representação inadequada para uma entidade ontológica procedendo como um computador. Uma vez mais, a via formal de argumentação parecia mais adequada e, conseqüentemente, decidimo-nos por um eixo argumentativo incidindo sobre o conceito de sistema formal, corrigindo-se outros aspectos que nos tinham sido apontados.

Chegámos, fundamentalmente, à mesma ideia: a de que é possível conceber a existência de volição, como propriedade metalógica na natureza, problematizando a viabilidade da simulação artificial da inteligência humana.

O presente Apêndice testemunha os raciocínios elaborados na primeira versão, e entretanto, pelas razões acabadas de aduzir, ultrapassados. Constitui, assim, um repositório útil das reflexões progressivamente percorridas e conducentes à versão actual da dissertação.

## **1. Princípio da Resolução Holográfica**

### **1.1. Entropia termodinâmica e informação binária.**

No que se segue, recorre-se à relação entre entropia termodinâmica e informação binária confrontando-a com a relação entrópica definida por Lee Smolin (Smolin, 2001a). Da consideração conjunta das duas relações surgirá uma contradição a que se chamará «paradoxo holográfico». Procurando-se ultrapassar o paradoxo holográfico será então proposto o Princípio da Resolução Holográfica, cujo enunciado será utilizado posteriormente para construir o «argumento da variabilidade nomotética».

No contexto da formulação do Princípio Holográfico, a quantidade de informação especificando um sistema físico está associada, por um lado, à quantidade de entropia

termodinâmica desse sistema e, por outro, ao número de dimensões do espaço de Hilbert que lhe corresponde. Procura-se uma descrição fundamental do sistema em que as propriedades atômicas da matéria descritas quanticamente (e.g. valores de spin de um conjunto de sistemas físicos componentes formando uma rede de Planck tridimensional) são relacionadas com propriedades macroscópicas mensuráveis, como a temperatura e a massa, através do tratamento termodinâmico dos buracos negros (Bardeen *et al.*, 1973).

A temperatura  $T$ , a variação da massa  $dM$  e a variação da entropia  $dS_{BN}$  de um buraco negro estão relacionadas através da expressão:

$$(1.1) \quad dM = TdS_{BN}$$

A variação da massa  $dM$  pode ser observada através dos efeitos gravitacionais que lhe estão associados. Por outro lado, a temperatura  $T$  do buraco negro pode ser medida por um observador acelerando no sentido inverso à atracção gravitacional perto do horizonte do buraco negro. Essa temperatura está relacionada com a energia da nuvem de fótons em que o observador se vê envolto. A temperatura é proporcional à aceleração, facto conhecido pelo «efeito de Unruh», descrito pelo físico canadiano Bill Unruh [1945 -] e relacionado com radiação de Hawking, cuja detecção proporcionaria também uma medição possível da temperatura do buraco negro.

A questão crucial, subentendendo o estabelecimento de relações entre propriedades atômicas e propriedades macroscópicas da matéria, consiste em saber qual é o número mínimo de estados independentes que descrevem completamente, ao nível fundamental da matéria, um sistema físico confinado a uma região limitada de espaço.

Este número deverá ser igual ao número de dimensões do espaço de Hilbert associado ao sistema, uma vez que este contém todos os estados distintos e independentes que aquele pode assumir, tal que cada uma das suas dimensões representa a possibilidade efectiva de um estado.

A definição clássica de entropia termodinâmica é a de que, a menos de uma constante, aquela é igual ao logaritmo natural do número de configurações ou de microestados distintos que um sistema físico pode assumir.

Desse modo, torna-se natural definir que, a nível fundamental, a entropia termodinâmica  $S$  de um sistema físico cumprirá o seguinte:

$$(1.2) \quad S = k \ln \dim(H)$$

Em que  $\dim(H)$  representa a dimensão do espaço de Hilbert e  $k = 1,38065 \times 10^{-23} \text{ J.K}^{-1}$  é a constante de Boltzmann.

No âmbito da computação quântica sabemos que um sistema atômico pode ser utilizado para codificar informação, utilizando-se o *qubit* em lugar do conceito clássico de *bit* para a representar. A dimensão de um espaço de Hilbert é definida considerando todos aqueles estados físicos cuja observação pode ser feita; tais que a sua representação no espaço de Hilbert não pode ser reduzida a um conjunto de outros estados independentes.

Um *qubit* representa os dois estados diferenciáveis, independentes e observáveis - nomeáveis por exemplo 1 e 0 ou «sim» e «não» - que permitem definir um *bit*, indexando-se a cada um deles uma dada amplitude de probabilidade.

A atribuição final do valor de probabilidade 1 a um dos estados, mediante observação nas condições adequadas, significa imediatamente a probabilidade nula do outro. A propriedade de sobreposição de estados de um sistema atômico significa que até que uma observação seja feita, os dois estados independentes possíveis definindo um *bit* podem contribuir simultaneamente para a dinâmica do sistema. Os dois estados permanecerão simultaneamente representados no espaço de Hilbert.

O ganho em termos de informação implementável é apreciável se considerarmos em vez de um, dois *qubit*. Teremos assim a possibilidade de obter quatro combinações possíveis de estados: 00; 01; 10 e 11. Até que uma observação se efective sobre o sistema, aqueles permanecerão simultaneamente presentes no espaço de Hilbert e na evolução dinâmica de qualquer computação quântica que os envolva. Num sistema clássico apenas uma e só uma dessas combinações pode ser implementável para efeitos computacionais.

Existe um duplo ganho em termos da quantidade de informação armazenável e da optimização temporal através do processamento paralelo numa computação quântica.

No contexto desta discussão, não serão considerados processos computacionais envolvendo a quantidade de informação representável num espaço de Hilbert, associado a um sistema físico. Pretende-se simplesmente determinar a quantidade de informação necessária para especificar um sistema físico, limitado espacialmente, num dado momento da sua existência.

Por outras palavras, é apenas relevante saber qual é a quantidade de informação necessária para descrever um sistema físico, num dado instante, restringindo a uma região espacial limitada. Na ausência de uma teoria unificando gravitação e fenómenos atômicos, podemos tentar estimar a resposta mediante uma aproximação envolvendo campos quânticos. Um campo quântico é constituído por um ou mais osciladores em cada ponto do espaço contínuo. Mesmo um oscilador harmónico apresenta um espaço de Hilbert cuja

dimensão é infinita. Por outro lado, num espaço contínuo, o número de pontos num dado volume é infinito, por mais pequeno que seja.

O número de estados parece pois ser infinito.

Esta linha de raciocínio, porém, ignora quaisquer efeitos gravitacionais.

Para alcançar uma estimativa finita, começamos por considerar que distâncias inferiores ao comprimento de Planck  $l_p = 1,61624 \times 10^{-35}$  m não podem ser observadas numa teoria unificada, de acordo com o que é proposto pela teoria da gravidade de loops quânticos ou ainda pelas teorias das cordas. Definimos uma rede tridimensional formada por volumes de Planck;  $l_p^3$ , e associamos um oscilador quântico a cada um daqueles volumes.

O espectro de cada um desses osciladores é discreto e limitado inferiormente em consequência do valor finito e não nulo de cada volume de Planck. Por outro lado, o mesmo espectro será limitado superiormente em consequência da energia de Planck  $E_p = 1,95609 \times 10^9$  J, a qual define, mediante a equivalência entre energia e massa, a quantidade de energia máxima que pode existir num volume de Planck sem que aí se forme um buraco negro.

Desse modo, o número de osciladores, num volume total formado por  $V$  volumes de Planck, é justamente  $V$ , visto que existe um oscilador por cada um daqueles volumes.

Cada oscilador poderá apresentar um espectro formado por  $n$  estados, pelo que a quantidade total de estados quânticos independentes  $N$  num volume de espaço  $V$ , deverá ser:

$$(1.3) \quad N = n^V$$

Identificando  $N$  com a dimensão do espaço de Hilbert associado ao sistema sob descrição, encontramos e recorrendo a (1.2), obtemos:

$$(1.4) \quad S = \ln N = V \ln n \geq V$$

Relação que é genérica para qualquer volume, se convencionarmos, em consequência do procedimento de adimensionalização adoptado, que para as unidades de Planck, a constante de Boltzmann é  $k = 1$ , resultando igualmente  $l_p^3 = 1$ .

A última equação substancia a nossa intuição de que a quantidade de estados necessários para especificar um sistema físico, numa dada região limitada de espaço, deve

umentar em função do volume ocupado pelo sistema (Bousso, 2002).

Em consequência dos trabalhos de Bekenstein, Hawking, Susskind, Bousso e Smolin, no domínio da termodinâmica dos buracos negros, de onde resulta um limite entrópico de características universais, essa intuição é falsa. Com efeito, a primeira forma do limite entrópico é a de que a entropia termodinâmica de qualquer sistema físico, confinado a uma região limitada de espaço, não deve exceder um quarto da área, medida em unidades de Planck, da menor superfície que o envolve completamente. Uma unidade de área de Planck sendo igual a  $l_p^2 = 2,61226 \times 10^{-70} \text{ m}^2$ , escrita em unidades SI ou igual a 1, escrita em unidades de área de Planck.

Analogamente, a forma do limite entrópico sugerida por Smolin, associada à teoria da gravidade de loops quânticos, estabelece que a entropia termodinâmica correspondente a qualquer sistema físico, cuja observação seja feita através de uma região bidimensional, não pode exceder um quarto da área dessa superfície, medida em unidades de Planck.

A proposta de Smolin, que visa evitar um conjunto de dificuldades apontadas às formas anteriores do limite entrópico, assume que a área de uma superfície através da qual se observa um sistema físico é uma medida da capacidade de transferência de informação física através do espaço (Smolin, 2001a). Esta será a interpretação do limite entrópico que adoptaremos aqui. A ideia de que a observação de fenómenos físicos envolve transferência de informação e - considerada uma entidade ontológica entre realidade e computação - o processamento da mesma, torna possível uma análise das condições lógico-formais que caracterizam o tratamento dessa informação e, desse modo, a análise da própria observação como fenómeno.

A nossa tarefa imediata consiste em recensear certos limites lógicos do pensamento aplicado à inteligibilidade da realidade e deduzir os limites fundamentais correspondentes que se impõem às nossas capacidades de observação. Este tipo de análises metateóricas foi explorado por Turing e Gödel, mas sem uma relação directa com a realidade física, permitindo resultados de natureza estritamente lógica. A tentativa aqui recai agora sobre uma teoria informal sobre a natureza.

O limite entrópico estabelece uma relação entre entropia termodinâmica e a área, uma grandeza física familiar. Para que possamos efectivar uma análise com consequências físicas previsíveis sobre o que é observável no mundo, procedemos agora ao estabelecimento de uma relação entre entropia termodinâmica e informação binária.

Desse modo passamos a dispor da possibilidade de relacionar informação com área, isto é; de relacionar um conceito abstracto, passível de integrar uma análise metateórica, com uma grandeza física mensurável.

Seguindo requisitos de rigor e completude na exposição, começaremos por

enunciar a definição de entropia  $H$  associada à transmissão de uma mensagem proposta por Claude Shannon [1916 - 2001] (Shannon, 1948), relacionando-a depois com a expressão para a entropia termodinâmica. Shannon estudou o caso em que uma fonte de informação produz mensagens formadas por um número inteiro de símbolos, através do que chamou «canais discretos de transmissão».

Num alfabeto de  $m$  símbolos  $S_i$ , admita-se uma sequência formada por  $n$  desses símbolos, a que se chamará mensagem de comprimento  $n$ , transmitida através de um canal de transmissão.

Considere-se ademais que cada símbolo  $S_i$  ocorre  $n_i$  vezes em cada uma das mensagens de comprimento  $n$  transmitidas, tal que se tem:

$$(1.5) \quad \sum_{i=1}^m n_i = n$$

Seja  $p_i = \frac{n_i}{n}$  a probabilidade de ocorrência do símbolo  $S_i$  na transmissão de uma mensagem de comprimento  $n$ . Nestas condições, a definição de entropia  $H$  associada à transmissão de uma mensagem de comprimento  $n$ , formulada por Shannon, escreve-se:

$$(1.6) \quad H = -K \sum_{i=1}^m p_i \ln p_i$$

Em que  $K$  é uma constante positiva que tem como objectivo definir as unidades em que a entropia  $H$  é medida, em função do contexto em que é utilizada. Se  $K$  for igual à constante de Boltzmann,  $H$  será a expressão no teorema- $H$  da termodinâmica, reportando-se a fenómenos estudados no âmbito daquela (Huang, 1988).

Se  $K$  for a unidade, com o logaritmo de base natural  $\ln$ ,  $H$  será medida em *nat* por símbolo do alfabeto utilizado (Comley, Dowe, 2005), associada à transmissão de mensagens de comprimento  $n$ . A presença das probabilidades  $p_i$  tem como objectivo introduzir na expressão da entropia o efeito de alguma regra que as altera e assim faz variar o grau de desordem medido por  $H$ .

Se  $m = 2$ , isto é, se forem utilizados apenas dois símbolos, e se a sua ocorrência na mensagem for independente e, desse modo, equiprovável, situação em que a entropia  $H$

será máxima; tem-se  $n = n_1 + n_2$  e  $n_1 = n_2 = \frac{n}{2}$ , pelo que  $p_i = \frac{n_i}{n} = \frac{(n/2)}{n} = \frac{1}{2}$  e a entropia  $H$  resulta:

$$(1.7) \quad H = -K \sum_{i=1}^n \frac{n_i}{n} \ln \left( \frac{n_i}{n} \right) = K \frac{n}{2} \ln 2 = \frac{K}{2} \ln 2^n$$

Este resultado mostra que a função  $H$ , neste caso, representa a média logarítmica referente da produção de mensagens de comprimento  $n$ , num espaço de  $2^n$  mensagens possíveis, por unidade do número de símbolos no alfabeto utilizado. Do texto de Shannon (1948) define-se quantidade de informação  $I$  transmitida por uma fonte discreta, numa mensagem de comprimento  $n$ , através da expressão:

$$(1.8) \quad I = K \ln M$$

Em que  $M$  representa o número de mensagens de comprimento  $n$  possíveis e  $K$  é uma constante positiva relacionada com as unidades de informação utilizadas.

Considerando um alfabeto com dois símbolos e fazendo  $M = 2^n$ , (1.8) pode ser escrita:

$$(1.9) \quad I = K \ln 2^n$$

Comparando com (1.7), vemos que:

$$(1.10) \quad I = 2.H$$

A informação associada à transmissão de uma mensagem de comprimento  $n$  é então o valor da entropia  $H$  multiplicada pelo número de símbolos no alfabeto. Para a definição de uma relação entre unidades de informação distintas que seja útil para os nossos propósitos, procedemos do seguinte modo. Consideramos genericamente que a informação, associada à transmissão de uma mensagem de comprimento  $n$  pode ser descrita adequadamente por qualquer expressão do tipo:

$$(1.11) \quad I_{u(b)}(M) = \log_b M \ u(b) ; b \in ]1, +\infty[$$

Tal que  $M$  é o número de mensagens de comprimento  $n$  possíveis e em que  $u(b)$  representa a notação das unidades em que a quantidade de informação é medida. Por razões que se tornarão claras adiante, definimos no presente contexto que o número mínimo  $M$  de mensagens de comprimento  $n$  necessárias para descrever, ao nível fundamental da matéria, qualquer sistema físico, é tal que:

$$(12) \quad M = 2^n$$

Desse modo tem-se:

$$(1.13) \quad I_{u(b)}(M) = \log_b 2^n ; b \in ]1, +\infty[$$

Verificamos em seguida que para cada uma das unidades  $u(b)$ , a qual depende da base  $b$  do logaritmo, se verifica:

$$(1.14) \quad I_{u(b)}(M) = \log_b M = 1 \ u(b) \Leftrightarrow M = b^1$$

E ainda que:

$$(1.15) \quad I_{u(b)}(M) = \log_b b^a = a \ u(b) ; M = b^a , b \in ]1, +\infty[ , a \in R^+$$

Em que o número  $a$  representa o valor da quantidade de informação para  $M = b^a$  nas unidades  $u(b)$  em que essa informação é expressa.

Definimos agora que a unidade de informação associada à transmissão de uma mensagem contendo um só símbolo, num alfabeto de dois símbolos, denominada «*bit*», sendo:

$$(1.16) \quad I_{bit}(2) = \log_2 2^1 = 1 \ bit \text{ em que } u(2) \equiv bit$$

Tal que  $M = 2^n$ .

A representação da quantidade de informação escrita *bit*, também chamada informação binária, associada à recepção de uma mensagem de comprimento  $n$ , resultará, com  $M = 2^n$ :

$$(1.17) \quad I_{bit}(2^n) = \log_2 2^n = n \log_2 2 = n \cdot 1 \text{ bit} = n \text{ bit}$$

Para efeitos de conversão entre unidades de quantidade de informação, comparamos no contexto deste trabalho o *bit* com outras unidades, para valores distintos da base  $b$  no logaritmo. Para tal notamos que o conjunto das mensagens de comprimento  $n$  possíveis deverá ser o mesmo, independentemente das unidades em que a quantidade de informação associada à mensagem é expressa. O valor de  $M$  será o mesmo.

Pretendemos uma expressão de conversão entre qualquer unidade de quantidade de informação e o *bit*. Dispomos então a seguinte sequência de identidades, válidas para qualquer base  $b$ :

$$(1.18) \quad \log_b M = \log_b M \Leftrightarrow \log_b 2^{\log_2 M} = \log_b M \Leftrightarrow \\ \log_2 M \cdot (\log_b 2) = \log_b M \Leftrightarrow \\ \log_2 M = \frac{\log_b M}{\log_b 2}$$

Introduzindo nesta expressão unidades em função das respectivas bases logarítmicas:

$$(1.19) \quad \log_2 M \text{ bit} = \frac{1}{\log_b 2} \log_b M u(b)$$

Ou seja:

$$(1.20) \quad I_{bit}(M) = \frac{1}{\log_b 2} I_{u(b)}(M)$$

Uma unidade de interesse corresponde ao caso em que  $b$  é igual à constante de Napier ou número de Euler;  $e \approx 2,71828$ .

A quantidade de informação  $I_{nat}(M) = \ln M$  diz-se expressa em unidades naturais e a respectiva unidade é designada por  $u(e) \equiv nat$ , tendo-se:

$$(1.21) \quad I_{bit}(M) = \frac{1}{\ln 2} I_{nat}(M)$$

Esta relação significa que para um conjunto de  $M$  mensagens possíveis de comprimento  $n$ , a cada *bit* de informação recebida corresponde necessariamente uma quantidade de informação de  $\ln 2$ , medida em *nat*.

Cumpre entretanto compreender em que é que consiste a entropia termodinâmica de um sistema físico, de modo a que a possamos relacionar com quantidade de informação, expressa em *bit*. A entropia termodinâmica de um sistema fechado é uma grandeza física que está associada à temperatura e à energia desse sistema e cuja propriedade fundamental, no âmbito da termodinâmica, é a de que consiste numa grandeza cujo valor aumentará, ao longo da evolução temporal do sistema, até atingir um valor máximo, consistente com as restrições físicas externas impostas ao sistema. Para efeitos desta discussão consideramos a formulação de entropia termodinâmica obtida no domínio da mecânica estatística, onde a mesma é conhecida por «entropia de Boltzmann».

Considere-se um sistema físico formado por átomos tal que cada átomo pode assumir um estado quântico, distinto e independente de todos estados restantes. Cada conjunto de estados na classe, resultante da combinação dos estados acessíveis a todos os átomos no sistema, define um «microestado quântico». Uma visão realista de um sistema macroscópico consiste em assumir que este efectua incessantemente transições extremamente rápidas entre os microestados quânticos que lhe estão associados e que de qualquer medição macroscópica sobre o mesmo resulta um valor médio das grandezas físicas que caracterizam esses estados (Huang, 1988). Uma vez que as transições ocorridas são a consequência de processos puramente aleatórios, é razoável supor que um sistema macroscópico assumirá cada microestado quântico com a mesma probabilidade, tal que a efectivação de cada um desses estados é consistente com as restrições físicas externas impostas ao sistema. O postulado fundamental da mecânica estatística reside na assumption de que cada microestado assumível por um sistema físico fechado apresenta a mesma probabilidade de se verificar. Quando uma restrição externa ao sistema é eliminada

– tal como o relaxamento de uma válvula, permitindo por exemplo que um gás se expanda e aumente de volume – o número permissível de microestados quânticos aumenta. Decorrido algum tempo, o sistema assumirá microestados quânticos, consistentes com as novas condições impostas, pertencentes à classe de microestados cujo número aumenta até ao número máximo de estados permitidos pelas novas restrições externas.

Esta situação apresenta uma analogia com a propriedade fundamental da entropia termodinâmica, nomeadamente, a de que se trata de uma grandeza cujo valor aumenta ao longo da evolução temporal do sistema, até atingir um valor máximo, consistente com as restrições físicas externas impostas ao sistema. Sugere que se identifique entropia termodinâmica com o número de microestados de um sistema, cuja ocorrência é permitida pelas restrições físicas externas impostas ao sistema.

A entropia termodinâmica é uma grandeza física que aumenta de forma aditiva. Por outro lado, o número de microestados aumenta de forma multiplicativa. Desse modo, a única maneira de identificar a entropia termodinâmica  $S$  com o número de microestados, tornando consistentes as duas formas de variação, consiste em considerar para a entropia uma função, cujo aumento seja aditivo sempre que o aumento do seu argumento corresponda a uma multiplicação, isto é; tal que:

$$(1.22) \quad S(n_1) + S(n_2) = S(n_1 \cdot n_2) \quad ; \quad n_1 \geq 2, n_2 \geq 2, n_1, n_2 \in \mathbb{N}$$

Este requisito é cumprido definindo entropia termodinâmica através da expressão:

$$(1.23) \quad S = k \ln \Omega$$

Onde  $\Omega$  representa o número de microestados quânticos consistentes com as restrições macroscópicas impostas ao sistema e  $k = 1,38065 \times 10^{-23} \text{ J.K}^{-1}$  é a constante de Boltzmann, de modo a que os valores obtidos para  $S$  são consistentes com a escala de temperatura kelvin (Huang, 1988). Considerando que  $\Omega = \dim(H)$ , em que  $H$  é o espaço de Hilbert associado ao sistema, obtemos a igualdade (1.2):

$$(1.2) \quad S = k \ln \dim(H)$$

Adoptando-se as unidades de Planck e considerando que, assumida a adimensionalização das constantes neste sistema de unidades, a constante de Boltzmann

resulta  $k = 1$ , obtém-se:

$$(1.24) \quad S = \ln \dim(H)$$

Justificamos agora a asserção de que número mínimo  $M$  de mensagens de comprimento  $n$  necessárias para descrever, ao nível fundamental da matéria, qualquer sistema físico, verifica:

$$(1.12) \quad M = 2^n$$

A dimensão de um espaço de Hilbert, associado a um sistema físico, mede o número de configurações possíveis que se podem obter a partir dos estados quânticos distintos e independentes que cada parte elementar nesse sistema pode assumir. No jargão das teorias de campos quânticos: a dimensão do espaço de Hilbert representa o número de possibilidades na combinatoria formada pelos estados nos espectros dos osciladores presentes no campo quântico. Cada uma dessas configurações corresponde a um microestado quântico. Um observador macroscópico perfaz uma observação sobre o sistema de interesse, cujos valores das propriedades resultam de uma média das transições desse sistema num conjunto de microestados quânticos possíveis.

Porém, se afirmamos que cada microestado quântico é composto por um conjunto de estados, fundamentais e irreduzíveis, então em princípio deverá ser possível preparar uma experiência, permitindo efectivar uma observação macroscópica associada à existência isolada de cada um dos estados quânticos no microestado quântico dado. Isto é; a consideração de um só desses microestados quânticos e a sua atomização nos respectivos estados componentes quânticos deverá ser verificável, empiricamente, pelo menos indirectamente através de algum processo causal.

De outro modo, a sua consideração teórica deixaria de fazer sentido, uma vez que é aceite como um aspecto descritivo da realidade. Que um observador se apropria do conhecimento da existência de um estado, através de uma observação, significa que o sujeito que observa consegue diferenciar esse estado de outros estados distintos. Para que essa diferenciação cognitiva se efective é necessário que, em relação ao estado de interesse, seja possível a ocorrência de pelo menos outro estado. Suponhamos que um microestado é composto por  $n$  estados quânticos independentes. Visto que cada um destes tem existência independente dos restantes, isso significa que os mesmos deverão ser diferenciáveis de alguma maneira, mediante algum conjunto de propriedades que os

descrevam. No âmbito de uma teoria quântica de campo, procedemos à escolha de subconjuntos de  $r$  estados, respeitando o facto de que os componentes fundamentais da matéria apresentam comportamentos e estruturas análogos e que a cada um deles se encontra associado um espectro com o mesmo número inteiro de estados quânticos  $r$ . Além disso, consideramos que cada conjunto de estados no espectro de um desses componentes fundamentais é tal que esses estados são mutuamente exclusivos. O número de microestados quânticos ou de configurações que contam para o cálculo da entropia termodinâmica será então:

$$(1.25) \quad \Omega = r^n ; r \geq 2 , n \in N$$

O requisito de que  $r \geq 2$  ilustra o facto de que para que um estado quântico se torne reconhecível é necessário que este seja diferenciável de pelo menos outro, cuja ocorrência é possível. Nestas condições resulta que o número  $\Omega$  de microestados quânticos ou configurações será sempre:

$$(1.26) \quad 2^n \leq \Omega \leq r^n$$

Tendo-se, por consequência, para a entropia termodinâmica:

$$(1.27) \quad \ln 2^n \leq S \leq \ln r^n$$

No espírito da teoria da informação de Shannon, gostaríamos entretanto de associar um microestado quântico, formado por  $n$  estados quânticos, um por oscilador, a uma mensagem de comprimento  $n$ , cujos símbolos são escolhidos a partir de um alfabeto com um número  $r$  desses símbolos. Tal que a cada símbolo corresponde um estado quântico no espectro comum a todos os componentes do sistema físico sob consideração.

Por outras palavras, entendemos que um microestado quântico corresponde a uma mensagem num conjunto de mensagens possíveis e que a existência desse microestado pode ser comprovada através de algum procedimento experimental, situação em que se afirma que a mensagem que lhe está associada foi recebida pelo observador.

O caso em que o número  $M$  de mensagens é mínimo é aquele em que justamente se tem  $r = 2$ , valor para o qual resulta:

$$(1.28) \quad M = \Omega = 2^n ; n \in N$$

Esta observação conclui a fundamentação de que o número mínimo  $M$  de mensagens de comprimento  $n$  necessárias para descrever, ao nível fundamental da matéria, qualquer sistema físico é  $2^n$ .

Resta finalmente relacionar entropia termodinâmica com informação representada em *bit*.

Para esse efeito, reparamos que para  $\Omega = M$  a entropia termodinâmica  $S$  expressa em unidades de entropia de Planck é numericamente igual à quantidade de informação medida em *nat*. Multiplicando a constante de Boltzmann  $k = 1,38065 \times 10^{-23} \text{ J.K}^{-1}$  pela quantidade de informação escrita em *nat*  $I_{nat}(M)$ , obtemos uma expressão que é dimensionalmente homogénea, envolvendo entropia termodinâmica e quantidade de informação, tal que os valores numéricos de  $S(\Omega)$  e de  $I_{nat}(M)$  diferem por um valor  $k$ .

$$(1.29) \quad S(\Omega) = k.I_{nat}(M) \Leftrightarrow k \ln \Omega = k \ln(M)$$

Porém; para  $k = 1$ , em que a constante de Boltzmann é escrita em unidades de Planck sucede que os dois valores numéricos produzidos pela entropia termodinâmica  $S(\Omega)$  e pela quantidade de informação medida em *nat*  $I_{nat}(M)$  são iguais.

Verificando-se:

$$(1.29.1) \quad |S(\Omega)| = |I_{nat}(M)| \Leftrightarrow |k \ln \Omega| = |\ln(M)|$$

Em que o sinal de módulo assim aplicado " $|$ " significa que comparamos apenas os valores numéricos das expressões nos dois lados da igualdade.

Uma vez que  $\Omega = M = \dim H = 2^n$  e que fizemos corresponder a cada microestado quântico no espaço de Hilbert dado uma mensagem de comprimento  $n$ , a igualdade expressa em (1.29.1) sugere que associemos entropia termodinâmica, medida em unidades de entropia de Planck, directamente à quantidade de informação, medida em *nat*. Desse modo, propomos que a nível fundamental da matéria, para o número mínimo de

mensagens  $M = 2^n$  necessárias para descrever um sistema físico, a entropia termodinâmica que lhe corresponde está associada à quantidade de informação escrita em *nat*, de que um observador se apropria quando observa esse sistema, tal que os valores numéricos das duas grandezas são iguais. Notamos entretanto, de (1.10), que a entropia  $H$  de Shannon e a entropia termodinâmica  $S$  são grandezas numericamente diferentes, tal que se tem:

$$(1.30) \quad S(2^n) = 2.H$$

De facto, nas presentes condições, as afirmações anteriores mostram que a entropia termodinâmica  $S$  está associada à quantidade total de informação transmitida, considerado um conjunto de mensagens equiprováveis. Por seu lado, a entropia de Shannon  $H$  representa uma média da quantidade de informação transmitida por símbolo, no alfabeto utilizado para as formar, em que a cada símbolo é atribuída uma probabilidade de ocorrência, caracterizando uma certa ordem na produção de mensagens.

No caso da entropia termodinâmica  $S$ , todos os microestados são equiprováveis, o que significa que não existe nenhuma regra que faça prevalecer tendencialmente a ocorrência de um subconjunto de mensagens sobre outro, no universo de mensagens possíveis. Deste modo, devemos distinguir entropia termodinâmica da entropia de Shannon.

Por outro lado, à luz da entropia de Shannon sucede que o caso da entropia termodinâmica é justamente aquele em que o grau de desordem de uma fonte de mensagens é máximo, uma vez que os símbolos no seu alfabeto têm a mesma probabilidade de ocorrer. Esta observação parece trazer controvérsia sobre se devemos ou não identificar entropia termodinâmica com um conceito intuitivamente clarificador como o de informação. Esta objecção pode ser ultrapassada afirmando que no caso em que um sistema se caracterize pelas condições iniciais que tornam maximamente desordenado o seu estado, é também máxima a quantidade de informação de que um sujeito cognoscente se apropriará quando observar esse sistema. Assim, podemos sempre invocar duas interpretações distintas, embora complementares, sobre o significado da entropia termodinâmica  $S$ . Por um lado,  $S$  mede o estado de desordem do sistema. Por outro, mede a quantidade de informação mínima necessária de que um sujeito cognoscente se tem de apropriar para que possa descrever efectivamente esse sistema.

Recorrendo a (1.21) e (1.29), obtemos sucessivamente:

$$(1.31) \quad I_{nat}(M) = \ln 2 \cdot I_{bit}(M)$$

$$(1.32) \quad |S(\Omega)| = |\ln 2 \cdot I_{bit}(M)| \Leftrightarrow |k \ln \Omega| = |\ln 2 \cdot \log_2(M)|$$

Isto é, ao nível fundamental da matéria, para o número mínimo de mensagens  $M = 2^n$  necessárias para descrever um sistema físico, a entropia termodinâmica que lhe corresponde está associada à quantidade de informação escrita em *bit*, de que um sujeito cognoscente se apropria quando observa esse sistema, tal que os valores numéricos das duas grandezas cumprem a relação (1.32). Esta afirmação define a relação pretendida entre entropia termodinâmica e quantidade de informação binária, isto é; de informação medida em *bit*. No espírito da correspondência estabelecida, chamaremos à informação binária associada à entropia termodinâmica de um sistema físico, a «informação física» de que um sujeito cognoscente se apropria quando observa esse sistema. Combinando (1.29) com (1.31) obtemos a seguinte sequência de igualdades:

$$(1.32.1) \quad S(\Omega) = k \ln \Omega = k \cdot I_{nat}(\Omega) = k \cdot \ln 2 \cdot I_{bit}(\Omega)$$

Isto é:

$$(1.32.2) \quad S(\Omega) = k \cdot \ln 2 \cdot I_{bit}(\Omega)$$

Expressão que escrita no sistema de unidades de Planck, com  $k = 1$ , significa que a  $I_{bit}(\Omega) = 1 \text{ bit}$  corresponde uma quantidade de entropia termodinâmica  $S(\Omega) = \ln 2$  em unidades de entropia de Planck. E que, escrita no sistema de unidades SI, com  $k = 1,38065 \times 10^{-23} \text{ J} \cdot \text{K}^{-1}$ , significa que a  $I_{bit}(\Omega) = 1 \text{ bit}$  corresponde uma quantidade de entropia termodinâmica  $S(\Omega) = k \ln 2 \text{ J} \cdot \text{K}^{-1}$ , que é justamente o limite de von Neumann - Landauer.

## 1.2. Limite entrópico, entropia termodinâmica mínima de um sistema físico e área mínima associável à observação de um sistema físico.

Tirando partido da relação entre entropia termodinâmica e informação binária, pretendemos agora saber qual é a quantidade mínima daquela grandeza, válida numa observação. Esse conhecimento permitirá deduzir, através da expressão para o limite entrópico, a quantidade de área mínima mensurável, associada à observação de um sistema físico. Da expressão (1.32) verificamos que à apropriação, por parte de um observador, de 1 *bit* de informação sobre um sistema físico, corresponde sempre uma quantidade de entropia termodinâmica  $S = \ln 2$ . Um *bit* é, doutro modo, a quantidade de informação binária que um observador recebe quando observa um sistema físico a que estão associados dois estados quânticos, formando um microestado quântico composto por um ou outro desses estados. Na teoria da informação, esta é a situação em que um receptor acede a uma mensagem de comprimento 1, formada a partir de um alfabeto com dois símbolos. O número de estados quânticos – no caso, dois – é igual ao número mínimo de estados necessários para que um ou outro possam ser cognitivamente diferenciáveis entre si e, desse modo, reconhecíveis. Isto é, a situação em causa ilustra as condições mínimas para as quais um sujeito cognoscente consegue observar um estado físico. Por estado físico entendemos qualquer aspecto observável na natureza, necessariamente diferenciável de outro. Noutros termos, o que não pode ser em absoluto diferenciável de outra coisa qualquer, não pode ser observado e, como tal, não existe.

Perante estas considerações, perguntamos qual é a quantidade mínima de entropia que um sistema físico pode ter ou, noutros termos, qual é a quantidade mínima de informação binária que um observador pode receber, para que possa formular uma proposição sobre um sistema físico. Dado que uma diferenciação entre estados físicos, por parte de um observador, é uma operação cognitiva que só se pode efectivar num universo de pelo menos dois estados físicos possíveis, concluímos que a quantidade mínima de informação binária nessas circunstâncias é, justamente, 1 *bit*.

Daqui decorre que a quantidade mínima de entropia termodinâmica num sistema físico observável é precisamente  $S_{\min} = \ln 2$ . Resultado é consistente com o limite de von Neumann-Landauer.

Neste ponto, retomamos a sugestão de Lee Smolin para a enunciação do Princípio Holográfico, envolvendo a expressão para o limite entrópico. Respeitando Smolin, a área de uma superfície através da qual se observa um sistema físico é uma medida da capacidade de transferência de informação física através do espaço (Smolin, 2001a).

Do ponto de vista de cada observador, informação física flui assim do passado para o futuro desse observador, através de um conjunto de superfícies bidimensionais, causalmente relacionadas entre si. No contexto da teoria da gravidade de loops quânticos, Smolin chamou a essas superfícies «ecrãs causais». Para efeitos deste trabalho, definimos em analogia o conceito de «ecrã holográfico»:

Definição de ecrã holográfico.

Dado um sistema físico  $F$ , chama-se ecrã holográfico  $A$  associado à observação do sistema, a uma superfície no espaço de área  $A(F)$ , através da qual é transferida a informação física sobre o sistema, no decurso da sua observação, e é tal que se verifica a seguinte relação - denominada limite entrópico - entre a entropia termodinâmica total  $S(F)$ , associada à informação física, e a área do ecrã e medidas em unidades de Planck :

$$(1.33) \quad S(F) = k \ln \dim(H) \leq k \frac{A(F)}{4}$$

Em que  $\dim(H)$  representa a dimensão de cada um dos espaços de Hilbert associados, respectivamente, a um e a outro lado do ecrã e  $k=1$  é a constante de Boltzmann, que se convencionou ser igual a 1 no sistema de unidades de Planck.

Fazendo  $\dim(H) = 2$ , obtemos o seguinte resultado:

$$(1.34) \quad S_{\min} \leq S(F) \leq k \frac{A(F)}{4}$$

Ou:

$$(1.35) \quad S_{\min} \leq k \frac{A(F)}{4}$$

Em que  $S_{\min} = \ln 2$  e  $k = 1$ , em unidades de Planck.

Esta forma do limite entrópico sendo válida para o número mínimo de mensagens

$M = 2^n$  necessárias para descrever um sistema físico, ao nível fundamental da matéria.

Representando (1.35) em unidades SI, obtemos, com  $k = 1,38065 \times 10^{-23} \text{ J.K}^{-1}$  :

$$(1.36) \quad k \ln 2 \leq k \frac{\xi \cdot l_p^2}{4 \cdot l_p^2} \quad (\text{J.K}^{-1})$$

De onde deduzimos que, numericamente se verifica para  $\xi$  :

$$(1.37) \quad \xi \geq \ln 16 \approx 2,77258$$

Esta relação significa que, de acordo com o limite entrópico, a uma quantidade mínima de entropia termodinâmica  $S_{\min} = k \ln 2 \text{ J.K}^{-1}$ , associada à observação de um sistema físico com dois estados possíveis, corresponde um ecrã holográfico cuja área mínima é:

$$(1.38) \quad A_{\min}(F) \approx 2,77258 \cdot l_p^2$$

Isto é; retornando a unidades de área de Planck, ter-se-á:

$$(1.39) \quad 2 < A_{\min}(F) < 3$$

As teorias da física que procuram unificar mecânica quântica e relatividade generalizada - nomeadamente; a teoria da gravidade de loops quânticos e as teorias das cordas - apresentam a área como uma grandeza quantificada, de valor mínimo igual ao quadrado do comprimento de Planck  $l_p^2$ . Isso significa que o valor de  $\xi$  em (1.37) deverá ser necessariamente um número inteiro positivo e que o seu valor mínimo, quando muito, deverá ser igual a 1. Existe uma razão para que comprimento e área sejam grandezas quantificadas, na ordem de grandeza, respectivamente, de um comprimento e de uma área de Planck. O comprimento de Planck torna-se importante quando se descreve o comportamento de sistemas físicos de dimensões na ordem de grandeza de um centímetro

dividido por  $10^{33}$  (1 seguido de 33 zeros). Nessas regiões do mundo, aspectos da realidade tratados pela mecânica quântica e pela gravidade restrita e generalizada devem ser considerados conjuntamente. Consequentemente, na descrição cabal do que sucede estarão necessariamente envolvidas as seguintes constantes da física (escritas em unidades SI).

A constante de Planck  $h$ :

$$h = 6,62607 \times 10^{-34} \text{ J} \cdot \text{s}$$

A constante da velocidade de propagação da luz no vácuo  $c$ :

$$c = 2,99792 \text{ m} \times 10^8 \text{ m} \cdot \text{s}^{-1}$$

A constante de gravitação universal  $G$ :

$$G = 6,67428 \times 10^{-11} \text{ m}^3 \cdot \text{kg}^{-1} \cdot \text{s}^{-2}$$

A toda e qualquer massa está associado um comprimento de onda de Compton  $\lambda_c = h/mc$ .

Este resulta da fórmula utilizada pelo físico francês Louis de Broglie [1892 - 1987] em 1924, relacionando o comprimento de onda  $\lambda$  da onda de probabilidades com a quantidade de movimento  $p$  do objecto atómico associado a essa onda;  $\lambda = h/p$ . O comprimento de Compton pode ser interpretado como um limite fundamental na medição da posição de uma partícula, para o qual o erro nessa medição é minimizado, quando se considera conjuntamente mecânica quântica e relatividade restrita. Para medirmos a posição de uma partícula torna-se necessário fazê-la interagir com fótons. A precisão dessa medição dependerá do comprimento de onda da radiação electromagnética associada aos fótons, tal que quanto menor for esse comprimento de onda, tanto maior será a precisão da medição. Isso significa que os fótons utilizados deverão ter associadas frequências elevadas do campo electromagnético associado que, portanto, deverão ser altamente energéticos. Se a energia do fóton que interage com a partícula em questão for superior à sua energia em repouso  $m \cdot c^2$ , existirá energia suficiente para originar outra partícula do mesmo tipo, perturbando a determinação da posição da partícula original. Podemos recorrer à primeira relação de incerteza de Heisenberg para determinar o limite mínimo, em função do comprimento de onda de Compton, para o erro na medição da posição de uma partícula.

A relação de incerteza referida relaciona o erro na determinação da posição  $\Delta x$  de uma partícula atômica com o erro na determinação da sua quantidade de movimento  $\Delta p$  :

$$(1.40) \quad \Delta x \cdot \Delta p \geq \frac{h}{4\pi}$$

Pelo que:

$$(1.41) \quad \Delta x \geq \frac{h}{4 \cdot \pi \cdot \Delta p}$$

Quando a quantidade de movimento da partícula excede  $m \cdot c$ , existe energia suficiente para que seja criada outra partícula idêntica, cuja energia em repouso deverá ser pelo menos  $m \cdot c^2$ . Desse modo, um limite mínimo para o erro  $\Delta x$  na determinação da posição da partícula, sem que se verifique uma perturbação desse tipo, é:

$$(1.42) \quad \Delta x \geq \frac{1}{4 \cdot \pi} \frac{h}{m \cdot c}$$

Ou

$$(1.42.1) \quad \Delta x \geq \frac{1}{4 \cdot \pi} \lambda_c$$

Porquanto:

$$(1.42.2) \quad \Delta p \leq m \cdot c$$

Esta expressão indica-nos que o erro  $\Delta x$  na medição da posição de uma partícula quântica, a menos de uma constante envolvendo  $\pi$  nunca será menor que o seu comprimento de onda de Compton. Procuramos agora um limite mínimo de comprimento, cujo conhecimento seja possível através de um procedimento empírico. Nas condições de

(1.42), se  $\lambda_c$  for tão pequeno quanto se deseje, o erro  $\Delta x$  poderá ser controlado de modo a aproximar-se de zero, caso em que a medição para a posição da partícula é absolutamente exacta, independentemente do valor da sua quantidade de movimento  $p$  se tornar totalmente indeterminado, como requer (1.40). Essa situação é, no entanto, inalcançável, existe um limite mínimo para o comprimento a partir do qual a medição da posição de um sistema físico não é mais possível. Para o determinarmos, utilizamos por simplicidade uma abordagem semi-clássica, envolvendo elementos da mecânica quântica e da relatividade generalizada.

Para um tratamento mais rigoroso, confirmando o mesmo resultado, indicamos (Garay, 1995), (Amelino-Camelia e Kowalski-Glikman, 2005) e (Calmet *et al.*, 2008).

Começamos por notar que toda e qualquer massa no universo é caracterizada pelo respectivo raio de Schwarzschild, o raio do horizonte de um buraco negro com a mesma massa do sistema físico dado. Desse modo, podemos considerar um buraco negro cujo raio de Schwarzschild  $r_{bn}$  se pode relacionar com o comprimento de onda de Compton  $\lambda_c$  do sistema. De (1.42), temos:

$$(1.43) \quad 2 \cdot \pi \cdot \Delta x \geq \frac{h}{2 \cdot m \cdot c}$$

Este é o valor do perímetro de uma circunferência de raio igual ao valor do erro  $\Delta x$  cometido na medição da posição de uma partícula. Consequentemente, essa partícula encontrar-se-á, para uma medição num plano, no interior do círculo definido por  $2\pi \cdot \Delta x$ .

A representação em unidades SI da métrica de Schwarzschild para um buraco negro, utilizando coordenadas polares, é:

$$(1.44) \quad ds^2 = \left(1 - \frac{2GM}{c^2 r}\right) c^2 dt^2 - \left(1 - \frac{2GM}{c^2 r}\right)^{-1} dr^2 - r^2 (d\theta^2 + \sin^2 \theta d\phi^2)$$

O valor de raio do horizonte  $r_{bn}$  do buraco negro é tal que:

$$(1.45) \quad \frac{2 \cdot G \cdot M}{c^2 \cdot r_{bn}} = 1 \Leftrightarrow r_{bn} = \frac{2 \cdot G \cdot M}{c^2}$$

Assumimos entretanto o valor do perímetro equatorial de um buraco negro, fazendo:

$$(1.46) \quad 2 \cdot \pi \cdot r_{bn} = 2 \cdot \pi \frac{2 \cdot G \cdot M}{c^2}$$

O perímetro em causa corresponde ao equador da menor das esferas envolvendo o buraco negro. Esfera a partir do interior da qual nenhum objecto físico (incluindo a luz) pode escapar; motivo pelo qual se chama à sua superfície o «horizonte do buraco negro». Consequentemente, nenhuma informação precisa sobre o destino do objecto será acessível a um observador, fora do horizonte. A única proposição formulável sobre a posição de um sistema físico no interior do horizonte é a de que aquele se encontra a uma distância indeterminada da singularidade do buraco negro, supostamente localizada no centro da circunferência, cujo valor máximo é  $r_{bn}$ .

Este estado de coisas é similar àquele em que uma medição sobre a posição de uma partícula atómica, comporta um erro  $\Delta x$ . Estendendo essa analogia, identificamos o caso em que o perímetro máximo de um buraco negro de Schwarzschild é igual ao perímetro do círculo, cujo raio é o erro mínimo  $\Delta x$  na medição da posição de uma partícula. Mais ainda, entendemos que a própria partícula se transformou num buraco negro, tal que o seu destino após o respectivo colapso é agora incerto. A massa da partícula é tal que produziu sobre si um buraco negro, a que está associada uma singularidade, em redor da qual existe um horizonte com um determinado raio  $r_{bn}$ , e nesta situação, esperamos então que duas condições se verifiquem. Por um lado, teremos atingido o limite mínimo de precisão para o qual é possível determinar a posição de qualquer objecto físico. Por outro lado, estaremos em presença de um sistema – a partícula que se transformou num buraco negro - cuja natureza é tal que qualquer outro objecto que com ele interaja, no sentido de providenciar informação a um observador externo, será absorvido, desaparecendo uma vez transposto o perímetro definido pelo raio do horizonte  $r_{bn}$ , justificando-se fisicamente essa indeterminação, de acordo com o que a primeira desigualdade de Heisenberg prevê.

Qualquer propriedade da partícula original estará agora afectada de uma indeterminação. Isso inclui a sua posição em relação ao centro da circunferência que define o horizonte do buraco negro em que a partícula se transformou ou, equivalentemente, a sua

posição em relação a um ponto no horizonte, segundo um raio imaginário que passa pelo centro daquela circunferência. Assim, escrevemos:

$$(1.47) \quad 2.\pi.r_{bn} = 2.\pi.\Delta x \Leftrightarrow 2.\pi \frac{2.G.M}{c^2} = \frac{h}{2.m.c}$$

Igualando as massas  $M$  e  $m$  nos dois lados da igualdade e fazendo  $\hbar = h/2\pi$ , resulta:

$$(1.48) \quad M = \frac{1}{2} \sqrt{\frac{\hbar.c}{G}}$$

Este é o valor da massa de um buraco negro de Schwarzschild, tal que a posição de qualquer partícula no interior do horizonte em relação ao respectivo centro se pode conhecer com o erro mínimo  $\Delta x = r_{bn} = \frac{\hbar}{2.M.c}$ .

Se substituirmos (1.48) na segunda igualdade de (1.45), vemos finalmente que:

$$(1.49) \quad r_{bn} = \sqrt{\frac{\hbar.G}{c^3}}$$

Que representa o raio de um buraco negro de Schwarzschild, tal que a posição de qualquer partícula no interior do horizonte em relação ao respectivo centro se pode conhecer com o erro mínimo  $\Delta x = r_{bn} = \frac{\hbar}{2.M.c}$ .

A expressão (1.47) mostra que se aumentarmos a massa  $M$  na tentativa de fazer baixar o erro  $\Delta x$ , o raio do horizonte aumentará  $r_{bn}$ , fazendo aumentar o intervalo de indeterminação da partícula em relação ao centro do horizonte. Doutra maneira, se diminuirmos a massa  $M$ , com o intuito de fazer baixar o raio do horizonte  $r_{bn}$ , o erro  $\Delta x$  aumentará. Este facto ilustra como a igualdade  $r_{bn} = \Delta x$  define uma condição limite em que teoria da relatividade generalizada e mecânica quântica se encontram, produzindo

consequências, num ou noutro caso, que sugerem a existência de um valor mínimo de indeterminação na medição do comprimento. A medição de um comprimento é uma operação cognitiva, utilizando objectos físicos e observações sobre eles, que requer que idealizemos no espaço - supostamente contínuo - uma grelha imaginária de pontos a que chamamos «referencial», projectada idealmente sobre o espaço. Admita-se uma recta subdividida em intervalos, em que cada um destes tem o valor de comprimento dado por (1.49), em que estabelecemos uma relação de um para um entre o conjunto de intervalos e o conjunto dos inteiros positivos. A recta diz-se «discretizada». Ao valor do comprimento de cada um dos intervalos chamamos «resolução do comprimento determinável», definida como o comprimento mínimo teoricamente atribuível a um segmento de recta onde se presume que um objecto possa estar localizado.

Qualquer medição de comprimento tem dois resultados possíveis. Ou o resultado dessa medição coincide com um ponto de subdivisão na recta; ou o resultado é tal que se encontra num interior de um dos intervalos. No primeiro caso, a medição terá um erro  $\Delta x$  relacionado. Isso significa que o objecto poderá estar algures à esquerda ou à direita do ponto de coincidência. Isto é, o resultado da medição pode ser associado ao intervalo imediatamente anterior ou imediatamente posterior ao ponto de coincidência. No segundo caso, o resultado da medição pode ser associado a um intervalo dado na recta, no sentido que o objecto cuja posição se mede se encontrará algures no interior desse intervalo.

Por consequência, uma coordenação em que se utiliza um referencial desta natureza reportará para a posição de um objecto físico a um ou dois intervalos, identificados por um número inteiro positivo. Por outro lado, a medição de um comprimento envolverá necessariamente um número inteiro de intervalos, razão pela qual se diz que o comprimento é uma grandeza quantificada. Afirmaremos que um objecto físico está localizado, por exemplo, no interior do décimo quinto intervalo na recta dada, contado a partir de uma origem convencionada, em que se associa algum aspecto observável desse objecto com a localização dada, coordenada da maneira referida na recta. O objecto dista entre catorze e quinze intervalos da origem. O valor do comprimento de cada intervalo - a resolução do comprimento determinável - define a forma e o limite quantitativo das proposições sobre comprimento e localização de um objecto que podem ser formuladas numa teoria física.

Qualquer medição de comprimento, envolvendo um valor menor que a resolução dada, implica a transferência de valores de energia superiores àqueles associados à massa  $M$  em (1.47), medição cuja efectivação é impossível porquanto, ainda antes, para  $M = 1/2\sqrt{\hbar c/G}$  a partícula colapsará num buraco negro, obviando essa tentativa de medição. Podemos inclusivamente assumir que para valores do comprimento de um

intervalo  $\Delta l \approx \sqrt{\hbar G / c^3}$ , superiores a este valor, a determinação desse comprimento é não só possível por via teórica, como recorrendo a meios empíricos, isto é; perfazendo uma experiência para o efeito, logo antes da formação de um micro-buraco negro.

O valor obtido em (1.49), correspondendo à resolução do comprimento determinável, é justamente o comprimento de Planck, notado  $l_p$ .

$$(1.50) \quad l_p = \sqrt{\frac{\hbar \cdot G}{c^3}} = 1,61625 \times 10^{-35} \text{ m}$$

Utilizando coordenação no plano e coordenação tridimensional, obtemos como valores mínimos para a área e para o volume, respectivamente:

A área de Planck  $l_p^2$ :

$$(1.51) \quad l_p^2 = \frac{\hbar \cdot G}{c^3} = 2,61226 \times 10^{-70} \text{ m}^2$$

E o volume de Planck  $l_p^3$ :

$$(1.52) \quad l_p^3 = \left( \frac{\hbar \cdot G}{c^3} \right)^{3/2} = 4,22207 \times 10^{-105} \text{ m}^3$$

Para efeitos de coordenação espacial, as três grandezas físicas, comprimento  $L$ , área  $A$  e volume  $V$ , deverão estar relacionadas através das expressões habituais:

$$(1.53) \quad A = L^2 \Leftrightarrow A = n^2 \cdot l_p^2 \quad ; \quad n \in N$$

Em que a área se encontra quantificada, associada a um sistema de dois eixos coordenados perpendiculares, válido para alguma região de espaço euclidiano, e:

$$(1.54) \quad V = L^3 \Leftrightarrow V = n^3 \cdot l_p^3 \quad ; \quad n \in N$$

O volume resultando quantificado, associado a um sistema de três eixos coordenados perpendiculares, válido para uma região de espaço euclidiano.

De acordo com (1.39), a área mínima de um ecrã holográfico, associado à observação de um sistema físico com dois estados possíveis, caracterizado por uma quantidade mínima de entropia termodinâmica  $S_{\min} = \ln 2$ , encontra-se entre duas e três áreas de Planck.

Testamos agora a validade desta asserção, face ao requisito de que comprimento e área são grandezas quantificadas, relacionáveis entre si. Para o efeito preservamos a seguinte propriedade geométrica, válida para qualquer região de espaço euclidiano:

Dado um comprimento  $L$ , existe sempre uma área  $A$ , tal que se verifica:

$$(1.55) \quad L = m \cdot l_p \Rightarrow A = L^2 = m^2 \cdot l_p^2 \quad ; \quad m \in \mathbb{N}$$

Esta condição afirma simplesmente que na natureza é sempre possível medir numa região de espaço euclidiano, comprimento e área, de modo a relacioná-las através de (1.55), tal que as duas grandezas se encontram quantificadas. Assumindo conjuntamente o Princípio Holográfico, na acepção de Smolin, e a propriedade anterior, tentamos agora construir o ecrã holográfico de menor área  $A_{\min}(F)$  verificando a saturação inferior do limite entrópico, dado por (1.35), para o caso em que se observa um sistema físico com dois estados possíveis. Ter-se-á, em unidades de área de Planck:

$$(1.56) \quad A_{\min}(F) \geq 4 \cdot S_{\min} / k \Leftrightarrow A_{\min}(F) \geq 4 \ln 2 \Leftrightarrow \\ A_{\min}(F) \geq 2,77258$$

Em virtude do facto de que a área é uma grandeza quantificada, resulta imediatamente que deve ser:

$$(1.57) \quad A_{\min}(F) \geq 3$$

Assumindo que o ecrã holográfico se construiu utilizando a propriedade geométrica considerada, obtemos da expressão anterior:

$$(1.58) \quad l_{\min}(F) = A_{\min}^{1/2}(F) \text{ e } l_{\min}(F) \geq \sqrt{3} \Leftrightarrow l_{\min}(F) \geq 1,73205$$

Em que  $l_{\min}(F)$  é um comprimento associado ao ecrã holográfico, verificando (1.55).

Uma vez que o comprimento é uma grandeza física quantificada, (1.58) implica que:

$$(1.59) \quad l_{\min}(F) \geq 2$$

Comparando (1.57) com (1.59) e recorrendo à relação entre comprimento e área dada por (1.55), verificamos que deverá ser necessariamente:

$$(1.60) \quad A_{\min}(F) \geq 4 \quad \text{e} \quad l_{\min}(F) \geq 2$$

Concluimos deste modo que qualquer ecrã holográfico associado à observação de um sistema físico com dois estados possíveis, tal que comprimento e área são grandezas físicas, quantificadas em unidades de Planck, e relacionáveis através de (1.55), tem como valor de área não menos de quatro áreas de Planck numa região de espaço euclidiano.

Vamos agora considerar um sistema físico com quantidade de entropia termodinâmica mínima  $S_{\min} = \ln 2$  que colapsou num buraco negro de Schwarzschild. A área mínima do seu horizonte em unidades de área de Planck, tratando-se de uma superfície esférica e fazendo de acordo com (1.49);  $r_{bn} = l_p$ , é dada por:

$$(1.61) \quad A_{horiz} = 4 \cdot \pi$$

A relação de Hawking entre entropia e área do horizonte estabelece que:

$$(1.62) \quad S_{bn} = \frac{A_{horiz}}{4} = \pi \approx 3,14159$$

Se considerarmos que a área está quantificada em áreas de Planck, mediante a definição de coordenação dada por (1.53) e o facto de que o comprimento deverá estar quantificado em comprimentos de Planck, obtemos:

$$(1.63) \quad S_{bn} < \frac{A_{horiz}}{4} = \frac{4}{4} = 1$$

Se  $S_{bn} = S_{\min} = \ln 2$  a relação anterior é consistente com o valor mínimo de entropia termodinâmica:

$$(1.64) \quad \ln 2 < 1 \Leftrightarrow 0,693157 < 1$$

Assumimos que a medição de um comprimento, nestas condições, deve estar necessariamente associado à medição da área  $A_{horiz} / 4$  na sua totalidade, porquanto é a quantidade mínima de área quantificada consistente com a transmissão da quantidade mínima de informação física e não pode existir nenhuma medição de um comprimento, interior a essa área, que permita identificar a sua subdivisão em regiões de menor área. Nesse caso, dada a geometria do problema, o mínimo comprimento mensurável deverá corresponder ao menor comprimento do perímetro que delimita um quarto da área do horizonte esférico considerado, ou seja:

$$(1.65) \quad l = \frac{\pi}{2} \approx 1,57079$$

Assumindo que o comprimento se encontra quantificado em comprimentos de Planck, ter-se-á para o comprimento mínimo e para a área mínima, associados à medição de da quantidade mínima de entropia termodinâmica, caracterizando um buraco negro de Schwarzschild:

$$(1.66) \quad l_{\min} = 2.l_p \quad A_{\min} = 4.l_p^2 \quad S_{\min} = \ln 2$$

Estes valores mínimos de comprimento e de área, consistentes com a quantidade mínima de entropia termodinâmica, são os mesmos que deduzimos para um ecrã holográfico situado numa região de espaço euclidiano.

Perante estas considerações, a saturação inferior do limite entrópico deverá ser agora, em unidades SI e com  $k = 1,38065 \times 10^{-23} \text{ J.K}^{-1}$  :

$$(1.67) \quad k \ln 2 < k \frac{4 \cdot l_p^2}{4 \cdot l_p^2} \quad (\text{J.K}^{-1})$$

Deste modo, teremos concluído que na natureza qualquer superfície associada à observação de um sistema físico com a quantidade mínima de entropia termodinâmica, tal que comprimento e área são grandezas físicas quantificadas em unidades de Planck, tem como valor da sua área não menos de quatro áreas de Planck.

Recorrendo a raios de luz, podemos estabelecer um mapeamento num ecrã holográfico, situado numa região euclidiana, de regiões de espaço não euclidianas, fortemente afectadas por efeitos gravitacionais. No limite, as regiões gravitacionalmente mais afectadas serão aquelas que correspondem a buracos negros. Um resultado conhecido da relatividade generalizada, recorrendo ao «Argumento da focagem», é o de que o mapeamento  $A_{map}$  de uma região do horizonte de um buraco negro num ecrã holográfico, situado numa região de espaço euclidiano, terá sempre uma área superior à área da região original  $A_{reg}$  (Susskind e Lindesay, 2005). Este facto sugere que podemos utilizar dois novos valores para a coordenação de comprimento e de espaço, enquanto grandezas físicas quantificadas, nomeadamente;  $l_{min} = 2l_p$  e  $A_{min} = 4l_p^2$ .

No âmbito do limite entrópico proposto por Smolin, área deve ser entendida como uma medida da capacidade de transferência de informação física no espaço. Uma vez que a informação deve ser uma grandeza física aditiva e que a mesma se encontra associada à quantificação em valores de entropia termodinâmica de  $S_{min} = \ln 2$ , os quanta de área  $A_{min}$  deverão também adicionar-se, à medida que a quantidade de informação aumenta.

Em regiões de espaço euclidiano teremos ecrãs holográficos formados por estes quanta de área adicionados, tal que a medição de um deles deverá produzir um micro-buraco negro de raio igual a um comprimento de Planck. Isto é, para uma quantidade apreciável de quanta de área  $A_{min}$ , o espaço permanecerá euclidiano, sendo afectado gravitacionalmente no limite, quando se efectiva a medição de um dos quanta de área.

Podemos, desta maneira, conceber como sistema de coordenação macroscópico, numa região de espaço euclidiano, um ecrã holográfico num plano, constituído por um conjunto de superfícies elementares planas, cujo valor de área de cada uma dessas superfícies corresponde a quatro áreas de Planck.

No espírito do domínio da física de onde derivámos as nossas ilações, envolvendo o Princípio Holográfico, cuja definição recorre ao limite entrópico, chamaremos ao tipo de coordenação descrita «coordenação holográfica». Um observador situado numa região de espaço euclidiano terá acesso a informação física oriunda de qualquer sistema físico, no seu passado, através da observação de eventos, num dos lados do ecrã holográfico, causalmente relacionados com o sistema físico original. Uma vez que a área deste ecrã se encontra quantificada em áreas elementares de  $A_{\min} = 4l_p^2$ , tal que a  $A_{\min}$  se encontra associada uma quantidade de entropia termodinâmica  $S_{\min} = \ln 2$  e ainda que a quantidade de informação física correspondente a esse valor é 1 *bit*, podemos afirmar que um observador através de um ecrã holográfico plano, de área  $A_H$  :

$$(1.68) \quad A_H = n \cdot A_{\min} \quad ; \quad A_{\min} = 4 \cdot l_p^2 \quad , \quad n \in \mathbb{N}$$

Recebe  $n$  *bit* de informação física.

Assim, a quantidade mínima de informação física, expressa em *bit* que um observador pode receber sobre o mundo é 1 *bit*, à qual se encontra associado um ecrã holográfico com uma área mínima de quatro áreas de Planck. Este resultado foi também sugerido pelo físico holandês Gerard 't Hooft [1946, ], conforme refere Smolin:

“For simple black holes, which do not rotate and have no electric charge, the values of the temperature and entropy can be expressed very simply. The area of the horizon of a simple black hole is proportional to the square of its mass, in Planck units. The entropy  $S$  is proportional to this quantity. In terms of Planck units, we have the simple formula

$$S = \frac{1}{4} \frac{A}{h \cdot G}$$

where  $A$  is the area of the horizon, and  $G$  is the gravitational constant.

There is a very simple way to interpret this equation which is due to Gerard 't Hooft, who did important work in elementary particle physics - for which he won the 1999 Nobel Prize for Physics — before turning his attention to the problem of quantum gravity. He suggests that the horizon of a black hole is like a computer screen, with one pixel for every four Planck areas. Each pixel can be on or off, which means that it codes one bit of information. The total number of bits of

information contained within a black hole is then equal to the total number of such pixels that it would take to cover the horizon. The Planck units are very, very tiny. It would take  $10^{66}$  Planck area pixels to cover a single square centimetre. So an astrophysical black hole whose horizon has a diameter of several kilometres can contain a stupendous amount of information”.

(Smolin, 2001b, p. 90)

Atendendo à definição de observação, convencionaremos que todo o objecto cuja observação seja impossível tem associada uma quantidade nula de informação física.

### 1.3. O paradoxo holográfico.

Na secção anterior deduzimos um limite mínimo para o comprimento, associado a uma forma de coordenação quantificada e descontínua do espaço, que denominámos resolução do comprimento determinável. Este comprimento foi obtido combinando, inicialmente, relatividade restrita com mecânica quântica, utilizando-se a primeira desigualdade de Heisenberg e o facto de que a partir de certa quantidade de energia seria criada uma nova partícula para além daquela cuja posição se mede. Recorrendo, em seguida, ao raio de um buraco negro de Schwarzschild, chegámos finalmente ao comprimento pretendido, cujo valor é o de um comprimento de Planck  $l_p$ . A resolução do comportamento determinável é um valor previsto teoricamente e corresponde à existência de um afastamento possível entre dois sistemas físicos que, por esse motivo, deveria poder ser observável.

Assumindo que, associado a esse comprimento, é válido na natureza um valor de área mínimo correspondendo ao quadrado daquela distância  $A_{\min} = l_p^2$ , também esta área deveria ser objecto de uma observação possível, permitindo a confirmação da sua existência. Recorrendo à ideia de que a observação, ao nível fundamental da descrição da matéria, corresponde a um processo cognitivo que envolve transferência de informação física, pressupondo necessariamente a efectivação de uma diferenciação entre estados distintos e independentes do sistema observado, concluímos a partir do limite entrópico que a quantidade de área mínima de um ecrã holográfico deve ser igual a quatro áreas de Planck  $A_H = 4.l_p^2$ . Este facto sugere que nenhuma observação pode ser efectuada, cuja área do ecrã holográfico associado e, desse modo, a área associada de qualquer superfície observável, possa ser inferior a quatro áreas de Planck. No sentido de Smolin, um ecrã

holográfico de área inferior a quatro áreas de Planck não teria capacidade suficiente para permitir sequer a transferência de 1 *bit* de informação física, tal que nessa situação nenhuma informação seria transferida e, conseqüentemente, nenhum conhecimento sobre o sistema físico acedível pelo observador.

Porém, se assim é, com que legitimidade é que se pode afirmar que na natureza é válido um valor de área igual a um quarto da menor área de um ecrã holográfico? De facto, uma vez que o valor que formalmente se atribui a uma área corresponde a uma extensão contínua entre zero e esse valor, se este for o valor mínimo mensurável, com que legitimidade é que se pode falar nos valores entre zero e esse valor mínimo, se a sua observação for impossível? Será a extensão entre zero e um dado valor mínimo observável contínua num sentido intuitivo do termo? E se nada ali for observável, o que restará da nossa intuição habitual de espaço, de possibilidade de diferenciação cognitiva entre dois pontos? Pontos considerados distintos mediante a presença de um afastamento entre eles, cuja realidade tem que estar para além do nada, sob risco de que os dois pontos coincidam e sejam o mesmo? Estas questões revelam que a continuidade e a sua relação com a descontinuidade, mediante a nossa própria cognição da realidade, é um problema não trivial. Dificuldades que parecem permear todos os formalismos quânticos em que, justamente, a quantificação das grandezas físicas presentes as promove uma incongruência entre os sistemas numéricos utilizados, reflectindo a forma de pensamento humano sobre o mundo, e o modo como a natureza é descontinuamente observável.

A nossa intuição habitual de espaço, em termos de uma extensão contínua, parece inadequada para lidar com estes problemas. Com o objectivo de obter uma afirmação coerente sobre estes aspectos da realidade, devemos proceder cautelosamente pela via lógica, começando por considerar o que se deverá entender por observação.

No contexto dos desenvolvimentos anteriores <sup>11</sup>, desejamos que a efectivação de uma observação se encontre associada à transferência de informação física. Desse modo, estabelecemos sucessivamente o seguinte.

#### Definição de estado físico de um objecto na natureza.

Um estado físico de um objecto na natureza é um aspecto desse sistema necessariamente diferenciável de pelo menos outro aspecto, por parte de um agente cognitivo chamado observador. Sempre que um observador perfaz essa diferenciação diz-se que aquele reconheceu o estado físico.

---

<sup>11</sup> Ver também a discussão sobre entropia de Von Neumann, associada à correlação quântica de um sistema, conducente a uma definição técnica de informação física (c.f. p. 76).

#### Definição de sistema físico.

Chama-se sistema físico a um conjunto não vazio de estados físicos, inteligivelmente relacionáveis pelo observador a um e um só objecto, tal que esse conjunto contém um ou mais subconjuntos não vazios de estados mutuamente exclusivos.

#### Definição de informação física.

Chama-se informação física sobre um sistema físico na natureza ao número inteiro de *bits* associados à quantidade de entropia termodinâmica do sistema  $S$ , tal que esta se encontra quantificada para um valor mínimo de  $S_{\min} = k \ln 2 \text{ J.K}^{-1}$  ( $S_{\min} = \ln 2$  no sistema de unidades de Planck), sendo que a esse valor corresponde 1 *bit* de informação física, transferido através de um ecrã holográfico cuja área mínima é  $A_H = 4.l_p^2$  ( $A_H = 4$  no sistema de unidades de Planck).

Em que  $l_p^2 = 2,61226 \times 10^{-70} \text{ m}^2$  é a área de Planck.

E em que  $k = 1,38065 \times 10^{-23} \text{ J.K}^{-1}$  é a constante de Boltzmann.

#### Definição de apreensão de informação física por parte de um observador.

Diz-se que foi feita a apreensão da informação física sobre um sistema físico por parte de um observador, sempre que a efectivação de uma experiência permitir ao observador reconhecer estados na natureza, em relação aos quais lhe é possível estabelecer uma relação de um para um entre um conjunto, com não menos de dois desses estados, e o conjunto dos números naturais.

#### Definição de observação.

Uma observação de um sistema físico consiste na apreensão por parte de um observador da informação física sobre o sistema observado, no processo da transferência dessa informação através do espaço e ao longo do tempo.

Sempre que um observador perfaz uma observação, aquele reconhece, pelo menos, um estado físico na natureza.

Definição de não existência ou de existência de um sistema físico.

Se no conjunto de estados físicos, definindo um sistema físico, existir pelo menos um desses estados que não possa ser observado por nenhum observador na natureza, o sistema físico em causa não pode ser observado e, como tal, não existe.

A informação física relativa a esse sistema não pode ser apreendida por nenhum observador e o seu valor, por definição, é zero *bit*.

Sempre que a disposição anterior não se verificar, diz-se que o sistema físico existe.

A consideração conjunta das definições anteriores, juntamente com a definição de ecrã holográfico, implica que não existe nenhum processo experimental através do qual seja possível observar qualquer região de espaço na natureza, cuja área seja menor que quatro áreas de Planck.

Admitamos uma região aproximadamente bidimensional, associada a um sistema físico, cuja área se pretende medir através de uma observação e convencionemos que a medição de um certo valor dessa área corresponde à situação em que a região é delimitável por processos físicos, tal que a extensão da respectiva delimitação tem o valor de área medido. A medição da área de uma região envolve necessariamente a presença do sistema físico a que a mesma está associada, uma vez que área é uma grandeza física relacional entre eventos, substancializados por sistemas físicos. Assim como não existe espaço sem a presença de eventos e dos sistemas físicos que os agenciam, também não pode existir uma superfície em si mesma, sem que diga respeito a alguma porção de matéria ou a algum processo dinâmico envolvendo energia, que seja observável.

Na secção anterior, definiu-se ecrã holográfico como qualquer superfície no espaço de área  $A$ , através da qual é transferida a informação física sobre o sistema que lhe está associado no decurso da sua observação. Desse modo, a região bidimensional associada ao sistema físico define um ecrã holográfico através do qual deverá fluir informação física sobre esse sistema, uma vez que só à custa da observação do mesmo é possível adquirir algum conhecimento sobre a área da região que lhe está associada. Consequentemente, através da região dada deve transferir-se pelo menos 1 *bit* de informação física, o valor mínimo de informação com base no qual se torna possível a um observador afirmar o que quer que seja sobre a realidade.

Porém, de acordo com a interpretação de área de Smolin, esta representa uma capacidade de transferência de informação física através do espaço e tem-se que uma

região bidimensional cuja área seja menor que  $A_H = 4.l_p^2$  representa um valor dessa capacidade que é inferior ao valor permitido para a transferência da quantidade mínima de informação física de 1 *bit*. Decorre então que qualquer superfície cuja área se pretenda delimitar a um valor inferior a quatro áreas de Planck, não pode ter nenhum sistema físico associado que seja observável. Este facto implica por sua vez que essa área, uma vez associada a um sistema físico cuja observação é impossível, não pode, por conseguinte, ser determinada. O resultado é consistente com o facto de que a área é uma grandeza física quantificada, apresentando um mínimo mensuravelmente intransponível. Poderíamos assumir que a existência de uma superfície com área superior à área considerada implicaria necessariamente a existência de áreas menores, isto é; que qualquer superfície de uma área dada é sempre decomponível em superfícies de áreas menores. No entanto, esta propriedade geométrica na descrição do mundo é falsificada mediante a impossibilidade empírica de proceder à decomposição hipostasiada.

Semelhante conclusão invoca imediatamente duas questões: o que existirá para além do limite mínimo de área observável e o que devemos afinal entender por área, em termos de extensão ou de possibilidade cognitiva de diferenciar alguma coisa no interior - hipoteticamente existente - de uma região bidimensional de área mínima. O problema torna-se ainda mais crítico se agora notarmos que é possível prever teoricamente a existência de uma área mínima  $l_p^2 < 4.l_p^2 = A_{\min}$ , cuja determinação empírica, através de uma observação é, por outro lado, impossível, uma vez que é inferior à menor das áreas de um ecrã holográfico  $A_H = 4.l_p^2$ . A esta contradição, que surge à luz da definição de observação e do facto de que o que não pode de todo ser observado, necessariamente não pode existir, chamaremos «paradoxo holográfico», cujo enunciado é o seguinte.

O paradoxo holográfico.

Chama-se paradoxo holográfico à contradição que resulta da comparação entre o valor mínimo de área existente na natureza;  $A_0 = l_p^2$  - previsto teoricamente através da consideração conjunta da relatividade generalizada e da mecânica quântica - e o valor da quantidade mínima de área associado à observação de um sistema físico, obténível a partir do limite entrópico e da quantidade mínima de entropia termodinâmica num sistema físico;  $A_{\min} = 4.l_p^2$ , tal que se tem  $A_0 < 4.A_0 = A_{\min}$ .

É entretanto notório que, do ponto de vista formal, o valor para  $A_0 = l_p^2$  surge na própria expressão do limite entrópico, a partir da qual é possível deduzir o valor para  $A_{\min} = 4.l_p^2$ . Este facto reforça o motivo pelo qual chamamos paradoxo à contradição obtida.

O comprimento de Planck, por seu lado, é um valor recorrente na física moderna que surge na teoria da gravidade dos loops quânticos e nas teorias das cordas; as tentativas contemporâneas de unificação entre gravidade e teorias de campo quânticas. Este facto sugere que a sua eliminação do discurso científico como via possível para a solução do problema não constitui uma hipótese viável. O paradoxo holográfico indicia assim a presença de uma contradição teórica que é necessário ultrapassar.

#### **1.4. A condição de não contradição na natureza e a relativização do observador.**

O problema colocado pelo paradoxo holográfico é o de que dada a impossibilidade - imputada por uma teoria sobre o mundo - sobre a observação de um dado aspecto da realidade, este não pode existir e, como tal, não pode ser simultaneamente representado por outra proposição formulável na teoria. Esse estado de coisas tornaria a teoria inconsistente. Justamente um dos requisitos que devemos exigir a qualquer teoria sobre a realidade é o de que seja consistente. O mundo deverá ser de tal modo que permita que sobre ele se produzam descrições não contraditórias. Caso contrário, qualquer juízo sobre a natureza resultará logicamente instável no que diz respeito ao seu valor lógico, tornando a prossecução do pensamento humano sobre o mundo uma tarefa penosa e inconclusiva.

Assumiremos, pois, válida a possibilidade de formulação de um discurso não contraditório sobre a natureza como primeiro postulado neste trabalho, cujo enunciado é o seguinte.

A condição de não contradição na natureza.

Todos os fenómenos na natureza são tais que podem ser descritos recorrendo a um discurso onde não existe contradição.

Concretizando este postulado em termos de um discurso formulável no âmbito de uma teoria informal sobre a natureza, associaremos provisoriamente este enunciado ao de outra proposição a que chamaremos a «condição de inteligibilidade lógica da natureza»,

pressupondo-se a validade da identidade ontológica entre o que é computável numa teoria informal e o que ocorre no mundo.

A proposição enuncia-se como segue.

A condição de inteligibilidade lógica da natureza.

Toda e qualquer descrição do mundo recorre teorias informais empíricas que são consistentes.

Assumiremos que a condição de inteligibilidade lógica da natureza é formalmente equivalente à condição de não contradição na natureza, entendendo-se por consistência de uma teoria informal a impossibilidade de deduzir, a partir dos postulados e princípios da teoria, duas proposições cujos enunciados sejam contraditórios.

Tendo em vista a condição de inteligibilidade lógica da natureza e, assumindo que as considerações teóricas desenvolvidas nas secções anteriores são válidas ou que, pelo menos, existem superfícies na natureza cuja área é igual a uma área de Planck, tal que  $l_p^2 < A_{\min}$ , devemos agora procurar uma solução para o paradoxo holográfico.

De acordo com a definição de observação dada, a condição de existência de qualquer coisa no mundo é a de que exista pelo menos uma situação em que essa objecto possa ser observado. Essa situação parece estar para além das possibilidades físicas de observação recenseadas. Por outro lado, o limite imposto à observação resultou da consideração conjunta de factos físicos e de uma propriedade da cognição que identificámos como a possibilidade de um sujeito cognoscente diferenciar aspectos distintos na realidade no decurso de uma observação. Assim sendo, na dedução da área mínima de um ecrã holográfico utilizou-se uma propriedade própria do sujeito cognoscente, isto é; do observador.

O limite para a área mínima de um ecrã holográfico, associado a um sistema físico observado, é por conseguinte uma propriedade que diz respeito a sujeitos cognoscentes e que resulta da combinação lógica entre propriedades do mundo físico e uma propriedade da cognição. Acrescendo a este facto, devemos ter em consideração que o conjunto de propriedades que atribuímos à realidade física, expressas num discurso teórico consistente, deriva de actividade cognitiva fundamentada na observação do mundo. Assim sendo, afigura-se plausível que o conjunto de propriedades físicas associadas à descrição da observação, enquanto fenómeno na realidade, embora caracterizando aspectos físicos do mundo como entropia termodinâmica e área, sejam factores correlativos que, juntamente com propriedades da cognição, caracterizam em suma o próprio acto de observar. Se assim

for, o limite entrópico e a proposição sobre a área mínima de um ecrã holográfico são afirmações que caracterizam o observador no universo que observa.

Esta assumção permite relativizar o nosso lugar como observadores no mundo, porquanto nada obsta então a que possam existir outros sistemas físicos capazes de proceder a uma observação, no sentido em que a mesma foi definida. Sucede apenas que todo o elemento na classe de sujeitos cognoscentes humanos, capazes de produzir um discurso teórico sobre a realidade com as características actuais, se caracteriza pelo facto de lhe ser impossível observar a realidade através de ecrãs holográficos de área inferior a quatro áreas de Planck.

Não significa isto que não existam na realidade as condições para ultrapassar esse limite, obtendo-se desse modo uma classe de observadores com outras propriedades físicas e, eventualmente, cognitivas. Nem tão pouco que um observador limitado por um determinado valor mínimo do ecrã holográfico não encontre um processo físico para ultrapassar esse limite. A situação é a de que essa variabilidade ainda não foi efectuada e que as propriedades que nos caracterizam, enquanto observadores, se mantêm sensivelmente as mesmas desde o início da aventura humana.

Estas afirmações sugerem fortemente que propriedades físicas atribuíveis ao acto de observar devem estar consistentemente relacionadas com o próprio comportamento observável das coisas e, conseqüentemente, com as teorias que os observadores produzem sobre o mundo. Teremos encontrado aqui um indício seguro para a presença de variabilidade nomotética? Não necessariamente.

A relativização da nossa presença física no mundo como observadores, como qualquer outro sistema caracterizado por propriedades cuja caracterização quantitativa pode variar, faz com que o entendimento teórico da realidade seja agora mais vasto e complexo. Continua a poder ser válida a tese da estabilidade nomotética, isto é; a afirmação de que numa teoria consistente sobre a natureza, estritamente fundamentada na observação, existem proposições chamadas leis da natureza que são válidas em todo o local; para qualquer tempo e em relação a qualquer observador no universo, cuja origem e valores lógicos não dependem de nenhum factor externo à teoria em que são formuladas. Apenas se incluiu o facto de que também o observador é uma entidade, como qualquer outra, na natureza, caracterizada por um conjunto de propriedades que determinam o modo como regista o comportamento dos restantes sistemas físicos.

Deveremos, pois, prepararmo-nos para descrições teóricas sobre o mundo que contemplem a possibilidade de visões distintas da realidade, consistentes entre si, e que dependem das propriedades dos observadores que as experimentam. Continua a ser válida a possibilidade de erigir uma teoria completa sobre a natureza que inclua todas essas

descrições possíveis, coerentes com a possibilidade de variação quantitativa de pelo menos uma propriedade de natureza física, atribuível às capacidades de observação de sujeitos cognoscentes num dado conjunto de observadores.

### 1.5. O critério lógico de realismo, o Princípio da Resolução Holográfica e a solução do paradoxo holográfico.

Uma solução possível para o paradoxo holográfico poderá ser a seguinte.

Seja uma teoria completa sobre o mundo formulável por observadores humanos. Isto é, uma teoria que descreve de forma adequada, se não todos, a maioria dos fenómenos observáveis. O limite inferior para o valor da área de um ecrã holográfico  $A_{\min} = 4.l_p^2$ , invocado nessa teoria, representa a quantidade mínima de área mensurável por observadores humanos, no decurso de todo e qualquer processo de observação ao seu alcance. Representa uma medida do limite inferior da capacidade cognitiva humana no universo, recorrendo à observação. A sua invocação é feita no contexto de uma proposição que podemos entender caracterizar as capacidades de observação de um conjunto particular de observadores no mundo e que, conseqüentemente, permitirá definir o que chamaremos «classe de observadores», associada a uma dada «resolução holográfica».

Começemos por tratar o conceito de «observador».

Definição de observador.

Chama-se observador a uma entidade cognoscente, isto é; que é capaz de produzir conhecimento.

Tendo em consideração o limite referido, convencionamos agora o seguinte:

Definição de observadores humanos.

Observadores humanos são entidades cognoscentes no universo, pertencentes à classe de entidades cognoscentes para as quais o valor mínimo de área mensurável  $A_{\min}$  é tal que, em relação ao valor de área de referência inferior,  $A_0 = l_p^2$ , de uma superfície que existe na natureza, se verifica  $A_{\min} = 4.A_0$ .

O problema agora surge mediante o facto de que o valor  $A_0 = l_p^2$  da área de referência, não sendo observável por observadores humanos, não pode ser logicamente invocável numa teoria consistente sobre o mundo. Esta dificuldade pode ser ultrapassada através do seguinte procedimento. Consideramos, no espírito do que foi afirmado, que no universo existem várias classes de observadores, caracterizados por valores mínimos mensuráveis de área, distintos entre si, a que chamaremos «resoluções holográficas», sucedendo que o conjunto dos observadores humanos se inclui numa dessas classes.

O valor da área de referência  $A_0 = l_p^2$  é uma consequência lógica dessa teoria mais vasta, formulável por todos os observadores, contemplando todas as classes de observadores no universo. A teoria deverá permitir que cada observador possa assumir logicamente a existência de outros observadores, com propriedades distintas das suas, assim como a existência de observações que estão para além dos limites físicos impostos às suas capacidades de observação. Estes requisitos sugerem que comecemos por estabelecer um critério que permita identificar por via lógica o que existe no universo, para além do que é observável por um dado sujeito cognoscente, sem que com isso se viole a possibilidade em absoluto da observação do que se afirma existir. Essa condição, chamada «critério lógico de realismo», enuncia-se da seguinte maneira.

#### Critério lógico de realismo.

Para que um sistema físico exista no universo é suficiente que se verifiquem, na teoria utilizada para o descrever, as condições lógicas para a sua observação por parte de pelo menos um observador nesse universo.

Por observação entende-se a recepção, por parte de um agente cognitivo, de uma quantidade não nula de informação física proveniente do sistema em causa. Por condições lógicas para a sua observação entende-se as proposições que determinam formalmente na teoria o conjunto de configurações de estados possíveis que podem ser observados para aquele sistema. A este respeito é conveniente notar que neste trabalho será adoptada a interpretação de Copenhaga, proposta pelo físico holandês Niels Bohr [1885 - 1962] para a mecânica quântica, em que o vector de onda associado a um objecto observável representa, não uma entidade física em si mesma, mas um componente lógico num procedimento formal, envolvendo todos os estados observáveis do sistema a que está associado. Neste sentido, o vector de onda não é um sistema físico observável, mas uma

ferramenta lógica que descreve o comportamento do objecto a que está associada.

É agora que, com base no critério lógico de realismo e a partir do conhecimento teórico de que dispomos, podemos procurar as condições lógicas que determinam a existência de uma entidade física que apenas se revelará inobservável relativamente a um subconjunto de observadores possíveis no mundo. O paradoxo holográfico refere dois limites mínimos para a área de um ecrã holográfico. O primeiro é a área de Planck. O segundo corresponde ao quadruplo dessa área, que identificamos como a resolução holográfica, notada  $R_H(0)$ , caracterizando a classe de observadores a que pertencemos.

Estes dois valores estão relacionados pela relação simples  $R_H(0) = 4.l_p^2$ , sendo que a área de Planck parece ser o menor valor de área, inferior à nossa resolução holográfica, que podemos obter teoricamente a partir do conjunto das nossas observações sobre o mundo, utilizando tudo o que sabemos sobre a realidade. Perante estas afirmações propomos um enunciado, que deverá ser considerado uma proposição de lei sobre a natureza, que representa justamente o conjunto de condições lógicas referidas no critério lógico de realismo, permitindo assegurar a existência de objectos no mundo para além das nossas capacidades de observação. Chamamos-lhe o «Princípio da Resolução Holográfica». A sua enunciação requer, de acordo com o que foi afirmado, que formalizemos o seguinte.

Definição de resolução holográfica e de classe de observadores.

Chama-se resolução holográfica  $R_H(i); i \in N_0$ , de um conjunto de observadores ao menor valor da área de um ecrã holográfico, através do qual todo e qualquer observador no conjunto dado recebe a quantidade mínima de informação física de 1 *bit*.

Ao conjunto de observadores cuja resolução holográfica é  $R_H(i)$  chama-se a classe de observadores  $C_O(i); i \in N_0$ , de resolução holográfica  $R_H(i)$ .

Definição de valor mínimo da área de referência.

Dada uma classe de observadores de resolução holográfica  $R_H(i); i \in N_0$ , chama-se valor mínimo da área de referência  $A_0(i); i \in N_0$ , de uma classe de observadores ao menor valor de área, inferior à resolução holográfica  $R_H(i)$  que pode ser obtido logicamente recorrendo ao conjunto das proposições que descrevem as observações

feitas pelos observadores naquela classe, no âmbito de uma teoria formulada pelos observadores na mesma classe.

Não vendo motivo para que se considere a existência de valores de resolução holográfica superiores a  $R_H(0) = 4.l_p^2$ , uma vez que não se conhece um limite superior para a medição de valores de área no universo. Em face desta asserção e das duas definições anteriores, enunciaremos finalmente o «Princípio da Resolução Holográfica», segundo postulado proposto nesta investigação.

Princípio da Resolução Holográfica.

Todo e qualquer observador no universo pertence a uma classe observadores  $C_o(i); i \in N_0$ , caracterizada por uma resolução holográfica  $R_H(i); i \in N_0$ , associada à quantidade mínima de informação física que cada observador nessa classe pode receber no decurso de uma observação.

Sendo que para toda e qualquer resolução holográfica  $R_H(i)$ , existe sempre outra,  $R_H(i+1)$ , associada a uma classe distinta de observadores  $C_o(i+1)$ , tendo-se:

$$(1.69) \quad R_H(i) = 4.R_H(i+1) \quad ; \quad R_H(i+1) = A_0(i) \quad , \quad i \in N_0$$

Em que  $A_0(i)$  é o valor mínimo da área de referência da classe de observadores  $C_o(i)$  de resolução holográfica  $R_H(i)$ , com  $R_H(0) = 4.A_H(0)$  e  $A_H(0) = l_p^2$ .

Mediante a consideração do Princípio da Resolução Holográfica, como representação do conjunto de condições lógicas para a existência de um objecto no universo, referidas no critério lógico de realismo, o paradoxo holográfico é imediatamente solucionado. A inconsistência que deriva da invocação lógica de dois limites mínimos para o valor possível da área - em que um deles é inferior à resolução holográfica dos observadores numa dada classe  $C_o(i)$  - é eliminada, porquanto um desses valores é justamente a resolução holográfica da classe de observadores  $C_o(i+1)$  que imediatamente se segue, em termos do próximo valor  $R_H(i+1)$ , da resolução holográfica. Este valor não pode ser directamente observado por um sujeito em  $C_o(i)$ . A nova

resolução, mais baixa, é logicamente dedutível a partir da informação empírica disponível, na teoria formulada pelos observadores na classe  $C_o(i)$ , providenciando a forma de relacionar o conjunto das suas observações com outras possíveis, para além das suas próprias capacidades de observação, efectivadas pelos observadores na classe  $C_o(i+1)$  e em consistência com o Princípio da Resolução Holográfica.

Noutros termos, a resolução holográfica de uma dada classe de observadores  $C_o(i+1)$ , que deverá estar associada a um sistema físico que existe, em consequência do critério lógico de realismo e do Princípio da Resolução Holográfica, é alcançável por via estritamente teórica a partir do discurso científico produzido por observadores na classe  $C_o(i+1)$ , verificando-se  $R_H(i+1) = 4 \cdot R_H(i)$ .

### 1.6. Informação e o Princípio da Resolução Holográfica.

De acordo com o Princípio da Resolução Holográfica, a resolução holográfica – aqui medida em unidades de área de Planck – é uma grandeza física quantificada recorrendo a um subconjunto dos números fracionários  $Q$ . Com efeito, representando por  $R_H(0) = 4 \cdot l_p^2$  a resolução holográfica da classe de observadores  $C_o(0)$  em que nos incluimos, temos sucessivamente:

$$(1.70) \quad R_H(1) = \frac{R_H(0)}{4}, R_H(2) = \frac{R_H(0)}{4^2}, \dots, R_H(n) = \frac{R_H(0)}{4^n} ; n \in N$$

O valor mínimo da área de referência sendo também uma grandeza física quantificada medida, de modo análogo, em unidades de área de Planck:

$$(3.71) \quad A_0(1) = \frac{A_0(0)}{4}, A_0(2) = \frac{A_0(0)}{4^2}, \dots, A_0(n) = \frac{A_0(0)}{4^n} ; n \in N$$

E em que  $A_0(0) = l_p^2$  é o valor mínimo da área de referência relativo à classe de observadores  $C_o(0)$  a que pertencem todos os observadores humanos.

Toda e qualquer medição de área efectivada por um observador numa classe de

observadores  $C_o(i)$  estará, em relação a esse observador, quantificada em unidades da sua resolução holográfica  $R_H(i)$ , o mesmo sucedendo com o comprimento, desde que se considere que para qualquer resolução holográfica é  $L_{\min}[R_H(i)] = [R_H(i)]^{1/2}$ , representando o comprimento mínimo mensurável por um observador em  $C_o(i)$ . Tem-se, assim, para um sujeito pertencente  $C_o(i)$ , com resolução holográfica  $R_H(i)$ :

$$(3.72) \quad A[R_H(i)] = n \cdot R_H(i) \quad ; \quad n \in N$$

$$(3.73) \quad L[R_H(i)] = m \cdot [R_H(i)]^{1/2} \quad ; \quad m \in N$$

Em que  $A[R_H(i)]$  representa uma área medida por um observador com resolução holográfica  $R_H(i)$  e  $L[R_H(i)]$  representa um comprimento medido pelo mesmo observador.

Seja agora  $I_f[F; R_H(i)]$  ;  $i \in N$ , a quantidade de informação física medida em *bit*, recebida por um observador, com resolução holográfica  $R_H(i)$ , quando observa o sistema físico  $F$ . Dois observadores  $O_1$  e  $O_2$  com resoluções holográficas,  $R_H(i)$  e  $R_H(i+1)$ , afirmarão, em consequência do critério lógico de realismo e do Princípio da Resolução Holográfica, que um determinado sistema físico  $F$  existe e que esse sistema é observável pelos dois sujeitos, se o observador com a resolução holográfica  $R_H(i)$  puder observar esse sistema. Nestas circunstâncias, perguntamos pela quantidade de informação física a que esses observadores terão acesso quando observam o mesmo sistema físico  $F$ .

De acordo com (1.72),  $O_1$  e  $O_2$  utilizarão ecrãs holográficos cujas áreas são, respectivamente;  $A[R_H(i)] = n \cdot R_H(i)$  e  $A[R_H(i+1)] = m \cdot R_H(i+1)$  ;  $n, m \in N$ . A cada uma das áreas elementares  $R_H(i)$  e  $R_H(i+1)$  encontra-se associada a transferência de 1 *bit* de informação física em relação, respectivamente, a  $O_1$  e  $O_2$ . Uma vez que  $O_1$  observa  $F$ , o ecrã holográfico que utiliza não pode ter área menor que  $R_H(i)$ . Este é o menor valor da área de uma superfície associável ao sistema físico  $F$  que o observador  $O_1$  pode conceber, quando observa  $F$ . Aquele valor de área deverá, portanto,

ser uma propriedade física do sistema, que não pode ser escamoteada por  $O_2$ , ainda que este o observe utilizando outra resolução holográfica.

Admitamos, entretanto, que a quantidade mínima de informação física de que  $O_1$  necessita para especificar o sistema físico  $F$  é superior a 1 *bit*. Daqui se segue que o número de áreas elementares  $R_H(i)$  que compõem o ecrã holográfico, utilizado pelo observador  $O_1$  para observar o sistema físico  $F$ , é igual ao número de bits de informação física que esse sujeito apreende sobre  $F$ , uma vez que é transferido 1 *bit* de informação física por área elementar. Dessa maneira, na observação do sistema físico  $F$ ,  $O_1$  considera uma superfície que lhe está associada, cuja área é  $A[R_H(i)] = n \cdot R_H(i)$ , tal que esta deve ser uma propriedade do sistema observado, que uma vez mais não pode ser escamoteada pelo observador  $O_2$ . Consequentemente, numa situação em que não sejam contemplados outros efeitos físicos - de natureza por exemplo relativística - os dois observadores concordarão em relação ao valor da área da superfície associada ao sistema físico  $F$ , no decurso da sua observação. Isto é; ter-se-á:

$$(3.74) \quad A[R_H(i)] = A[R_H(i+1)] \quad ; \quad n \in N$$

Esta relação equivale a escrever:

$$(3.75) \quad n \cdot R_H(i) = m \cdot R_H(i+1) \quad ; \quad n, m \in N$$

Que recorrendo à igualdade (1.69) no enunciado do Princípio da Resolução Holográfica, se transforma na expressão:

$$(3.76) \quad n \cdot 4 \cdot R_H(i+1) = m \cdot R_H(i+1) \quad ; \quad n, m \in N$$

Este facto significa que:

$$(3.77) \quad m = 4 \cdot n \quad ; \quad n, m \in N$$

Neste ponto, reparamos que  $m$  representa a quantidade de informação física que

$O_2$  recebe quando observa o sistema físico  $F$  e que, analogamente,  $n$  representa a quantidade de informação física de que  $O_1$  se apropria quando observa o mesmo sistema. A unidade de informação, nos dois casos, é o *bit* e está associada à capacidade comum que os observadores dispõem de diferenciar estados físicos, distintos e independentes, no decurso de uma observação. Convencionando que  $I_f[F; R_H(i)]$  representa a quantidade de informação física que um observador recebe, quando observa um sistema físico  $F$  utilizando a resolução holográfica  $R_H(i)$ , obtemos finalmente a relação procurada:

$$(3.78) \quad I_f[F; R_H(i+1)] = 4 \cdot I_f[F; R_H(i)] \text{ (bit)}$$

Que significa que observadores com resoluções holográficas mais baixas observarão o mundo com grau de precisão superior, apropriando-se de maiores quantidades de informação física sobre os sistemas físicos que observam.

Intuitivamente, este resultado sugere a presença de uma forma de variabilidade nomotética, no sentido de que a quantidade de entropia termodinâmica e outras grandezas associadas, como a massa e a temperatura, devem variar em função da resolução holográfica com que se observa a natureza. Mais ainda, de (1.71), relativizando o valor mínimo da área de referência, concluímos que no argumento que nos conduziu à previsão teórica de  $A_0(0) = l_p^2$ , um ou mais dos resultados utilizados nessa dedução devem variar em função do valor da resolução holográfica. Eventualmente, o valor da constante de Planck na primeira relação de incerteza de Heisenberg, a forma da métrica de Schwarzschild para um buraco negro ou alguma das constantes na expressão para o raio do horizonte do buraco negro;  $r_{bn} = \frac{2 \cdot G \cdot M}{c^2}$ .

A variação possível das constantes da física é um tópico actual no qual se inscreve a tentativa do físico português António Magueijo [1968 -] para ultrapassar o «problema do horizonte» na cosmologia do Big Bang, como alternativa à «teoria da inflação cósmica», proposta por Alan Guth [1947 -] (Magueijo, 2003). O problema do horizonte diz respeito ao facto de que o universo é homogéneo e que esse facto requer que, imediatamente após o Big Bang, se tenham verificado interacções físicas entre zonas distantes do universo. Algumas destas, porém, estariam causalmente desconectadas, no sentido de que a distância entre elas não poderia ser percorrida por um raio de luz no tempo adequado para que uma interacção ocorresse. Desse modo, Magueijo sugeriu que a

velocidade da luz seria então superior ao seu valor presente, formulando uma «teoria cosmológica de velocidade da luz variável» («VSL; variable speed of light theory of cosmology»). A constante de estrutura fina, notada pela letra grega  $\alpha$ , caracteriza a intensidade da interacção electromagnética. A constante é adimensional e a sua expressão definítoria e respectivo valor, de acordo com o sistema de unidades SI, são dados por:

$$(3.79) \quad \alpha = \frac{e^2}{\hbar \cdot c \cdot 4 \cdot \pi \cdot \epsilon_0} = 7,29735 \times 10^{-3}$$

Onde  $e$  é a carga eléctrica elementar,  $e = 1,60218 \times 10^{-19} \text{ C}$ ;  $\hbar$  é a constante reduzida de Planck,  $\hbar = 1,05457 \times 10^{-34} \text{ J} \cdot \text{s}$ ;  $c$  é a velocidade de propagação da luz no vazio,  $c = 2,99792 \times 10^8 \text{ m} \cdot \text{s}^{-1}$  e  $\epsilon_0$  é a permitividade eléctrica no vazio,  $\epsilon_0 = 8,85418 \times 10^{-12} \text{ F} \cdot \text{m}^{-1}$ .

Em 1999, uma equipa liderada pelo astrónomo australiano John Webb verificou experimentalmente a existência de indícios da variação temporal da constante de estrutura fina (Webb *et al.*, 1999, 2001). Sendo a constante adimensional, a sua variação significa que uma das constantes em (1.79) poderá ter variado ao longo do tempo. Este facto embora relevante para o sistema de unidades SI, é pelo contrário relativamente inócuo para o sistema de unidades de Planck, em consequência do procedimento de adimensionalização das constantes, que pode ser reafirmado permanecendo ainda válidas as relações entre todas as constantes. Razão pela qual semelhante variação não teria consequências de maior para as ideias propostas neste trabalho. O facto de que a resolução holográfica pode influir na descrição física do mundo invoca um problema crítico de objectividade no mundo, no sentido de que, perante um mesmo fenómeno físico, observadores hipotéticos, com resoluções holográficas distintas, interagirão de forma diferente com os mesmos sistemas físicos, experimentando situações físicas díspares e produzindo discursos teóricos discordantes.

Essa pode ser a situação, relativamente, ao raio de um buraco negro de Schwarzschild.

Admitam-se dois observadores  $O_1$  e  $O_2$ , tais que o segundo utiliza uma resolução holográfica inferior à de  $O_1$  para observar o mundo. Seja um sistema físico, inicialmente observável pelos dois sujeitos, que se aproxima de um buraco negro de Schwarzschild.

Existirá um caso em que  $O_2$  verificará a presença do sistema físico situado fora do horizonte do buraco negro que observa, tal que o sistema ainda se poderá afastar desse horizonte. Pelo contrário, o observador  $O_1$  afirmará que o mesmo sistema físico colapsou no buraco negro, sendo-lhe impossível observá-lo no futuro. Existe pois aparente quebra de objectividade. As físicas dos dois observadores são inconsistentes e, aparentemente, a única solução para o problema é a de que observadores com resoluções holográficas diferentes habitarão universos distintos e incomunicáveis. Este facto não torna inconsistente que cada observador, a partir do seu ponto de vista, possa prever a possibilidade da existência de outros universos, tal como se sugeriu aqui, no decurso da solução do paradoxo holográfico, e tal como sucede com a «interpretação dos universos paralelos» («many world interpretation») para a mecânica quântica, proposta pelo físico americano Hugh Everett III [1930 - 1982] (Everett, 1957). Porém, lança dúvidas sobre a origem da possibilidade dessa capacidade de previsão. Afinal, qual poderá ser o significado de um universo alternativo, cuja existência é possível, mas do qual os indícios empíricos são, por definição, nulos. A questão desse modo coloca-se; poderão observadores com resoluções holográficas diferentes interagir, coexistindo no mesmo universo, tal que é possível formular uma teoria completa consistente, envolvendo todos as visões epistémicas possíveis?

Uma resposta possível pode ser equacionada da seguinte maneira.

Suponhamos um observador  $O_2$  que utiliza um ecrã holográfico cuja área é  $A_H = 4 \cdot R_H(i)$ , associado à observação do sistema físico  $F$ . Admita-se outro observador  $O_1$  que observa o mesmo sistema, utilizando um ecrã holográfico de área igual à sua resolução holográfica  $R_H(i-1)$ , tal que se tem  $R_H(i-1) = A_H$ . O observador  $O_2$  tem à sua disposição uma quantidade de informação física de 4 *bit*, correspondendo a um universo de  $2^4$  estados físicos distintos e independentes para o sistema. O observador  $O_1$ , por seu lado, recebe de  $F$  uma quantidade de informação física de 1 *bit*, correspondendo a  $2^1$  estados físicos distintos e independentes observáveis no mesmo sistema.

De que modo podemos relacionar estes dois factos, intentando uma descrição objectiva do mundo, comum aos dois observadores? Uma possibilidade será considerar que os dois estados físicos que  $O_1$  observa correspondem a dois ou mais subconjuntos distintos do conjunto de estados físicos observáveis por  $O_2$ . Mais exactamente, nos termos de uma teoria que envolva vários observadores com resoluções holográficas diferentes, que dado o mesmo sistema físico  $F$ , a observação de um estado físico  $E_f[R_H(i-1)]$  de  $F$ , a

uma resolução holográfica  $R_H(i-1)$ , é equivalente à observação de qualquer estado físico  $E_f[R_H(i)]$  do mesmo sistema num subconjunto do conjunto de estados, observáveis a uma resolução holográfica  $R_H(i)$ , desde que esse facto não produza uma inconsistência na teoria informal.

Teríamos, assim, duas situações possíveis:

a) Aquela em que dois observadores  $O_1$  e  $O_2$ , com resoluções holográficas diferentes, observariam estados físicos equivalentes - no sentido de que um estado físico, observável por  $O_1$ , é equivalente à observação de qualquer estado num subconjunto de estados observáveis por  $O_2$  -, ocorrendo a possibilidade de interacção entre observadores e de uma descrição objectiva do mundo, mediante a equivalência entre estados estabelecida;

b) A situação em que essa equivalência não se verifica, implicando a impossibilidade de interacção entre os observadores.

Esta abordagem do problema da consistência entre duas visões da natureza, sob resoluções holográficas distintas, e conseqüente concordância objectiva sobre a realidade, sugere que o contacto entre universos díspares poderia suceder, ainda que esporadicamente, desde que esse facto não implicasse inconsistência numa teoria completa sobre o mundo e, por conseguinte, violação do requisito de objectividade. Os estados físicos, por outro lado, não necessitam de ser os mesmos, correspondendo às mesmas propriedades físicas atribuíveis a um sistema. De forma intuitiva, esperaríamos então que físicas diferentes, correspondendo a resoluções holográficas diferentes, apresentassem alguns aspectos logicamente consistentes, permitindo, nuns casos, interacção entre universos, e ainda outros aspectos logicamente inconsistentes, proibindo noutros casos essa interacção. Dessa forma, a consistência da teoria informal manter-se-ia sempre salvaguardada. Mais ainda, esperaríamos que a observação de uma dada propriedade física, utilizando uma resolução holográfica, correspondesse à observação, não necessariamente da mesma propriedade, mas de outra propriedade física, utilizando uma resolução holográfica distinta, tal que a equivalência formal referida entre estados físicos se verificasse válida.

Por exemplo, na situação anterior, envolvendo um buraco negro de Schwarzschild,

poderíamos hipostasiar que a informação que o observador  $O_2$  recebe de uma zona no espaço - que  $O_1$  identifica com o interior do buraco negro - é equivalente à radiação de Hawking, do ponto de vista de  $O_1$ . Se do ponto de vista de  $O_2$ , o sistema físico se afastasse do buraco negro, essa situação seria então ainda vista como uma forma de radiação por parte de  $O_1$ , tal que o observador  $O_2$ , do ponto de vista de  $O_1$ , seria também discernível como tal. Por fim, esperaríamos que observadores com resoluções holográficas sucessivamente mais baixas, recenseassem novas propriedades sobre o mundo, em função da maior quantidade de estados físicos que as suas capacidades de observação permitem referenciar.

O que foi afirmado sugere assim que a assumpção do Princípio da Resolução Holográfica requer que se tenha em consideração a possibilidade da existência de vários universos exibindo comportamentos físicos diferentes, dependendo das resoluções holográficas dos observadores, que são consistentemente relacionáveis entre si. Essa variedade de possibilidades de comportamento físico e respectivas equivalências - entre o que é observável para o mesmo sistema físico por observadores com resoluções holográficas distintas - não deixa, em princípio, de poder ser consistentemente descrito por uma só teoria informal sobre a natureza. Continua a poder ser válida a tese da estabilidade nomotética, em que as proposições de lei são agora outras, mais complexas e abrangentes, tão só permitindo uma fenoménica mais extensa. A especificação teórica das relações entre as físicas formuláveis por observadores com resoluções holográficas diferentes, e respectivos casos em que a interacção entre estes é ou não possível, está para além do âmbito deste trabalho.

O fundamento para a evocação do Princípio da Resolução Holográfica reside no facto de que este se revelará útil na análise da tese da estabilidade nomotética. Recorrendo a uma forma de coordenação do espaço consistente com aquele princípio, construiremos um argumento que sugere a possibilidade de ocorrência de variabilidade nomotética na natureza. Munidos deste resultado, confrontaremos posteriormente a tese forte da inteligência artificial, sustentando que o pensamento humano é equivalente a uma computação numa teoria informal sobre a natureza, e a tese fisicalista da inteligência artificial, mantendo que a simulação artificial da inteligência humana é possível, mediante intervenção técnica na natureza, recorrendo a processos físicos.

## **2. Argumento da Variabilidade Nomotética (versão fisicalista).**

### **2.1. Coordenação holográfica de um evento numa região de espaço.**

Na secção anterior analisamos a relação entre informação e observação de um sistema físico. Definimos que um observador utiliza um ecrã holográfico sempre que observa o mundo, cuja área representa uma medida da capacidade de transferência de informação física sobre o sistema, através do espaço. Identificando a ocorrência de uma contradição, relativa à invocação de dois valores mínimos de área na realidade, denominada paradoxo holográfico, sugerimos que o menor dos valores da área de um ecrã holográfico poderia caracterizar as capacidades de observação de uma classe observadores, propriedade a que chamámos resolução holográfica. Como solução para o paradoxo holográfico e subentendendo a possibilidade de classes distintas de observadores no universo, sugerimos, finalmente, o Princípio da Resolução Holográfica como lei da natureza.

Nesta secção construiremos um argumento indiciando a existência de variabilidade nomotética no mundo. A estratégia que adoptaremos será a seguinte.

Respeitando o Princípio da Resolução Holográfica, começaremos por definir uma forma de coordenação de um evento numa região de espaço euclidiano. Essa coordenação, que denominaremos «coordenação holográfica», será feita sobre um ecrã holográfico associado a uma superfície num plano. Caracterizaremos então um conjunto de propriedades que uma teoria informal sobre a natureza deverá exhibir. Em seguida formaremos a lista de todos os valores de abcissa relativos à coordenação holográfica sobre o ecrã dado, reconhecendo que no âmbito dessa teoria, qualquer um desses valores deverá ser mensurável utilizando a resolução holográfica adequada.

Respeitando o conjunto de propriedades que uma teoria sobre a natureza deve cumprir, definiremos o que se deve entender por «proposição contingente» e «proposição universal» no contexto da mesma. Finalmente, sobre a lista de valores de abcissa definida - e assumindo a validade da tese da estabilidade nomotética - aplicaremos um procedimento matemático conhecido por «diagonalização». Em decurso, identificaremos uma inconsistência a partir da qual será finalmente deduzido o «argumento da variabilidade nomotética».

Iniciamos o nosso trajecto invocando o conceito de ecrã holográfico associado a uma observação e respectiva relação com o Princípio Holográfico. Um observador terá, em princípio, acesso a eventos num ecrã holográfico que estão causalmente ligados a outros

eventos no passado desse observador. O conjunto desses eventos define toda a informação física sobre todos os sistemas físicos que podem ser observados no universo através do ecrã. É com base nos eventos observáveis no mesmo e na informação física que os mesmos definem que um observador pode formular proposições sobre o mundo e construir uma teoria descrevendo o comportamento das coisas. As proposições devem então ser formuladas no contexto de uma determinada correspondência. Os eventos observáveis num ecrã holográfico são tema de afirmações que caracterizam o seu comportamento. Este, por outro lado, deve ter alguma forma de correspondência, quer causal, com eventos distantes do ecrã, cuja informação física se diz transferida através daquele. Noutros termos, a teoria física que descreve o comportamento dos eventos - observáveis no ecrã holográfico por um observador próximo - deve cumprir uma correspondência lógica com a teoria física utilizada para descrever o comportamento dos objectos afastados do ecrã, causalmente conectados aos sistemas físicos cujo comportamento é observável ao nível da superfície bidimensional dada. Esta correspondência é afinal o que estabelece o Princípio Holográfico que, na sua forma mais geral, propõe que num universo de dimensão  $D$ , a descrição completa de uma região  $R$  que tem fronteira  $B = \partial R$  de dimensão  $D-1$ , pode ser obtida através da descrição de todos os processos físicos que ocorrem em  $B$ .

Quando se afirma que são feitas observações sobre um ecrã holográfico, isso não significa que as mesmas se efectuem mediante algum processo experimental capaz de produzir cortes bidimensionais no espaço, anulando de alguma forma a sua tridimensionalidade. O significado daquela asserção é o de que são possíveis observações que dizem respeito a sistemas físicos, cuja posição e dimensões espaciais podem ser consideradas localizáveis em regiões tridimensionais, numa vizinhança muito próxima de uma superfície bidimensional imaginária chamada ecrã holográfico. Do ponto de vista de um observador, utilizando uma determinada resolução holográfica  $R_H(i)$ , a amplitude da vizinhança referida, pode-se convencionar definir um paralelepípedo cuja área é  $A = n \cdot R_H(i)$ ;  $n \in \mathbb{N}$  e tal que a sua espessura é dada por  $[R_H(i)]^{1/2}$ , correspondendo a um ecrã holográfico.

O Princípio da Resolução Holográfica garante que podemos considerar espessuras progressivamente menores, aproximando-se infinitamente de zero, sem contudo atingir esse limite, porquanto uma resolução holográfica nula não pode ser considerada. De facto, se assim fosse, a relação (1.69) não seria válida, o que é consistente com a circunstância de que através de um ecrã holográfico de área nula não pode ser transferida informação física e, conseqüentemente, não pode ser feita qualquer observação.

Estas afirmações conferem precisão ao conceito de ecrã holográfico.

O Princípio Holográfico, por seu lado, assegura que os eventos que ocorrem no interior do paralelepípedo ideal definido - observáveis a resoluções holográficas diferentes - permitem a qualquer observador descrever o comportamento físico dos sistemas físicos distantes do ecrã holográfico e produzir uma teoria sobre a realidade. Uma teoria informal, que descreve a maioria dos fenómenos deverá considerar todas as classes de observadores no universo, associadas a todas as resoluções holográficas possíveis. Esta teoria pode, em princípio, ser formulada por qualquer observador em qualquer resolução holográfica, desde que seja respeitado o critério lógico de realismo. Dessa maneira, podemos considerar do ponto de vista de cada observador, em qualquer classe de observadores, um ecrã holográfico composto por todas as resoluções holográficas menores que a sua área, tal que esta pode ser arbitrária.

Seja um observador que coordena o espaço a partir de um referencial tridimensional ortonormado numa região de espaço-tempo euclidiano, onde se situa um ecrã holográfico, perpendicular a um dos eixos coordenados, no qual o observador regista todos os eventos, simultâneos do seu ponto de vista, causalmente relacionados com eventos situados no seu passado, obtendo assim informação para enunciar proposições sobre o mundo e respectivos valores lógicos; «1» indicando que uma proposição é verdadeira ou «0», indicando que a mesma é falsa. A teoria em causa será uma teoria empírica sobre a natureza e, em vista da tese da estabilidade nomotética, será a mesma do ponto de vista de todos os observadores, permitindo prever o que é observável, utilizando qualquer resolução holográfica. Definimos agora uma forma de coordenação de eventos ocorrendo sobre um ecrã holográfico que cumpra o Princípio da Resolução Holográfica.

Definição de sistema de coordenação holográfica de eventos num ecrã holográfico.

Chame-se região holográfica inicial a um ecrã holográfico que é um quadrado, cuja área é  $\mathcal{R}_H$ , verificando a relação  $A(\mathcal{R}_H) = 4^j \cdot R_H(0)$  ;  $j \in N_0$ , onde  $R_H(0) = 4 \cdot l_p^2$ , contido num plano e situada numa região de espaço euclidiano.

Chame-se  $\mathcal{L}_H$  ao comprimento holográfico da região holográfica inicial, tal que se tem  $A(\mathcal{R}_H) = [\mathcal{L}_H]^2$ , isto é;  $\mathcal{L}_H = 2^{j+1} \cdot l_p$  ;  $j \in N_0$ .

Seja um observador que utiliza um referencial tridimensional ortonormado e considere-se o plano  $P$  definido pelos eixos  $xx$  e  $yy$  como a superfície bidimensional que

contém a região holográfica inicial  $\mathcal{R}_H$ .

Convencione-se que aquela se situa no primeiro quadrante de  $P$ , com o vértice inferior esquerdo coincidindo com a origem das coordenadas.

Associe-se a cada região  $R_H[x, y, R_H(i)]$  em  $\mathcal{R}_H$ , região que é um quadrado e cuja área é igual a uma das resoluções holográficas  $R_H(i)$  (ver fig. 4.2. p. 238), o trio ordenado de números  $[x, y, R_H(i)]$  coordenando o vértice inferior direito daquele quadrado e verificando as seguintes relações:

$$(2.1) \quad x = r.[R_H(i)]^{1/2} \quad ; \quad r \in N$$

Em que  $r$  representa o número de regiões com resolução holográfica  $R_H(i)$ , incluindo a região holográfica considerada, contadas até ao vértice inferior direito dessa região, situadas entre duas rectas paralelas ao eixo dos  $xx$ , contendo as arestas superior e inferior de  $R_H[x, y, R_H(i)]$  e distando  $[R_H(i)]^{1/2}$  entre si, segundo uma recta paralela ao eixo dos  $yy$ .

E:

$$(2.2) \quad y = s.[R_H(i)]^{1/2} \quad ; \quad s \in N$$

Tal que  $s$  representa o número de regiões com resolução holográfica  $R_H(i)$ , incluindo a região holográfica considerada, contadas até ao vértice superior direito dessa região, situadas abaixo dessa região entre duas rectas paralelas ao eixo dos  $yy$ , contendo as arestas esquerda e direita de  $R_H[x, y, R_H(i)]$  e distando  $[R_H(i)]^{1/2}$  entre si, segundo uma recta paralela ao eixo dos  $xx$ .

Os valores máximos de  $r$  e de  $s$  devem ser tais que  $x \leq \mathcal{L}_H$  e  $y \leq \mathcal{L}_H$ .

Chamaremos a  $x$  o valor da abcissa da coordenação holográfica da região  $R_H[x, y, R_H(i)]$  e a  $y$  o respectivo valor da ordenada.

Para cada resolução holográfica  $R_H(i) \leq R_H(0)$ , seja  $\mathcal{E}_H(i)$  o conjunto de eventos em  $\mathcal{R}_H$ , observáveis a essa resolução, em que a extensão de cada um desses

eventos se confina a  $R_H[x, y, R_H(i)]$ .

Sempre que um evento em  $\mathcal{E}_H(i)$  for observável numa região  $R_H[x, y, R_H(i)]$  em  $\mathcal{R}_H$ , utilizando a resolução holográfica  $R_H(i)$ , diz-se que esse evento é coordenável holograficamente, à resolução holográfica  $R_H(i)$ , sendo-lhe atribuído o trio de coordenadas  $[x, y, R_H(i)]$  de acordo com (2.1) e (2.2).

A região holográfica inicial  $\mathcal{R}_H$  pode ter uma área superior ou igual à resolução holográfica  $R_H(0)$ , de acordo com o requisito de que seja um quadrado cujo lado é  $\mathcal{L}_H = 2^{j+1} \cdot l_p$ ;  $j \in N_0$ . A resolução holográfica  $R_H(0) = 4 \cdot l_p^2$  é aquela que caracteriza a classe de observadores humanos, em consistência com o ponto de vista holográfico a partir do qual podemos formular teorias sobre a natureza. Na perspectiva de cada observador - utilizando a sua resolução holográfica própria - os eventos escolhidos para serem coordenados holograficamente são justamente aqueles que permitem a descrição do mundo ao nível fundamental da matéria.

Consideramos em decurso uma forma comum de coordenação holográfica a partir da relação entre resoluções holográficas e utilizando para unidade de coordenação o comprimento holográfico  $\mathcal{L}_H$  da região holográfica inicial  $\mathcal{R}_H$ .

Admita-se que se verifica  $A(\mathcal{R}_H) = 4^j \cdot R_H(0)$ ;  $j \in N_0$ ,  $j \neq 0$ .

Ter-se-á então o seguinte:

$$(2.3) \quad \mathcal{L}_H = 2^j [R_H(0)]^{1/2} = 2^j \cdot L_H(0)$$

Onde  $L_H(0) = 2 \cdot l_p$  representa o menor comprimento que um observador na classe de observadores humanos pode medir, utilizando a resolução holográfica  $R_H(0)$ .

Da relação anterior conclui-se que:

$$(2.4) \quad L_H(0) = 2^{-j} \cdot \mathcal{L}_H$$

Utilizando as relações (2.1) e (2.2), verificamos que para coordenações holográficas feitas usando  $R_H(0)$ , é:

$$(2.5) \quad x = r \cdot 2^{-j} \cdot \mathcal{L}_H \quad ; \quad r, j \in N, \quad 1 \leq r \leq 2^j$$

Em que para  $r = 2^j$ , se tem  $x = \mathcal{L}_H$ , e:

$$(2.6) \quad y = s \cdot 2^{-j} \cdot \mathcal{L}_H \quad ; \quad s, j \in N, \quad 1 \leq s \leq 2^j$$

Tal que para  $s = 2^j$ , se obtém  $y = \mathcal{L}_H$ .

Os valores de  $r$  e de  $s$  correspondem a números inteiros positivos. Desse modo,  $x$  e  $y$  podem facilmente ser representados num sistema numérico de base 2, tal que resulta:

$$(2.7) \quad x = (a_0 \cdot 2^0 + a_1 \cdot 2^{-1} + \dots + a_k \cdot 2^{-k} + \dots + a_j \cdot 2^{-j}) \cdot \mathcal{L}_H \quad ; \quad a_k \in \{0,1\}$$

$$(2.8) \quad y = (b_0 \cdot 2^0 + b_1 \cdot 2^{-1} + \dots + b_k \cdot 2^{-k} + \dots + b_j \cdot 2^{-j}) \cdot \mathcal{L}_H \quad ; \quad b_k \in \{0,1\}$$

Expressões válidas desde que para  $a_0 = 1$  ou  $b_0 = 1$ , os restantes factores, respectivamente;  $a_k$  e  $b_k$ , sejam nulos.

De acordo com (1.70) é válida a seguinte relação:

$$(2.9) \quad L_H(i) = 2^{-i} \cdot L_H(0) \quad ; \quad i \in N_0$$

Ou seja, usando (2.4);

$$(2.10) \quad L_H(i) = 2^{-(j+i)} \cdot \mathcal{L}_H \quad ; \quad j \in N, \quad i \in N_0$$

Daqui resulta:

$$(2.11) \quad x = r \cdot 2^{-(j+i)} \cdot \mathcal{L}_H \quad ; \quad r, j \in N, \quad i \in N_0, \quad 1 \leq r \leq 2^{j+i}$$

Em que para  $r = 2^{j+i}$ , se tem  $x = \mathcal{L}_H$ , e:

$$(2.12) \quad y = s \cdot 2^{-(j+i)} \cdot \mathcal{L}_H \quad ; \quad s, j \in N \quad , \quad i \in N_0 \quad , \quad 1 \leq s \leq 2^{j+i}$$

Em que para  $s = 2^{j+i}$ , se obtém  $y = \mathcal{L}_H$ .

Fazendo  $n = j+i$  e adoptando o mesmo raciocínio de há pouco, obtemos como resultado que os valores  $x$  e  $y$ , relativos a uma coordenação holográfica, utilizando qualquer resolução holográfica  $R_H(i)$ ;  $i \in N_0$ , podem ser expressos na forma:

$$(2.13) \quad x = (a_0 \cdot 2^0 + a_1 \cdot 2^{-1} + \dots + a_k \cdot 2^{-k} + \dots + a_n \cdot 2^{-n}) \cdot \mathcal{L}_H \quad ; \quad n \in N_0 \quad , \quad a_k \in \{0,1\}$$

$$(2.14) \quad y = (b_0 \cdot 2^0 + b_1 \cdot 2^{-1} + \dots + b_k \cdot 2^{-k} + \dots + b_n \cdot 2^{-n}) \cdot \mathcal{L}_H \quad ; \quad n \in N_0 \quad , \quad b_k \in \{0,1\}$$

Uma vez mais, desde que, para  $a_0 = 1$  ou  $b_0 = 1$ , os restantes factores, respectivamente;  $a_k$  e  $b_k$ , sejam nulos.

A figura seguinte ilustra o que foi definido para a coordenação holográfica, recorrendo ao comprimento holográfico  $\mathcal{L}_H$  da região holográfica inicial  $\mathcal{R}_H$  como unidade de comprimento.

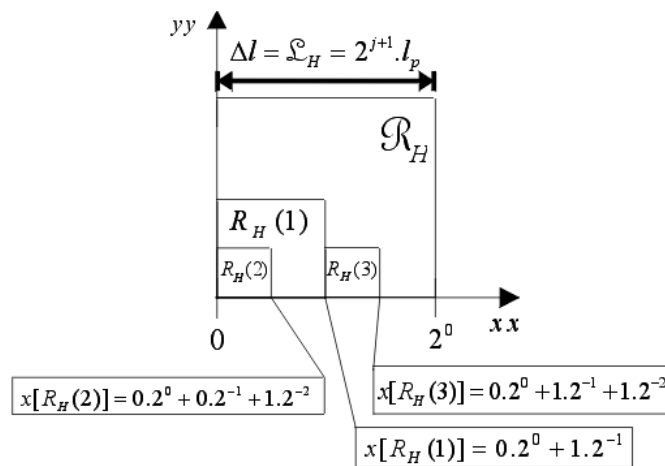


Figura A2.1 - Coordenação holográfica de eventos em regiões holográficas, utilizando o comprimento holográfico  $\mathcal{L}_H$  como unidade de comprimento. Representou-se apenas o valor de abscissa  $x$  na coordenação.

Por uma questão de clareza, indicou-se apenas a coordenação relativa aos valores de abcissa. As ordenadas são obtidas recorrendo a representação idêntica. Na figura adoptou-se ainda uma notação sintética em que, por exemplo,  $R_H(1)$  representa a região holográfica  $R_H[x, y, R_H(1)]$ , com coordenação holográfica obtida utilizando a resolução holográfica  $R_H(1)$  e em que  $x[R_H(1)]$  representa o valor de  $x$  para a respectiva abcissa da região holográfica coordenada.

Recorrendo ao critério lógico de realismo, qualquer observador pode considerar teoricamente qualquer coordenação holográfica, utilizando qualquer valor da resolução holográfica, sucedendo no entanto que o mesmo estará impedido de observar eventos, cuja observação só se possa efectivar a resoluções holográficas inferiores à sua.

## 2.2. Propriedades de uma teoria sobre a natureza.

Definida uma forma de coordenação holográfica, compatível com o Princípio da Resolução Holográfica, procedemos à consideração do conjunto mínimo de propriedades que uma teoria informal sobre a natureza deverá cumprir.

A teoria em causa deverá ser tal que em seu âmbito se podem definir variáveis cujo domínio são classes de conjuntos. A teoria incluirá um conjunto denumerável e finito de proposições fundamentais verdadeiras sobre a natureza, onde se incluem os observadores, chamadas postulados, identificáveis com as leis da natureza.

Sempre que possível, a cada proposição na teoria será atribuído um valor lógico que, de acordo com o Princípio do Terceiro Excluído, deverá ser um de dois valores possíveis: aquele que indica que a proposição é verdadeira - facto representado pelo símbolo « $V$ » - ou aquele que indica que a mesma é falsa - facto representado pelo símbolo « $F$ ». Sempre que não seja possível atribuir, no âmbito da teoria, um valor lógico a uma proposição verdadeira, esta dir-se-á uma «proposição indecidível».

Se for possível deduzir na teoria que o valor lógico de uma determinada proposição existe, isto é; que uma proposição na teoria é necessariamente verdadeira ou falsa, independentemente do facto desse valor não poder ser atribuído à proposição através de uma dedução, a proposição dir-se-á «não computável na teoria dada». Na teoria será possível formular, a partir de qualquer proposição, outra proposição chamada negação da proposição dada, representada pelo símbolo « $\sim$ » seguido da proposição que se pretende negar. Será possível agrupar várias proposições em conjuntos ou classes de proposições.

A teoria terá a possibilidade de considerar proposições pertencentes a conjuntos

finitos ou infinitos delas mediante a utilização livre de símbolos de quantificação, em que o símbolo « $\forall$ » representa o quantificador universal, referindo-se a todas as proposições no conjunto e o símbolo « $\exists$ » representa o quantificador existencial, referindo-se a pelo menos uma proposição no conjunto dado. O símbolo « $\in$ » será utilizado para expressar a relação de pertença de uma proposição a um conjunto de proposições. Para além destes símbolos serão ainda utilizados na teoria aqueles associados a operações habituais do cálculo proposicional, com a notação seguinte: « $\Rightarrow$ » para implicação material entre duas proposições; « $\Leftrightarrow$ » para equivalência material entre duas proposições; « $\wedge$ » para conjunção entre duas proposições e « $\vee$ » para disjunção entre duas proposições.

Na teoria sobre a natureza será possível contar objectos.

Desse modo, estarão ali definidos os números naturais, formando o conjunto  $N_0$ .

Sobre este conjunto estarão definidas as quatro operações aritméticas elementares: adição, subtração, multiplicação e divisão. Recorrendo às mesmas será ainda possível definir a operação potenciação, permitindo a definição do conjunto dos números fracionários positivos  $Q^+$  e, em particular, permitindo o desenvolvimento na base 2 dos valores de abcissa na coordenação holográfica.

Por consistência de uma teoria entenderemos o facto de que não é possível deduzir nessa teoria uma proposição e a sua negação. Requeremos que a teoria seja consistente. Sobre a relação entre aspectos observáveis no mundo e proposições representáveis na teoria são válidas as seguintes afirmações:

a) Todos os aspectos observáveis na natureza correspondem a proposições na teoria;

b) Nenhuma proposição indecível ou não computável na teoria poderá corresponder a aspectos observáveis na natureza;

c) Todas as proposições que se refiram a aspectos observáveis na natureza deverão ser falsificáveis ou cumprir o critério lógico-formal para a validade de uma proposição sobre a natureza.

d) Todo o aspecto observável no mundo será chamado «evento».

Consideramos que um evento será computável na teoria, se nela existir um processo lógico através do qual é possível obter a proposição que corresponde à sua

ocorrência ou à negação da sua ocorrência. Se não existir tal processo, o evento dir-se-á não computável na teoria considerada. Um evento denominar-se-á absolutamente não computável, se não existir nenhuma teoria onde esse evento seja computável.

Numa teoria sobre o mundo não existem referências a eventos não computáveis e, por conseguinte, a eventos absolutamente não computáveis. De acordo com o requisito b), esses eventos não são observáveis na natureza. A teoria considera-se completa. Descreve, por hipótese, com absoluta inteligibilidade tudo o que sucede no mundo, descrevendo o que é observável no contexto das suas capacidades de previsão. Poderão existir proposições indecidíveis e proposições não computáveis na teoria, porém, estas dizem respeito a aspectos que inerem à própria teoria, mas não ao comportamento actual dos sistemas físicos no mundo. Colocamos como princípio que a teoria é suficientemente ampla para que tudo o que for previsto pela teoria, referente a aspectos sobre a natureza, é necessariamente computável e, como tal, observável. Estas afirmações pretendem invocar a existência de uma identidade ontológica entre o que é observável e o que é computável na teoria. Finalmente, numa teoria completa sobre a natureza deverão ser válidos o critério lógico de realismo, o Princípio Holográfico e o Princípio da Resolução Holográfica, enunciados que caracterizam sistemas físicos específicos chamados observadores e que devem ser entendidos como leis da natureza, pertencendo, desse modo, ao conjunto de postulados da teoria.

### **2.3. A lista de todos os valores de abcissa numa coordenação holográfica de eventos.**

Enunciadas as propriedades que pretendemos que uma teoria informal sobre a natureza cumpra para os nossos propósitos, retornamos agora à coordenação holográfica de eventos e procedemos à consideração da lista contendo todos os valores que a abcissa  $x$  pode assumir nessa coordenação. A teoria sobre a natureza inclui o Princípio da Resolução Holográfica e em vista da tese da estabilidade nomotética deverá ser a mesma para todas as classes de observadores no mundo.

Além disso, assumindo a identidade ontológica entre o que existe e o que é computável na teoria, tem-se que a lista em questão deverá ser computável e completa, no sentido que deverá representar extensamente o aspecto escolhido no conjunto de todas as observações possíveis - no caso; o valor de abcissa na coordenação holográfica. Noutros termos; a teoria deverá ser capaz de prever consistentemente a existência de todo e qualquer valor de abcissa, associada à coordenação de uma região holográfica, na qual se

encontra confinado um evento observável. A teoria em causa deverá ser suficientemente profícua em termos computacionais para que a produção de semelhante lista seja possível, sem que daí resulte qualquer inconsistência.

Só nesta situação é possível conceber que computabilidade simbólica num discurso universal sobre a natureza e comportamento das coisas, num universo inteligível, assumam uma congruência absoluta e final. Trata-se afinal da condição extrema e desejável que se pretende atingir e apenas a sua assumpção permitirá retirar conclusões epistemológicas decisivas relativas à sua viabilidade. O Princípio da Resolução Holográfica, além disso, permite considerar todas as classes de observadores na natureza, perfazendo a identificação sistemática de todas as regiões holográficas numa região holográfica inicial.

Este facto indicia a possibilidade de se obter a lista pretendida, recorrendo a um processo empírico, substanciável do ponto de vista de qualquer observador, através do critério lógico de realismo. Cada observador, caracterizado por uma resolução holográfica  $R_H(i)$  ;  $i \in N$  , pode recensear todas as regiões holográficas possíveis  $R_H$  na região holográfica inicial  $\mathcal{R}_H$  . Desse modo, a construção de uma lista contendo todos os valores de abcissa, correspondendo a todas as coordenações holográficas feitas por observadores em todas as classes de observadores no universo é, em princípio, possível. A lista é computável e todos valores de abcissa nela representados correspondem a proposições computáveis que, por sua vez, referenciam aspectos observáveis na natureza.

A forma geral do valor de abcissa de uma coordenação holográfica, tal como foi definida, é dada pela expressão (2.13). Apelando à simplicidade e sem perda de generalidade em relação ao que se segue, considerem-se todos os valores de abcissa, à excepção de  $x = \mathcal{L}_H$  , tal que  $\mathcal{L}_H = 2^{j+1} \cdot l_p$  ;  $j \in N$  , isto é; tal que a região holográfica inicial  $\mathcal{R}_H$  tem sempre um valor de área superior à resolução holográfica  $R_H(0) = 4 \cdot l_p^2$  .

A forma geral do valor de abcissa na coordenação holográfica de eventos passará a ser:

$$(2.15) \quad x = (a_1 \cdot 2^{-1} + \dots + a_k \cdot 2^{-k} + \dots + a_n \cdot 2^{-n}) \cdot \mathcal{L}_H \quad ; \quad n \in N \quad , \quad a_k \in \{0,1\}$$

Assuma-se pois que a lista de todos os valores de abcissa relativos à coordenação holográfica foi produzida e ordenada de alguma maneira. Essa produção deverá corresponder a um processo computável, no contexto da teoria sobre a natureza, associada ao conjunto de todas as observações holograficamente coordenáveis por todas as classes

de observadores no universo. A forma geral da lista  $\mathcal{Q}_H$  de valores de abcissa, considerando apenas os factores «0» ou «1» nas parcelas escritas na base 2, corresponderá então à seguinte tabela:

$$(4.16) \quad \mathcal{Q}_H \quad \begin{array}{c|cccccc} & 2^{-1} & 2^{-2} & \dots & 2^{-k} & \dots & 2^{-n} \\ \hline x_1[R_H(q)] & a_1(1) & a_1(2) & \dots & a_1(k) & \dots & a_1(n) \\ x_2[R_H(q)] & a_2(1) & a_2(2) & \dots & a_2(k) & \dots & a_2(n) \\ \vdots & \vdots & \vdots & & \vdots & & \vdots \\ x_j[R_H(q)] & a_j(1) & a_j(2) & \dots & a_j(k) & \dots & a_j(n) \\ \vdots & \vdots & \vdots & & \vdots & & \vdots \\ x_n[R_H(q)] & a_n(1) & a_n(2) & \dots & a_n(k) & \dots & a_n(n) \end{array}$$

$$q, n \in N \quad ; \quad a_j(k) \in \{0,1\}$$

Na primeira linha da tabela estão escritos os índices da expressão, na base 2, para o primeiro valor de abcissa possível  $x_1[R_H(q)]$  na lista  $\mathcal{Q}_H$ , associado à coordenação holográfica. Na  $j$ -ésima linha da tabela encontram-se escritos os índices da expressão, na base 2, para o  $j$ -ésimo valor de abcissa possível  $x_j[R_H(q)]$  na lista  $\mathcal{Q}_H$ , associado à coordenação holográfica. Na  $n$ -ésima linha da tabela encontram-se escritos os índices da expressão, na base 2, para o  $n$ -ésimo valor de abcissa possível  $x_n[R_H(q)]$  na lista  $\mathcal{Q}_H$ , associado à coordenação holográfica. Na notação utilizada, a letra  $q$  indica que existem  $q \in N$  regiões holográficas com coordenação holográfica  $R_H[x_j, y, R_H(i)]$  com o mesmo valor de abcissa  $x_j$ . Facto elucidado graficamente pela figura seguinte:

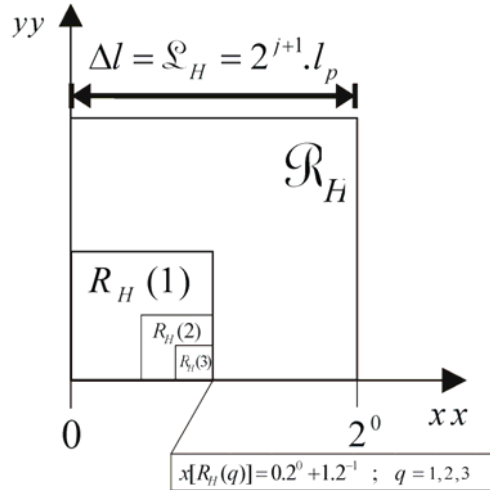


Figura A2.2 - Três regiões holográficas; exibindo o mesmo valor de abcissa, na coordenação holográfica.

A situação em causa resultando da omissão do valor de ordenada  $y$ , para efeitos da formulação da lista  $\mathcal{Q}_H$ . Cada uma das linhas na tabela corresponde a uma sequência de zeros e uns. Um exemplo possível é:

$$(2.17) \quad x_j[R_H(q)] = (0, 1, 0, 1, 0, 0, 1, 1, 0, 1, 0, 0, 1, \dots)$$

Em que as reticências no interior dos parênteses indicam que a sequência pode continuar indefinidamente, desde que à direita do último dígito exista pelo menos outro dígito diferente de zero. Desse modo, a notação  $a_j(k)$  na lista representa a  $k$ -ésima entrada da  $j$ -ésima expressão correspondente ao valor de abcissa  $x_j = x_j[R_H(q)]$  relativo à coordenação holográfica de todas as regiões holográficas, com coordenação holográfica  $R_H[x_j, y, R_H(i)]$ , para as quais é válido aquele valor. Esta observação conclui a especificação da lista  $\mathcal{Q}_H$  de todos os valores de abcissa possíveis - à exceção de  $x = \mathcal{L}_H$  - relativos à coordenação holográfica de eventos na região holográfica inicial  $\mathcal{R}_H$ .

## 2.4. Proposições universais e contingentes numa teoria sobre a natureza.

Um observador utilizando a resolução holográfica  $R_H(i); i \in N_0$ , através de um ecrã holográfico - correspondendo a uma região holográfica  $R_H$  na região holográfica inicial  $\mathcal{R}_H$  - observa e coordena holograficamente eventos nesse ecrã. De acordo com o Princípio Holográfico, os eventos observados definem toda a informação física sobre todos os eventos distantes do ecrã, em relação aos quais os primeiros se encontram causalmente ligados. A partir da informação física recebida sobre o mundo através do ecrã holográfico, o observador pode então formular proposições sobre aquilo que observa, construindo uma teoria sobre o mundo. Desejavelmente e ao fim de algum tempo, aquele sujeito cognoscente produzirá uma teoria completa sobre a natureza. No contexto da mesma, e de acordo com o critério lógico de realismo e com o Princípio da Resolução Holográfica, o observador formula consistentemente proposições sobre o que é observável por outros observadores, utilizando resoluções holográficas distintas da sua.

No conjunto das sentenças formuladas no âmbito na teoria - que são computáveis e decidíveis - existirão necessariamente dois subconjuntos de proposições; o conjunto das proposições que se dizem «universais» e o conjunto das proposições chamadas «contingentes». As proposições nos dois conjuntos poderão dizer respeito a aspectos da própria teoria ou a eventos observáveis no mundo. Em face da tese da estabilidade nomotética, todos os observadores em questão dispõem do conhecimento da mesma teoria sobre o que é observável, utilizando qualquer resolução holográfica. No contexto desse corpo de conhecimentos, por uma «proposição universal» entende-se uma afirmação enunciada na teoria sobre a natureza, cuja forma, significado e valor lógico são os mesmos em relação a todos os observadores no mundo, em qualquer local e tempo do universo, coordenados por cada um desses observadores. Inversamente, por uma «proposição contingente» entende-se uma afirmação formulada na teoria, cuja forma e significado são os mesmos em relação a todos os observadores no mundo, em qualquer local e tempo do universo, coordenados por cada um desses observadores, mas cujo valor lógico pode variar de observador para observador.

O Princípio da Resolução Holográfica é um exemplo de uma proposição universal na teoria. A afirmação de que «um dado valor da resolução holográfica, característico de uma classe específica de observadores no universo  $C_o(i); i \in N_0$ , é válido, para efeitos de observação, numa classe de observadores utilizando uma teoria onde esse valor é formulável», é verdadeira para qualquer observador em  $C_o(i); i \in N_0$  e falsa para os

restantes observadores. A afirmação em causa é, por isso, um exemplo de uma proposição contingente.

Relativamente ao que será tratado na secção seguinte, interessa-nos considerar as proposições universais na teoria informal sobre a natureza e relacioná-las com a coordenação holográfica que um observador efectiva sobre um ecrã holográfico. Sempre que um observador coordena um evento confinado a uma região holográfica  $R_H$ , a teoria que utiliza deverá ser válida para a descrição desse evento e respectivas relações do sistema físico, a que respeita o evento, com os restantes sistemas físicos no universo.

Essa região holográfica, que confina o sistema físico associado ao evento observado, corresponde justamente a um dos aspectos do sistema sob observação. Desse modo, podemos associar a cada uma das regiões holográficas  $R_H$ , na região holográfica inicial  $\mathcal{R}_H$  o conjunto das proposições universais na teoria e, conseqüentemente, os respectivos valores lógicos. Mais ainda, visto que se tratam de proposições universais, podemos definir na teoria uma função  $\mathcal{L}$  que descreve um determinado comportamento associável a todos os observadores no universo, quando perfazem uma observação sobre uma dada região holográfica, utilizando a resolução holográfica adequada.

Esse comportamento é o seguinte. Suponhamos que um observador caracterizado por uma resolução holográfica  $R_H(i)$ ;  $i \in N_0$  observa um evento  $\xi$  confinado a uma região holográfica, cuja coordenação holográfica é  $R_H[x, y, R_H(i)]$ . Admita-se que esse observador detém o conhecimento cabal de uma teoria sobre a natureza  $\mathcal{T}$  e, conseqüentemente, a possibilidade de obter todas as proposições que podem ali ser formuláveis. Consideremos também que o mesmo observador tem à sua disposição um conjunto de proposições que representam outros eventos observados, relacionáveis com o evento em questão.

Podemos agora assumir que a informação empírica e teórica utilizável pelo observador, possa ser suficientemente vasta para permitir a realização de uma experiência cujo intuito consiste em determinar o valor lógico de uma ou mais proposições universais em  $\mathcal{T}$ , no decurso da observação do evento  $\xi$ . Este processo subentende duas *assumpções* fundamentais:

a) Que numa teoria sobre a natureza  $\mathcal{T}$  toda e qualquer proposição universal sobre eventos observáveis possa ser relacionável, em termos da sua consistência, com pelo menos uma proposição representando o evento observado;

b) Que para toda a relação considerada em a) existe uma experiência exequível substanciando essa dependência.

Uma proposição universal é um postulado na teoria e, como tal, o valor lógico da proposição que representa o evento  $\mathcal{E}$  é obtível, através de uma dedução, a partir dessa proposição universal ou, doutro modo; esse evento pode ser causalmente relacionado com outro evento que o origina, tal que a proposição que o representa resulta, mediante dedução, da proposição universal considerada. A validade da segunda condição resulta do facto de que, dada a validade da primeira, a proposição em  $\mathcal{T}$  que representa a dependência entre eventos hipostasiada, sendo uma sentença sobre a realidade, deverá em princípio ser falsificável. O observador deverá realizar várias experiências, utilizando regiões holográficas distintas daquela a que se confina o evento  $\mathcal{E}$ , até que consiga chegar a uma situação física permitindo testar o valor lógico da proposição universal escolhida; ou até que consiga reunir informação física suficiente para o efeito. No entanto, numa situação ideal é admissível que o valor lógico de toda e qualquer proposição universal em  $\mathcal{T}$  possa ser determinado pelo observador, através da observação do evento  $\mathcal{E}$ .

Nessa situação e assumindo que o procedimento descrito é possível e que o mesmo conclui com a afirmação por parte do observador do valor lógico da proposição, definimos que a função  $\mathcal{L}$  descreve, de modo geral, o resultado do conjunto dos procedimentos agenciados por todo e qualquer observador, utilizando a resolução holográfica  $R_H(i)$ ;  $i \in N_0$ , na determinação do valor lógico de qualquer proposição  $p$  na teoria  $\mathcal{T}$ , associada à observação de um evento  $\mathcal{E}$ , confinado a uma região holográfica  $R_H$ . Esta possibilidade pode inclusivamente estender-se a proposições que não representem eventos na natureza. Bastará para tanto que o observador utilize o conceito de informação física para codificar a informação que diz respeito à teoria que utiliza, recorrendo ao evento observado.

Assim, chamando a  $\mathcal{L}$  «função nomotética», definimos o seguinte.

Definição de função nomotética.

Seja  $p$  uma proposição no conjunto de proposições  $\mathcal{P}$  formuláveis na teoria informal sobre a natureza  $\mathcal{T}$ .

Seja um observador que observa um evento  $\mathcal{E}$ , confinado a uma região holográfica  $R_H$ , cuja coordenação holográfica é  $R_H[x, y, R_H(i)]$ , utilizando a resolução

holográfica  $R_H(i)$  ;  $i \in N_0$ .

Assuma-se que o evento  $\mathcal{E}$  se encontra associado à proposição  $p$ , de tal modo que a sua observação permite, através de um procedimento que pode ser descrito na teoria  $\mathcal{T}$ , determinar o valor lógico de  $p$ , «V» ou «F».

Seja  $\{R_H[x, y, R_H(i)]\}$  o conjunto de todas as regiões holográficas pertencentes à região holográfica inicial  $\mathcal{R}_H$ , representadas pelas suas coordenações holográficas.

Chama-se função nomotética  $\mathcal{L}$  à aplicação de  $\mathcal{P} \times \{R_H[x, y, R_H(i)]\}$  sobre  $\{0,1\}$ , em que a cada proposição  $p \in \mathcal{P}$ , cujo valor lógico é determinado mediante a observação de um evento, confinado à região  $R_H$  em  $\mathcal{R}_H$  com coordenação holográfica  $R_H[x, y, R_H(i)]$ , associa um e um só valor em  $\{0,1\}$ , tal que:

$$\mathcal{L} : \mathcal{P} \times \{R_H[x, y, R_H(i)]\} \rightarrow \{0,1\} ;$$

(2.18)

$$\mathcal{L} \left[ p, R_H[x, y, R_H(i)] \right] = \begin{cases} 1 ; & \text{se e só se o valor lógico de } p \text{ é } V \\ 0 ; & \text{se e só se o valor lógico de } p \text{ é } F \end{cases}$$

Com o auxílio da função nomotética, podemos agora explicitar o que se entende por proposição universal e proposição contingente de outra forma.

Seja  $\{R_H(i)\}$  ;  $i \in N_0$  o conjunto de todas as resoluções holográficas e  $\{R_H[x, y, R_H(i)]\}$  ;  $i \in N_0$  o conjunto de todas as regiões holográficas na região holográfica inicial  $\mathcal{R}_H$ , representadas pelas respectivas coordenações holográficas.

Definição de proposição universal numa teoria  $\mathcal{T}$ , em relação a uma região holográfica inicial  $\mathcal{R}_H$ .

Uma proposição  $p \in \mathcal{P}$  em  $\mathcal{T}$  diz-se universal em  $\mathcal{R}_H$  se e só se:

$$(2.19) \quad \forall R_H(i) \in \{R_H(i)\}, \forall R_H \in \{R_H[x, y, R_H(i)]\} :$$

$$\mathcal{L} \left[ p, R_H[x, y, R_H(i)] \right] = 1$$

Definição de proposição contingente numa teoria  $\mathcal{T}$ , em relação a uma região holográfica inicial  $\mathcal{R}_H$ .

Uma proposição  $p \in \mathcal{P}$  em  $\mathcal{T}$  diz-se contingente em  $\mathcal{R}_H$  se e só se:

$$(2.20) \quad \exists R_H(i) \in \{R_H(i)\}, \exists R_H \in \{R_H[x, y, R_H(i)]\} : \\ \mathcal{L} [ p, R_H[x, y, R_H(i)] ] = 0$$

Deste modo, associámos o valor lógico de cada uma das proposições numa teoria sobre a natureza  $\mathcal{T}$  a todas as regiões holográficas numa região holográfica inicial  $\mathcal{R}_H$ .

As figuras seguintes ilustram o resultado dessa associação para uma proposição universal e para uma proposição contingente.

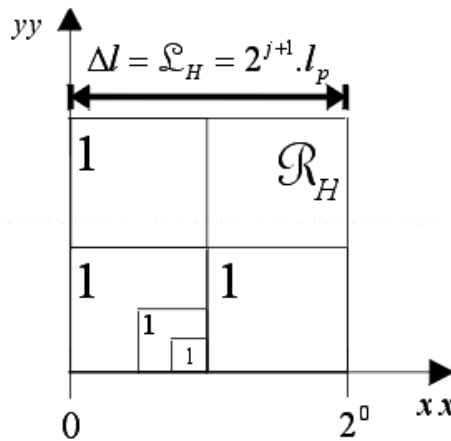


Figura A2.3 - O valor lógico de uma proposição universal associado a cinco regiões holográficas, através do respectivo valor da função nomotética  $\mathcal{L}$ .

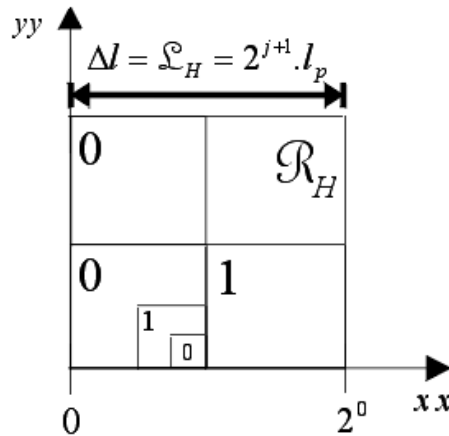


Figura A2.4 - Os valores lógicos de uma proposição contingente associados a cinco regiões holográficas, através do respectivo valor da função nomotética  $\mathcal{L}$ .

A partir das definições anteriores podemos entretanto definir proposição universal e proposição contingente numa teoria sobre o mundo  $\mathcal{T}$ .

Definição de proposição universal numa teoria sobre o mundo  $\mathcal{T}$ .

Uma proposição  $p \in \mathcal{P}$  em  $\mathcal{T}$  diz-se universal se e só se a proposição for universal em relação a todas as regiões holográficas no universo.

Definição de proposição contingente numa teoria sobre o mundo  $\mathcal{T}$ .

Uma proposição  $p \in \mathcal{P}$  em  $\mathcal{T}$  diz-se contingente se e só se a proposição for contingente em relação a pelo menos uma região holográfica no universo.

Definidas proposição universal e proposição contingente, numa teoria sobre a natureza, procedemos à dedução do argumento da variabilidade nomotética.

## 2.5. Dedução do Argumento da Variabilidade Nomotética.

O argumento da variabilidade nomotética produz um resultado que se deduz da consideração conjunta da lista de todos os valores de abcissa na coordenação holográfica e do que é preconizado pela tese da estabilidade nomotética.

Esta propõe que numa teoria consistente sobre a natureza, estritamente

fundamentada na observação, existem proposições chamadas leis da natureza que são válidas em todo o local; para qualquer tempo e em relação a qualquer observador no universo, cuja origem e valores lógicos não dependem de nenhum factor externo à teoria em que são formuladas. Noutros termos, e utilizando o que foi definido na secção anterior, a tese da estabilidade nomotética sustenta que existem proposições universais cujo valor lógico é verdadeiro e independente quer do local e do tempo, coordenáveis por qualquer observador; quer ainda de qualquer factor externo à teoria em que são formuladas.

Sendo esse o caso, a tese da estabilidade nomotética permite pois afirmar que existem sentenças numa teoria sobre a natureza  $\mathcal{T}$  que são proposições universais em relação a uma região holográfica inicial  $\mathcal{R}_H$ . É justamente desta asserção que nos iremos ocupar.

Para o efeito retomamos a definição de proposição universal e relacionamo-la com a lista de todos os valores de abcissa na coordenação holográfica de  $\mathcal{R}_H$ . Seja então a lista  $\mathcal{A}_H$  considerada. A cada valor de abcissa  $x_j[\mathcal{R}_H(q)]$  na coordenação holográfica corresponde um número infinito e denumerável de regiões holográficas  $R_H$ , indexadas pela letra  $q$  ( $q \in N$ ). Dessa maneira, a cada valor de abcissa encontra-se associado um conjunto de regiões holográficas  $R_H$ , que é distinto de todo e qualquer outro conjunto associado a qualquer outro valor da abcissa.

$$(2.16) \quad \mathcal{A}_H \quad \begin{array}{c|cccccc} & 2^{-1} & 2^{-2} & \dots & 2^{-k} & \dots & 2^{-n} \\ \hline x_1[\mathcal{R}_H(q)] & a_1(1) & a_1(2) & \dots & a_1(k) & \dots & a_1(n) \\ x_2[\mathcal{R}_H(q)] & a_2(1) & a_2(2) & \dots & a_2(k) & \dots & a_2(n) \\ \vdots & \vdots & \vdots & \ddots & \vdots & & \vdots \\ x_j[\mathcal{R}_H(q)] & a_j(1) & a_j(2) & \dots & a_j(k) & \dots & a_j(n) \\ \vdots & \vdots & \vdots & & \vdots & \ddots & \vdots \\ x_n[\mathcal{R}_H(q)] & a_n(1) & a_n(2) & \dots & a_n(k) & \dots & a_n(n) \end{array}$$

$$q, n \in N \quad ; \quad a_j(k) \in \{0,1\}$$

Uma proposição universal  $p$  é tal que para toda e qualquer região holográfica em  $\mathcal{R}_H$ , cuja coordenação holográfica é  $R_H[x, y, R_H(i)]$ ;  $i \in N_0$ , se verifica:

$$(2.21) \quad \mathcal{L} \left[ p, R_H \left[ x, y, R_H(i) \right] \right] = 1$$

Isto é; o respectivo valor da função nomotética é sempre igual a 1. Seja  $\{R_H; x_j[R_H(q)]\}$  o conjunto de todas as regiões holográficas em  $\mathcal{R}_H$ , cujo valor de abcissa é  $x_j[R_H(q)]$  e  $\{R_H(i)\}$  o conjunto de todas as resoluções holográficas. De acordo com a definição de proposição universal numa teoria  $\mathcal{T}$ , em relação a uma região holográfica inicial  $\mathcal{R}_H$ , podemos então escrever:

$$(2.22) \quad \forall R_H \in \{R_H; x_j[R_H(q)]\}, \forall R_H(i) \in \{R_H(i)\}, j, q \in N, i \in N_0:$$

$$\mathcal{L} \left[ p, R_H \left[ x_j, y, R_H(i) \right] \right] = 1$$

Uma vez que a proposição  $p$  é universal, a afirmação expressa na afirmação (2.22) não se altera com a variação do valor de abcissa  $x_j[R_H(q)]$ , indexado à letra  $j$ . Esse facto significa que podemos afirmar que para todo e qualquer valor de abcissa  $x_j$  na coordenação holográfica, uma proposição universal  $p$  cumpre (2.21), para toda a região holográfica coordenável, em qualquer resolução holográfica. Isto é, podemos afirmar que:

$$(2.23) \quad \forall x_j[R_H(q)] \in \mathcal{Q}_H, \forall R_H(i) \in \{R_H(i)\}, \forall R_H(q) \in \mathcal{R}_H, j, q \in N, i \in N_0:$$

$$\mathcal{L} \left[ p, R_H \left[ x_j, y, R_H(i) \right] \right] = 1$$

Em que  $R_H(q)$  representa qualquer região holográfica em  $\mathcal{R}_H$  cujo valor de abcissa é  $x_j$  e onde  $R_H(i)$  é a resolução holográfica utilizada pelo observador na coordenação da região holográfica  $R_H(q)$ . Desde que se preserve o sentido prescrito a (2.23), esta expressão pode ser condensada noutra:

$$(2.24) \quad \forall x_j \in \mathcal{Q}_H, \forall R_H(i) \in \{R_H(i)\}, \forall R_H(q) \in \mathcal{R}_H:$$

$$\mathcal{L} \left[ p, R_H \left[ x_j, R_H(i) \right] \right] = 1$$

Afirmação que se mantém válida para uma proposição universal  $p$ , onde omitimos o valor de ordenada  $y$  na coordenação holográfica. Associámos assim os valores de abcissa  $x_j$ , na coordenação holográfica, ao valor lógico de uma proposição universal  $p$ , em relação à região holográfica inicial  $\mathcal{R}_H$  onde aquela coordenação é feita. Com este resultado, encontramos-nos preparados para proceder à conclusão do argumento da variabilidade nomotética.

A forma geral de um valor de abcissa, correspondente à sua representação na lista  $\mathcal{A}_H$ , é dada por:

$$(2.25) \quad x_j = \left[ a_j(1).2^{-1} + \dots + a_j(k).2^{-k} + \dots + a_j(n).2^{-n} \right]. \mathcal{L}_H ;$$

$$n \in N, a_j(k) \in \{0,1\}$$

A lista  $\mathcal{A}_H$  contém todos os valores de abcissa na coordenação holográfica. Esse facto é conjuntamente assegurado pelo Princípio da Resolução Holográfica - que garante a existência de todos os valores na forma considerada - e pela identidade ontológica estabelecida entre o que é computável numa teoria sobre a natureza  $\mathcal{T}$  e o que é real no mundo. De facto, se no âmbito da teoria  $\mathcal{T}$ , chegássemos à conclusão de que existe um valor de abcissa, cuja medição fosse impossível no decurso de uma coordenação holográfica, isso significaria que a teoria teria previsto a existência de um aspecto no mundo não observável. Essa afirmação é inconsistente nos termos da teoria, em relação à realidade que descreve, porquanto todas as suas previsões que se referem a aspectos no mundo devem necessariamente ser observáveis. É por isso que considerámos semelhante teoria completa.

Tendo estas afirmações em mente, construa-se um valor de abcissa  $x_l$  da seguinte maneira. Para a primeira parcela  $a_l(1)$  na sua expressão fazemos  $a_l(1) = a_1(1)$ , isto é; igualamo-la à primeira parcela no primeiro valor  $x_1$  na lista. Para a segunda parcela  $a_l(2)$ , fazemos  $a_l(2) = a_2(2)$ , igualando-a à segunda parcela do segundo valor de abcissa na lista  $\mathcal{A}_H$ . Procedendo dessa maneira, até que se tenha, sucessivamente:

$$(2.26) \quad a_l(1) = a_1(1) ; a_l(2) = a_2(2) ; \dots ; a_l(k) = a_k(k) ; \dots$$

$$a_l(l) = a_l(l) ; \dots ; a_l(n) = a_n(n)$$

A situação é ilustrada em (2.27), onde se omitiram as entradas que não se encontram sobre a diagonal, se retirou a linha para o valor genérico  $x_k$  e onde se incluíram as entradas correspondentes a  $x_l$ . As setas indicam a correspondência entre as entradas de  $x_l$  e as restantes.

$$(2.27) \quad \mathcal{Q}_H \begin{array}{c|cccccc} & 2^{-1} & 2^{-2} & \dots & 2^{-k} & \dots & 2^{-n} \\ \hline x_1 & a_1(1) & & & & & \\ x_2 & \downarrow & a_2(2) & & & & \\ \vdots & & \downarrow & \ddots & & & \\ x_l & a_1(1) & a_2(2) & \dots & a_l(l) & \dots & a_l(n) \\ \vdots & & & & & \ddots & \uparrow \\ x_n & & & & & & a_n(n) \end{array}$$

$$n \in N ; a_j(k) \in \{0,1\}$$

O valor de abcissa  $x_l$  cumpre a forma genérica dada por (2.25), existe e é representável na lista  $\mathcal{Q}_H$ . A sua expressão geral é:

$$(2.28) \quad x_l = \left[ a_1(1).2^{-1} + a_2(2).2^{-2} + \dots + a_l(l).2^{-l} + \dots + a_n(n).2^{-n} \right]. \mathcal{Q}_H ;$$

$$n \in N, a_j(k) \in \{0,1\}$$

A lista  $\mathcal{Q}_H$  é um resultado na teoria sobre a natureza  $\mathcal{J}$ . Consequentemente, todo e qualquer observador pode formulá-la nessa teoria e obter o valor  $x_l$ .

Concebemos agora outro valor de abcissa  $x_m$ , calculável a partir do anterior, construído de acordo com um método que inclui informação sobre o valor lógico de uma

proposição universal  $p$ , na teoria  $\mathcal{T}$ , em relação à região holográfica inicial  $\mathcal{R}_H$ . Começamos por considerar um observador que utiliza a resolução holográfica  $R_H(i)$  para determinar o valor lógico de uma proposição  $p$ , formulada em  $\mathcal{T}$ , em todas as regiões holográficas  $R_H(q)$  em  $\mathcal{R}_H$ , coordenáveis por esse observador, cujo valor de abcissa é  $x_j$ . Com a resolução holográfica  $R_H(i)$  que utiliza, esse observador determinará o valor lógico de  $p$  num número finito  $Q(j,i) \equiv Q[x_j, R_H(i)]$  de regiões holográficas  $R_H(q)$ . Obterá, portanto, um número  $Q(j,i)$  de valores lógicos para a proposição  $p$ , em relação à região holográfica inicial  $\mathcal{R}_H$ .

Definimos agora a função  $\mathcal{F}$  que corresponde ao produto dos valores da função nomotética  $\mathcal{L}$ , que resultam da determinação anterior dos valores lógicos de  $p$  em  $\mathcal{R}_H$ .

Mais exactamente; definimos que:

$$(2.29) \quad \mathcal{F} : \{0,1\} \rightarrow \{0,1\} ; \mathcal{F} = \prod_{q=1}^{Q(j,i)} \mathcal{L} \left[ p, R_H \left[ x_j, R_H(i) \right] \right]_q$$

Onde indexámos o índice  $q$  na factorização ao resultado da função nomotética para cada região holográfica  $R_H(q)$ , cujo valor de abcissa é  $x_j$ , e em relação à qual o observador determina o valor lógico da proposição  $p$ . Se a proposição  $p$  for universal, verifica-se naturalmente que  $\mathcal{F} = 1$ , pois todos os valores de  $\mathcal{L}$  são iguais a 1.

Prosseguimos assumindo que o mesmo processo se verifica para, pelo menos, um observador em cada classe de observadores no universo, isto é; para cada resolução holográfica  $R_H(i)$ ;  $i \in N_0$ . Indexamos a função  $\mathcal{F}$  ao índice  $i$  associado à resolução holográfica e definimos uma nova função  $\mathcal{H}$ , tal que:

$$(2.30) \quad \mathcal{H} : \{0,1\} \rightarrow \{0,1\} ; \mathcal{H} = \prod_{i=0}^n \mathcal{F}(i)$$

Se a proposição  $p$  for universal, resultará  $\mathcal{H} = 1$ . A função  $\mathcal{H}$  depende de  $\mathcal{F}$ , que por sua vez depende do valor de abcissa  $x_j$  e da proposição  $p$ .

Podemos notar essa dependência escrevendo  $\mathcal{H} \equiv \mathcal{H}(x_j, p)$ .

Retornamos então ao valor de abcissa  $x_m$  cuja forma se pretende obter. Definimos que para cada uma das suas entradas, representadas na lista  $\mathcal{A}_H$ , se verifica o seguinte:

$$(2.31) \quad a_m(k) = [a_k(k) - \mathcal{H}(x_k, p)]^2 ; k \in N$$

Fazendo sucessivamente:

$$(2.32) \quad \begin{aligned} a_m(1) &= [a_1(1) - \mathcal{H}(x_1, p)]^2 \\ a_m(2) &= [a_2(2) - \mathcal{H}(x_2, p)]^2 \\ &\vdots \\ a_m(k) &= [a_k(k) - \mathcal{H}(x_k, p)]^2 \\ &\vdots \\ a_m(n) &= [a_n(n) - \mathcal{H}(x_n, p)]^2 ; n \in N \end{aligned}$$

Em virtude do facto de que  $p$  é uma proposição universal, obtemos:

$$(2.33) \quad a_m(k) = [a_k(k) - 1]^2 ; k \in N$$

Utilizando a representação (2.27), ilustramos o procedimento substituindo as entradas sobre a diagonal com a respectiva aplicação de (2.33), representando-se o valor de abcissa  $x_m$ .

$$(2.34) \quad \mathcal{Q}_H \quad \begin{array}{c|cccccc} & 2^{-1} & 2^{-2} & \dots & 2^{-k} & \dots & 2^{-n} \\ \hline x_1 & [a_1(1) - 1]^2 & & & & & \\ x_2 & \downarrow & [a_2(2) - 1]^2 & & & & \\ \vdots & & \downarrow & \ddots & & & \\ x_m & a_m(1) & a_m(2) & \dots & a_m(m) & \dots & a_m(n) \\ \vdots & & & & & & \uparrow \\ x_n & & & & & \ddots & [a_n(n) - 1]^2 \end{array}$$

$$n \in N ; a_j(k) \in \{0,1\}$$

O valor de abcissa  $x_m$  tem a forma geral:

$$(2.35) \quad x_m = [[a_1(1) - 1]^2 \cdot 2^{-1} + [a_2(2) - 1]^2 \cdot 2^{-2} + \dots + a_m(m) \cdot 2^{-l} + \dots + [a_n(n) - 1]^2 \cdot 2^{-n}] \cdot \mathcal{Q}_H ;$$

$$n \in N, a_j(k) \in \{0,1\}$$

E cumpre a forma geral (2.25) de qualquer valor de abcissa na coordenação holográfica, porquanto de (2.33) verificamos que:

- a) Sempre que for  $a_k(k) = 1$ , tem-se  $a_m(k) = 0$ ;
- b) Sempre que for  $a_k(k) = 0$ , tem-se  $a_m(k) = 1$ .

O valor de abcissa  $x_m$  é construído utilizando uma diagonalização sobre a lista  $\mathcal{Q}_H$ , contendo todos os valores de abcissa na coordenação holográfica, e respeitando o facto de que o valor da função  $\mathcal{H}$  para uma proposição universal  $p$  é sempre 1. A possibilidade de formulação da lista  $\mathcal{Q}_H$  fundamenta-se no facto de que o Princípio da Resolução Holográfica prevê a possibilidade de valores distintos de resolução holográfica no universo, relacionáveis entre si de uma certa maneira. Essa relação permite definir uma forma de coordenação holográfica e, de acordo com esta, proceder à medição de todos os valores na lista, de modo a cumprir-se a forma geral da sua representação dada por:

$$(2.25) \quad x_j = \left[ a_j(1).2^{-1} + \dots + a_j(k).2^{-k} + \dots + a_j(n).2^{-n} \right]. \mathcal{L}_H ;$$

$$n \in N, a_j(k) \in \{0,1\}$$

Isto é; em princípio, a lista pode ser obtida empiricamente em consequência do enunciado de uma proposição universal – o Princípio da Resolução Holográfica – que se refere ao conjunto de todas as regiões holográficas no universo, assegurando a possibilidade da sua observação. A lista tem, portanto, de poder ser obtida por um processo dedutivo na teoria, consistente com a possibilidade de observação dos dados que a compõem, cumprindo-se a identidade ontológica entre o que é computável e o que pode ocorrer no mundo.

O valor de abcissa  $x_m$ , por seu lado, resulta de um processo que é formulável no contexto da teoria dada, fundamentado na representação de aspectos observáveis na natureza, nomeadamente; na lista de todos os valores de abcissa e no facto de que  $p$  é uma proposição universal. A sua forma cumpre (2.25).

Mostraremos, em seguida, que o valor de abcissa  $x_m$  é uma «proposição não computável» na teoria  $\mathcal{T}$  e que, por esse motivo, não pode estar na lista  $\mathcal{L}_H$ , que é computável. Para o efeito, partimos da assumpção de que  $x_m$  se encontra numa dada posição da lista, tal como foi representado em (2.34).

Invocamos agora o procedimento da sua construção, tendo em mente que a entrada  $a_m(m)$  de  $x_m$  na lista se encontra sobre a diagonal da mesma. Este procedimento revela um problema. Quando substituímos o valor de  $a_m(m)$  na expressão (2.33), utilizada para construir  $x_m$ , obtemos:

$$(2.36) \quad a_m(m) = [a_m(m) - 1]^2$$

Esta expressão é falsa para qualquer valor de  $a_m(m)$  em  $\{0,1\}$ .

Com efeito:

a) Quando  $a_m(m) = 1$ , a expressão (2.36) transforma-se na proposição « $1 = 0$ »

b) Quando  $a_m(m) = 0$ , a expressão (2.36) transforma-se na proposição «  $0 = 1$  ».

O problema não ocorre, se contrariamente ao que foi suposto, assumirmos que o valor de abcissa  $x_m$  não se encontra na lista  $\mathcal{A}_H$ . Porém, isso significa que é possível deduzir um valor de abcissa  $x_m$  na teoria  $\mathcal{T}$ , cumprindo a forma geral (2.25), que se refere a um aspecto na natureza e que não se encontra na lista de todos os resultados que permitem descrever esse aspecto.

A definição de uma proposição não computável é a de que se for possível deduzir na teoria uma determinada proposição, embora à mesma possa ser atribuível um valor lógico, a proposição dir-se-á não computável na teoria dada.

Essa parece ser, justamente, a situação do valor de abcissa  $x_m$ .

A forma da proposição que lhe corresponde, de acordo com (2.25) e (2.35), é tal que esse valor existe, mas não temos meio de produzir a dedução da sua existência, uma vez que esta envolveria conhecer todas as entradas de  $x_m$  na lista e não temos como atribuir-lhe um desses valores. Visto que a proposição que corresponde ao valor de abcissa  $x_m$  se refere a um aspecto na natureza observável e, simultaneamente, é uma proposição não computável, obtemos uma contradição. Invocando a exigência de que numa teoria sobre a natureza nenhuma proposição indecível ou não computável pode corresponder a aspectos observáveis na natureza, concluímos que:

(2.36.1)  $x_m$  é um valor de abcissa observável na coordenação holográfica da região holográfica inicial  $\mathcal{R}_H$ , que não pode ser observado.

Em face desta contradição, a teoria sobre a natureza  $\mathcal{T}$  torna-se inconsistente.

Tendo como objectivo remover a inconsistência, revemos agora as suposições consideradas na construção do valor de abcissa  $x_m$ . Pretendemos determinar uma situação alternativa em que se cumpram os seguintes requisitos:

$\alpha$ ) A lista  $\mathcal{A}_H$  de todos os valores de abcissa, na coordenação holográfica de uma região holográfica inicial  $\mathcal{R}_H$ , pode ser obtida computacionalmente numa teoria sobre a natureza  $\mathcal{T}$  que é, por hipótese, completa;

$\beta)$  O valor de abcissa  $x_m$  pode ser construído através do processo descrito, recorrendo em particular à expressão (2.31);

$\chi)$  O valor de abcissa  $x_m$  corresponde a uma proposição computável e encontra-se na lista  $\mathcal{Q}_H$ , sendo mensurável na natureza, utilizando-se a resolução holográfica adequada.

O primeiro requisito só pode ser válido se o Princípio da Resolução Holográfica for verdadeiro. Este enunciado foi sugerido no decurso da resolução do paradoxo holográfico, razão pela qual é um postulado em  $\mathcal{J}$ . As duas condições seguintes só podem cumprir-se se a situação decorrente de (2.36) não se verificar. Essa possibilidade pode ser obtida adoptando o procedimento seguinte.

Retome-se a expressão geral, utilizada para construir  $x_m$ :

$$(2.31) \quad a_m(k) = [a_k(k) - \mathcal{H}(x_k, p)]^2 ; k \in N$$

Para  $k = m$ , obtemos:

$$(2.37) \quad a_m(m) = [a_m(m) - \mathcal{H}(x_m, p)]^2$$

Esta expressão só é verdadeira para qualquer valor de  $a_m(m)$  em  $\{0,1\}$ , se for:

$$(2.38) \quad \mathcal{H}(x_m, p) = 0$$

Atendendo ao modo como foram definidas as funções  $\mathcal{H}$  e  $\mathcal{F}$  e tendo em consideração as expressões (2.29) e (2.30), significa isso que deve existir, pelo menos, um observador com resolução holográfica  $R_H(i)$ ;  $i \in N_0$ , para o qual se verifica:

$$(2.39) \quad \mathcal{F}(i) = \prod_{q=1}^{Q(j,i)} \mathcal{L} \left[ p, R_H \left[ x_m, R_H(i) \right] \right]_q = 0$$

Por sua vez, (2.39) só pode ocorrer, se existir pelo menos uma região holográfica  $R_H$  em  $\mathcal{R}_H$ , cujo valor de abcissa é  $x_m$ , coordenada por esse observador e para a qual se tem:

$$(2.40) \quad \mathcal{L}\left[p, R_H\left[x_m, R_H(i)\right]\right] = 0$$

Deste modo, deixamos cair a imposição de que  $p$  tenha de ser necessariamente universal. Uma vez que a proposição  $p$  pode agora ser qualquer na teoria  $\mathcal{T}$ , deverá ser verdadeira a seguinte afirmação:

$$(2.41) \quad \forall p \in \mathcal{P}, \exists x_m [R_H(q)] \in \mathcal{A}_H, \exists R_H(i) \in \{R_H(i)\}, \exists R_H(q) \in \mathcal{R}_H, i \in N_0:$$

$$\mathcal{L}\left[p, R_H\left[x_m, y, R_H(i)\right]\right] = 0$$

O facto desta afirmação se referir explicitamente a  $x_m$ , não torna este valor de abcissa especial em relação aos restantes. Se a lista  $\mathcal{A}_H$  se assume computável, então é sempre possível alterar a ordem dos valores de abcissa  $x_j$  ali representados. Desse modo, o índice  $m$  refere-se a uma sequência de zeros e uns na lista, que pode representar qualquer valor de abcissa  $x_j$  que se pretenda, no conjunto de todos os valores de abcissa observáveis. No contexto do que se acaba de afirmar, a proposição (2.41) significa então que para alguma resolução holográfica  $R_H(i); i \in N_0$ , existe pelo menos uma região holográfica no conjunto de todas as regiões holográficas em  $\mathcal{R}_H$ , representadas pelas respectivas coordenações holográficas  $\{R_H[x, y, R_H(i)]\}; i \in N_0$ , para a qual o valor da função nomotética  $\mathcal{L}$ , relativa a  $p$ , é 0.

Uma vez que não colocámos qualquer imposição sobre o valor de área da região holográfica inicial  $\mathcal{R}_H$  e que  $p$  pode - por força da remoção da contradição apontada - ser qualquer proposição na teoria  $\mathcal{T}$ , escrevemos o seguinte:

Seja uma região holográfica  $\mathcal{R}_H$  qualquer, pertencente ao conjunto  $\mathcal{V}$  de todas as regiões holográficas no universo;

Seja uma teoria  $\mathcal{T}$  qualquer, pertencente à classe  $\mathcal{Y}$  de teorias que cumprem o seguinte conjunto  $\mathcal{W}$  de propriedades:

1) As quatro operações algébricas elementares – adição, subtração, produto e divisão – estão definidas sobre o conjunto dos números naturais com o zero  $N_0$  e sobre o conjunto dos fracionários positivos  $Q^+$ .

2) A potenciação de base 2 é uma operação definida em  $\mathcal{T}$ .

3) A teoria  $\mathcal{T}$  é consistente.

4) Sobre a relação entre aspectos observáveis no mundo e proposições representáveis na teoria são válidas as seguintes afirmações:

4.1) Todos os aspectos observáveis na natureza correspondem a proposições formuláveis na teoria.

4.2) Nenhuma proposição indecidível ou não computável na teoria pode corresponder a aspectos observáveis na natureza.

4.3) Todas as proposições que se refiram a aspectos observáveis na natureza deverão ser falsificáveis ou cumprir o critério lógico-formal para a validade de uma proposição sobre a natureza.

5) O critério lógico de realismo, o Princípio Holográfico e o Princípio da Resolução Holográfica são postulados na teoria.

Seja  $p$  uma proposição qualquer, pertencente ao conjunto  $\mathcal{P}$  de todas as proposições formuláveis em  $\mathcal{T}$ ;

Verifica-se o seguinte:

$$(2.42) \quad \forall \mathcal{R}_H \in \mathcal{V}, \forall \mathcal{T} \in \mathcal{Y}, \forall p \in \mathcal{P} : \\
\left[ \exists R_H(i) \in \{R_H(i)\}, \exists R_H \in \{R_H[x, y, R_H(i)]\}, i \in N_0 : \right. \\
\left. \mathcal{L}[p, R_H[x, y, R_H(i)]] = 0 \right]$$

Ou seja,

$$(2.43) \quad \forall \mathcal{R}_H \in \mathcal{V}, \forall \mathcal{T} \in \mathcal{Y}, \forall p \in \mathcal{P} : p \text{ é uma proposição contingente em } \mathcal{R}_H$$

Para que as condições  $\beta)$  e  $\chi)$  se cumpram é necessário agora que qualquer proposição  $p$  na teoria sobre a natureza  $\mathcal{T}$  seja contingente na região holográfica  $\mathcal{R}_H$ . Mais ainda, as mesmas condições asseguram a consistência da teoria impedem que nela se verifique uma forma de não computabilidade, relativa à obtenção do valor de abcissa  $x_m$ , tornando consistente o seu processo de construção e possível a sua mensurabilidade, através de uma observação.

Reconhecidas as propriedades que caracterizam a teoria  $\mathcal{T}$ , a afirmação (2.43) permite finalmente enunciar o argumento da variabilidade nomotética, referindo as consequências de uma proposição universal em  $\mathcal{T}$  para a consistência da teoria.

Seja uma teoria sobre a natureza  $\mathcal{T}$ , cumprindo o conjunto  $\mathcal{W}$  de propriedades.

Seja  $p$  uma proposição qualquer, formulada em  $\mathcal{T}$ .

É então válida a seguinte afirmação:

Argumento da variabilidade nomotética.

Para toda e qualquer região holográfica  $\mathcal{R}_H$ , para a qual é válida qualquer teoria  $\mathcal{T}$  sobre o mundo, cumprindo o conjunto de propriedades  $\mathcal{W}$ , e para qualquer proposição  $p$  formulável em  $\mathcal{T}$ , é verdadeira a seguinte proposição em  $\mathcal{T}$ :

Se  $p$  for uma proposição universal em  $\mathcal{R}_H$ , formulada em  $\mathcal{T}$ ,  $\mathcal{T}$  é inconsistente.

O enunciado do argumento da variabilidade nomotética parece assim sugerir que qualquer descrição consistente do mundo requer que toda a proposição no formalismo utilizado seja contingente. Noutros termos, que de alguma maneira ocorra necessariamente

variabilidade nomotética na natureza e, em particular, que as leis que pensávamos reger o mundo sejam variáveis. Esta interpretação do resultado revelar-se-á, no entanto, provisória perante análise posterior e cabal das suas consequências para a consistência da teoria em que é formulada. Uma vez que a proposição que corresponde ao argumento deduzido é uma proposição na teoria informal, sucede que a respectiva validade se aplica ao seu próprio enunciado, tornando-a assim uma proposição contingente. Desse facto resulta uma contradição que torna a teoria inconsistente. Uma vez que essa é considerada uma teoria completa sobre a natureza, o argumento da variabilidade nomotética torna a realidade ininteligível ou, pior, contraditória.

Sugere-se, então, que a solução para situação seja presumir a existência de uma propriedade no mundo, cuja expressão não tem cabimento nos limites explicativos de uma teoria informal sobre a realidade, mesmo que essa teoria possa ser completa.

Caracteriza-se essa situação dizendo que há certos eventos no mundo, cuja descrição é metalógica, sugerindo-se que a origem desses eventos, apresentando estrutura e regularidade, se deve à expressão de volição na natureza.

