

UNIVERSIDADE DE SÃO PAULO
FACULDADE DE FILOSOFIA, LETRAS E CIÊNCIAS HUMANAS
DEPARTAMENTO DE FILOSOFIA
PROGRAMA DE PÓS-GRADUAÇÃO EM FILOSOFIA

William Ananias Vallério Dias

O Problema do Reduccionismo no Pensamento de
Edward Fredkin

São Paulo
2017

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

Catálogo na Publicação
Serviço de Biblioteca e Documentação
Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo

D541p Dias, William
 O Problema do Reduccionismo no Pensamento de
Edward Fredkin / William Dias ; orientador Osvaldo
Pessoa Jr.. - São Paulo, 2017.
 126 f.

Dissertação (Mestrado)- Faculdade de Filosofia,
Letras e Ciências Humanas da Universidade de São
Paulo. Departamento de Filosofia. Área de
concentração: Filosofia.

1. FILOSOFIA DA CIÊNCIA. 2. REDUCCIONISMO. I.
Pessoa Jr., Osvaldo, orient. II. Título.

William Ananias Vallério Dias

O Problema do Reduccionismo no Pensamento de
Edward Fredkin

Dissertação apresentada ao
Programa de Pós-Graduação em
Filosofia do Departamento de
Filosofia da Faculdade de Filosofia,
Letras e Ciências Humanas da
Universidade de São Paulo, para
obtenção do título de Mestre em
Filosofia sob a orientação do Prof. Dr.
Osvaldo Frota Pessoa Jr.

São Paulo
2017

AGRADECIMENTOS

Não sei se o universo em que vivemos é ou não um programa de computador, mas se for acredito que exista um Grande Programador responsável pelo sistema ter alcançado o estado em que terminei este trabalho e agradeço a ele por isso. Mas também há muitos elementos “dentro da máquina” a quem devo minha gratidão.

Agradeço ao meu orientador, professor Osvaldo, pelo apoio à ideia do projeto, pela liberdade oferecida na escolha dos temas trabalhados, pelas sugestões valiosas para a redação do texto e pela paciência com todas as mudanças que fiz no projeto, desde a proposta de trabalho até o final.

Agradeço aos professores Ricardo Tassinari e Anderson Beraldo-de-Araújo pelas sugestões dadas no processo de qualificação que foram de grande ajuda para o término do trabalho.

Agradeço aos funcionários da secretaria do Departamento de Filosofia da Universidade de São Paulo pelo auxílio com minha total falta de aptidão com os processos burocráticos da pós-graduação.

Agradeço à minha família, especialmente minha mãe Neuza e meu padrasto Horácio, pelo constante apoio e carinho.

Agradeço a todos os colegas e amigos pelos bons momentos passados juntos que tornaram o desenvolvimento deste trabalho mais leve.

Agradeço a todos os professores que tive a oportunidade de conhecer durante todos esses anos de estudo na Universidade de São Paulo, tanto no Instituto de Física quanto no Departamento de Filosofia, pelo valioso conhecimento transmitido que contribuíram para minha formação acadêmica, assim como agradeço a todos os professores que passaram por minha vida e contribuíram para minha formação pessoal.

Agradeço a todos os alunos que tive a oportunidade de lecionar física, matemática ou filosofia por me inspirarem a melhorar cada vez mais.

Por fim, agradeço ao CNPq por conceder uma bolsa de dois anos que permitiu a conclusão deste trabalho com maior tranquilidade.

RESUMO

DIAS, William. O Problema do Reduccionismo no Pensamento de Edward Fredkin. 2017. 128f. Dissertação (Mestrado) – Faculdade de Filosofia, Letras e Ciências Humanas. Departamento de Filosofia, Universidade de São Paulo, São Paulo.

O estadunidense Edward Fredkin, um pioneiro na área de computação, é conhecido por defender a hipótese de que o mundo natural é fundamentalmente um sistema de computação digital se partirmos do princípio de que todas as grandezas físicas são discretas, de modo que cada unidade mínima de espaço e tempo pode assumir apenas uma quantidade finita de estados possíveis. Nesse cenário, as transições de estado do universo nas escalas mais elementares poderiam ser representadas por modelos de autômatos celulares, sistemas computacionais formados de unidades espaciais básicas (células) que modificam seus estados em dependência de uma regra de transição que toma o próprio estado da célula com relação às unidades vizinhas. Quando as mudanças de estados das células são consideradas em escalas maiores é possível notar um comportamento coletivo que parece seguir uma regra própria, não contemplada na programação básica atuando no nível das células. Fredkin acredita que o nível mais microscópico de nosso universo funcione como um autômato celular e, quando sua computação é tomada em maiores escalas, o padrão coletivo é identificado com os elementos que definimos em nossa física atual como elétrons, moléculas, pedras, pessoas e galáxias, ainda que todos esses elementos macroscópicos sejam apenas o resultado de uma computação alterando estados presentes em unidades mínimas de espaço. Diante disso, a intenção deste trabalho é mostrar que a conjectura de Fredkin pode ser interpretada como uma hipótese reducionista, uma vez que todo sistema explicado por nossas teorias físicas podem ser completamente definidos em termos de uma estrutura computacional.

Palavras-chave: Edward Fredkin, ontologia digital, autômatos celulares, reduccionismo, metafísica.

ABSTRACT

DIAS, William. The Problem of Reductionism in Edward Fredkin's Thought. Thesis (Master Degree). 2017. 128f. – Faculdade de Filosofia, Letras e Ciências Humanas. Departamento de Filosofia, Universidade de São Paulo, São Paulo.

Edward Fredkin, an American computer pioneer, is known for defending that the natural world is fundamentally a digital computing system, assuming that all physical quantities are discrete, in a way that each unit of space and time can only attain a finite number of possible states. In this scenario, the state transitions of the universe, taking place in the most elementary scales, could be represented by cellular automata models, computer systems formed by basic space units (cells) that modify their states in dependence on a transition rule that takes the state of the cell itself with respect to neighboring units. When cell state changes are considered on larger scales, it is possible to notice a collective behavior that seems to follow a rule of its own, not contemplated in basic programming at the cell level. Fredkin believes that the most microscopic level of our universe works as a cellular automaton and when its computation is taken at larger scales, the collective pattern is identified with the elements we define in our current physics as electrons, molecules, stones, people and galaxies, although all these macroscopic elements are only the result of a computation altering the states in minimum space units. The purpose of this work is to show that Fredkin's conjecture can be interpreted as a reductionist hypothesis, since every system explained by our physical theories can be completely defined in terms of a computational structure.

Keywords: Edward Fredkin, digital ontology, cellular automata, reductionism, metaphysics.

SUMÁRIO

Introdução.....	7
Capítulo 1: Universo Digital.....	11
1.1. Edward Fredkin.....	12
1.2. Natureza Finita.....	16
1.3. Informação e Digitalidade.....	21
1.4. Autômatos Celulares.....	25
1.4.1. Autômato Celular Unidimensional Elementar.....	27
1.4.2. O Jogo da Vida.....	31
1.5. Computação Universal.....	36
1.6. Computação Reversível.....	41
1.7. Modelos de Mecânica Digital.....	44
1.8. Implicações da Hipótese de Fredkin.....	49
1.8.1. Simetrias Contínuas.....	49
1.8.2. Sistema de Referência Absoluto.....	50
1.8.3. Mecânica Quântica e Determinismo.....	52
1.8.4. O "Outro".....	55
1.8.5. Reduccionismo.....	57
Capítulo 2: Reduccionismo.....	58
2.1. Conceito de Reduccionismo.....	58
2.2. Reduccionismo como Visão de Mundo.....	62
2.2.1. Reduccionismo Significativo.....	63
2.2.3. Reduccionismo Trivial.....	71
2.3. Sondas Epistemológicas.....	76
2.3.1. Ontologia Geral.....	77
2.3.2. Determinação e Separabilidade Espacial com Sondas Epistemológicas.....	79
2.4. Objeções ao Reduccionismo.....	82
2.4.1. Realização Múltipla.....	83
2.4.2. Tiquismo.....	85
2.4.3. Emergentismo.....	85
2.4.4. Holismo e Não-Separabilidade.....	88
2.5. Considerações Gerais sobre Reduccionismo.....	90
Capítulo 3: Reduccionismo em um Universo Digital.....	92
3.1. Ontologia Geral Digital.....	92
3.2. Autômatos Celulares e Níveis.....	94
3.3. Sondando um Mundo Digital.....	99
3.4. Copiando um Mundo Digital.....	103
3.5. Autômatos Celulares e Anti-Reduccionismo.....	105
3.5.1. Mundos Digitais Multiplamente Realizáveis.....	106
3.5.2. Mundos Digitais Não-Deterministas.....	108
3.5.3. Computação Emergente.....	108
3.5.4. Limite de Informação.....	112
3.5.5. O Problema da Consciência.....	117

Considerações Finais.....	118
Referências.....	121

Introdução

Esta dissertação pretende explorar algumas implicações filosóficas de uma tese defendida pelo especialista em computação Edward Fredkin, mais precisamente, a tese que declara nossa realidade física como equivalente a um sistema computacional digital. Resumidamente, de acordo com Fredkin, partindo do princípio de que a quantidade de informação contida em uma porção finita de espaço-tempo seja limitada, seria possível associar mudanças nos estados microscópicos da natureza com transições de estados correspondentes a processos computacionais não muito diferentes do que ocorre em um computador digital, como o *laptop* utilizado para escrever este texto.

A proposta de Fredkin é pouco ortodoxa, mas não é absurda. Basta notar que a realidade física é computável – nossa tecnologia atual permite gerar simulações de cenários naturais com um grau considerável de fidelidade, sugerindo assim a conjectura de que pelo menos alguns componentes de nosso universo atuam de acordo com uma dinâmica computacional. Essa ideia não é uma exclusividade de Fredkin, sendo estudada em diferentes perspectivas por vários estudiosos, principalmente físicos, matemáticos e cientistas da computação. Alguns nomes relevantes de pesquisadores nessa área incluem Konrad Zuse, Stephen Wolfram, Gregory Chaitin, Jürgen Schmidhuber, Gerard 't Hooft, Seth Lloyd e David Deutsch. Todos eles concordam com uma ontologia natural envolvendo algum aspecto computacional, embora discordem em que sentido a computação da natureza é realizada. Por exemplo, alguns autores favoráveis à natureza computacional (e.g. Deutsch, 1985; Lloyd, 2006) preferem não considerar a computação do universo como primitiva, assumindo que tais efeitos computacionais são resultantes de princípios quânticos mais fundamentais. Nesse caso, a computação é dependente da física e não o oposto. Entretanto, Fredkin discorda desse caminho. Para ele, a computação que regula a realidade natural é mais básica em relação ao que é descrito pela teoria quântica. Desse modo, peculiaridades quânticas seriam apenas aparentes, tendo em vista que estão apoiados em um algoritmo processado em escala mais elementar. Assim, no sistema de Fredkin, a base do mundo natural não é definida por efeitos fundamentalmente quânticos, mas por processos fundamentalmente computacionais.

Diante disso, a única opção para uma descrição realista do universo seria um modelo computacional baseado em variáveis discretas. Outros formalismos, como as equações diferenciais utilizadas nas teorias físicas atuais, seriam apenas boas aproximações para

fenômenos derivados de um comportamento digital mais primitivo. Por outro lado, Fredkin acredita que esse novo modelo computacional seria caracterizado por regras simples. Sendo assim, uma descrição física da realidade seria simplesmente uma descrição do comportamento geral de estados mais elementares que permitem definir um algoritmo único para todas as leis naturais. Partindo desse algoritmo, seria possível explicar qualquer situação física já prevista por nossas teorias.

A conjectura de Fredkin está apoiada principalmente na tendência da ciência em descobrir que grandezas antes pensadas como contínuas são coerentes com a experiência apenas em uma descrição discreta, como ocorreu no caso de grandezas como carga elétrica e energia. O autor acredita que, com o avanço das descobertas científicas, a discretude se aplicaria até mesmo ao espaço e ao tempo. Se o espaço e tempo são discretos, então cada volume de espaço-tempo só pode assumir uma quantidade finita de informação, de modo que um processo de computação digital poderia representar alterações nos estados de cada unidade mínima espacial em cada instante mínimo de tempo. Haveria, portanto, uma dinâmica computacional em baixa escala responsável por determinar os acontecimentos em escalas maiores.

Em suma, uma computação básica seria capaz de definir aspectos menos básicos. Segundo Fredkin, isso é possível se entendermos a computação do mundo natural como um autômato celular. Para fins introdutórios, em uma descrição bastante informal, um autômato celular seria um programa constituído de partes mínimas (células) em que cada uma dessas partes pode modificar seu estado em cada etapa da computação por conta de uma regra que leve em conta seu próprio estado e o estado de células próximas, suas vizinhas. Assim, uma única regra agindo em cada célula com base na vizinhança permite uma computação paralela poderosa: em cada passo do algoritmo, todas as células podem se modificar simultaneamente. Ocorre que, quando essas alterações celulares são tomadas em conjunto, o comportamento coletivo das células pode seguir um padrão tão organizado que poderia ser interpretado como uma regra nova, ainda que a computação só envolva as regras básicas mencionadas. Em autômatos celulares, simplicidade pode formar complexidade quando considerada em escalas maiores.

Um exemplo famoso de autômato celular é o Jogo da Vida (ou simplesmente *Life*), idealizado em 1970 pelo matemático John Conway. Embora o Jogo da Vida seja extremamente simples no nível das células (admite apenas dois estados possíveis para cada

célula: viva ou morta), os coletivos de células no estado viva parecem seguir um comportamento próprio que vai além da simples mudança entre estados de viva e morta. Considere, por exemplo, a figura abaixo:

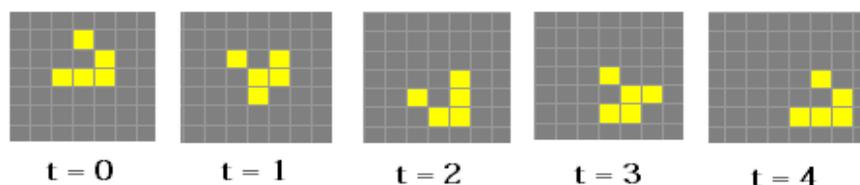


Figura 0: Exemplo de um padrão específico de células vivas no Jogo da Vida

Quando tomamos cinco células vivas na configuração mais à esquerda da figura, as regras do Jogo da Vida (basicamente, deixar a célula em seu estado se possuir exatamente duas vizinhas vivas, torná-la viva caso tenha exatamente três vizinhas vivas e torná-la morta caso tenha qualquer outro número de vizinhas) fazem com que esse agrupamento pareça formar um objeto que se movimenta rotacionalmente pelo espaço, ainda que nenhuma célula se movimente – elas apenas alteram seus estados conforme as regras básicas. Autômatos celulares, portanto, estimulam a intuição de que aspectos novos surjam de leis simples. Assim, quando a ideia de autômato celular é aplicada à natureza, as aparentes leis de escalas mais altas seriam justificadas como consequências imediatas das leis elementares ocorrendo em escalas menores.

Mas, diante desse universo definido por uma dinâmica computacional, seria possível defender que elementos básicos interagindo de acordo com um algoritmo são tudo que existe. Entidades como elétrons, átomos, moléculas, organismos e galáxias seriam apenas manifestações da computação que ocorre por conta do processamento dos dados mínimos contidos no espaço e no tempo. Ainda que comportamentos complexos venham a surgir em escalas maiores, eles não seriam nada além do processo computacional mais básico em funcionamento. Pelo que se pode inferir da obra de Fredkin, ele parece concordar com essas declarações e isso permite classificá-lo como um reducionista.

Uma posição reducionista consiste na defesa de alguma forma de redução, isto é, a alegação de que um dado elemento seja totalmente definido em termos de outros. O termo redução pode ser aplicado aos mais diversos contextos, sejam entidades (o mosquito completamente definido por suas moléculas), propriedades (propriedades mentais totalmente

determinadas por propriedades cerebrais) ou mesmo conceitos (o conceito de temperatura totalmente determinado pelo conceito da média da energia cinética molecular). Na perspectiva fredkiniana, o mundo macroscópico seria totalmente definido por uma dinâmica digital microscópica e isso configura um reducionismo nos termos acima definidos. Diante disso, o propósito desta dissertação será analisar a perspectiva de Fredkin em termos dessa concepção de reducionismo.

As intenções do presente texto incluem tanto um aspecto expositivo, visando divulgar a visão de Fredkin em território nacional, quanto um aspecto crítico, visando desenvolver uma análise mais aprofundada do significado filosófico dessa perspectiva.

Este trabalho está dividido em três capítulos:

O primeiro capítulo apresenta a hipótese de Fredkin por meio de uma descrição detalhada de sua conjectura e tomando várias de suas obras como referência. Iniciando com uma contextualização apresentando a pessoa de Edward Fredkin, partimos para suas premissas básicas, analisando suas propostas para a construção de um modelo baseado na conjectura e indicando algumas implicações derivadas dessa perspectiva.

O segundo capítulo muda o foco para um exame da noção de reducionismo, esclarecendo o que se pretende entender com uma posição reducionista e como ela aparece em contexto científico.

Por fim, o terceiro capítulo aplica as noções apresentadas no segundo capítulo na perspectiva defendida por Fredkin. Diante das noções de reducionismo adotadas, será discutido de que maneira o pensamento de Fredkin se enquadra nela.

Basicamente, o roteiro deste trabalho consiste em oferecer a descrição de uma posição específica no primeiro capítulo, apresentar o pano de fundo filosófico mais geral no segundo capítulo e discutir como a hipótese do autor dialoga com esse conceito mais geral no terceiro capítulo.

Capítulo 1: Universo Digital

O protagonista deste trabalho, o estadunidense Edward Fredkin (1934-), não é exatamente um filósofo de formação, mas sim um dos pioneiros no ramo da computação. Não obstante, ele se comprometeu com a defesa de uma tese tipicamente filosófica ao declarar que o universo em que vivemos poderia simplesmente ser um grande computador. Em outras palavras, Fredkin atribui propriedades computacionais à realidade física¹. Essa metafísica computacional também pode ser encontrada em outros autores, mas Fredkin é um dos estudiosos que mais levam esse ponto de vista a sério, já que, para ele, não se trata apenas de afirmar que o mundo natural pode ser interpretado como um computador, utilizando o aspecto computacional como uma metáfora, mas sim entender esse mundo como um computador em si mesmo. Para Fredkin, as leis da natureza consistem literalmente na execução de um programa.

Vale ressaltar que a relação entre física e computação não é estranha. É um fato indiscutível que a natureza é pelo menos parcialmente computável, caso contrário não seria possível simular computacionalmente aspectos do mundo natural. Não é difícil desenvolver simulações computacionais de certos sistemas físicos: um aparelho de *videogame*, por exemplo, pode exibir facilmente na tela de um monitor uma série de imagens representando um personagem caindo de acordo com a aceleração da gravidade ou colidindo com outros elementos do cenário seguindo as leis de conservação da quantidade de movimento. Uma dificuldade maior surge ao se afirmar que a natureza é totalmente computável. Embora seja possível desenvolver simulações muito próximas do universo real, isso não significa que todos os aspectos da realidade natural sejam contemplados na simulação². Na ausência de uma descrição última da natureza, nada parece garantir que essa natureza não venha a apresentar

1 Em Fredkin (2003), essa perspectiva é chamada de Filosofia Digital (*Digital Philosophy* ou, abreviadamente, DP).

2 É o que acontece nos projetos de realidade virtual. Via de regra, o objetivo de gerar uma realidade virtual é construir, por meio simulações computacionais, elementos que sejam indistinguíveis de suas versões originais. Isso não exige uma reprodução exata de todos os detalhes físicos do elemento em questão, mas apenas algum modo de causar uma impressão de realidade para aqueles que experimentarem a simulação. Obras de ficção científica que levam situações de realidades simuladas por computadores ao extremo (como o popular filme *Matrix*, de 1999) geralmente imaginam alguma tecnologia capaz de modificar os estados cerebrais dos personagens para gerar experiências sensoriais programadas. No caso da perspectiva de Fredkin temos menos a ideia de vivermos em um ambiente simulado e mais a ideia de que o próprio universo seja, em si mesmo, um sistema computacional. Talvez o universo físico seja, de fato, a simulação do que ocorre em um universo maior, mas o que está em jogo é o fato do nosso mundo natural apresentar um comportamento computacional. Ver Bostrom (2003) para um estudo mais aprofundado sobre a possibilidade de vivermos em uma realidade simulada.

indecidibilidades que não possam ser capturadas por um processo computacional. Fredkin, no entanto, declara que o mundo natural é totalmente computável. Para ele, tudo que ocorre na natureza é derivado de uma estrutura computacional básica responsável por definir todas as demais estruturas observadas em nosso domínio físico, das menores às maiores escalas. Isso é possível desde que certas suposições ontológicas sejam consideradas. No caso, o autor assume que toda grandeza fisicamente significativa na natureza seria discreta. A partir disso, seria possível pensar em uma natureza capaz de atuar como um computador.

Neste capítulo, essa proposta de universo computacional será apresentada em maiores detalhes. Antes disso, porém, seria conveniente oferecer uma contextualização por meio de uma apresentação da pessoa de Edward Fredkin, a mente por trás dessa ideia.

1.1. Edward Fredkin

De acordo com dados biográficos registrados em Wright (1988) e Hagar (2016), Fredkin nasceu em Los Angeles, Califórnia, em 1934. Foi o último dos três filhos de imigrantes russos que vieram para a América pouco tempo depois da Revolução de 1917. Não teve um início de vida muito fácil – sua mãe foi vítima de câncer quando ele tinha apenas onze anos e não teve muitos amigos durante a infância³. Apesar disso, por ser interessado em matemática e ciências físicas desde criança, Fredkin ingressou no Caltech (*California Institute of Technology*) em 1951 para se tornar físico. Diante das dificuldades por ter que trabalhar enquanto estudava, preferiu interromper seu curso logo no segundo ano de graduação para servir as Forças Aéreas Americanas, se formando como piloto em 1955. Iniciou seu contato com o mundo da computação em 1956, quando foi chamado para testar o SAGE (abreviação de *Semi Automatic Ground Environment*), um sistema com fins militares de defesa aérea, nos laboratórios Lincoln, no MIT (*Massachusetts Institute of Technology*). Como o SAGE ainda demoraria cerca de um ano para ser testado, aproveitou esse período para aprender tudo que era preciso saber sobre computadores até aquela época. Seu talento para lidar com essas novas máquinas era notável e não demorou muito para que se tornasse um especialista no assunto. Em 1958, Edward saiu das Forças Aéreas para trabalhar na BBN (*Bolt, Beranek and Newman*), uma empresa de consultoria em Boston, sob a supervisão do psicólogo e cientista da computação Joseph Carl Robnett Licklider (1915-1990), que não

³ De acordo com o que revelou no artigo de Wright (1988), os garotos preferiam jogar futebol com uma pessoa a menos no time do que escolher Fredkin para fazer parte da equipe.

demorou para reconhecer as habilidades do recém-chegado, um reconhecimento que se mostraria valioso alguns anos mais tarde. O trabalho de Fredkin envolvia, principalmente, auxiliar os membros da equipe em assuntos ligados à computação. Em 1959, ele tem contato com um PDP-1 (*Programmed Data Processor-1*), um computador relativamente poderoso na época⁴, convencendo a BBN a adquirir a primeira máquina lançada. Fredkin havia sugerido algumas alterações de *hardware* para o desenvolvedor do equipamento, Ben Gurley (1926-1963), que, para sua surpresa (cf. Hagar, 2016, p. 422), foram aceitas e implementadas no produto final.

Em 1962, Fredkin deixou a BBN para montar sua própria empresa de tecnologia, a *Information International Incorporated* ou simplesmente *Triple-I*. Um trabalho particularmente importante dessa empresa foi desenvolvido para a *Woods Holes*, um instituto oceanográfico que precisava de um sistema capaz de traduzir certos dados obtidos por equipamentos de pesquisas submarinas para um computador (cf. Wright, 1988, online). O sucesso desse sistema permitiu um contrato com os laboratórios Lincoln, popularizando a empresa em 1968 e tornando Fredkin um milionário com apenas trinta e quatro anos de idade (Hagar, 2016, p. 423).

Em 1960, o MIT havia iniciado o Projeto MAC (*Multi Access Computer* ou *Machine Aided Cognition*), um projeto de pesquisa em computação e inteligência artificial apoiado pela DARPA (*Defense Department's Advanced Research Projects Agency*), onde Licklider, o supervisor de Fredkin na época da BBN, estava trabalhando. Robert Fano (1917-2016), o primeiro diretor do projeto, chamou Fredkin para participar da equipe juntamente com outros nomes fortes no cenário da tecnologia computacional, como o matemático Marvin Minsky (1927-2016). Em 1968, Licklider assumiu a direção do projeto e, juntamente com Minsky, convenceu o chefe do Departamento de Engenharia Elétrica do MIT, Louis Smullin (1916-2009), a aceitar Fredkin como professor visitante por um ano, lecionando uma disciplina relacionada à resolução de problemas (cf. Hagar, 2016, p. 423). Ao fim daquele ano, em vista do sucesso de seu trabalho, Fredkin foi aceito oficialmente como professor, vindo a se tornar diretor do Projeto MAC entre 1971 e 1974. Foi uma contratação bastante atípica na história do MIT, já que Fredkin não possuía uma especialização acadêmica superior completa (Hagar, 2016, p. 423). Não obstante, o bom trabalho apresentado por Fredkin em seu ano como visitante mostrou que ele era perfeitamente capacitado para assumir esse cargo. Além disso,

⁴ O PDP-1 é famoso no meio *hacker* por ter sido o primeiro *hardware* a rodar o jogo de batalha de navas *Spacewar!*, desenvolvido pelo cientista norte-americano Steve Russell.

como a computação era uma área muito nova, havia uma grande falta de profissionais familiarizados com computadores e o conhecimento prático de Fredkin com essas máquinas era inegável.

Em 1974, Fredkin retorna para o Caltech, passando um ano sob a tutela do físico Richard Feynman (1918-1988), vencedor do prêmio Nobel de física de 1965, a quem Fredkin conheceu em 1961, em uma visita ao Caltech. Esse período permitiu tanto a Fredkin ensinar as novidades da computação para Feynman, quanto a Feynman ensinar os avanços da física para Fredkin. Feynman reconheceu Fredkin como um pensador original e promissor, capaz de trazer um novo olhar para a física (cf. Wright, 1988, online).

Edward retornaria ao MIT em 1975, inspirado com sua atualização de conhecimentos em física, para trabalhar em suas ideias a respeito dos fundamentos da realidade natural: a intuição de que o mundo se comportava como um grande computador. Muito desse pensamento foi fortalecido em seus testes com autômatos celulares no PDP-1 quando trabalhava na BBN, mas essa intuição já o acompanhava pelo menos desde seus primeiros contatos com máquinas computacionais na segunda metade dos anos 1950 (Hagar, 2016, pp. 422-423). Vale notar que seu contato com máquinas de computação durante o treinamento nos laboratórios Lincoln despertou um forte sentimento de que um computador era o retrato da perfeição:

“Meu estilo de vida inteiro estava apenas esperando o computador chegar”, diz ele. “O computador era em essência simplesmente a coisa perfeita.” Em alguns aspectos, um computador era preferível a qualquer outra aglomeração de matéria que ele já havia encontrado – mais sofisticado e flexível do que outras máquinas inorgânicas, e mais lógico do que as orgânicas. “Veja, quando eu escrevo um programa, se eu o escrevo corretamente, ele vai funcionar. Se estou lidando com uma pessoa, e eu digo a ela alguma coisa, e digo a ela corretamente, isso pode ou não funcionar.”(Wright, 1988, online).

A citação acima também revela um aspecto da personalidade de Edward: seu forte apreço pela racionalidade. Wright descreve Fredkin como um homem extremamente racional, raramente esboçando emoções, e que acredita haver uma resposta perfeitamente lógica para qualquer problema:

Fredkin exala racionalidade. Sua voz não é tão uniforme e precisa quanto a do Sr. Spock, mas está bem perto, e os paralelos não terminam por aí. Ele raramente expressa emoção - exceto, talvez, por algum sinal mínimo de irritação diante das circunstâncias mais difíceis. Ele nunca viu um problema que não tivesse uma solução perfeitamente lógica, e acredita fortemente que a inteligência pode ser mecanizada sem limites. (Wright, 1988, online)

Uma amostra dessa peculiaridade de Edward seria o fato de ele ter oferecido o “*Fredkin Prize*”, uma alta recompensa em dinheiro destinada a qualquer um que conseguisse elaborar um programa de computador capaz de derrotar um campeão de xadrez (cf. Wright, 1988, online). Tal proeza foi conseguida em 1997, quando o programa *Deep Blue*, criado por Feng-Hsiung Hsu, Thomas Anantharaman e Murray Campbell, venceu o campeão mundial russo Garry Kasparov⁵. Diante desse jeito de ser, não soaria tão surpreendente que Fredkin desenvolvesse uma visão de mundo pautada em um algoritmo bem definido.

As teses pouco ortodoxas de Fredkin nunca foram muito bem recebidas pela maioria dos estudiosos do MIT. Apesar disso, ele contou com alguns colaboradores, incluindo Marvin Minsky, uma das pessoas mais influentes em sua vida intelectual (Hagar, 2016, p. 424). Além de Minsky, Fredkin também contou com o apoio de alunos interessados na relação entre física e modelos computacionais como o canadense Norman Margolus (1955-) e o italiano Tomaso Toffoli (1943-). Juntamente com Margolus e Toffoli, Fredkin formou no MIT um grupo de estudos denominado Mecânica da Informação (*Information Mechanics*), onde realizavam pesquisas sobre esses assuntos e se mantinham atentos para novas ideias relacionadas ao tema que surgissem no círculo acadêmico (Longo & Vaccaro, 2013, p. 24). Por sinal, um grupo razoável de acadêmicos também tinha algum interesse nessa linha de pesquisa, com um destaque especial para o engenheiro alemão Konrad Zuse (1910-1995), responsável por desenvolver o Z-3, um dos primeiros projetos de computador programável, e autor do texto *Calculating Space*⁶ (Zuse, 1970/1969), um dos primeiros trabalhos apresentando explicitamente uma descrição do universo físico com base em sistemas computacionais digitais. O trabalho de Zuse serviu como uma das mais fortes influências na sofisticação das ideias de Fredkin para uma digitalização do mundo natural e uma representação das leis físicas por meio de um sistema de autômatos celulares. Em 1981, Fredkin liderou a organização da primeira Conferência de Física da Computação, um evento que contou com a participação de diversos pesquisadores envolvidos ou interessados, em algum grau, com as relações entre física, computação e informação, incluindo nomes como Konrad Zuse, John Wheeler, Richard Feynman, Freeman Dyson, Carl Adam Petri, Arthur Burks, Rolf Landauer, Hans Moravec, Norman Packard, Danny Hillis e Gregory Chaitin entre os convidados (cf.

5 <http://www-03.ibm.com/ibm/history/ibm100/us/en/icons/deepblue/>. Data de acesso: 03 de agosto de 2016.

6 *Rechnender Raum*, no original alemão.

Longo & Vaccaro, 2013, p. 22)⁷. Nessa conferência, muito foi discutido sobre a possibilidade de descrever a natureza por meio de computação. Segundo Zuse, foi a primeira vez que suas ideias a respeito de um universo digital foram levadas a sério (cf. Fredkin, 2003, p. 194; Longo & Vaccaro, 2013, p. 25).

Fredkin deixou o MIT em 1986, mas não abandonou completamente a vida acadêmica. Nos anos 90, ele lecionou por seis anos na Universidade de Boston e, mais recentemente, já nos anos 2000, se tornou professor na Universidade Carnegie Mellon (Hagar, 2016, p. 424). Embora suas ideias não tenham tanto apoio quanto gostaria, Fredkin não desiste de defendê-las e a maioria de seus textos sobre o assunto está disponível gratuitamente ao público por meio da Internet⁸. Na sequência, esses pensamentos serão explorados com maior profundidade.

1.2. Natureza Finita

O ponto de partida do pensamento de Fredkin envolve a adoção de uma caracterização específica da natureza que permita a descrição exata dos sistemas físicos por meio de um modelo computacional. Essa característica básica é denominada Natureza Finita⁹ e pode ser definida como a hipótese de que todas as grandezas trabalhadas pela física são discretas e finitas. Nas palavras do autor:

Natureza Finita é a hipótese de que fundamentalmente toda grandeza da física, incluindo espaço e tempo, se revelará discreta e finita; de que a quantidade de informação em qualquer volume pequeno de espaço-tempo será finito e igual a um pequeno número de possibilidades. (Fredkin, 1993, p. 116)

Uma estrutura pode ser considerada discreta se todos os pontos de seu domínio forem isolados, sendo que um ponto x será isolado em uma estrutura A se existir uma vizinhança de x na topologia de A que não tenha nenhum ponto pertencente a A com exceção de x (cf. Beraldo-de-Araújo & Baravalle, 2016, p. 3). Nesse sentido, números naturais constituem uma estrutura discreta, mas o mesmo não ocorre com números reais: se representarmos todos os números reais como pontos em uma reta nunca encontraremos pontos isolados¹⁰.

7 Os trabalhos apresentados no evento foram arquivados no volume 21 do periódico *International Journal of Theoretical Physics*, em 1982.

8 Ver o endereço <http://www.digitalphilosophy.org>. Data de acesso: 28 de agosto de 2017.

9 No original, *Finite Nature*.

10 Note que no caso dos números naturais temos uma estrutura infinita e discreta, ou seja, infinitude permite

Por muito tempo, principalmente com o advento do cálculo diferencial e integral desenvolvido no século XVII, os físicos consideraram a natureza como fundamentalmente não-discreta, utilizando equações diferenciais de variáveis contínuas para descrever os fenômenos naturais. Isso é bem intuitivo quando grandezas fundamentais como espaço e tempo são levadas em conta. De acordo com a física clássica, a trajetória de uma partícula no espaço pode ser representada por uma linha com tantos pontos quanto números reais, assim como o tempo decorrido no deslocamento da partícula pode ser entendido como um fluxo ininterrupto de uma quantidade não-enumerável de instantes não isolados. Consequentemente, grandezas derivadas da mudança de posição no espaço em um certo intervalo de tempo, como velocidade e aceleração, também serão funções contínuas de suas variáveis dependentes. Se espaço e tempo são mesmo grandezas físicas fundamentalmente contínuas, as teorias físicas definidas por equações diferenciais são, de fato, uma descrição correta da natureza. Entretanto, a continuidade do espaço e do tempo físico não é uma necessidade lógica: não podemos definir, *a priori*, que essas grandezas não demonstrem um caráter discreto.

O debate da continuidade do espaço e do tempo existe, pelo menos, desde os famosos paradoxos de Zenão de Eleia (c. 490-430 AC), crítico tanto de uma visão contínua quanto discreta de grandezas espaciais e temporais. Como um discípulo de Parmênides (c. 530-460 AC), Zenão era contrário à realidade da pluralidade, do movimento e da mudança. Assim sendo, apresentou diversas ilustrações em que alguma conclusão absurda surgia quando espaço e tempo eram pensados como constituídos de partes finitas ou infinitas (Aristóteles, 1991a, 231a18-241b32). O desenvolvimento da matemática e da filosofia dos séculos seguintes forneceu soluções a muitos dos problemas indicados por Zenão, mas ainda não há um posicionamento definitivo sobre a natureza da continuidade do espaço e do tempo. Por mais que equações diferenciais funcionem bem aos nossos propósitos científicos, não há garantia alguma de que descrevam realmente a natureza e não sejam apenas boas aproximações. Para o filósofo da ciência Amit Hagar, essa não é uma questão que podemos definir apenas pela lógica:

A questão em jogo, portanto, não é a consistência dessas estruturas, mas sim sua aplicabilidade ao mundo físico: que descrição matemática, a discreta ou a contínua, é uma

discretude, mas uma estrutura finita necessariamente será discreta, pois se essa estrutura admite apenas uma quantidade finita de pontos, então não poderemos encontrar infinitos pontos entre dois pontos quaisquer e, consequentemente, seria possível isolar qualquer ponto dessa ontologia em alguma vizinhança, mesmo que seja muito pequena. Desse modo, definir uma “Natureza Finita” já implica uma “Natureza Discreta”. Não obstante, Fredkin possui uma definição própria de discreto que será apresentada na próxima seção.

descrição adequada da realidade física? E se ambas são descrições adequadas, então qual delas é mais fundamental? (...) A notável aplicabilidade do contínuo em descrever o espaço macroscópico não pode ser tomada como uma evidência garantida por conta de sua aparente singularidade; como no caso das geometrias não-Euclidianas, a natureza fundamental do espaço físico pode muito bem ser diferente, e claramente não precisa ser decidida por meros argumentos *a priori*. (Hagar, 2014, p. 35)

Assim como a geometria do espaço não pode ser definida como euclidiana por argumentos *a priori*, o caráter discreto ou contínuo da natureza também não pode ser definido desse modo. Se não for logicamente necessário que a realidade natural seja fundamentalmente contínua, então é perfeitamente possível que representações discretas sejam a descrição adequada da natureza fundamental. Seguindo esse raciocínio, não há impedimentos puramente lógicos para trabalhar com base na Natureza Finita.

Mesmo que Fredkin esteja logicamente autorizado a assumir a Natureza Finita como uma possibilidade, isso não significa que ela seja verdadeira. Diante disso, é preciso mais alguma justificativa para escolher essa hipótese em detrimento da noção contínua de natureza. Como já mencionado, a história mostrou, ao longo dos séculos, que equações diferenciais são muito úteis para descrever variações nos sistemas físicos. No entanto, a tecnologia do século XX permitiu o desenvolvimento de sistemas computacionais que permitem representar os mesmos sistemas físicos com um grau considerável de aproximação por meio de simulações. Tanto simulações computacionais quanto equações diferenciais podem ser excelentes ferramentas para descrever a realidade física. Porém, se a natureza for discreta, equações diferenciais não retratam fielmente a realidade, enquanto se a natureza for contínua, nossas simulações de computador não são reproduções completas do que ocorre no mundo natural. Caso a natureza seja mesmo contínua, seria preciso uma quantidade infinita de passos computacionais para processar com exatidão todos os dados presentes em uma porção extremamente pequena de espaço e tempo. A existência de tantos dados em uma região espaço-temporal tão minúscula causou espanto no físico Richard Feynman, levando-o a questionar a legitimidade desse fato. Em seu livro *The Character of Physical Law* (Feynman, 1965), lemos o seguinte:

Sempre me incomodou que, de acordo como as leis são entendidas hoje, uma máquina computacional tome um número infinito de operações lógicas para revelar o que ocorre em uma região arbitrariamente pequena de espaço em um intervalo arbitrariamente pequeno de tempo. Como tudo isso pode se dar em um espaço tão pequeno? Por que deveria levar uma quantidade infinita de lógica para revelar o que acontece em uma porção tão pequena de

espaço-tempo? Então eu tenho frequentemente elaborado a hipótese de que fundamentalmente a física não requeira uma sentença matemática, de que no fim a maquinaria será revelada e as leis se tornarão simples, como um tabuleiro de damas com todas as suas aparentes complexidades. (Feynman, 1965, pp. 57-58).

Feynman suspeita que exista uma limitação para os dados contidos em um volume finito de espaço-tempo, motivado principalmente por um estranhamento diante da complexidade computacional que uma quantidade infinita de dados exigiria. Embora simplicidade seja desejável do ponto de vista científico, ela não garante uma certeza a respeito do comportamento real da natureza. Talvez o mundo seja mesmo complexo, no sentido de apresentar uma infinidade de dados em cada porção espaço-temporal, independentemente de nossa preferência por simplicidade. Por outro lado, também não há garantia alguma de que o mundo natural não seja fundamentalmente simples e os dados contidos em cada porção espaço-temporal sejam finitos.

O autor compreende que a hipótese de Natureza Finita não é indubitavelmente correta, mas a assume como verdadeira, como um ponto de partida necessário para defender seu ponto de vista. Em seu modo de pensar, não há como definir um sistema computacional para a natureza sem se comprometer com a Natureza Finita. Ao revisar suas ideias em um artigo de 2004, Fredkin diz o seguinte:

Não podemos progredir nessa discussão sem assumir que a física é sempre localmente finita. Não pretendemos argumentar nesse ponto, muito embora não saibamos a natureza de como ou porque isso deve ser verdadeiro; nós assumimos que é verdadeiro. (Fredkin, 2004, pp. 35-36).

Na opinião de Fredkin, uma motivação para apostar na Natureza Finita seria a tendência da atividade científica em adotar representações discretas. Muitos conceitos da física, antes pensados como contínuos, agora são fisicamente significativos apenas quando pensados como discretos. Fredkin apresenta alguns exemplos disso (Fredkin, 1993, p. 116-117). Um primeiro exemplo seria a própria constituição da matéria. Por muito tempo, pela dificuldade em observar o que ocorre em escalas muito pequenas, muitos cientistas duvidavam da coerência da teoria atômica. Com novas descobertas, como a explicação do movimento browniano, e novas tecnologias, como o microscópio de tunelamento, um modelo de matéria constituída de partículas mínimas é tido como a melhor explicação atualmente. Um

segundo caso seria a eletricidade: se antes a eletricidade era pensada como um fluido, hoje é entendido que cada partícula elementar contribui com uma quantidade mínima de carga elétrica. Um terceiro exemplo seria a energia. Max Planck (1858-1947) inaugurou o campo da física quântica ao explicar a incompatibilidade entre as previsões teóricas e os dados experimentais da radiação emitida por um corpo negro. O desenvolvimento da mecânica quântica levou à explicação de que essa energia não era irradiada pelo corpo de modo contínuo, mas em parcelas mínimas (os *quanta*). Além disso, os avanços nessa área mostraram que muitas outras propriedades surgiram no escopo da teoria assumindo apenas uma quantidade finita de valores bem definidos (como o momento angular intrínseco dos elétrons, o *spin*, que só pode assumir dois estados). Todos esses exemplos demonstram que não há como garantir que as observações empíricas futuras não exigirão alguma representação discreta para uma física coerente. Além disso, Fredkin salienta que, em termos experimentais, a física discreta parece mais acessível do que a física contínua. Esse argumento é claramente mais pragmático do que realista, mas vale a pena destacá-lo. A quantização de conceitos pode ser aceita quando um fenômeno só é explicado satisfatoriamente em um dado grau de observação por meio de grandezas discretas (como no caso da radiação do corpo negro). Uma continuidade, por sua vez, exigiria uma fragmentação infinita por definição, tornando muito difícil pensar em um experimento concreto capaz de comprovar definitivamente a existência de uma grandeza contínua. Nas palavras do autor:

Até agora não há um argumento convincente baseado em evidências experimentais que aponte para qualquer grandeza da física como definitivamente contínua. O que frequentemente dizemos é “Se é discreta, então a quantização deve ocorrer abaixo de algum nível.” É difícil propor até mesmo um teste que poderia verificar que alguma grandeza da física fosse de fato contínua. (Fredkin, 1993, p. 117)

Mais uma vez: não é possível provar definitivamente se a hipótese de Natureza Finita é mesmo verdadeira, mas Fredkin oferece razões para tomá-la como uma possibilidade ao se orientar pela própria história da ciência. Um raciocínio indutivo não pode garantir a certeza de uma conclusão, mas oferece plausibilidade para a aceitação de uma proposição. Se for minimamente razoável se guiar pela tendência da história, então também será minimamente razoável acreditar que a tendência de grandezas antes pensadas como contínuas se revelem como discretas no decorrer do trabalho científico. Para a sorte de Fredkin, muitas teorias físicas contemporâneas, como a gravidade quântica em laço (cf. Smolin, 2001), defendem que

uma saída para a unificação de todas as forças da natureza seja justamente um corpo teórico onde espaço e tempo sejam pensados como discretos. Há também esforços experimentais para encontrar evidências favoráveis à discretude do espaço e do tempo (e.g. Hogan, 2012). Diante disso, a hipótese de Natureza Finita deve ser encarada, no mínimo, como uma conjectura não imediatamente descartável.

O fato é que uma perspectiva computacional da realidade entra em voga se a Natureza Finita for tomada como verdade (Fredkin, 2003, p. 201). Se há apenas uma quantidade limitada de estados possíveis em uma porção de espaço-tempo, então é perfeitamente possível que exista um isomorfismo entre as transições desses estados e as transições de um algoritmo computacional. Aceitando a Natureza Finita, também damos abertura para os modelos previstos em uma conjectura digital, onde cada estado bem definido no sistema computacional poderia ser mapeado diretamente a um estado bem definido da natureza. Segundo Fredkin:

Uma consequência da natureza finita é que todo volume finito de espaço-tempo pode estar em apenas um de um número finito de estados (...). Dada a suposição da NF (Natureza Finita), é tautologicamente verdadeiro que um computador digital finito possa executar um modelo exato. Claro, há uma contingência quanto às condições iniciais e condições de contorno, juntamente com um programa correto para o computador. Além disso, eventualmente deve ser possível encontrar um modelo de processos informacionais que mapeie univocamente seus bits ou variáveis de n-estados de modo direto nas quantidades correspondentes da física que está modelando. (Fredkin, 2004, p. 36)

1.3. Informação e Digitalidade

Os termos “informação” e “digital” aparecem com frequência nos textos de Fredkin. Diante disso, antes de prosseguir com os pormenores do pensamento do autor, é importante apresentar com uma maior clareza como esses conceitos são definidos em seus trabalhos.

Em um dos artigos mais recentes do autor (Fredkin, 2013), ele declara que, em sua perspectiva, toda informação deve ser representada de modo discreto por uma configuração, uma coleção espacialmente distribuída de estados que podem ser interpretados de maneira significativa por algum processo. Assim, a noção de informação para Fredkin envolve conceitos como “discreto”, “estados”, “processo”, “configuração” e “significativo”.

Para começar, o autor define “discreto” (*discrete*) como a propriedade de um sistema finito fechado capaz de assumir N estados possíveis distintos que possam entrar em uma relação de um-para-um com N números inteiros distintos (cf. Fredkin, 2013, p. 370). Assim

sendo, uma informação será representada de modo discreto quando assumir uma quantidade finita de N estados possíveis diferentes mapeáveis com N inteiros diferentes.

Já um “estado” (*state*) pode ser definido como tudo aquilo que possui um “significado” (*meaning*) determinado por algum “processo” (*process*). Diversos estados organizados espacialmente constituem uma “configuração” (*configuration*). Assim, dependendo do processo em jogo, significados distintos podem ser atribuídos para certas configurações. Os conceitos de processo e significado não são muito óbvios, mas o entendimento se torna mais claro com exemplos dados pelo próprio autor (cf. Fredkin, 2013, p. 374): em um computador digital comum, diferentes regiões magnéticas de um disco rígido representam diferentes estados individuais e, em conjunto, formam configurações que possuem significados dependendo do processo. Considere um arquivo de vídeo .MPEG gravado no disco rígido de um computador qualquer. O modo com que as orientações magnéticas nas diferentes regiões do disco estão organizadas determinam uma configuração correspondente ao arquivo. O que essa configuração significa dependerá do processo executado. Caso o processo seja a abertura do arquivo por um programa de execução de vídeos, então o significado dessa configuração envolverá algo como a transmissão de dados para o sistema de vídeo do computador, de tal maneira que uma série de imagens apareça em sequência na tela do monitor. Caso o processo seja a gravação do vídeo em um *pendrive*, então o significado da configuração seria um conjunto de dados que deve ser copiado para a memória do dispositivo. Configurações, portanto, possuem significados atribuídos por processos. Uma configuração pode não ter significado algum para um dado processo, embora tenha muito significado para outros, como um arquivo de som que não é aberto por um programa de edição de imagens, mas pode ser reproduzido por um *software* multimídia que processa tanto dados de áudio quanto de vídeo.

Em resumo, informação é um conceito associado a uma configuração dotada de significado em termos de algum processo¹¹. No caso do universo digital, centro da filosofia do autor, a configuração que particularmente nos interessa diz respeito aos estados mínimos dotados de significado para algum processo no espaço e no tempo. Como já vimos, de acordo com a Natureza Finita, esses estados são finitos e contáveis, portanto, podem ser associados aos números inteiros, representando a informação de modo discreto. Mas representação de informação discreta também poderia ser entendida como informação digital. Resta ainda

¹¹ Nesse sentido, a concepção de informação de Fredkin parece depender mais de semântica, ao contrário de definições de informação independentes de significado, como é o caso de Shannon (1948).

esclarecer o conceito do termo “digital”.

Um exemplo imediato de dispositivo que trabalha com informação digital são os computadores digitais. Ao contrário de seus antepassados analógicos, computadores digitais conseguem diferenciar configurações específicas de um sistema e associá-las a dígitos (daí a palavra “digital”, entendendo “digital” como uma associação com dígitos). A palavra “dígito” pode ser entendida em dois sentidos diferentes: em um sentido mais abstrato, dígito seria um elemento de um conjunto contável de símbolos, enquanto em um sentido mais concreto, dígito pode ser definido como o estado macroscópico de um sistema capaz de ser diferenciado sem ambiguidade com relação a outros estados (cf. Piccinini & Scarantino, 2011, pp.7-8). Desse modo, um computador digital opera de acordo com o reconhecimento de configurações macroscópicas específicas em seu sistema. Assim, a passagem de corrente elétrica pelos terminais do circuito de um processador em uma dada configuração do sistema pode ser associada ao dígito “1” e a ausência de corrente elétrica, na mesma configuração, poderia ser associada ao dígito “0”. Se a configuração física desses estados pode ser associada a dígitos, então podemos concluir que a informação atribuída a esses estados é digital. Informação digital, nesse entendimento, pode ser definida como uma configuração associável a dígitos¹².

Repetimos mais uma vez que partimos da suposição de Natureza Finita, onde um volume de espaço-tempo apresenta uma quantidade finita de estados distintos possíveis. Se esses estados podem ser mapeados com um conjunto finito de números inteiros, então podem ser representados por dígitos. Se podem ser representados por dígitos, então os estados presentes em um volume de espaço-tempo de nosso universo podem exibir o que definimos como informação digital¹³.

Ocorre que, de acordo com a visão de Fredkin apresentada acima, informação implica em estados com significado para algum processo. Se afirmamos que os estados finitos de nosso espaço-tempo exibem informação digital, então também devemos supor que existe um processo capaz de apontar algum significado para esses estados. Ora, sabemos que os

12 O modo mais simples de representar dígitos é por meio de um sistema binário, ou seja, cada dígito sendo associado a um entre dois símbolos (via de regra, 0 ou 1). Dígitos representados dessa maneira são chamados de “dígitos binários” ou simplesmente *bits* (“*binary digits*”), termo cunhado por John Tuckey (1915-2000).

13 A discussão com respeito à descrição digital de nossa realidade pode tomar um caminho bem mais complicado. Para alguns autores, “digital” e “analógico” são apenas modos de representação que dependem do grau de abstração do observador em relação à realidade (Floridi, 2009). Assim, não faria sentido admitir uma postura realista quanto à natureza digital ou analógica do mundo. Outros autores (e.g. Berto & Tagliabue, 2014) discordam dessa posição. Neste trabalho não aprofundaremos essa discussão. Para os nossos propósitos, consideramos possível a ideia de Fredkin a respeito da digitalidade presente no mundo e, a partir disso, analisamos as implicações filosóficas de nosso interesse.

fenômenos naturais obedecem às leis da física. Duas partículas colidindo obedecem à conservação da quantidade de movimento, por exemplo. Essas leis indicam que deve existir algum significado atribuído aos sistemas físicos, de modo que, dadas certas condições, certos comportamentos são observados. Assim, as leis físicas devem cumprir o papel de processos que dão algum significado aos estados físicos. O que Fredkin propõe é que, uma vez que a informação dos estados físicos seja entendida como digital, então os processos definidos pelo que chamamos de leis físicas envolvem manipulação de dados digitais, o que não seria diferente do funcionamento de um computador digital.

A conclusão parece plausível teoricamente, mas, na prática, não é tão simples. O desafio de Fredkin é oferecer um modelo computacional que atenda seus pressupostos e também seja capaz de explicar as leis físicas conhecidas com o mesmo grau de sucesso empírico do presente. Até a data em que este texto foi escrito, nenhum modelo definitivo foi apresentado. O próprio Fredkin admite que sua conjectura se trata de uma área muito nova e ainda há muito a ser estudado antes de indicar uma solução capaz de atender todas as expectativas científicas (Fredkin, 2003, p. 189). Enquanto essa meta não é atingida, tudo que é afirmado nessa linha de pesquisa está no território das conjecturas teóricas. Entretanto, Fredkin tem seus palpites: ele possui alguma ideia de como poderia ser um programa capaz de representar a realidade em alguma escala fundamental com precisão. O estudo da elaboração dos detalhes referentes ao *software* que descreve nosso universo é chamado pelo autor de Mecânica Digital (abreviadamente MD)¹⁴ (cf. Fredkin, 1990; Fredkin, 2003).

Cabe ressaltar que o programa da natureza atua em uma escala ainda menor do que a física atual está acostumada. Em escalas maiores, nossas já conhecidas teorias físicas baseadas em equações diferenciais funcionam muito bem, mas a Mecânica Digital rege uma escala inferior à conhecida, onde as unidades de espaço e tempo assumem os estados associados à informação digital interpretada como significativa pelo programa. Fredkin não dá uma confirmação do nível dessa escala. É possível que ela seja algo extremamente pequeno, ao nível da escala de Planck (cerca de 10^{-35} m), mas o autor não vê motivos para que essa escala mínima não seja algo mais próximo de nossos recursos, na ordem de 10^{-17} m (Fredkin, 1993, p. 116). A condição é que o programa consiga descrever os mesmos fenômenos descritos pelas outras teorias, seja a escala muito pequena ou não muito distante das escalas que já temos acesso. Existindo uma escala de tamanho em que a informação contida ali possa ser

14 No original, *Digital Mechanics*.

identificada por um programa e permita derivar as mesmas previsões de nossa física já conhecida, não haveria motivos para rejeitar a perspectiva digital de natureza como mera especulação.

De acordo com Fredkin (2003, p. 190), a elaboração de um modelo de universo digital deve ser pensada levando em conta três princípios: simplicidade, economia e navalha de Occam¹⁵. Em outras palavras, a Mecânica Digital trabalha para encontrar o algoritmo mais simples e eficiente possível. Vale reforçar mais uma vez que simplicidade é um valor cognitivo e não precisa necessariamente ser um aspecto da natureza. No entanto, se existir um algoritmo simples, capaz de emular um programa complexo exibindo os mesmos resultados a partir dos mesmos dados de entrada, seria perfeitamente razoável do ponto de vista prático utilizar este programa em nosso trabalho científico. As suspeitas de Fredkin com relação ao sistema computacional responsável pela representação do universo apontam para um autômato celular universal reversível (ACUR¹⁶), em outras palavras, um autômato celular capaz de executar computação universal e reversível. Diante disso, seria conveniente discutir o que são autômatos celulares, computação universal e computação reversível para uma melhor compreensão das razões para esse tipo de sistema ser escolhido. Esses serão os temas das próximas seções.

1.4. Autômatos Celulares

De acordo com Poundstone (1985, pp.14-15), as primeiras ideias de autômatos celulares surgiram por volta da década de 1950, com o matemático húngaro naturalizado estadunidense John von Neumann (1903-1957) e o matemático polonês Stanislaw Ulam (1909-1984). Havia um interesse de von Neumann em representar matematicamente a reprodução de seres vivos para mostrar que aspectos reprodutivos não são uma exclusividade de sistemas biológicos. Para tanto, ele aplica ideias de Ulam que envolviam programas de computador capazes de gerar padrões geométricos complexos por repetições de instruções. O resultado foi um sistema dividido em unidades mínimas capazes de assumir diversos estados dependendo de regras bem definidas. Grosso modo, o sistema concebido por von Neumann seria como um enorme tabuleiro dividido em várias casas, onde cada uma delas poderia

15 “Navalha de Occam” é um termo que faz referência ao pensamento do filósofo medieval William Occam (1285 – 1347) de sempre dar preferência à explicação mais simples possível (cf. Fredkin, 2003, p. 190, Nota 2).

16 No texto original, *Reversible Universal Cellular Automata* (RUCA) (Fredkin, 2003).

assumir 29 estados diferentes, contando com o estado vazio. Dada uma configuração inicial, o estado global do tabuleiro mudaria de acordo com regras bem estabelecidas. Poundstone resume bem essa explicação:

O espaço celular de Von Neumann pode ser pensado como uma exótica e solitária forma de xadrez. O tabuleiro é ilimitado, e cada quadrado pode ser vazio ou conter um dos 28 tipos de peças do jogo. O jogador solitário arranja as peças do jogo em um padrão inicial. A partir daí, regras estritas determinam todas as configurações sucessivas do tabuleiro. (Poundstone, 1985, p. 16)

Com o passar dos anos, o interesse por autômatos celulares cresceu consideravelmente, tendo em vista a versatilidade desses sistemas na modelagem de padrões complexos a partir de regras de transição simples. Exemplos de situações simuladas por autômatos celulares incluem turbulência de fluidos, crescimento de plantas, teoria ecológica, evolução de DNA, propagação de doenças contagiosas, dinâmica social, incêndio de florestas e padrões de atividade elétrica em redes neurais (cf. Illachinski, 2001).

Há uma grande variedade de autômatos celulares possíveis, mas, em geral, eles possuem cinco características básicas (cf. Illachinski, 2001, p. 5), a saber:

Distribuição de células: a base do sistema envolve uma distribuição de objetos mínimos denominados células em uma, duas, três até n dimensões.

Homogeneidade: todas as células são equivalentes em seu formato. Não há restrições quanto à forma geométrica das células (em um autômato celular de duas dimensões cada célula poderia ter formato triangular, retangular ou hexagonal, por exemplo).

Estados discretos: cada célula pode assumir um estado entre uma quantidade finita de estados possíveis.

Interações locais: cada célula possui uma relação de vizinhança local com uma ou mais de outras células do sistema de modo que cada célula só poderá interagir com sua vizinhas.

Dinâmica discreta: em cada unidade discreta de tempo nesse sistema, cada célula atualiza seu estado atual de acordo com regras de transição que levam em conta os estados da célula e de suas vizinhas.

Por mais simples que essas características sejam, autômatos celulares podem exibir um comportamento altamente complexo a partir dessa simplicidade. Mais do que isso: autômatos celulares são modelos computacionais. Dada uma configuração inicial das células como

entrada (*input*) em um dado instante de tempo, a regra do autômato permite gerar uma nova configuração para as células como saída (*output*) no próximo instante de tempo. Isso pode simular diferentes algoritmos computacionais. Para um maior esclarecimento, estudaremos brevemente dois exemplos de autômatos celulares: o autômato celular de uma dimensão elementar e o mais famoso autômato celular em duas dimensões, o Jogo da Vida.

1.4.1. Autômato Celular Unidimensional Elementar

Consideremos um autômato celular definido com um conjunto enumerável $C = \{c_1, c_2, \dots, c_n\}$ de células em algum formato (por exemplo, quadrado) distribuídas em uma única dimensão (uma linha ou uma coluna) de modo que cada célula possa assumir um dos k estados associados ao conjunto $\Sigma = \{0, 1, \dots, k - 1\}$ em um instante de tempo t , onde $t \in \mathbb{N}$. Assim, o estado de uma célula c_i no instante t pode ser dado por uma função $\sigma_i(t) \in \Sigma$. Em uma dimensão, o estado de cada célula no instante $t + 1$ (ou seja, $\sigma_i(t + 1)$) dependerá do estado da própria célula e de suas vizinhas de acordo com uma função ϕ tal que:

$$\sigma_i(t + 1) = \phi(\sigma_{i-r}(t), \sigma_{i-r+1}(t), \dots, \sigma_{i+r-1}(t), \sigma_{i+r}(t))$$

Aqui, r especifica a quantidade de células em cada lado que contam como vizinhas para a célula i (em um autômato celular unidimensional em uma linha, cada célula teria até r vizinhas do lado direito e até r vizinhas do lado esquerdo¹⁷). A função ϕ dependerá de uma $(2r + 1)$ -upla indicando um padrão de estados com k^{2r+1} possibilidades de vizinhança. Para cada vizinhança possível, ϕ oferece um estado específico para a célula i . No caso mais simples, o chamado autômato celular elementar¹⁸, onde Σ conta com apenas dois estados (0 e 1, portanto $k = 2$) e cada célula só tenha o máximo de uma vizinha de cada lado ($r = 1$), teremos ϕ em função de uma tripla dada pelos estados da célula i , $i + 1$ e $i - 1$, gerando 8 padrões de vizinhança. Nessas condições, com 8 padrões de vizinhança que podem levar a dois estados possíveis cada um, haverá um total de 256 maneiras de definir $\sigma_i(t)$ com base em ϕ , ou seja, 256 regras possíveis ditando o comportamento do autômato. As regras para autômatos

17 Aqui imaginamos uma linha ou coluna infinita de células, mas é perfeitamente possível imaginar um autômato celular unidimensional finito. Nesse caso, poderíamos definir alguma regra especial para o comportamento das células nas bordas (talvez uma indicação para considerar a ausência de uma célula na vizinhança como uma célula em estado 0) ou desenvolver uma topologia específica em que a primeira célula seja vizinha da última célula, como se as células estivessem distribuídas em uma superfície cilíndrica ou toroidal.

18 No original, *Elementary Cellular Automata*.

celulares nessas condições foram enumeradas de 0 a 255 e estudadas detalhadamente por Stephen Wolfram (cf. Wolfram 1983, 1984, 2002), responsável por catalogar diferentes padrões apresentados por autômatos celulares unidimensionais quando os estados das células em cada instante de tempo são analisados em paralelo para cada regra aplicada. Tomaremos a Regra 90 como exemplo base.

De acordo com a Regra 90, a célula i sempre assume o estado 0 quando ambas as suas células vizinhas tomam o estado 0 ou quando ambas tomam o estado 1 e assume o estado 1 caso uma das vizinhas esteja no estado 0 e a outra vizinha esteja no estado 1. A figura abaixo¹⁹ representa todos os casos possíveis para esta regra: tomamos a célula central como nossa célula i de referência e, na linha logo abaixo, é indicado qual será seu estado com base nas duas vizinhas. Adaptando a figura às definições apresentadas acima, uma célula em branco é equivalente a uma célula no estado 0 e uma célula preenchida é equivalente a uma célula no estado 1.



Figura 1: Representação da Regra 90, mostrando como a célula central variará em relação às vizinhas.

Adotando esta regra, vejamos o que ocorre quando a mesma é aplicada, partindo de um instante inicial ($t = 0$), onde apenas uma célula da linha esteja preenchida. A Figura 2 mostra essa situação:



Figura 2: Estado do conjunto de células em $t = 0$ com uma única célula preenchida.

No próximo instante, a regra é aplicada. A única célula preenchida é esvaziada (pois tem duas células vizinhas em branco), mas suas duas vizinhas são preenchidas (pois ambas possuíam uma vizinha em branco e uma vizinha preenchida). Assim, teremos o que é mostrado na Figura 3:

¹⁹ As figuras exibidas nesta seção foram retiradas do aplicativo *NetLogo*, disponível em www.netlogoweb.org.
Data de acesso: 22 de agosto de 2016.

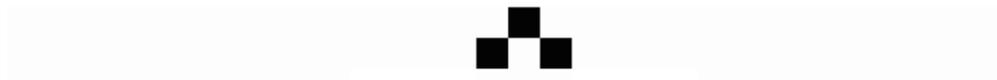


Figura 3: Evolução do autômato celular unidimensional de $t = 0$ (primeira linha) para $t = 1$ (segunda linha) de acordo com a regra 90.

Se a regra continuar a ser aplicada mais nove vezes, veremos o padrão formado pelo autômato no decorrer do tempo até $t = 10$. Isso é o que a Figura 4 apresenta:

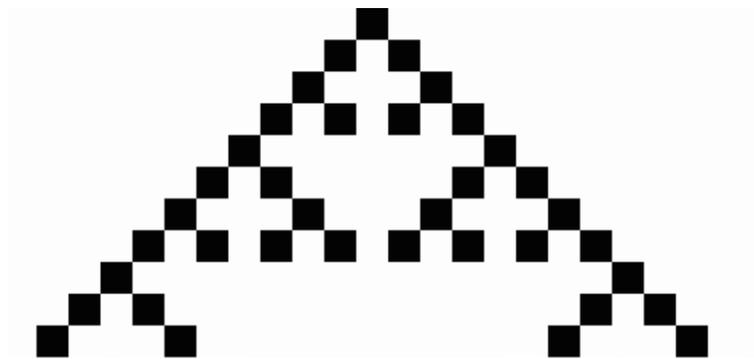


Figura 4: Evolução do autômato celular unidimensional de $t = 0$ (primeira linha) até $t = 10$ (décima primeira linha) de acordo com a regra 90.

Assim, prosseguindo na aplicação da mesma regra até $t = 128$, teremos um padrão como o apresentado na Figura 5 a seguir:

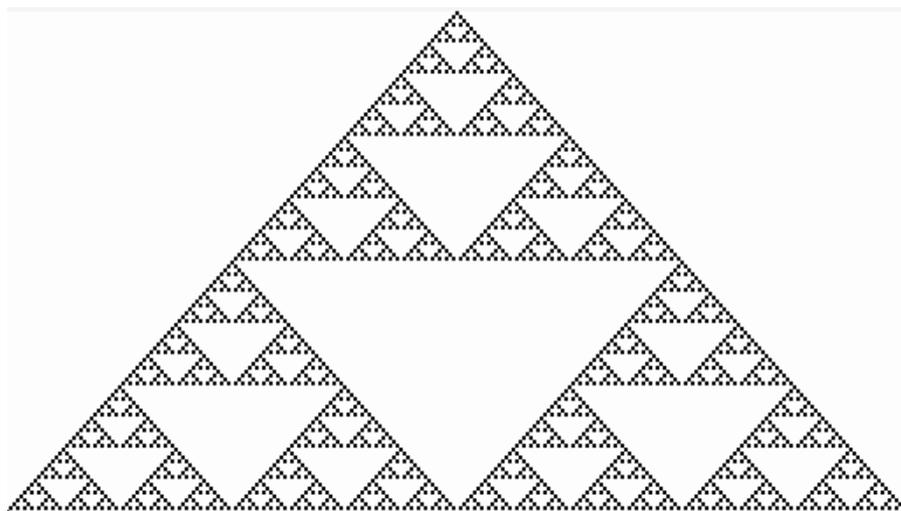


Figura 5: Evolução do autômato celular unidimensional de $t = 0$ a $t = 128$ de acordo com a regra 90.

Notamos na Figura 5 que, quando a Regra 90 é aplicada por uma quantidade apreciável de tempo, um padrão interessante começa a transparecer quando todas as linhas são analisadas simultaneamente. Mais precisamente, encontramos diversos agrupamentos triangulares de células preenchidas que obedecem um padrão próximo de uma dimensão fractal dada por um fator $\log_2 3$ (cf. Wolfram, 1983, p. 607). Esse comportamento não é uma exclusividade da Regra 90. Diversas outras regras permitem o surgimento de padrões qualitativamente complexos. Os diferentes tipos desses autômatos celulares unidimensionais básicos foram separados em Wolfram (1984) de acordo com a seguinte classificação:

Classe 1 ($c1^{20}$): Regras que levam a padrões homogêneos, onde todas as células terminam preenchidas ou em branco. Exemplo: Regra 254 (a célula só assume estado em branco se estiver em branco e suas vizinhas também estiverem em branco).



Figura 6: Padrão de evolução do autômato regido pela Regra 254 com uma célula inicial preenchida. A tendência é que todas as células terminem preenchidas.

Classe 2 ($c2$): Regras que geram padrões de evolução estáveis simples ou que variem de modo periódico. Exemplo: Regra 4 (a célula só assume estado preenchido se estiver preenchida e suas vizinhas estiverem em branco).



Figura 7: Padrão de evolução do autômato regido pela Regra 4 com uma célula inicial preenchida. A tendência é que apenas a célula inicial continue preenchida durante todos os instantes de tempo.

²⁰ Abreviação utilizada em Illachinski (2001).

Classe 3 (c3): Regras que geram padrões aparentemente caóticos e não-periódicos. Exemplo: Regra 30 (se a vizinha da esquerda estiver preenchida, a célula se tornará preenchida apenas se estiver em branco e sua vizinha da direita também estiver em branco, mas se a vizinha da esquerda estiver em branco, a célula ficará em branco apenas se estiver em branco e sua vizinha da direita também estiver em branco) e a já explorada Regra 90.



Figura 8: Padrão de evolução regido pela Regra 30 com uma célula inicial preenchida.

Classe 4 (c4): Regras que levam a padrões complexos, mas com algumas estruturas se propagando localmente. Exemplo: Regra 110 (a célula fica em branco apenas se estava em branco e sua vizinha da direita também estava em branco ou se estava preenchida e ambas as vizinhas estavam preenchidas, caso contrário, a célula fica preenchida).



Figura 9: Padrão de evolução regido pela Regra 110 com uma célula inicial preenchida.

Veremos que esse comportamento peculiar no decorrer do tempo permite uma aproximação com modelos computacionais. Antes disso, porém, consideremos nosso segundo exemplo de autômato celular: o Jogo da Vida (ou simplesmente *Life*).

1.4.2. O Jogo da Vida

O Jogo da Vida (*Game of Life* ou apenas *Life*) é um autômato celular desenvolvido pelo matemático John Conway e popularizado pelo escritor Martin Gardner em 1970 (Poundstone, 1985). Inspirado pela imprevisibilidade dos padrões encontrados nos trabalhos de Stanislaw Ulam, Conway queria desenvolver um jogo que seria “tão imprevisível quanto possível, ainda que com as regras mais simples possíveis” (cf. Poundstone, 1985, p. 25). Embora seja chamado de jogo, não há um vencedor ou perdedor. O único papel do jogador é

estabelecer alguma condição inicial para as células. Uma vez estabelecido o padrão inicial, as regras do programa se encarregarão de definir os padrões futuros. Do nosso ponto de vista, interagir com o Jogo da Vida estaria mais próximo de assistir a uma animação do que jogar um *videogame* (Poundstone, 1985, p. 24).

O Jogo da Vida obedece as já mencionadas características mais básicas de um autômato celular. No caso, podemos considerá-lo como um sistema contendo um conjunto $C = \{c_1, c_2, \dots, c_n\}$ de células quadradas organizadas em um plano bidimensional. Isso pode ser ilustrado como um grande tabuleiro quadriculado. Cada célula pode assumir um de dois estados de um conjunto $\Sigma = \{\text{Viva}, \text{Morta}\}$. Em uma representação gráfica, podemos ilustrar uma célula viva como um quadrado preenchido e uma célula morta como um quadrado em branco. Em cada instante de tempo discreto t (onde $t \in \mathbb{N}$) o autômato se encontra em uma geração (*generation*), onde cada célula se encontra em um dos dois estados possíveis. As regras definirão qual será o estado das células na próxima geração em $t + 1$. Essas regras são estabelecidas de acordo com a situação da vizinhança de cada célula. No presente caso, cada célula terá um máximo de oito vizinhas e as regras são as seguintes: (i) se a célula possuir exatamente duas vizinhas vivas, então ela manterá seu estado atual na próxima geração (se estiver viva, ela continuará viva e se estiver morta, continuará morta), (ii) se a célula possuir exatamente três vizinhas vivas, então ela assumirá o estado de viva na próxima geração e (iii) se a célula possuir qualquer número de vizinhas diferente de dois ou três (zero, um, quatro, cinco, seis, sete ou oito), então ela assumirá o estado de morta na próxima geração. E essas são as únicas regras para o Jogo da Vida. Embora essa estrutura não corresponda ao modelo da reprodução de qualquer ser vivo em especial, é possível imaginar essas regras em termos de uma analogia com colônias de organismos (Poundstone, 1985, p. 26). Com base nessa analogia, as criaturas do jogo morreriam por isolamento (quando possuem apenas uma ou nenhuma vizinha) ou por superpopulação (quando possuem mais do que três vizinhas), no entanto, podem sobreviver com exatamente duas vizinhas e se existir a relação entre três vizinhas permitiria o surgimento de uma nova criatura. Note mais uma vez que o papel do jogador neste jogo é praticamente inexistente. Sua função se resume apenas a decidir qual será a configuração inicial de células vivas em $t = 0$. Todas as gerações posteriores serão definidas pelas regras²¹. Em certo sentido, poderíamos dizer que as próprias células é que são as

21 Embora seja importante destacar a existência de certas configurações que nunca poderão ser alcançadas pela mera aplicação das regras do autômato celular, mas podem ser definidas como estado inicial do autômato. Configurações possíveis apenas como configurações iniciais estabelecidas pelo jogador são chamadas de Jardim do Éden (*Garden of Eden*) (Poundstone, 1985).

verdadeiras “jogadoras”. Não obstante, os resultados podem ser surpreendentes. Vejamos, a seguir, alguns exemplos de padrões formados pelo Jogo da Vida²².

Um primeiro padrão peculiar é o *blinker*²³. Iniciando em $t = 0$ com três células vivas em sequência na horizontal²⁴ teremos três células vivas em sequência na vertical em $t = 1$, voltando a ter três células vivas seguidas na horizontal em $t = 2$, como pode ser verificado na Figura 10²⁵. Esse padrão tende a se repetir indefinidamente.

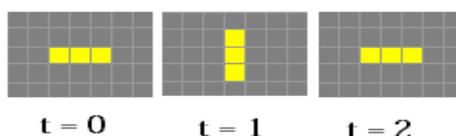


Figura 10: Padrão das três primeiras gerações de um *blinker* iniciado com três células vivas em sequência na horizontal.

Um padrão não tão interessante, mas relevante, é o *block*. Devido às regras do Jogo da Vida é possível construir um agrupamento de quatro células vivas que nunca morrerão, não importa quantas gerações se passem. A figura 11 mostra esse padrão:

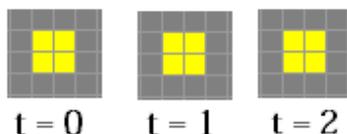


Figura 11: Padrão *block*. Pelas regras, as quatro células nunca poderão morrer.

Um outro padrão bastante popular é o *glider*²⁶. Formado por um agrupamento de cinco células vivas, um *glider* aparece no tabuleiro do jogo seguindo um padrão próprio de movimento diagonal no espaço. A Figura 12 mostra as cinco primeiras gerações de um *glider*.

22 Todos os exemplos apresentados foram retirados do que foi apresentado em Poundstone (1985).

23 Em inglês, seria algo como “piscador”.

24 O efeito será o mesmo iniciando com três células vivas em sequência na vertical.

25 As imagens do Jogo da Vida foram retiradas do software *Game of Life* de Edwin Martin (<http://www.bitstorm.org/gameoflife/>). Data de acesso: 25 de agosto de 2016.

26 Uma tradução literal seria “planador”.

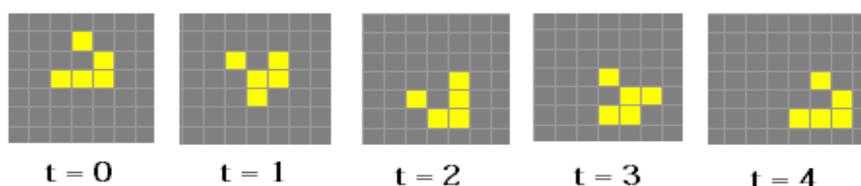


Figura 12: Padrão das cinco primeiras gerações iniciadas com um *glider*. As cinco células vivas formam um aglomerado que aparenta se mover de modo rotacionado através do tabuleiro.

Embora o movimento do *glider* exiba um padrão bem organizado, tal padrão não estava de modo algum contemplado quando as regras do Jogo da Vida foram criadas. Isso mostra que, assim como ocorre nos autômatos celulares unidimensionais, o comportamento das células do Jogo da Vida pode ser bastante complexo quando analisado em uma perspectiva maior no decorrer do tempo. Mesmo uma configuração inicial simples com poucas células vivas no tabuleiro pode levar a um padrão geral mais sofisticado, difícil de prever a partir da simplicidade do início. Considere a Figura 13 a seguir:

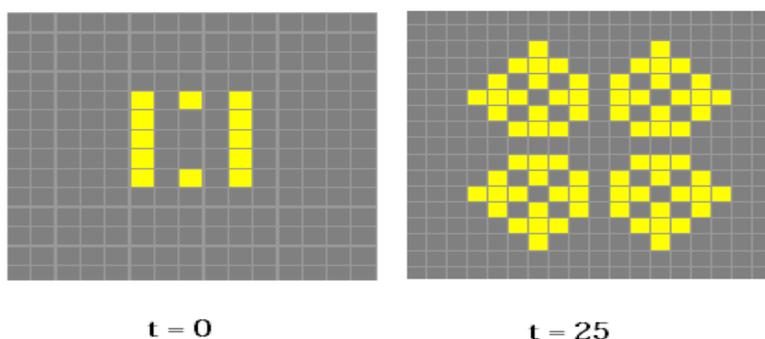


Figura 13: Um padrão com poucas células vivas em $t=0$ pode levar a um padrão mais complexo com diversas células espalhadas em $t=25$.

O tabuleiro é iniciado com um total de 12 células vivas agrupadas em um arranjo simples com duas sequências verticais de células vivas e duas células vivas sem vizinhas próximas entre essas duas sequências. Após 25 gerações, a configuração global muda para quatro agrupamentos mais complexos de células vivas com uma perceptível simetria espacial.

Conforme testamos diferentes arranjos iniciais no Jogo da Vida não é muito difícil encontrar agrupamentos de células que, juntos, formam uma estrutura estável e conseguem até mesmo formar padrões como *blinkers* e *gliders* automaticamente. Tome como exemplo o arranjo inicial apresentado na Figura 14 a seguir:

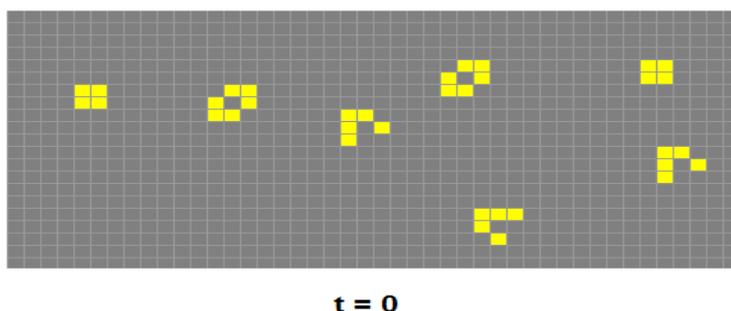


Figura 14: Estado inicial de um arranjo de células capaz de gerar *gliders* indefinidamente

Iniciando uma partida do Jogo da Vida com um arranjo inicial especificado na Figura 14 será perceptível o aparecimento periódico de padrões *glider* em consequência do comportamento de certos agrupamentos de células após um certo número de gerações. Em um certo sentido, esse padrão inicial de células vivas como um todo poderia ser interpretado como uma “fábrica de *gliders*”, como se seguisse uma regra própria para formar esses padrões²⁷. No entanto, a única computação atuando no sistema é a execução das regras do Jogo da Vida, regras estas que não incluem uma direção para formar *gliders* de maneira ordenada.

Talvez esses agrupamentos devam ser encarados como objetos especiais únicos constituídos de células vivas, assim como encaramos um ser vivo pluricelular como um organismo constituído por várias células trabalhando em conjunto. Não obstante, mesmo uma afirmação simples como o movimento de um *glider* é questionável: tudo que ocorre são células mudando seu estado de viva para morta. Nenhuma célula se desloca efetivamente em relação ao tabuleiro, embora o aglomerado de células vivas identificado como *glider* pareça se movimentar do ponto de vista de alguém visualizando o desenrolar de gerações do Jogo da Vida em um monitor. Não parece ser uma questão tão simples afirmar que exista o movimento de um *glider* no tabuleiro ou mesmo que ele seja, de fato, um objeto único e não apenas um aglomerado de células vivas, afinal, tudo que está acontecendo envolve apenas as regras básicas do Jogo da Vida. Esse detalhe será de grande importância para as discussões posteriores deste trabalho. Mas suponhamos que essas estruturas estáveis possam mesmo ser tratadas como “organismos”. Nesse caso, é tentador pensar que, dado um tabuleiro de extensão infinita, o Jogo da Vida poderia formar gigantescas estruturas estáveis assim como seres humanos são organismos complexos em um mundo regido por leis microscópicas

²⁷ O padrão da figura 14 também é conhecido como *Glider Gun* (revólver de gliders), descoberto por um grupo de pesquisa do MIT liderado por William Gosper (cf. Poundstone, 1985, p. 105).

básicas. Illachinski, ao abordar o Jogo da Vida em seu livro, faz um comentário com esse teor na seguinte citação:

Ao se observar a complexidade aparentemente ilimitada e a variedade dos padrões evolutivos do Jogo da Vida, se torna quase impossível deixar de imaginar, juntamente com Conway, que, se o jogo fosse executado em um tabuleiro infinito, deveriam surgir verdadeiras 'formas de vida', e talvez elas mesmas evoluíssem para 'organismos' mais complexos, possivelmente sencientes. (Illachinski, 2001, p. 132)²⁸

Os dois exemplos de autômatos celulares apresentados ilustram a possibilidade de gerar complexidade a partir da simplicidade. A regra de transição de estado para cada célula é extremamente simples, dependendo apenas das células vizinhas, mas os resultados globais estão longe da trivialidade. Se a hipótese de Natureza Finita de Fredkin estiver correta e existir uma quantidade limitada de informação no espaço, então é razoável supor que exista uma unidade mínima de espaço comparável à uma célula de autômato celular, onde as leis da física seriam, no final das contas, as transições de estado dessas células, gerando a complexidade observável na natureza. Obviamente, é necessário especificar qual autômato celular seria equivalente ao nosso universo. Precisamos de mais alguns detalhes. Como mencionado anteriormente, Fredkin ainda considera outras duas condições importantes na computação da realidade natural: universalidade e reversibilidade.

1.5. Computação Universal

O princípio de computação universal foi introduzido pelo matemático britânico Alan Turing (1912-1954). Em um famoso artigo da década de 30, Turing discutiu o conceito de número computável, um número real cuja representação como decimal pode ser calculada por meios finitos ou, em uma outra definição, um número real cuja representação decimal pode ser escrita por uma máquina (cf. Turing, 1936, p. 230). Turing pensou em um tipo específico de máquina abstrata que ficou conhecido na literatura especializada como “máquina de Turing”.

Uma máquina de Turing pode ser imaginada como um dispositivo capaz de assumir uma certa quantidade de configurações (ou estados). Essa máquina é alimentada por uma fita (*tape*) dividida em seções (ou quadrados), de tal modo que cada seção pode estar vazia ou

28 Vale comentar que Illachinski parece compreender o surgimento de aspectos como vida e consciência dentro de uma visão emergentista (cf. Illachinski, 2001, p. 629-632), enquanto Fredkin, como veremos nos próximos capítulos, parece trabalhar esses conceitos em uma linha mais próxima do reducionismo.

conter um símbolo pertencente a algum alfabeto finito. Em qualquer momento, há apenas uma seção sendo examinada (*scanned*) pela máquina. O dispositivo deve ser capaz de verificar o conteúdo da seção (se está vazia ou se contém algum símbolo) e assumir um determinado comportamento, como apagar o símbolo (caso a seção esteja preenchida) ou inserir um símbolo novo (caso a seção esteja vazia). A máquina também pode mudar a seção, movendo a seção examinada para a esquerda ou para a direita. Basicamente, a partir de uma configuração inicial, a rotina da máquina de Turing consiste em examinar o símbolo contido na seção, decidir o que fazer na seção (apagar, inserir um símbolo novo ou simplesmente manter o símbolo como está), mover a fita em uma seção (para a direita ou para a esquerda) e entrar em uma nova configuração. Como resultado, a sequência de símbolos escritos na fita corresponderão aos dígitos de um número real. Se cada estágio do movimento da máquina for completamente definido em termos de suas configurações, temos uma máquina automática (*automatic machine*). Caso uma máquina automática verifique apenas dois tipos (*kinds*) de símbolos, onde símbolos de primeiro tipo envolvem apenas dois casos (0 e 1, por exemplo) e os outros casos são chamados simplesmente símbolos de segundo tipo, temos uma máquina de computação (*computing machine*). Se a máquina é iniciada com uma fita em branco em uma determinada configuração, a subsequência impressa de símbolos de primeiro tipo constituirá a sequência computada pela máquina. O número real cuja representação decimal binária corresponda à sequência computada será o número computado pela máquina. Uma máquina pode ser circular (*circular*) ou não-circular (*circle-free*). A máquina é considerada circular se chegar a uma configuração na qual não consegue mais se mover ou se move apenas imprimindo símbolos de segundo tipo e não símbolos de primeiro tipo, caso contrário, será considerada não-circular. Assim, uma sequência será computável se puder ser computada por uma máquina não-circular (Turing, 1936, p. 232)²⁹. Como dispositivos abstratos, as máquinas de Turing não estão limitadas ao espaço ou ao tempo. Em princípio, podemos imaginar uma máquina de Turing trabalhando com uma fita de extensão infinita por quanto tempo se desejar.

Alguns exemplos simples de máquinas de Turing podem ser dados. Um primeiro exemplo (chamarei de máquina A) consiste em imprimir apenas 1s e uma outra máquina (chamarei de máquina B) imprime apenas 0s. Ambas possuem apenas uma configuração. No caso de A, se a máquina examinar uma seção com um símbolo diferente de 1, deve apagar o símbolo, inserir 1 e mover a fita, caso contrário, deve deixar como está e mover a fita. No

²⁹ Diversos textos oferecem uma definição mais formal de máquina de Turing (e.g. Sipser, 2013, pp. 167-170). Para os propósitos deste trabalho, uma definição informal será suficiente.

caso de B, se a máquina examinar uma seção com um símbolo diferente de 0, deve apagar o símbolo, inserir 0 e mover a fita, caso contrário, deve deixar como está e mover a fita. Assim, A imprimirá um número real com uma representação binária correspondente a uma sequência infinita de 1s e B imprimirá um número real com uma representação binária correspondente a uma sequência infinita de 0s.

Não é difícil notar que há uma quantidade infinita de máquinas de Turing possíveis, mas podemos imaginar uma única máquina capaz de imprimir qualquer sequência computável (Turing, 1936, p. 242). Entra aqui o princípio de computação universal. Podemos, então, definir uma máquina de computação universal U. Nessa máquina, a configuração inicial identifica na fita qual máquina ela deve imitar. A partir disso, U pode se comportar exatamente como a máquina identificada. Considere as duas máquinas definidas no parágrafo anterior. Podemos preparar a configuração inicial de U de tal modo que, ao ser iniciada, U deve verificar se a primeira seção da fita em análise possui um símbolo de primeiro tipo ou de segundo tipo, por exemplo. Caso seja um símbolo de primeiro tipo, deve se comportar como A, imprimindo apenas 1s. Caso seja um símbolo de segundo tipo, deve se comportar como B, imprimindo apenas 0s. Isso pode ser expandido para quantas máquinas se desejar. Assim, poderíamos ter uma única máquina capaz de atuar como qualquer outra, desde que se tenha disponibilidade ilimitada de espaço e tempo.

Ainda na década de 30, o matemático Alonzo Church (1903-1955) utilizou um formalismo chamado cálculo- λ que alcançou resultados equivalentes à teoria da computação de Turing e com isso a literatura em lógica matemática passou a utilizar o termo “Tese de Church-Turing” (cf. Kleene, 1967, p. 232). Essa tese declara que uma função é computável se e somente se pode ter seus valores calculados por uma máquina de Turing ou, equivalentemente, por funções do cálculo- λ ³⁰. Desse modo, qualquer sistema que exiba um comportamento equivalente a uma máquina de Turing também seria capaz de definir valores para uma função computável. Esse sistema sequer precisa ser uma máquina - um funcionário lendo, apagando e escrevendo símbolos em uma folha de papel de acordo com instruções pré-definidas poderia cumprir o mesmo papel de uma máquina de Turing (supondo que seja possível encontrar uma pessoa com papel e tempo o suficiente para executar qualquer função compatível com as máquinas).

Note que nossos computadores eletrônicos atuais seguem um comportamento

30 A tese de Church-Turing não possui uma prova, mas até o presente momento nunca foi refutada.

equivalente às máquinas de Turing. Ainda que não tenham fitas, nossos *laptops*, *desktops* e *smartphones* utilizam outros recursos de memória, como regiões magnéticas em um disco rígido ou cartão de memória que podem ser modificados de modo equivalente às operações de apagar e substituir símbolos nas seções da fita de uma máquina de Turing. Um programa rodando em um *laptop* nada mais é do que um sistema eletrônico realizando operações correspondentes às instruções identificadas nos registros armazenados na memória do disco. A memória desses computadores também pode ser ampliada: podemos expandir a memória de um *laptop* adicionando mais pentes de memória ao *hardware* do sistema, por exemplo. Desse modo, o princípio de computação universal também se torna válido. Sem limites de memória e tempo, um computador eletrônico poderia simular qualquer outro, seguindo a mesma ideia de computação universal já mencionada. Um exemplo prático da imitação de vários sistemas computacionais em um único computador seria a emulação de *videogames* – um computador com memória suficiente pode realizar as mesmas operações de um console de *videogame*. Assim, jogos que só funcionam em um console A e jogos que só funcionam em um console B podem todos funcionar em um único computador que rode programas emuladores dos dois consoles.

O fato de nossa realidade física permitir a construção de computadores universais macroscópicos é, para Fredkin, uma evidência de que o comportamento da realidade microscópica também obedece o princípio de computação universal. Nas palavras do autor:

Se a física microscópica (assumindo Natureza Finita) não fosse universal, então seria tautologicamente verdadeiro que a construção macroscópica de um computador comum não seria possível; mas a natureza nos permite construir computadores! (Fredkin, 1990, p. 257)

O argumento de Fredkin consiste em admitir que se todo computador macroscópico funciona por conta do que ocorre em escala microscópica e computadores macroscópicos realizam computação universal, então a dinâmica microscópica realiza computação universal. Uma vez que, de acordo com a ideia de Fredkin, a dinâmica microscópica limite de nosso mundo físico pode ser comparada a um autômato celular, então o autômato celular subjacente à realidade natural realizará computação universal, já que permitiria a formação de sistemas equivalentes a máquinas de Turing em uma escala mais ampla. Encontrar um autômato celular capaz de reproduzir efeitos em alta escala equivalentes ao funcionamento de máquinas de Turing universais é uma das tarefas da Mecânica Digital³¹.

31 Note que a dinâmica de autômatos celulares no nível das células é diferente das máquinas de Turing. Não é

Wolfram, por exemplo, sugeriu que autômatos celulares unidimensionais de classe 4 podem executar computações universalmente (Wolfram, 1984, p. 2). Esses autômatos exibem estruturas estáveis que se propagam de modo periódico e poderiam ser interpretadas como detentores de pacotes que preservam informação e se deslocam no decorrer do tempo. Diante disso, é possível defender a hipótese de que exista um autômato celular de classe 4 que apresente um comportamento capaz de imitar uma máquina de Turing universal. Nas palavras de Illachinski a respeito da universalidade computacional de autômatos celulares em uma dimensão:

Não obstante, muitos dos ingredientes chave necessários para provar universalidade, tais como conjuntos de estruturas de propagação a partir de análogos de componentes convencionais de 'hardware' (i.e. fios, portões e memória) podem ser explicitamente construídos, aparecem, pelo menos em princípio, como suportados por certas regras unidimensionais. O componente mais básico requerido é um mecanismo para transportar pacotes de informação de uma parte do sistema a outra, i.e., padrões de propagação persistentes de partículas, cuja presença é usualmente um indicativo de comportamento classe c4. (Illachinski, 2001, p. 89)

Em uma obra que escreveu em parceria com os matemáticos Elwyn Berkelamp e Richard Guy (Berkelamp et. al., 2004/1982), John Conway mostrou que o Jogo da Vida equivale a uma máquina de Turing universal, pois pode ter uma configuração inicial definida para realizar qualquer operação computável em termos de *gliders* e outros elementos de maior escala definidos pelo comportamento do autômato (cf. Berkelamp et. al., 2004/1982, pp. 940-957)³². Isso mostra que existem autômatos celulares capazes de exercer computação universal e, seguindo a ideia de Fredkin, nosso mundo seria definido por um autômato celular desse tipo.

Se o programa fundamental da natureza for um autômato celular capaz de realizar computação universal, uma boa notícia é que esse autômato celular pode ser extremamente simples, lembrando que, nessa hipótese digital, simplicidade é tomada como valor. De fato, Edwin Banks, um aluno de Fredkin, mostrou em sua tese de doutorado que um autômato celular de duas dimensões, onde cada célula pode assumir apenas dois estados e se comunicar

difícil construir máquinas de Turing capazes de simular autômatos celulares (afinal, se nossos computadores pessoais são equivalentes a máquinas de Turing e simulam autômatos celulares, então máquinas de Turing podem simular esses autômatos), mas não podemos dizer que todo autômato celular seja capaz de simular uma máquina de Turing. A proposta de Fredkin, portanto, depende de autômatos celulares que simulem máquinas de Turing ou nosso mundo macroscópico (que inclui máquinas de Turing!) não seria possível.

32 Uma demonstração detalhada da computação universal no Jogo da Vida pode ser encontrada em Rendell (2016).

com apenas quatro vizinhas, pode simular qualquer computação realizada por qualquer outro autômato celular (Banks, 1971). Assim, ainda que o autômato celular responsável pela dinâmica de nosso universo seja relativamente complexo, talvez com muitos estados (ainda que finitos) por célula, de acordo com a tese de Banks, poderíamos representá-lo por meio de um formalismo equivalente, porém mais simples, com apenas dois estados por célula.

1.6. Computação Reversível

Muitas computações são encaradas como irreversíveis. Isso ocorre quando os dados de saída (*output*) apresentam menos informação do que os dados de entrada (*input*). Por exemplo, a partir de uma entrada “ $2 + 2$ ” para um programa que calcule uma soma de dois números naturais há apenas uma única saída: “4”. Porém, partindo apenas de “4”, não é possível computar se tal número é um resultado de uma operação “ $2 + 2$ ”, “ $3 + 1$ ” ou “ $0 + 4$ ”. Entretanto, a irreversibilidade da computação não é uma necessidade lógica. Considerando algum dispositivo de memória extra onde todos os passos da computação sejam guardados, um computador poderia utilizar esses dados armazenados para retornar ao estado inicial a partir do dado de saída, permitindo assim um modo reversível de computação (Bennett, 1973)³³. Desse modo, a computação pode ser revertida, pois a informação é conservada: o sistema consegue processar como retornar para a entrada exata a partir da saída.

Estudos em computação reversível aplicada foram parte dos esforços de Fredkin e seus companheiros nos anos 1970 (Fredkin, 1990, pp. 256-257), discordando de pesquisas da época que consideravam uma computação reversível fisicamente impossível (e.g. Landauer, 1961). Juntamente com Tomaso Toffoli, Fredkin desenvolveu uma computação baseada em um princípio chamado de lógica conservativa³⁴. Para os autores, lógica conservativa pode ser entendida como um modelo computacional em que os processos podem ser revertidos justamente pela possibilidade de conservar todos os dados envolvidos nas operações. Em particular, Fredkin e Toffoli desenvolveram um modelo baseado em colisões elásticas de esferas rígidas denominado modelo das bolas de bilhar (*Billiard Ball Model*), onde a informação é conservada e as computações associadas às colisões podem ser revertidas (Fredkin & Toffoli, 1982). Não muito tempo depois, Norman Margolus mostrou que esse

33 Isso é particularmente simples para computações de funções computáveis inversíveis. Se há registro da função computada pela máquina, bastaria que ela assumisse uma configuração capaz de operar na função inversa.

34 A palavra “lógica” deve ser entendida aqui mais como um modelo de computação do que um formalismo como a lógica clássica ou a lógica modal.

modelo de bolas de bilhar poderia ser simulado por um autômato celular (Margolus, 1984). Sendo assim, é perfeitamente possível pensar em autômatos celulares capazes de exercer computação reversível.

A reversibilidade é exigida na Mecânica Digital para atender à representação da reversibilidade observável em fenômenos microscópicos mais fundamentais. Em escalas menores, de acordo com a física contemporânea, as leis físicas apresentam simetria em relação à conjugação de carga, transformação de paridade e reversibilidade temporal denominada simetria CPT (Carga, Paridade e Tempo). Simetria, nesse sentido, significa que uma partícula elementar em movimento ainda estaria submetida às mesmas leis, se fossem invertidos o sinal de sua carga elétrica, o sinal de uma de suas coordenadas espaciais e o sentido do fluxo do tempo. Em respeito à tal simetria, um modelo de Mecânica Digital também exigiria que as transformações de estado do sistema nas menores escalas fossem reversíveis. De acordo com o autor:

(...) as leis que governam a evolução do estado de um modelo da MD devem ser invariantes em CPT [Carga-Paridade-Tempo]. É claro; a física é invariante em CPT. Mas no modelo da MD, é claro e cristalino que a invariância em CPT implica um tipo muito forte de conservação de informação. (Fredkin, 2003, p. 196)

A reversibilidade se tornaria possível no programa da natureza por meio da conservação da informação utilizada na computação: toda informação recebida (ou perdida) em um volume de espaço-tempo específico deve ser compensada pela perda (ou ganho) de informação em outro volume vizinho. Citando Fredkin mais uma vez:

As leis da física tornam a reversibilidade microscópica uma parte intrínseca de todos os processos físicos microscópicos. Nós somos levados a acreditar que o custo de seguir em frente a partir de um estado deve ser o mesmo custo de voltar para trás a partir daquele estado. Isso significa que informação é microscopicamente e localmente conservada. Nós postulamos que em qualquer volume de espaço-tempo, informação que é recebida ou perdida daquele volume deve ser perdida ou recebida daquelas regiões que são vizinhas espaço-temporais do volume. Nesse caso, conservação de informação é como a conservação de energia. (Fredkin, 2003, p. 206)

Vale notar que nem todos os sistemas de autômatos celulares admitem reversibilidade. O próprio Jogo da Vida não é reversível. Por exemplo, um *block*, padrão estável de quatro células vivas (ver Seção 1.4.2), pode se originar de um aglomerado de três células vivas vizinhas entre si conforme o arranjo da figura abaixo:



Figura 15: Aglomerado de três células vivas do Jogo da Vida, vizinhas entre si e que geram um block na próxima geração.

Se iniciarmos o Jogo da Vida apenas com um *block*, é sabido que ele se manterá assim por todas as demais gerações. No entanto, se iniciarmos com as três células na configuração da Figura 15, a próxima geração formará um *block* e continuará assim por todas as gerações subsequentes. A Figura 16 ilustra os dois casos.

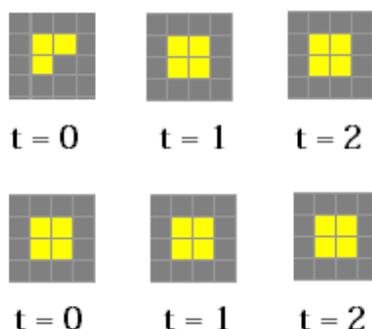


Figura 16: Formação de um block no Jogo da Vida por dois meios possíveis.

Podemos notar que um *block* pode se originar de dois *inputs* diferentes. Porém, revertendo a computação não podemos determinar se o *block* era formado por três células na geração anterior ou se já era um block, tornando o *output* da computação inversa indeterminado. Sendo assim, seria preciso algum recurso extra onde a informação de qual *input* originou o *block* fosse armazenada para permitir a reversibilidade do processo. Isso mostra que o Jogo da Vida não pode ser escolhido como um sistema responsável por representar a realidade³⁵. Embora o jogo de Conway seja equivalente a um computador universal, como mencionado na seção anterior, não é o suficiente para o tipo de sistema desejado por Fredkin: um autômato celular universal reversível (ACUR).

³⁵ A não ser que recursos de conservação de informação sejam adicionados à dinâmica do Jogo da Vida para permitir reversibilidade.

1.7. Modelos de Mecânica Digital

Fredkin não declara o conhecimento de um modelo computacional definitivo para o mundo natural em nenhum momento. A princípio, tudo apresentado aqui até agora é apenas especulação. Não obstante, com base no conhecimento físico atual, o autor apresenta uma ideia de como o sistema que deseja poderia parecer em seus aspectos mais básicos e como ele poderia se aplicar à natureza fundamental (Fredkin; 1990, 2003, 2004).

Já sabemos que a Mecânica Digital procura um modelo de natureza correspondente a um autômato celular. Isso exige pelo menos duas grandezas em comum entre natureza e programa: grandezas espaciais e grandezas temporais. Em um autômato celular, as células se encontram em uma relação de vizinhança entre si (o que pode ser entendido como uma relação espacial) e podem alterar seus estados dependendo dos estados vizinhos (havendo mudança, concluímos que há tempo). Assim, Fredkin já apresenta duas unidades básicas para modelos de Mecânica Digital: L (unidade básica de espaço ou simplesmente o comprimento de separação entre as células) e T (unidade básica de tempo, indicando cada ciclo de operação da regra de transição do autômato celular). Além delas, há uma terceira unidade básica: a unidade de informação B. Sendo guiado por uma preferência pela simplicidade (cf. Fredkin, 2003, p. 191), a unidade de informação mais simples corresponderia a um *bit* (ou seja, existiriam dois estados possíveis para cada portador mínimo de informação). Em termos físicos, no sistema sugerido por Fredkin essa unidade de informação carregaria a mesma unidade de medida do momento angular, assim como a constante de Planck. Ele apresenta algumas razões para essa escolha: em primeiro lugar, o sentido do momento angular só pode assumir dois valores (um associado ao sentido horário de giro e outro para o sentido anti-horário). Isso também se aplica ao momento angular intrínseco das partículas (o chamado *spin* das partículas elementares, com apenas dois valores possíveis em cada direção). Além disso, como a reversibilidade é necessária, adotar essa unidade de medida pode ser bastante útil, pois o momento angular inverteria seu valor quando a direção do tempo fosse invertida (a inversão de um giro em sentido horário necessariamente resultará em um giro em sentido anti-horário). Nas palavras do autor:

No Sistema Internacional de unidades, B tem as mesmas dimensões de momento angular ou ação, ML^2T^{-1} [M para unidade de massa, L para unidade de comprimento e T para unidade de

tempo]. O valor de B é o mesmo valor de \hbar – a constante de Planck reduzida. Por conta da constante B ter momento angular, quando a direção do tempo é revertida o momento angular de B deve mudar de orientação (do sentido anti-horário para o sentido horário e vice-versa). Nós representamos isso pelo sinal de B, no qual é positivo (+) para o tempo seguindo em uma direção e negativo (-) para o tempo seguindo na direção oposta. (Fredkin, 2003, p. 212)

Como já mencionado (Seção 1.3), para Fredkin, informação envolve uma configuração significativa para algum processo (no caso, o autômato celular da natureza). Não há uma obrigatoriedade em relação ao modo como esses estados são representados, desde que o sistema seja capaz de identificá-los e diferenciá-los. Na concepção de Fredkin, uma unidade de informação em termos de momento angular é perfeitamente razoável, embora não garanta uma certeza de que realmente seja a unidade correta, como podemos ler na seguinte passagem:

Um sistema físico de dois estados, tais como um férmion que está ou com spin para cima ou com spin para baixo, pode representar um bit de informação. Nós podemos, com um grande risco de possível confusão, chamar tais sistemas de estados físicos com dois estados de bits. $\text{Spin} \pm \frac{1}{2} \hbar$ é uma escolha bastante razoável para tais bits físicos. Isso significa que o bit físico é representado por quantidade conservada, o momento angular. Isso não coloca nenhum problema na computação. O conceito de bit é independente de como ele é representado; informação digital tem a ver com os significados inerentes no arranjo de coisas e não é relacionado com quais são essas coisas. (Fredkin, 2004, p. 35)

A unidade básica de espaço será simplesmente a distância efetiva entre duas células do autômato. Já a unidade básica de tempo envolve um tratamento um pouco mais complexo: ela não seria uma unidade mínima de tempo, mas corresponderia a um ciclo de subunidades de tempo ou fases de tempo indicadas por P (cf. Fredkin, 2003, p. 212), de modo que, em cada uma dessas subunidades de tempo, o autômato processe a informação contida em cada vizinha de uma dada célula. Isso é necessário para que o tempo exigido pela regra de transição do programa para verificar o estado da vizinhança seja contemplado fisicamente no sistema. A transição efetiva de estado ocorre em um tempo T, mas isso só pode acontecer após P subinstantes onde o processamento é efetuado. De acordo com Fredkin:

A regra é, com efeito, uma lei fundamental da física que pode tanto ser pensada como uma regra que tenha P como um parâmetro e o microtempo como um argumento como pode ser pensada como tendo P como microregras aplicadas de novo e de novo em rotação. (Fredkin, 2003, p. 213)

Além dessas três unidades básicas (B, L e T), mais a variável com respeito às subfases de tempo (P), Fredkin também leva em conta o número de dimensões espaciais (D), um valor correspondente à regra do algoritmo (R), outro relativo à idade do universo computado em unidades de T (A) e um valor referente às condições iniciais do universo (I). D considera simplesmente as nossas três dimensões espaciais. R é um número referente à tabela de regras aplicável ao autômato entre as várias listas de instruções possíveis para um dado autômato celular. R assume um valor específico para nosso mundo. Se fosse diferente, o comportamento da natureza também seria diferente em algum aspecto. Como todo algoritmo computacional pode ser representado por uma sequência de números inteiros, não é difícil associar uma das regras possíveis do autômato com um número em especial. Nas palavras de Fredkin:

Programas de computador, algoritmos, e autômatos são algo novo e não há formas canônicas definidas para expressar suas representações. No entanto, o que sabemos com certeza é que qualquer algoritmo ou autômato celular pode ser representado por uma tabela de transição ou por uma cadeia de instruções para qualquer computador universal e que podemos fazer com que cada representação seja equivalente a um inteiro. Assim, todo algoritmo computacional pode ser representado por um conjunto de inteiros (palavras de computador) ou pode ser representado como um inteiro muito grande (um bloco inteiro de palavras de computador). Tais inteiros são conjuntos concatenados de instruções para algum modelo canônico de computador. No modelo da MD, uma tabela simples e curta dá a definição da regra. R define o processo que cobre o presente estado do universo no próximo estado. (Fredkin, 2003, p. 214)

As quantidades restantes, A e I, são chamadas por Fredkin de “constantes cosmológicas” (cf. Fredkin, 2003, p. 214). A seria simplesmente a contagem de ciclos desde o início do funcionamento do programa até o estado presente e I seria um número indicando o estado inicial global do autômato entre todos os arranjos iniciais possíveis.

Com apenas essas unidades básicas seria possível derivar por análise dimensional diversas grandezas físicas associadas a elas. A Figura 17 a seguir, retirada e traduzida de Fredkin (2003, p. 216), mostra uma comparação entre grandezas físicas definidas em unidades do Sistema Internacional e grandezas físicas definidas em unidades da Mecânica Digital.

Dimensão	Unidades SI (m ^a , kg ^b , s ^c)	Unidades BLT (B ^d , L ^e , T ^f)
Comprimento	m	L
Massa	kg	BL ⁻² T
Tempo	s	T
Momento Angular	m ² kg s ⁻¹	B
Energia	m ² kg s ⁻²	BT ⁻¹
Momento	m kg s ⁻¹	BL ⁻¹
Força	m kg s ⁻²	BL ⁻¹ T ⁻¹
Potência	m ² kg s ⁻³	BT ⁻²
Pressão	m ⁻¹ kg s ⁻²	BL ⁻³ T ⁻¹
Momento de Inércia	m ² kg	BT
Carga ao quadrado	C ²	BLT ⁻¹
Viscosidade	m ² s ⁻¹	BL ⁻³

a Distância que a luz viaja em 1/299792458 s

b Massa padrão de 1901.

c 9,192,631,770 ciclos da transição hiperfina do Césio¹³³.

d Igual a \hbar , a constante de Planck reduzida.

e c vezes T

f Unidade natural de tempo.

Figura 17: Tabela comparativa entre unidades de grandezas físicas adotadas pelo Sistema Internacional e as unidades adotadas na Mecânica Digital (Fredkin, 2003, p. 217).

A missão da Mecânica Digital é encontrar um modelo computacional isomórfico ao comportamento do mundo natural, ou seja, os mesmos resultados obtidos pelos experimentos da física contemporânea devem ser derivados pela Mecânica Digital por meio do algoritmo correspondente ao modelo computacional. Se, ao tomarmos três unidades para informação, distância e tempo, conseguimos gerar unidades equivalentes ao que é trabalhado nas medições da física atual, é um sinal de que o modelo desejado por Fredkin apresenta alguma chance de ser bem sucedido. Note que é perfeitamente possível comparar as unidades de Fredkin com muitas das constantes aceitas pela física. Se a velocidade da luz é expressa em termos de distância e tempo, também podemos definir uma constante da Mecânica Digital em termos de L e T correspondente à velocidade de uma partícula do autômato celular equivalente à luz. Se há uma equivalência entre massa e energia, além do fato de que energia pode ser definida em

termos de momento angular e o inverso do tempo (uma vez que a constante de Planck é medida com unidades de momento angular, e frequência é definida como o inverso do período), as unidades adotadas no modelo digital também já fornecem dados para definir massa e energia no contexto do programa. O desafio é trabalhar essas unidades em um conjunto de relações que seja compatível com os frutos de nossa física já conhecida.

Embora essas unidades de medida básicas sejam direcionamentos úteis para pensar um programa fundamental do universo, nada disso fornece qualquer comprovação definitiva de uma descrição digital da realidade. As indicações mostradas nos parágrafos acima apenas mostram um esforço de Fredkin em procurar a representação computacional mais adequada para o mundo natural. Segundo Fredkin, a natureza pode ser descrita por um autômato celular universal reversível, embora ainda seja um trabalho em progresso definir os detalhes deste autômato. Uma das opções trabalhadas pelo autor para determinar o programa da natureza seria um modelo *Salt* de autômato celular (cf. Fredkin, 2003, 2004). O nome significa “sal” em inglês, chamado assim pelo fato das células do autômato se agruparem como no arranjo dos átomos de sódio e cloro nas moléculas de sal. Nesse modelo, as células se encontram em uma distribuição espacial cartesiana de três dimensões, onde o estado inicial de cada célula é definido pela soma dos números inteiros associados às suas coordenadas espaciais x , y e z de modo que se $x + y + z$ for um número par, a célula assume um estado A, enquanto se $x + y + z$ for um número ímpar, ela assume um estado B. O processamento dos estados efetuado pelo programa ocorre em ciclos, onde cada ciclo depende de seis subinstantes de tempo (rotulados de 0 a 5, por exemplo) de modo que em subinstantes pares, o programa lida apenas com células no estado A e em subinstantes ímpares, o programa lida apenas com células no estado B. Essa construção é feita visando as condições de reversibilidade, como é esperado de um modelo compatível com a Mecânica Digital. Detalhes mais técnicos sobre o estudo do modelo *Salt* podem ser encontrados em Miller & Fredkin (2005) e alguns resultados interessantes (como a simulação de movimentos próximos de uma trajetória circular por meio de um autômato celular tipo *Salt*) podem ser conferidos em Miller & Fredkin (2012)³⁶. Contudo, nunca é demais repetir que o modelo *Salt* não é nada definitivo, mas apenas um dos possíveis candidatos aos primeiros modos de representação do mundo natural digital. Talvez o programa por trás da realidade seja bem diferente, algo que o próprio autor reconhece. Citando Fredkin mais uma vez:

36 Simulações de um autômato celular aos moldes do Salt podem ser encontradas no endereço <http://busyboxes.org>. Data de acesso: 21 de setembro de 2016.

Tenha em mente que todo detalhe aqui é simplesmente um exemplo de como a MD deve funcionar. Há pouca dúvida de que os detalhes verdadeiros, quando nós os conhecermos, serão diferentes; mas o que é oferecido aqui é um quadro conceitual que deve, com muitos detalhes diferentes, ser um modelo dos processos fundamentais na física. (Fredkin, 2004, p. 37)

Para nossos propósitos nesta dissertação não estaremos tão preocupados com o fato de ainda não possuímos um programa definitivo, uma vez que nosso interesse estará centrado nos aspectos mais filosóficos da questão e não nos detalhes técnicos. Sabemos que Fredkin deseja representar a natureza por meio de um autômato celular capaz de executar computação universal e reversível, partindo do pressuposto de que exista um limite para a quantidade de informação contida em cada volume de espaço-tempo. Diante disso, em uma descrição geral, o pensamento do autor pode ser resumido em quatro leis fundamentais (cf. Fredkin, 2003, p. 206), a saber: (i) toda informação é conservada; (ii) os processos fundamentais da natureza devem ser processos de computação universal; (iii) o estado de qualquer sistema físico deve apresentar algum meio digital de representação; e (iv) o único tipo de mudança é aquela causada por algum processo informacional digital. Mesmo que ainda não tenhamos um modelo específico pronto, a simples aceitação dessas leis fundamentais nos leva a implicações conceituais de interesse filosófico, previstas pelo próprio autor em seus trabalhos. Não temos condições de entrar nos pormenores de todas essas consequências, mas é válido traçar alguns comentários a respeito delas em nossa próxima seção.

1.8. Implicações da Hipótese de Fredkin

1.8.1. Simetrias Contínuas

O autor é ciente de que nossa física atual considera o conceito de simetria (Fredkin, 2004, p. 38). No presente contexto, simetria é entendida no sentido físico, como a ideia de que certas propriedades de um sistema físico não se alteram quando o mesmo sistema é submetido a determinadas transformações como rotação, translação no espaço e translação no tempo. Simetrias podem ser discretas (propriedades do sistema não se alteram diante de mudanças discretas, como a alteração do valor da carga elétrica) ou contínuas (propriedades do sistema não se alteram diante de mudanças contínuas, como uma rotação ou translação contínua). Conciliar simetrias contínuas com a Mecânica Digital é um problema.

Ao entrar nessa discussão, Fredkin invoca o chamado Teorema de Noether (Noether, 1918), um enunciado que permite derivar leis de conservação a partir da noção de simetria em um sistema físico (assim, a lei de conservação do momento angular, por exemplo, pode ser entendida a partir de uma simetria rotacional), bem como derivar noções de simetria com base em princípios de conservação (cf. Fredkin, 2003, p. 205; 2004, pp. 38-39). No caso da Mecânica Digital, a conservação de uma informação discreta (como a conservação do momento angular) pode levar a uma simetria assintoticamente contínua em maiores escalas. Para Fredkin, o mundo físico é fundamentalmente discreto, mas em uma escala macroscópica haveria uma aparência de continuidade, ou seja, quanto mais se amplia a escala, mais contínuo o mundo parece aos nossos olhos. Nas palavras do autor:

Se nossa física microscopicamente discreta e finita conserva exatamente o momento angular, momento, e energia, então em uma escala suficientemente grande a granularidade do espaço e tempo discretos, a simetria da rotação (isotropia angular), simetria de translação espacial, e simetria de translação temporal serão assintoticamente contínuas! Então a MD [Mecânica Digital] retira essas simetrias de sua posição de mais fundamentais e as substitui com leis de conservação mais fundamentais. As simetrias aparentemente contínuas surgem assintoticamente – em escalas acima do arranjo de espaço-tempo – e elas têm nos enganado ao nos fazer acreditar na realidade da continuidade física. (Fredkin, 2004, p. 38).

Resumidamente, haveria, de fato, uma relação entre conservação e simetria, mas não uma simetria perfeitamente contínua. As conservações da natureza fundamental (garantidas pela primeira lei da MD em conservar a informação) se aproximam cada vez mais de simetrias rotacionais, espaciais e temporais³⁷, mas essas simetrias seriam apenas aproximadamente contínuas.

1.8.2. Sistema de Referência Absoluto

A Mecânica Digital é baseada em um programa capaz de identificar a informação ligada à cada partícula mínima do sistema. Mas isso implica em um processamento de informação associado em última instância a um sistema de referência privilegiado, afinal, toda informação, lembramos, é uma configuração significativa para um determinado contexto e, no caso, o contexto seria o ACUR da natureza como um todo. Fredkin também está ciente disso.

³⁷ Mais precisamente, conservação da informação de momento angular levaria à simetria rotacional, conservação da informação de momento linear levaria à simetria translacional e conservação da informação de energia levaria à simetria temporal (cf. Fredkin, 2004, pp. 38-39).

Para ele, contudo, isso não significa uma violação de nenhuma lei experimental específica, mas sim a violação do conceito teórico de que não existe um referencial privilegiado. Na visão do autor, não há problema algum em assumir um sistema de referência absoluto:

Não conhecemos nada nos dados experimentais (ou em modelos matemáticos que correspondam a eles) que nos obrigue a rejeitar o conceito de um único e fixo sistema de referência. Todavia a comunidade da física tem escolhido fazê-lo por motivos heurísticos. Isso acarretou em um conceito útil como a ideia de que não há um sistema de referência fixo se tornando um dogma na física contemporânea. A evidência experimental é meramente que temos tentado detectar um sistema de referência fixo e temos falhado. As leis matemáticas da física certamente não proibem tal sistema de referência fixo; pelo contrário, elas admitem que todo sistema de referência irrotacional e não acelerado é consistente com as leis da física. (Fredkin, 2003, p. 208)

Assim sendo, embora não possamos detectar um sistema de referência absoluto, isso não implica sua inexistência. Veremos nessa mesma seção que Fredkin não considera nosso universo físico como um sistema completamente isolado de qualquer outra coisa, ou seja, na visão do autor seria perfeitamente possível imaginar um referencial exterior ao próprio universo.

Uma outra consequência ao admitir um referencial privilegiado seria a conclusão de que as simetrias mencionadas na subseção anterior não seriam absolutas. Se há um sistema de referência absoluto, conseqüentemente não seria possível admitir simetrias translacionais e rotacionais perfeitas. Essa perda de isotropia é um preço a se pagar pelas leis informacionais e computacionais adotadas pela hipótese digital em um nível elementar, embora essas mesmas simetrias ainda possam ser consideradas como boas aproximações em escalas maiores. É o que vemos na citação a seguir:

Não há dúvidas de que a simetria translacional absoluta e a simetria rotacional absoluta, como as propriedades mais básicas da física, são uma impossibilidade informacional. Essas duas simetrias foram desenvolvidas e se tornaram uma fixação da física, começando com Newton e continuando com Poincaré e Einstein. Modelos matemáticos que assumem essas simetrias são essencialmente corretos. Entretanto há leis sobre processos informacionais que são desconhecidas para a física contemporânea. Essas novas leis tornam claro que, no nível mais microscópico da física, tanto a simetria translacional contínua quanto a simetria rotacional contínua devem ser violadas. Além do mais, simetria contínua no tempo também deve ser violada no nível microscópico mais básico. Todavia ainda é razoável e conveniente fazer uso dessas simetrias contínuas para todos os processos físicos acima dos níveis mais microscópicos. (Fredkin, 2003, p. 209).

Em resumo, se são aceitas as duas últimas leis fundamentais exibidas na seção 1.7, a

saber, todo sistema admite representação digital e toda mudança é causada por um processamento de informação, então também se faz necessário aceitar um sistema de referência privilegiado que defina essas transformações informacionais digitais, mesmo que algumas simetrias sejam violadas em níveis de escala mais fundamentais (embora ainda seja matematicamente aceitável usá-las em uma escala mais ampla, mas não como correspondência fiel à realidade elementar).

Quanto à dificuldade de conciliar um único sistema de referência com os resultados da relatividade geral, onde espaço e tempo podem ser “deformados”, Fredkin lembra que o movimento de uma partícula do ACUR é totalmente definido pela programação³⁸. Nesse sentido, não é o espaço-tempo absoluto do programa que se modifica, mas sim o modo com que a regra por trás do autômato lida com a movimentação. Citando Fredkin mais uma vez:

Deve-se lembrar que as propriedades físicas do nosso espaço-tempo não são diretamente determinadas pela estrutura; ao invés disso, elas são determinadas pelos caminhos que as partículas tomam conforme elas se movimentam pela estrutura; é o movimento delas que se conformam às leis da física. Uma vez que o AC [Autômato Celular] Cartesiano é universalmente computável, é um fato tautológico que a estrutura do sistema pode suportar regras que produzem partículas e campos para os quais os movimentos das partículas estão de acordo tanto com a relatividade especial quanto com a relatividade geral. Naturalmente, o mesmo é verdadeiro para partículas relativísticas para as quais as massas e relógios internos devem obedecer as leis da relatividade especial. (Fredkin, 2004, p. 41).

1.8.3. Mecânica Quântica e Determinismo

Talvez a consequência mais polêmica da hipótese computacional da natureza seja a controvérsia gerada com os resultados não-determinísticos da mecânica quântica. Se o mundo físico é mesmo equivalente a um programa de computação reversível, onde toda a informação é conservada, isso significa que não há espaço para dúvidas quanto às transições de estado ocorridas. Um valor de *input* terá um valor de *output* bem definido e vice-versa graças à reversibilidade da computação. Nesse sistema, o determinismo parece evidente. Contudo, a história revela que as descobertas da mecânica quântica mostram uma grande dificuldade em encontrar uma explicação determinista para os resultados de experimentos em escala atômica e subatômica. Mesmo insinuando que a mecânica quântica seja uma teoria incompleta (cf.

³⁸ Na verdade, as chamadas “partículas em movimento” seriam alterações de estados das células em uma sequência bem ordenada. Repetimos que no Jogo da Vida, por exemplo, temos a impressão do movimento de uma célula para outra quando duas células consecutivas na mesma linha ou coluna trocam seus estados.

Einstein et. al., 1935) e sugerindo a existência de variáveis ainda desconhecidas nessa teoria que justifiquem um determinismo, como ocorre em algumas interpretações da teoria quântica (e.g. Bohm, 1952), seria necessário admitir que a natureza se comporta de modo não-local, ou seja, uma ação poderia causar um efeito à longa distância de modo instantâneo, transgredindo a impossibilidade de transmissão mais rápida que a luz adotada pela relatividade especial. Segundo o físico John Bell, variáveis ocultas são incompatíveis com qualquer teoria da física que envolva localidade (Bell, 1966). Isso mostra que adequar as excentricidades quânticas ao modo de pensar da física clássica é uma tarefa bastante complicada. Como a computação adotada por Fredkin em sua perspectiva digital parece adotar uma visão de mundo clássica, conciliá-la com a mecânica quântica seria tão problemático quanto.

Muitos autores simpáticos à concepção de um universo baseado em computação, mas cientes da dificuldade em conciliar determinismo computacional com a teoria quântica, preferem admitir que o universo é um computador fundamentalmente quântico (e.g. Deutsch, 1985; Lloyd, 2006). Nesse outro ponto de vista, se existir um programa por trás da realidade natural, tal programa deveria processar informação quântica, de tal modo que os *bits* em jogo seriam definidos por estados em superposição. Nessa abordagem, a mecânica quântica seria mais fundamental do que a computação da natureza e já estaria incluída na interpretação computacional do universo. Essa é uma opinião aceita por Richard Feynman, por exemplo, como podemos ler na seguinte citação:

O programa que Fredkin tem sempre apoiado sobre tentar encontrar uma simulação computacional da física, me parece um excelente programa a seguir. Ele e eu temos tido maravilhosos, intensos e intermináveis argumentos, e meu argumento é sempre que o real uso dessa ideia envolveria a mecânica quântica, e portanto total atenção e aceitação dos fenômenos da mecânica quântica – o desafio de explicar fenômenos quânticos – tem sido colocados em argumentação, e assim esses fenômenos tem sido entendidos muito bem ao analisar a situação. E não estou feliz com todas essas análises que levam em conta apenas a teoria clássica, porque a natureza não é clássica, diabos, e se você quer fazer uma simulação da natureza, é melhor torná-la quântica, e, por Deus, isso é um problema maravilhoso, pois não parece nada fácil. Obrigado. (Feynman, 1982, p. 486)

Fredkin, por outro lado, não considera a mecânica quântica como uma teoria fundamental. Ele defende que os fenômenos chamados quânticos não são os mais básicos, mas sim resultados aparentes de operações ainda mais elementares. Para Fredkin, existiria apenas um sistema computacional básico moldado pela Mecânica Digital, não fazendo sentido em falar na separação entre mundo clássico e mundo quântico, como podemos ver na citação

a seguir:

A Filosofia Digital toma uma visão definitiva da mecânica quântica que não é ligada com a interpretação de Copenhague ou mesmo com qualquer outra interpretação contemporânea. Nós rejeitamos a ideia de que há algo como um ato irreversível de observação ou que há um mundo clássico e um mundo quântico. A Filosofia Digital assume que algum sistema de Mecânica Digital é tudo que há, enquanto impõe leis informacionais incluindo conservação de informação. (Fredkin, 2003, p. 202).

Essa visão de Fredkin recebe apoio do físico Gerard 't Hooft, para quem a ontologia do universo também não está sedimentada nas leis quânticas: a mecânica quântica seria apenas uma ferramenta para descrever estatisticamente fenômenos de uma escala menor em uma escala maior, assim como, por exemplo, a termodinâmica é uma descrição macroscópica do que é trabalhado microscopicamente na mecânica estatística. Em escalas ainda menores, onde nossa experiência ainda não consegue alcançar, as leis físicas podem ser muito diferentes do que é apresentado pela mecânica quântica. Nas palavras de 't Hooft:

É aqui que eu gostaria de defender uma aproximação diferente. Não devemos nos esquecer que a mecânica quântica não descreve realmente que tipo de fenômeno dinâmico está acontecendo de fato, mas apenas nos dá resultados probabilísticos. Para mim, parece extremamente implausível que qualquer teoria razoável para a dinâmica na escala de Planck levaria a processos que sejam tão complicados de descrever, que devêssemos esperar flutuações estocásticas em qualquer teoria de aproximação descrevendo efeitos desse nível em escalas maiores. Parece muito mais razoável primeiramente tentar uma teoria clássica e determinística para o domínio na escala de Planck. Poderia se especular então que o que chamamos de mecânica quântica hoje, possa não ser nada mais do que uma técnica engenhosa para tratar essa dinâmica estatisticamente. ('t Hooft, 2001, p. 308-309).

Se existe uma possibilidade lógica em tomar os fenômenos quânticos como não-básicos e não existe nenhuma comprovação experimental que diga o contrário, então, mais uma vez, esforços como os de 't Hooft e Fredkin não podem ser descartados de imediato. Além disso, a ideia de um mundo descrito por um autômato celular bem definido envolve um determinismo completo, o que 't Hooft chama de “superdeterminismo” ('t Hooft, p. 10). Se for o caso, mesmo as ações dos observadores envolvidos já seriam pré-definidas por leis deterministas e isso poderia contornar os impedimentos propostos por Bell ('t Hooft, 2015, pp. 73-76). Já existem até mesmo tentativas de explorar experimentalmente essa possibilidade (e.g. Hossenfelder, 2011).

Na conjectura de Fredkin, portanto, o determinismo é bastante evidente. A partir de estados iniciais, o estado do próximo instante será definido pelas regras do autômato. Ocorre,

porém, que essa determinação não pode ser oferecida por qualquer um “dentro da máquina”. Uma vez que fazemos parte do ACUR, nossa capacidade de raciocínio também está limitada pelas capacidade do autômato e, portanto, seria impossível para qualquer um de nós apresentar um *output* mais rapidamente do que o próprio programa é capaz. Segundo Fredkin:

Se há uma questão cuja resposta depende da evolução futura exata de parte do universo então, em geral, não há um meio mais rápido de conseguir essa resposta exata além de deixar que essa parte do universo continue sua evolução. Essa observação basicamente matemática tem um certo apoio no entendimento da natureza da Filosofia Digital. Enquanto a Filosofia Digital é absolutamente determinista, nós decidimos chamar isso de “determinismo incognível”. O conhecimento de que há determinismo não permite que nós, vivendo no universo, façamos qualquer previsão exata do futuro. Nesse sentido a Mecânica Digital é como a Mecânica Quântica. (Fredkin, 2003, p. 210)

De acordo com a citação acima, nossa realidade física conta com um determinismo intrínseco, mas não se segue que os estados futuros sejam previsíveis por qualquer habitante do mundo: não temos a capacidade de determinar com precisão os estados futuros do universo por conta de nossas limitações cognitivas. Nesse sentido, a Mecânica Digital é tão imprevisível quanto a mecânica quântica, não obstante seja objetivamente determinista.

1.8.4. O “Outro”

Vimos que Fredkin define nosso universo físico como um grande autômato celular, cujas leis são determinadas por certas regras bem definidas. Mas se o universo físico é um sistema computacional, ele seria um sistema completamente isolado ou um programa funcionando em alguma máquina? As duas posições são defensáveis, mas nos limitaremos à visão que Fredkin defende: nosso mundo é um sistema computacional rodando em algum outro nível de realidade maior. Fredkin chama esse outro nível apenas de “*Other*” (Outro):

Se espaço, tempo, matéria e energia são todos consequências do processo informacional acontecendo no Computador Supremo, então tudo em nosso universo é representado por esse processo informacional. O local onde o computador está, o dispositivo que executa esse processo, escolhemos chamar de “Outro”. (Fredkin, 1993, p. 119)

A justificativa que Fredkin encontra para postular a existência do “Outro” é a contingência de nosso universo. Nossa realidade física se originou de certas condições iniciais seguindo determinadas regras, mas não há razão para dizer que essas condições e essas regras

sejam necessárias. O universo poderia ter sido um autômato diferente, com regras diferentes. Basta lembrar que o princípio de computação universal permite que diferentes algoritmos sejam executados e as leis de nosso mundo físico constituem apenas um desses infinitos algoritmos. Isso é um indício de que, ao menos em tese, poderia haver uma realidade de segunda ordem, onde nosso universo é apenas uma instância de todos os programas possíveis. Podemos ler isso na citação a seguir:

De onde o Outro veio? Na verdade essa questão é bem fácil de lidar. A natureza dos sistemas de leis que podem suportar computação é muito mais ampla do que a natureza dos sistemas que estão limitados à física do nosso universo. Em outras palavras, muitas das propriedades de nosso mundo que são necessárias para que ele tome a forma que possui não são necessárias para outros tipos de mundos que podem suportar computação universal. E computação universal, do tipo que pode simular outros computadores de propósito geral, já é uma propriedade de todos os computadores comerciais comuns. (Fredkin, 1993, p. 119)

Fredkin não se compromete com nenhuma propriedade específica do “Outro”. Ele apenas afirma que essa realidade não estaria necessariamente limitada às mesmas regras do nosso universo. Se nossa realidade possui 3 dimensões espaciais, o “Outro” pode conter ou não essas três dimensões: talvez ele tenha muito mais dimensões. Se nossa realidade física teve um início, o “Outro” pode ou não ser eterno. O “Outro” pode ser muitas outras coisas que podemos apenas especular. Alguém com uma veia mais religiosa poderia imaginar o “Outro” como uma indicação da existência de um Deus criador além do universo (na linha de raciocínio de Fredkin, um “Deus programador”), enquanto outros podem imaginar o “Outro” como um mundo muito parecido com o nosso, mas com maiores poderes computacionais, suficientes para nos gerar como uma simulação de si mesmos³⁹. Citando Fredkin mais uma vez, em um artigo mais recente:

As leis da física no Outro não precisam ter muito em comum com as leis da física em nosso universo. Em particular, a necessidade de um início (como o Big Bang) não precisa ser uma propriedade do Outro. O Outro pode ou não ter leis de conservação. Por outro lado, Outro também deve ter um espaço-tempo 3 + 1 dimensional, mas ele poderia ter mais ou menos dimensões. O Outro teria que ser maior (quanto ao número de estados possíveis) do que nosso Universo. Muito maior. Poderia se imaginar que o Outro existe ou tenha sido construído em um Grande Autômato Celular que implementa a física de nosso universo e uma vez que as condições iniciais fossem estabelecidas no Grande Autômato Celular o processo seria colocado em movimento e deixado para evoluir sem perturbações. (Fredkin, 2013, p.380)

Embora continue no campo da especulação, Fredkin comenta que poderia haver um

³⁹ Semelhantemente ao que é discutido em Bostrom (2003).

propósito para o desenvolvimento de nosso universo: talvez a computação de toda a nossa realidade esteja ocorrendo com o objetivo de encontrar um *output* em especial (Fredkin, 1993, p. 120), mas não é possível ter qualquer certeza quanto a isso. Também não é impossível concluir que o “Outro” seja uma realidade computada em outra realidade de ordem superior, que, por sua vez, pode ser o programa de outra realidade e assim por diante, levando facilmente para um cenário de “bonecas russas” com uma regressão infinita de realidades computacionais. Não obstante, também poderíamos viver no caso de realidade computacional mais simples de todos: o “Outro” poderia ser uma única realidade constituída de uma gigantesca máquina de Turing universal computando todos os algoritmos possíveis. Essa é uma opinião compartilhada pelo cientista da computação Jürgen Schmidhuber, ao defender que é mais simples projetar uma única máquina com a única instrução de executar todas as computações possíveis do que programar uma máquina específica para cada realidade (Schmidhuber, 1997). Há, portanto, várias especulações, mas nenhuma certeza.

Em resumo, Fredkin expressa um realismo metafísico quanto à existência objetiva de um domínio além de nossas experiências observáveis, onde as leis da computação universal seriam executáveis e nosso universo físico formaria apenas uma das várias computações possíveis nessa realidade de segunda ordem com leis físicas não necessariamente iguais às nossas. Graças ao princípio de computação universal que permite uma realização múltipla de algoritmos em sistemas de naturezas diferentes, como já vimos na seção 1.5, não é necessário conhecer os detalhes desse outro nível de realidade: é suficiente compreender que estados específicos desse domínio constituem os dados digitais de um programa computacional.

1.8.5. Reduccionismo

Por fim, a questão filosófica que mais irá nos interessar neste trabalho é que Fredkin parece defender uma visão reducionista.

Nosso objetivo no restante desta dissertação será justamente analisar o pensamento de Fredkin em termos do pensamento reducionista, em particular, uma visão reducionista aplicada ao mundo físico, onde propriedades e leis consideradas em nossas teorias científicas podem ser completamente definidas por propriedades e leis mais fundamentais.

Exploraremos o conceito de reduccionismo no segundo capítulo e, no terceiro capítulo, ele será aplicado explicitamente ao trabalho de Fredkin.

Capítulo 2: Reduccionismo

O segundo tópico de interesse deste trabalho é o tema do reduccionismo. Uma intuição inicial é que reduccionismo pode ser compreendido como a visão de que certos aspectos da realidade são totalmente determinados por outros – conhecer todos os fatores determinantes seria suficiente para conhecer tudo a respeito do que é determinado. Nesse sentido, um castelo de areia seria completamente determinado por grãos de areia organizados em uma configuração específica, aspectos mentais seriam completamente determinados por interações neurais ocorridas no cérebro, um ser vivo seria inteiramente definido pelo arranjo de suas moléculas e fenômenos explicados por teorias da biologia poderiam ser totalmente explicados por teorias físicas. Todos esses exemplos expressam casos de redução, onde um elemento específico é totalmente explorado em termos de elementos mais básicos e gerais⁴⁰. A visão reducionista, portanto, não se dá por satisfeita enquanto não esgotar as explicações a respeito de um dado fenômeno a partir daquilo a que se reduz (Jones, 2000, p. 13).

O objetivo deste capítulo é apresentar um panorama geral da filosofia reducionista, principalmente no contexto científico. A partir disso, será possível examinar como o pensamento de Fredkin se adequa nessa perspectiva filosófica no próximo capítulo.

2.1. Conceito de Reduccionismo

Ideias reducionistas são bastante antigas. A filosofia do pré-socrático Parmênides, já no século V AC, anunciava que a realidade era uma e toda diversidade, movimento e mudança observável era apenas ilusória. O mecanicismo dos séculos XVII e XVIII encarava o mundo como uma grande máquina totalmente determinada por interações locais de suas partes básicas. Mais recentemente, no século XX, houve uma tentativa de reduzir toda a matemática à lógica por parte de intelectuais como Gottlob Frege (1848-1925) e Bertrand Russell (1872-1970). Na mesma época, no contexto das ciências empíricas, os positivistas lógicos desejavam

40 Não há uma definição única para redução. Searle (1992, p. 113), por exemplo, define redução da seguinte maneira: dados os elementos A e B, ocorre redução de A para B se e somente se A não for nada além de B. A definição de Searle parece invocar uma noção de identidade. No caso deste texto, entenderemos que A se reduz a B quando A pode ser completamente definido em termos de B. Isso não envolve necessariamente uma identidade – a mente não precisa ser idêntica ao cérebro para ser reduzida a ele, bastaria ser encarada como uma realização da estrutura neural. Nesse sentido, a mente é algo distinto do cérebro, ainda que seja completamente determinada pela atividade cerebral e, portanto, se reduza a ele. O conceito de redução poderia ser analisado em detalhes, mas não é o propósito deste trabalho. Para um estudo elaborado do conceito de redução ver van Riel (2014).

reduzir todas as sentenças sobre experiências e todo o conhecimento científico a um único corpo teórico por meio de um único método. Nas palavras de Rudolf Carnap (1891-1970), um dos principais nomes do positivismo lógico:

[...] ciência é uma unidade, de modo que todas as sentenças empíricas podem ser exprimidas em uma linguagem única, todos os estados de coisas são de um único tipo e são conhecidos por um mesmo método. (Carnap, 1934, p. 32)

A visão de Carnap revela claramente um desejo de unificar a ciência em três aspectos: epistemológico (uma única linguagem ou modo de representação), ontológico (um único tipo de elemento existente) e metodológico (um único método para alcançar a unidade). São justamente nessas três frentes (epistemológica, ontológica e metodológica) que o pensamento reducionista é costumeiramente classificado no contexto das ciências da natureza (a área que particularmente nos interessa neste trabalho), sendo Ruse (2005, p. 793) um exemplo de autor que utiliza essa separação. A princípio parece uma divisão razoável, uma vez que o processo de redução pode ocorrer com representações, elementos representados e métodos. Vejamos como essa classificação é feita em Ruse (2005).

Em um reducionismo epistemológico é aceito que uma teoria pode ser apenas um caso especial de outra, de modo que seja possível derivar logicamente a teoria reduzida a partir da teoria a que se reduz, como ocorre na termodinâmica, cujas teorias podem se reduzir à mecânica estatística. Isso significa que mudanças de teoria no decorrer da história da ciência não precisam significar um completo abandono da teoria anterior. Se há uma redução de teorias no processo de mudança, então a teoria nova absorve a teoria antiga de algum modo, ainda que seja necessário explicitar as condições que permitam a derivação lógica dos mesmos aspectos da teoria antiga a partir da teoria nova.

Já um reducionismo ontológico admite a crença de que o todo da realidade apresenta um número mínimo de entidades. Desse modo, o objetivo de um reducionista ontológico será identificar quais são essas entidades e excluir os elementos restantes da categoria de real e tomá-los apenas como representações. Essa posição pode ser exemplificada pelo materialismo, a visão de que apenas entidades materiais existem efetivamente, negando assim a existência objetiva de entidades imateriais, como uma consciência distinta e separada do cérebro. Assim, apenas processos efetuados pelos elementos materiais que constituem nosso cérebro seriam responsáveis por aspectos como comportamentos, emoções e sensações, pois a

realidade só poderia abrigar entidades materiais.

Por fim, o reducionismo metodológico indica simplesmente que um método suficiente para adquirir conhecimento completo de um todo é compreender totalmente suas partes. Independentemente do que é tomado como real ou irreal no reducionismo ontológico, um reducionista metodológico assume que um cientista deve buscar compreender um dado sistema com base em seus componentes. Assim, o químico poderia compreender todos os detalhes de uma reação química estudando os aspectos físicos envolvidos.

O físico Steven Weinberg critica essa classificação tripla por não ser muito clara – a diferença entre as três categorias não é muito exata, segundo ele – e, além disso, ela se apóia mais no que cientistas fazem, fizeram ou poderiam fazer do que na estrutura da própria natureza (Weinberg, 1996, nota 4, p. 51)⁴¹. A perspectiva reducionista de Weinberg é mais objetiva: relações de redução seriam verdadeiras independentemente da existência dos cientistas. Ainda que os cientistas não consigam explicar reações químicas complexas em termos de interações físicas, as teorias químicas não teriam princípios autônomos separados de princípios mais profundos da física. Podemos dizer, portanto, que o reducionismo de Weinberg segue um pano de fundo filosófico realista.

Em contexto científico, o realismo pode ser entendido basicamente como a ideia de que certas entidades, estados e processos definidos por teorias científicas existem objetivamente (Hacking, 1983, p. 21). Nesse sentido, teorias científicas são verdadeiras ou falsas, retratando ou não retratando fielmente elementos de uma realidade independente da existência de seres humanos. Adaptando esse conceito para a visão reducionista, a realidade seria composta de entidades, estados e processos básicos descritos por alguma teoria fundamental que determinam todos os outros elementos representados por teorias menos fundamentais, de modo que teorias científicas verdadeiras seriam aquelas que fizessem referência a esses elementos básicos.

Uma vez que a proposta de Fredkin é bastante próxima do realismo - afinal, vivemos em um mundo computacional seguindo um programa definido no “Outro”, ou seja, há claramente elementos de realidade independentes de nossa existência - será conveniente adotar um conceito de reducionismo em uma linha mais realista. Sendo assim, o estudo do pensamento reducionista no presente capítulo será feito tomando a visão de Weinberg como referencial teórico.

41 No caso, Weinberg está falando de Ernst Mayr, um biólogo evolucionista que classifica o reducionismo nessas três categorias (cf. Weinberg, 1996, pp. 50-51).

Ao aprofundar suas reflexões em uma obra posterior (Weinberg, 2001), Weinberg aborda a questão do reducionismo em duas distinções principais.

A primeira delas seria a classificação entre um reducionismo significativo (*grand reductionism*) e um reducionismo trivial (*petty reductionism*). O primeiro deles seria a ideia de que toda a natureza é como é por conta de leis universais simples nas quais todas as explicações científicas acabam se reduzindo de alguma maneira. Já o segundo seria a visão de que as coisas são como são por conta de propriedades de seus constituintes, como no exemplo da dureza do diamante, justificada pelo arranjo dos átomos de carbono (Weinberg, 2001, p. 111).

Uma segunda distinção proposta por Weinberg que, segundo ele, quase nunca é lembrada, seria a compreensão do reducionismo como programa de pesquisa de um lado e a interpretação do reducionismo como visão da natureza de outro (Weinberg, 2001, p. 112). Essa diferença pode ser entendida tomando o estudo meteorológico como pano de fundo. Os fenômenos atmosféricos são definidos em termos de princípios gerais da aerodinâmica e leis de radiação, mas para fins de previsão de chuva para o dia seguinte, talvez seja mais viável pensar em termos de noções menos básicas, como frentes frias e nuvens carregadas. Mas ainda que o reducionismo não proporcione uma linha de pesquisa satisfatória para os meteorologistas, ele sempre carregará a intuição geral de que certas noções utilizadas no estudo da atmosfera não são independentes de princípios da física – frentes frias são como são por conta de propriedades do ar e do vapor de água que por sua vez são como são por conta de princípios físicos. Podemos não saber quais são as leis últimas da natureza, mas sabemos que elas não são dadas em termos de frentes frias e nuvens (cf. Weinberg, 2001, p. 112).

A investigação do pensamento reducionista será feita nas próximas seções tomando essas distinções como base. Na seção 2.2 é discutido como um mundo reducionista deve ser estruturado, mostrando como os diferentes níveis estão relacionados, permitindo tanto uma compreensão de reducionismo significativo (leis básicas que determinam leis menos básicas) quanto uma compreensão de reducionismo trivial (propriedades básicas determinam propriedades menos básicas). A seção 2.3 muda o foco para o papel do conhecimento, mostrando como, pelo menos a princípio, níveis menos elementares podem ser conhecidos a partir dos níveis mais elementares – para tanto, serão utilizados recursos abstratos de sondas epistemológicas (Pessoa, 2005) que permitem compreender em que sentido essas determinações ocorrem. Por fim, a seção 2.4 apresenta algumas das principais objeções que

podem ser desenvolvidas contra o pensamento reducionista.

2.2. Reduccionismo como Visão de Mundo

A visão de mundo reducionista, segundo Weinberg (2001, p. 112), pode ser imaginada como um esquema de todas as áreas da ciência representadas por pontos ligados entre si por setas. Cada seta saindo de um dos pontos aponta para uma área com princípios capazes de explicar a área de origem da seta. A intuição básica é que esse esquema não terá duas regiões isoladas de pontos, o que indicaria disciplinas científicas independentes. Ao invés disso, considerando que Weinberg assume uma postura realista, todas as disciplinas estarão conectadas seguindo um padrão que tende a unir todas as ciências em um único ponto representando o conjunto de leis e princípios mais fundamentais da natureza - a tão sonhada, mas ainda desconhecida, teoria unificada dos físicos. A Figura 18 apresenta um esboço de como esse esquema pode ser ilustrado utilizando algumas áreas científicas como exemplo.

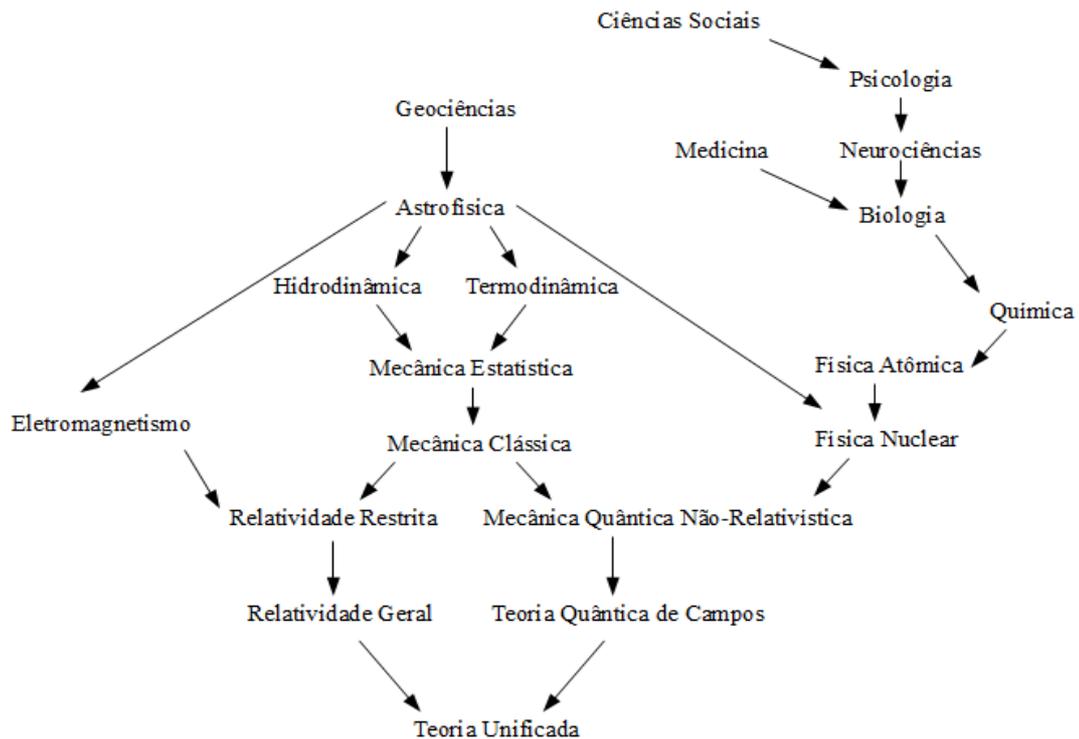


Figura 18: Esquema de uma visão reducionista com algumas áreas da ciência representadas. As setas partem de uma dada área para a área responsável por explicá-la. Como exemplo, conceitos de hidrodinâmica e termodinâmica podem ser explicados por noções da mecânica estatística que, por sua vez, é dependente de conceitos gerais da mecânica clássica. A tendência é uma convergência para uma suposta teoria unificada ainda não definida.

Uma vez que cada área da ciência atua em um determinado domínio, é possível pensar esse esquema redutivo em termos de níveis. Para duas áreas da ciência ligadas por setas, a área de origem da seta estará em um nível menos fundamental em relação à área apontada pela seta (Weinberg, 2001, pp. 112-113). Tomando mais uma vez o exemplo da meteorologia: o nível em que nuvens carregadas e frentes frias são estudadas é menos fundamental do que o nível em que partículas e campos são estudados. No entanto, o nível das partículas e campos é mais fundamental e mais geral, pois não permite em princípio explicar apenas nuvens, mas qualquer sistema que dependa de partículas e campos, o que abarca uma grande área de nosso universo material. Seguindo esse raciocínio, o nível mais fundamental da natureza seria aquele que não poderia ser explicado por mais nenhum outro⁴².

Se uma realidade reducionista segue essa estrutura de níveis, então é preciso esclarecer de que maneira esses níveis estão conectados em uma relação de redução. Lembrando que Weinberg (2001, p. 111) definiu dois tipos de reducionismo, temos uma redução baseada em leis e princípios causais (para reducionismo significativo) e uma redução baseada em estados e propriedades (para reducionismo trivial). As relações entre níveis seguindo esses dois pontos serão examinadas nas subseções a seguir.

2.2.1. Reduccionismo Significativo

O reducionismo considerado mais interessante por Weinberg (cf. Weinberg, 2001, p. 111) é o reducionismo significativo (*grand reductionism*), onde as leis dos níveis mais básicos definem as leis dos níveis menos básicos. O caso limite (e desejável por Weinberg) seria que todas as leis específicas se reduzissem a um único sistema de leis gerais, que determinassem o funcionamento de todos os níveis superiores na hierarquia de níveis da realidade natural. No contexto da física teórica, por exemplo, é recorrente aplicar esforços para encontrar uma suposta teoria unificada, capaz de conciliar dois ramos da física ainda desvinculados - a saber, a teoria quântica de campos e a relatividade geral. Mas esse tipo de reducionismo traz também a dúvida sobre como níveis diferentes podem se conectar por meio de suas leis. Uma primeira intuição seria afirmar que as ciências aplicáveis aos níveis mais básicos permitem explicar as ciências menos elementares. Podemos formalizar uma ciência a partir de suas teorias. Assim, poderíamos dizer que quando uma ciência é totalmente explicada por outra dizemos que

42 Alguns autores (e.g. Schaffer, 2003) acreditam que é impossível encontrar um nível fundamental, ainda que seja possível separar a realidade em níveis e pensar em relações redutivas entre esses níveis.

ocorre uma redução inter-teórica ou redução entre teorias, ou seja, teorias de um dado nível científico se reduzem a teorias de outro nível.

A maioria dos filósofos envolvidos com a questão da redução se interessa por esse aspecto de conexão entre teorias e todos eles, em maior ou menor grau, sempre citam o filósofo Ernest Nagel (1901-1985), um dos primeiros estudiosos a explorar profundamente o problema da redução inter-teórica durante os anos 1940. A ideia básica por trás do pensamento de Nagel é que níveis científicos são explicados por outros pelo fato de níveis menos básicos constituírem derivações lógicas formuladas a partir dos níveis mais básicos.

Em sua obra *A Estrutura da Ciência* (Nagel, 2006), lançada originalmente em 1961, Nagel apresenta a noção de redução de teorias nos seguintes termos:

Uma redução, no sentido em que empregamos a palavra aqui, é a explicação de uma teoria ou de um conjunto de leis experimentais estabelecidas em um campo de investigação por outra teoria formulada habitualmente, embora não invariavelmente, para outro domínio. Para maior brevidade, chamaremos o conjunto de teorias ou leis experimentais que são reduzidas a outra teoria de ciência secundária e a teoria a qual se efetua ou se propõe a redução de ciência primária. (Nagel, 2006, pp. 445-6).

De acordo com o entendimento apresentado na citação acima, Nagel define uma ciência como um conjunto que pode englobar teorias e leis experimentais⁴³ de modo que seja possível estabelecer uma relação entre ciências na qual uma ciência secundária seja explicada por uma ciência primária. Isso é análogo à nossa concepção de ciências organizadas em níveis, onde a ciência do nível menos fundamental é explicada pela ciência do nível mais fundamental - adaptando para a terminologia de Nagel, ciências em níveis mais básicos seriam primárias e ciências em níveis menos básicos seriam secundárias. Já o conceito de explicação pode ser entendido aqui em termos de uma dedução, na qual o elemento explicado surge como uma consequência lógica de um conjunto de sentenças tomadas como premissas (Nagel, 2006, p. 42).

Nagel considera duas maneiras de reduzir uma ciência à outra. A primeira delas envolve um caso de redução em que as leis da ciência secundária utilizam termos descritivos que não mudam de significado quando tomados pela ciência primária. Nesse caso, as relações de dedução entre ciência primária e secundária utilizariam um vocabulário homogêneo e por isso também são chamadas de reduções homogêneas (cf. Nagel, 2006, p. 447). Alguns

⁴³ Leis experimentais são entendidas aqui como regularidades observáveis no domínio de uma ciência, enquanto teorias, na acepção de Nagel, envolvem leis científicas sobre aspectos inobserváveis (cf. Nagel, 2006, p. 117).

exemplos, dados pelo próprio Nagel, seriam a redução da mecânica clássica de partículas aplicada a massas pontuais para uma mecânica de corpos extensos aplicada a objetos rígidos ou a redução da mecânica exclusivamente terrestre de Galileu para uma mecânica newtoniana de gravitação que englobe corpos celestes (cf. Nagel, 2006, pp. 446-447). Nesses dois exemplos, os termos utilizados nos sistemas científicos envolvidos na ciência secundária são os mesmos termos envolvidos na ciência primária – uma mecânica de corpos rígidos ainda utilizará os mesmos conceitos aplicados na mecânica de partículas e a gravitação de Newton ainda emprega termos também utilizados na mecânica de Galileu.

Contudo, há casos de redução em que o vocabulário da ciência secundária parece tratar de noções que não estão contempladas no vocabulário da ciência primária. Tais casos são chamados por Nagel de reduções não-homogêneas ou heterogêneas (Nagel, 2006, p. 450). Um exemplo clássico de redução heterogênea seria a redução da termodinâmica à mecânica estatística. A termodinâmica utiliza algumas noções já conhecidas da mecânica como pressão, volume e peso, mas também apresenta conceitos próprios como temperatura, calor e entropia. Essas noções aparentemente exclusivas levaram ao pensamento inicial de que a termodinâmica seria uma ciência independente que não poderia ser reduzida à pura mecânica, mas as descobertas do século XIX, como a interpretação microscópica da entropia por Ludwig Boltzmann (1804-1906), mostraram que muitos desses conceitos diferenciados poderiam receber uma conceitualização equivalente na mecânica (cf. Nagel, 2006, pp. 451-2). Um exemplo simples seria o conceito de temperatura utilizado em termodinâmica para sistemas macroscópicos que pode ser reduzido ao conceito de energia cinética média de moléculas tomadas em escala microscópica (Nagel, 2006, pp. 452-3). O desafio está em esclarecer como os vocabulários heterogêneos dessas duas áreas podem se interligar.

Se a redução for entendida como uma conexão lógica entre dois corpos teóricos, então tal conexão nunca poderá ocorrer se os termos adotados na ciência reduzida não aparecem na ciência a que se reduz – se a disciplina secundária contém termos que não aparecem nas suposições teóricas da disciplina primária, uma derivação lógica seria, *prima facie*, impossível (cf. Nagel, 2006, p. 463). Diante disso, o autor apresenta duas condições necessárias para tornar reduções do tipo heterogênea possíveis: conectabilidade e derivabilidade. Dado um termo X que apareça na disciplina secundária, mas não no corpo teórico da primária, a condição de conectabilidade estipula suposições de algum tipo que permitam ligar o significado de X com as suposições teóricas da ciência primária. Essas suposições constituem

o que aparece na literatura da filosofia da ciência como leis de ponte (*bridge laws*). Graças a essas suposições auxiliares, a derivabilidade, a segunda das condições estipuladas, faz com que todas as leis e teorias da ciência secundária, incluindo as leis e teorias que utilizem X, sejam derivadas logicamente dos pressupostos que aparecem na disciplina primária (Nagel, 2006, pp. 464-5). Segundo o autor, os vínculos postulados por essas conexões não precisam ser puramente sintáticos e podem ser caracterizados em contexto metateórico, indo além da organização formal da linguagem das disciplinas, incluindo associações de significado, convenções previamente deliberadas para traçar as ligações entre os corpos teóricos ou hipóteses que relacionem os termos teóricos com fatos empíricos comuns às duas ciências (cf. Nagel, 2006, pp. 465-6). Seja como for, as leis de ponte devem possibilitar que a ciência secundária seja consequência lógica de uma ciência primária, mostrando como derivar os termos da ciência secundária a partir da ciência primária.

Diante de tudo isso, uma visão de mundo reducionista apoiada em um modelo de redução teórica aos moldes do que é sugerido por Nagel pode ser resumida nos seguintes termos: os níveis menos fundamentais são completamente definidos por níveis mais fundamentais por conta das ciências em níveis mais básicos permitirem derivar logicamente tudo que pode ser dito nas ciências de níveis menos básicos. Desse modo, o conceito de temperatura na termodinâmica é uma consequência lógica do conceito de energia cinética média da mecânica estatística. Os comentadores de Nagel, no entanto, questionam se de fato podemos interpretar reduções como consequências lógicas. Para alguns autores, a ciência primária não permite uma dedução da ciência secundária – partir da ciência primária pode levar a uma ciência rigorosamente distinta da ciência secundária.

Lawrence Sklar, por exemplo, argumenta em um de seus artigos (Sklar, 1967) que a noção de redução entre teorias não é tão simples quanto Nagel imaginava. A história mostra que casos de redução inter-teórica como uma derivação direta são bem raros, mesmo em situações mais simples como as reduções homogêneas (Sklar, 1967, p. 110). Tome o caso da redução da mecânica de Galileu para queda dos corpos para a teoria gravitacional de Newton, um caso de redução homogênea na terminologia de Nagel. De acordo com Nagel, a mecânica de Newton permite derivar a mecânica de Galileu sem o auxílio de leis de ponte, uma vez que o vocabulário utilizado na mecânica galileana é apenas um subconjunto do vocabulário da mecânica newtoniana. Sklar, no entanto, chama a atenção para o fato de que a teoria de Newton não permite inferir exatamente a teoria de Galileu – tudo que a ciência primária pode

deduzir é uma aproximação da ciência secundária (Sklar, 1967, p. 111). A ideia de que temos uma dedução aproximada ao invés de uma derivação exata é melhor entendida se considerarmos a crítica ao modelo redutivo de Nagel feita por Paul Feyerabend (1924-1994). Em particular, considere a seguinte passagem de Feyerabend (1962):

O que acontece aqui quando uma transição é feita de uma teoria T' para uma teoria mais ampla T (a qual assumimos ser capaz de cobrir todos os fenômenos cobertos por T') é algo muito mais radical do que a incorporação de uma teoria inalterada T' (isto é, inalterada com respeito aos significados de seus principais termos descritivos assim como aos significados dos termos de sua linguagem de observação) no contexto de T. Ao invés disso, o que acontece é uma completa substituição da ontologia (e talvez mesmo do formalismo) de T' pela ontologia (e formalismo) de T e uma correspondente mudança nos significados dos elementos descritivos do formalismo de T' (dado que esses elementos e esse formalismo ainda são usados). Essa substituição afeta não apenas os termos teóricos de T' mas também ao menos alguns dos termos observacionais que ocorriam em suas sentenças de teste. Isto é, não apenas a descrição de coisas e processos no domínio ao qual T' havia sido aplicado será infiltrada ou com o formalismo e termos de T ou, se os termos de T' ainda estiverem em uso, com os significados dos termos de T, como também as sentenças expressando o que é acessível à observação direta nesse domínio agora irão significar algo totalmente diferente. Em suma: introduzir uma nova teoria envolve mudanças de perspectiva com respeito aos aspectos observáveis e inobserváveis do mundo, bem como mudanças correspondentes até mesmo nos significados dos termos mais fundamentais da linguagem empregada. (Feyerabend, 1962, pp. 28-29)

De acordo com Feyerabend, a passagem de uma teoria a outra não é inofensiva em termos de significado. O nível mais fundamental não incorpora uma disciplina de nível menos fundamental sem nenhuma alteração nos significados atribuídos aos termos utilizados na ciência secundária, pois a ontologia ou mesmo o formalismo muda completamente com a chegada da ciência primária. Assim, retomando o exemplo da redução da mecânica de Galileu à mecânica de Newton, a visão de mundo presente na teoria newtoniana se torna diferente da visão de mundo presente na teoria galileana. Isso significa que qualquer tentativa de derivar a mecânica galileana partindo do contexto da mecânica newtoniana não poderá resultar em uma teoria galileana pura – se a perspectiva da teoria nova é tomada como premissa, é impossível derivar uma teoria idêntica à teoria antiga sem carregar nenhum aspecto, por mínimo que seja, da teoria nova. Para todos os efeitos, a ciência primária não permite deduzir a ciência secundária, mas alguma versão diferenciada dessa ciência secundária. Diante disso, é necessário repensar a noção de redução proposta por Nagel.

Essa ideia de mudança de perspectiva não é muito distante da filosofia defendida por Thomas Kuhn (1922-1996), apresentada bem resumidamente a seguir. Em seu conhecido livro *A Estrutura das Revoluções Científicas* (Kuhn, 2011), publicado originalmente no início da

década de 1960, Kuhn interpreta a história da ciência como uma sequência de ciclos. A princípio, cientistas tentam resolver seus problemas dentro de um determinado paradigma, um conjunto de métodos, teorizações e valores adotados pela maioria da comunidade científica. A ciência se encontra em seu período normal enquanto os problemas podem ser resolvidos nesse paradigma. Ocorre que, em algum momento da história, surgem anomalias, problemas que não podem ser solucionados por nenhum elemento do paradigma atual. Na persistência dessas anomalias, a ciência começa a entrar em um período de crise. Ocorrerá a chamada revolução científica quando a crise for solucionada com uma mudança de paradigma. Isso pode significar uma alteração extremamente radical no modo como antigos conceitos eram apresentados – uma mudança de ponto de vista parecida com o que se tem ao percebermos uma nova figura em uma ilusão de óptica⁴⁴. Nas palavras de Kuhn:

Guiados por um novo paradigma, os cientistas adotam novos instrumentos e orientam seu olhar em novas direções. E o que é ainda mais importante: durante as revoluções, os cientistas vêem coisas novas e diferentes quando, empregando instrumentos familiares, olham para os mesmos pontos já examinados anteriormente. É como se a comunidade profissional tivesse sido subitamente transportada para um novo planeta, onde objetos familiares são vistos sob uma luz diferente e a eles se apegam objetos desconhecidos. (Kuhn, 2011, p. 147)

O raciocínio de Kuhn reforça a ideia de Feyerabend. Se a teoria da relatividade, por exemplo, for pensada como uma mudança de paradigma no contexto da física, então há um referencial totalmente novo, a partir do qual o mundo é visto de modo muito diferente quando comparado à visão newtoniana – no terreno da relatividade, o módulo da velocidade da luz deve sempre ser o mesmo para qualquer observador, medidas de intervalo de tempo para um mesmo sistema podem mudar para diferentes observadores, etc. Após uma revolução, entramos em um paradigma que associa significados diferentes para conceitos do paradigma anterior. Desse modo, comparando com o comentário de Feyerabend, a tentativa de derivar uma teoria anterior dentro do paradigma atual não poderá resultar em uma versão idêntica da teoria antiga, mas uma teoria que inevitavelmente será vislumbrada com os óculos do novo paradigma.

Uma tentativa interessante de contornar essa crítica da mudança de significados apontada por Feyerabend e Kuhn foi proposta por Kenneth Schaffner (e.g. Schaffner, 1967, 1993)⁴⁵. Resumidamente, Schaffner concorda que, em um processo de redução entre dois

44 A famosa imagem pato-lebre desenvolvida pelo psicólogo Joseph Jastrow (1863-1944) seria um exemplo típico desse tipo de ilusão.

45 Schaffner (1967) apresenta uma elaboração inicial, enquanto Schaffner (1993) apresenta uma formulação

corpos teóricos, nem sempre é possível derivar logicamente a teoria reduzida a partir da teoria redutora. Não obstante, é possível derivar logicamente uma versão corrigida da teoria reduzida com base no que a teoria redutora apresenta de novo (cf. Schaffner, 1993, p. 429). Em outras palavras, se tomamos as proposições da mecânica estatística, por exemplo, não poderemos concluir proposições correspondentes à termodinâmica original, mas podemos derivar uma termodinâmica corrigida que leva em conta fatores da mecânica estatística. Ocorre que é possível comparar a termodinâmica derivada com sua versão original por meio de uma relação de analogia: se, quando comparada com a versão original, a versão corrigida não apresentar absolutamente nada em comum com ela, então a teoria redutora não absorveu nada da teoria reduzida e o processo é melhor entendido como uma substituição de teorias. Por outro lado, caso a versão corrigida apresente algum grau de semelhança com a versão original, então é possível falar em algum processo redutivo: quanto maior a semelhança, mais a teoria redutora absorveu a teoria reduzida. A Figura 19 abaixo esquematiza o modelo de Schaffner.

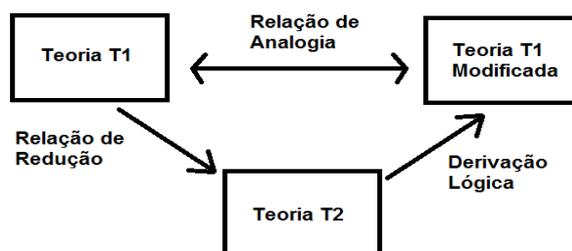


Figura 19: Diagrama do modelo de redução teórica proposto por Schaffner. Quando T1 se reduz a T2 só é possível derivar logicamente versão modificada de T1 que pode ser comparada com sua versão original por meio de analogia.

Autores posteriores influenciados pelo trabalho de Schaffner (e.g. Bickle, 1998; Churchland, 1986; Hooker, 1981), conhecidos como reducionistas *New Wave* (van Riel, 2014, p. 173), aprofundaram a análise da relação de analogia entre teoria reduzida corrigida e teoria reduzida original, aplicando essa ideia principalmente no contexto da filosofia da mente. Foge ao escopo deste trabalho entrar nos detalhes dessa análise, mas é suficiente para nossos propósitos entender que a ideia geral do modelo de Schaffner não resgata a teoria reduzida original da redutora, mas oferece um modo de avaliar o quanto da teoria reduzida foi absorvido pela redutora por meio de uma comparação entre a teoria reduzida original e o que é possível derivar a partir da redutora.

mais desenvolvida.

A vantagem do modelo de Schaffner é justamente evitar as objeções feitas a Nagel quanto à impossibilidade de uma derivação lógica da ciência secundária a partir da ciência primária. Schaffner deixa claro que a partir da ciência primária só podemos derivar logicamente uma versão modificada da ciência secundária, mas ela pode ser comparada com a versão original para avaliação das semelhanças e diferenças. A desvantagem desse modelo é a dificuldade em oferecer uma maior clarificação lógica do que seria uma relação de analogia, como o próprio Schaffner comenta (Schaffner, 1967, p. 146).

Weinberg, por sua vez, não encara as mudanças de significado como uma ameaça à sua posição reducionista. Ainda que os significados se alterem, o autor não considera que seja impossível fazer uma comparação entre teorias antes e depois das revoluções kuhnianas (Weinberg, 2001, p. 196). Ele concorda que há mudanças no modo em que pensamos no decorrer da história da ciência, mas essas mudanças são evolucionárias e não revolucionárias (Weinberg, 2001, p. 197). Nesse sentido, os novos paradigmas funcionam mais como um complemento dos antigos do que uma desconstrução. De fato, alguém que conhece a teoria da relatividade não enxergará a mecânica clássica com os mesmos olhos de alguém que não conhece o novo paradigma, mas dentro da perspectiva relativística é possível entender os pontos em que a teoria da relatividade se torna mais satisfatória do que a mecânica clássica. As revoluções mencionadas por Kuhn não mudam o esquema geral de mundo reduzido de Weinberg. O princípio básico continua o mesmo, ou seja, as áreas da ciência continuam sendo explicadas por áreas mais fundamentais, ainda que essas áreas mais básicas sejam desconhecidas. A própria ideia de teoria unificada, por exemplo, ainda é uma interrogação, mas parece razoável supor que nossas esferas de conhecimento científico atual possam ser explicadas por um campo ainda mais básico – as setas de explicação ainda podem convergir mais.

Nesse ponto, talvez seja conveniente melhorar a definição de explicação⁴⁶, indo além do que Nagel propõe como derivação lógica. Como Sklar comenta, raramente encontramos derivações lógicas diretas de uma teoria para outra (Sklar, 1967, p. 110), mas esse mesmo autor sugere uma ideia de explicação que pode ser útil aqui – a teoria redutora não explica a teoria reduzida no sentido de derivá-la diretamente a partir de seu formalismo, mas pode explicar no sentido de ajudar a compreender as razões da teoria reduzida ser bem sucedida (Sklar, 1967, p. 112). Novas descobertas no trabalho científico refinam o nosso conhecimento

46 Uma análise do conceito de explicação é um problema filosófico difícil que não convém aprofundar neste texto. Para mais detalhes sobre esse tema, ver Salmon (1989).

do mundo natural e permitem entender com mais profundidade o que de fato estava acontecendo no escopo das teorias mais antigas. Descobrir que existem partículas ainda menores constituindo os prótons ampliou nosso conhecimento de física atômica com o auxílio de uma novidade presente em uma teoria mais básica, a física nuclear. Essa noção de explicação parece compatível com a visão reducionista de Weinberg apresentada acima.

Concluimos assim que um reducionismo significativo aos moldes de Weinberg admite que as teorias dos níveis mais elementares estão vinculadas com os níveis menos elementares por conexões explicativas, no sentido de leis mais básicas explicitarem as razões de leis menos básicas funcionarem tão bem em seu papel de representar os elementos de seu domínio de atuação.

2.2.3. Reduccionismo Trivial

O segundo tipo de reducionismo considerado por Weinberg é chamado trivial (*petty*). Nessa variação, propriedades de alto nível são diretamente definidas a partir das propriedades de baixo nível. Um exemplo seria a liquidez da água, justificada pela maior liberdade de movimento das moléculas de H₂O – propriedades microscópicas (moléculas mais livres) definem um determinado aspecto macroscópico (água em estado líquido). Aqui, não consideramos relações entre teorias a respeito do domínio microscópico e macroscópico, mas apenas uma constatação a partir da constituição dos elementos envolvidos – o fato de propriedades das partes definirem totalmente alguma propriedade do todo. Weinberg considera esse tipo de reducionismo bem mais desinteressante quando comparado com o reducionismo significativo pelo fato de não ser tão geral, já que nem sempre as coisas podem ser explicadas apenas com base em seus constituintes – por exemplo, quando Einstein explicou as leis de Newton em termos da teoria da relatividade, ele não fez isso pensando nos componentes de alguma coisa, mas sim buscando princípios mais gerais que explicassem os princípios menos gerais, independente da constituição dos corpos envolvidos (Weinberg, 2001, p. 111). Não obstante, esse tipo de reducionismo merece atenção quando levamos em conta certas questões metafísicas – entender as relações entre parte e todo é uma preocupação ontológica, afinal. Com efeito, em uma relação redutiva, propriedades em níveis mais fundamentais definem totalmente propriedades menos fundamentais e qualquer mudança nas propriedades menos básicas significa que necessariamente houve alguma mudança nas

propriedades mais básicas. Esse tipo de relação é chamado superveniência e podemos aprofundar o conceito de reducionismo trivial por meio de um estudo dessa noção.

O conceito de superveniência assume um sentido específico em contexto filosófico. Na língua inglesa, o termo “*supervenies*” pode ser empregado no mesmo sentido que, em português, utilizamos a palavra “sobrevir”. Assim, a frase “*Upon a sudden supervened the death of the king*” (que, em uma tradução livre, significaria algo como “Repentinamente sobreveio a morte do rei”) utiliza a ideia de superveniência de um modo diferente do que aparece nos textos filosóficos (Kim, 1990, pp. 1-2). A origem desse termo no sentido filosófico contemporâneo não é muito clara⁴⁷, mas é seguro afirmar sua aparição no contexto filosofia moral do século passado, mais precisamente na obra de Richard M. Hare (1919-2002) (cf. Kim, 1990, p. 3). Segundo Jaegwon Kim (1990, p. 6), Hare introduz o termo superveniência nas discussões éticas na seguinte passagem:

Primeiro, vamos tomar aquela característica de “bom” que tem sido chamada de sua superveniência. Suponha que dizemos, “São Francisco foi um bom homem”. É logicamente impossível dizer isso e manter ao mesmo tempo que poderia ter havido outro homem colocado exatamente nas mesmas circunstâncias de São Francisco, que se comportou exatamente da mesma forma, mas que se diferencia de São Francisco apenas com respeito ao fato de que ele não era um homem bom. (Hare, 1953, p. 145)

Em resumo, Hare está dizendo que é logicamente impossível duas pessoas compartilharem exatamente os mesmos aspectos, mas se diferenciarem unicamente no que diz respeito a ser bom (Kim, 1984, p. 155). Há aqui uma covariação necessária entre propriedades morais (como ser bom) com propriedades não-morais descritivas. Covariação de propriedades é a intuição geral por trás da ideia de superveniência: propriedades de um tipo devem covariar com propriedades de outro tipo em algum sentido – uma mudança nas propriedades supervenientes não podem ocorrer sem uma mudança em propriedades de base (Kim, 1990, p. 7).

Embora na citação anterior Hare tenha pensado na ideia de superveniência como uma característica, Kim prefere pensar no conceito como uma relação – afinal, o que a citação de Hare mostra é que há superveniência de predicados éticos em relação a predicados não-éticos (Kim, 1990, p. 7). Podemos, então, definir superveniência como uma relação tal que se A é

47 Kim (1990, p. 5) conta que uma das primeiras aparições da palavra “*supervenies*” em um texto filosófico ocorreu em uma passagem da obra de Gottfried Leibniz (1646-1716) tratando sobre relações, mas não encontrou mais nenhum outro uso da palavra até o vigésimo século.

superveniente com respeito a B, então não é possível alterar A sem alterar B⁴⁸. Essa relação pode ocorrer com os mais diversos elementos, como sentenças, eventos, proposições, etc (Kim, 1984, p. 155). Contudo, considerando a noção de reducionismo trivial trabalhada aqui, focaremos na superveniência entre propriedades. Definiremos assim um conjunto A de propriedades supervenientes e um conjunto B tomado como base de superveniência, de modo que uma propriedade pertencente a A é superveniente a uma propriedade pertencente a B.

Um exemplo típico de superveniência pode ser pensado no âmbito da filosofia da mente. Pelo senso comum, propriedades mentais (digamos, a propriedade de possuir um estado mental de dor ou prazer) dependem diretamente de propriedades materiais (no caso, as interações ocorridas no cérebro), de modo que alterar um estado mental (digamos, passar de uma sensação de dor para uma sensação de prazer) só é possível com a alteração de um estado material (modificar o arranjo cerebral responsável pela sensação de dor para um arranjo que produza a sensação de prazer). Sendo assim, é possível imaginar uma relação de superveniência entre mente e cérebro. Propriedades mentais seriam propriedades supervenientes e propriedades cerebrais seriam propriedades da base de superveniência.

Mas se uma superveniência de A em relação a B significa que não é possível termos modificações em A sem modificações em B, então é preciso esclarecer exatamente o sentido de “não é possível”. Algo pode não ser possível em virtude de ser logicamente impossível ou pode ser impossível por conta de restrições próprias de nosso universo, ainda que pudesse haver possibilidade caso essas restrições fossem diferentes. Nessa linha de raciocínio, Kim (1984) diferencia uma noção fraca e uma noção forte de superveniência.

Começemos pela superveniência fraca. Considere dois conjuntos A e B de propriedades, sendo A um conjunto de propriedades supervenientes e B um conjunto de propriedades de base. Considere também que essas propriedades possam ser atribuídas a elementos x e y. Haverá superveniência fraca do conjunto A em relação ao conjunto B se, para quaisquer elementos x e y, se x e y compartilharem exatamente as mesmas propriedades em B, então compartilharão exatamente as mesmas propriedades em A (cf. Kim, 1984, p. 158).

Um exemplo utilizado por Kim seria a superveniência da propriedade de ser bom com respeito a propriedades de ser corajoso, honesto e benevolente⁴⁹ (Kim, 1984, p. 159). Seja o

48 Nada impede que se A for superveniente a B, B também seja superveniente a A. Podemos alterar a área de uma esfera sem alterar seu volume, mas também não podemos alterar o volume sem alterar a área. Simetria, portanto, não é necessária. A relação também é reflexiva (A é superveniente a A) e transitiva (se A é superveniente a B e B é superveniente a C, então A é superveniente a C). Não demonstraremos essas propriedades aqui, mas uma discussão mais completa pode ser encontrada em McLaughlin & Bennett (2014).

49 Certamente deve haver mais condições que determinem a bondade de um indivíduo, mas, para fins didáticos,

conjunto de propriedades supervenientes A formado por duas propriedades: “ser bom” (G) e “não ser bom” (\neg G). Seja o conjunto de base B formado por conjunções das três propriedades corajoso (C), honesto (H) e benevolente (V)⁵⁰. Em outras palavras, um indivíduo predicado de acordo com as propriedades de B pode apresentar essas três virtudes, nenhuma delas ou apenas alguma delas, havendo um total de oito conjunções possíveis entre as virtudes: $C\&H\&V$, $C\&H\&\neg V$, $C\&\neg H\&V$, $C\&\neg H\&\neg V$, $\neg C\&H\&V$, $\neg C\&H\&\neg V$, $\neg C\&\neg H\&V$, $\neg C\&\neg H\&\neg V$ ⁵¹. A partir disso, definimos que indivíduos considerados bons e indivíduos não considerados bons podem ser examinados de acordo com sua predicação com respeito a coragem, honestidade e benevolência. Assim, haverá superveniência fraca diante da impossibilidade de dois indivíduos compartilharem exatamente a mesma combinação de virtudes das combinações possíveis em B e, ainda assim, apresentarem propriedades distintas em A. Se A é superveniente a B e dois indivíduos, digamos, Fabiano e Renato, forem ambos corajosos, honestos e benevolentes, então necessariamente Fabiano e Renato terão a mesma predicação em A: ou tanto Renato e Fabiano são classificados como G ou ambos são classificados como \neg G.

Essa noção de superveniência, porém, carece de força modal. Tome mais uma vez o exemplo acima, relacionando um conjunto A de propriedades supervenientes referentes à bondade e um conjunto B de conjunções de virtudes. Tudo que a relação de superveniência fraca nos diz é que se dois indivíduos compartilharem as mesmas propriedades em B necessariamente compartilharão as mesmas propriedades em A. Porém, a superveniência fraca não nos diz quais propriedades em B devem corresponder às propriedades em A. Em nosso mundo não parece haver muita discordância em admitir que uma pessoa corajosa, honesta e benevolente seja boa. Assim, não teríamos muitos problemas em admitir uma superveniência de G com respeito a $C\&H\&V$ – se Fabiano e Renato forem corajosos, honestos e benevolentes em nosso mundo, então ambos devem ser bons em nosso mundo. O problema é que a superveniência fraca torna a superveniência de G com relação a $C\&H\&V$ apenas contingente. No contexto da superveniência fraca seria admissível um mundo em que uma pessoa corajosa, honesta e benevolente não fosse boa ou um mundo em que uma pessoa que não seja nem corajosa, nem honesta e nem benevolente fosse boa (Kim, 1984, p. 159). Se certas relações

bondade será avaliada em termos dessas três características.

50 Utilizando o sinal \neg para negação teremos que $\neg C$ significa não-corajoso, $\neg V$ significa não-benevolente e $\neg H$ significa não-honesto. Além disso, utilizamos o sinal $\&$ para expressar a operação lógica de conjunção.

51 Novamente, por razões de simplicidade do argumento, não estamos admitindo neutralidade. Aqui, ou o indivíduo possui a virtude ou não possui.

entre propriedades são necessárias, válidas para qualquer mundo possível, então precisamos de uma definição mais forte de superveniência.

Tomando mais uma vez um conjunto de propriedades de superveniência A e um conjunto de propriedades de base B, haverá superveniência forte de A com relação a B quando, necessariamente, para cada x e cada propriedade F em A, se x possuir F, então haverá uma propriedade G em B tal que x possui G e, necessariamente, se qualquer y possuir G, ele também possuirá F (Kim, 1984, p. 165).

Recorrendo ao exemplo da superveniência da bondade com respeito a combinações de virtudes (coragem, honestidade e benevolência), podemos entender a superveniência forte da seguinte maneira: necessariamente, se Fabiano é considerado bom por ser corajoso, honesto e benevolente, então, necessariamente, qualquer indivíduo (como Renato) que seja corajoso, honesto e benevolente também será bom. Em outras palavras, necessariamente, ao se admitir que a bondade de Fabiano é superveniente com relação à sua coragem, honestidade e benevolência, então é necessário que ela seja superveniente a essas virtudes e não a outras.

Kim também apresenta a superveniência forte em termos de superveniência global (Kim, 1984, pp. 167-168)⁵². Nesse caso, ele utiliza a linguagem de mundo possíveis⁵³. Dados um conjunto de propriedades de superveniência A e um conjunto de propriedades de base B, A está em uma relação de superveniência global com respeito a B se, para um mundo m_1 indistinguível de um mundo m_2 com respeito às propriedades de B, então m_1 também será indistinguível de m_2 com respeito às propriedades de A. Por exemplo, se propriedades mentais são supervenientes a propriedades físicas, então um mundo fisicamente idêntico ao nosso terá exatamente as mesmas propriedades mentais. Considere que, em um mundo m_1 , Eduardo ame sua colega de classe Mônica, ou seja, assume um estado mental de afeto com relação à garota. Havendo superveniência global do físico ao mental não seria possível conceber um mundo m_2 exatamente como o mundo m_1 em todos os seus aspectos físicos, com um Eduardo e uma Mônica fisicamente idênticos às suas contrapartes de m_1 , e, ainda assim, o Eduardo de m_2 não amar a Mônica de m_2 . De modo semelhante, podemos lembrar do exemplo da bondade. Haverá superveniência global de bondade (ou ausência de bondade) com as combinações de virtudes (no exemplo, coragem, honestidade e benevolência) se, dados dois mundos

52 Superveniência forte é equivalente a superveniência global. Isso está demonstrado em Kim (1984, p. 168).

53 Vivemos em uma realidade que constitui nosso mundo atual. Um mundo possível pode ser entendido como o mundo que nossa realidade poderia ser. Por exemplo, em nosso mundo atual, o piloto brasileiro Ayrton Senna morreu em 1994, mas podemos conceber um mundo possível em que a morte de Senna não tenha ocorrido nesse ano.

indistinguíveis com respeito ao modo como as combinações de virtudes são distribuídas entre seus habitantes, esses dois mundos também sejam indistinguíveis com respeito à distribuição da bondade.

Note que o conceito de superveniência indica apenas que certas propriedades covariam, não implicando que uma propriedade A se reduza a uma propriedade B. Em uma situação de superveniência de A com relação a B, nem sempre A será completamente determinado por B, ainda que mudar A implique mudar B. A noção de emergência (discutida com mais detalhes nas seções subsequentes) trabalha essa ideia: uma propriedade emergente depende das propriedades de nível inferior, mas não pode ser reduzida a elas – pelo menos não em uma interpretação totalmente metafísica de emergência.

Podemos, contudo, clarificar o reducionismo trivial de Weinberg em termos de superveniência. Propriedades de alto nível (supervenientes) são determinadas por propriedades de baixo nível (propriedades de base), de modo que mudanças nas propriedades de alto nível impliquem em mudanças de baixo nível – se a água deixa de ser líquida, então necessariamente ocorreu algo em suas moléculas. Assim, em um reducionismo trivial, o todo é completamente determinado por seus componentes, tornando impossível mudar o todo sem mudar as partes, o que pode ser explicado por meio do conceito de superveniência.

2.3. Sondas Epistemológicas

Até agora a noção de reducionismo adotada forneceu uma descrição de nossa realidade natural. Em linhas gerais, o mundo reducionista é dividido em níveis, de tal modo que há níveis mais fundamentais do que outros. Nessa perspectiva, conhecer todas as leis do nível mais fundamental permite compreender todas as leis do nível menos fundamental. Em adição a isso, propriedades de nível mais fundamental definem todas propriedades de nível menos fundamental e, portanto, o conhecimento de propriedades fundamentais permitiria o conhecimento de propriedades menos fundamentais. Assim, se conheço todos os princípios por trás da física de partículas elementares, poderia conhecer todos os princípios da física atômica e da química, uma vez que propriedades químicas são totalmente definidas com base em propriedades de física elementar. Semelhantemente, se conheço todos os princípios por trás da neurologia, poderia conhecer todos os princípios por trás da psicologia, uma vez que comportamentos são totalmente definidos por interações neurais. Apesar disso, ainda não foi

discutido de que maneira seria possível conhecer a determinação dos níveis menos básicos em termos dos níveis mais básicos.

Alguns autores (e.g. Fodor, 1974) defendem que o fato de o mundo ser fundamentalmente físico não implica que todas as leis de disciplinas mais específicas devam ser apresentadas em termos de leis físicas – afinal, o vocabulário da física de partículas é tão distante do vocabulário da economia, por exemplo, que expressar uma lei da economia em termos de uma lei da física nuclear seria inviável para qualquer padrão humano (cf. Fodor, 1974, pp. 102-4). De fato, como já comentado anteriormente, Weinberg diferencia entre adotar o reducionismo como visão de mundo e o reducionismo como método (Weinberg, 2001, p. 112). O fato de o mundo natural apresentar um nível básico não significa que precisamos lançar fora todas as disciplinas de níveis menos básicos e buscar um único vocabulário pautado nas ciências básicas, como queriam os positivistas lógicos. Parece evidente que seres humanos possuem uma capacidade muito limitada para descrever uma transação econômica em termos de interações entre partículas elementares, tornando impraticável qualquer teoria econômica apresentada no vocabulário da física nuclear. No entanto, podemos questionar se essa descrição é possível em princípio. Ainda que humanos como nós não consigam fazer isso, um ser sem as mesmas limitações de conhecimento poderia fazê-lo? Se sim, isso mostraria que o reducionismo também é defensável em termos epistemológicos, ainda que esse conhecimento não seja alcançável com nossa capacidade atual. Um modo de analisar essa questão seria a aplicação de recursos abstratos que podem ser chamados de sondas epistemológicas (Pessoa, 2005, 2011).

2.3.1. Ontologia Geral

Antes de definirmos essas sondas epistemológicas é importante esclarecer onde exatamente elas estão atuando. Basicamente, queremos descrever nossa realidade de modo sistemático ou científico. No entanto, uma descrição científica da realidade pode ser feita em uma linha realista ou anti-realista. Como já mencionado, em uma vertente realista, nossas representações fazem referência a um mundo com uma existência independente de seres humanos, enquanto uma vertente anti-realista não assume qualquer compromisso com a existência dessa realidade objetiva. Desse modo, se adotamos um posicionamento realista, então nossas representações podem ou não tratar de uma realidade objetiva. Nesse ponto, é

conveniente fazer uma distinção entre os termos “ôntico” e “ontológico”, inspirados na divisão feita por Heidegger (1971, Nota 14, p. 106). Separamos o termo “ôntico” para tratar de elementos da realidade objetiva, independentes da existência humana, a “coisa-em-si” kantiana, enquanto o termo “ontológico” é empregado para cuidar do que nossas representações e teorias retratam. Assim, haverá uma descrição fiel da realidade se o elemento ontológico de uma representação também corresponder a um elemento ôntico.

O problema é que não temos garantias de que nossas teorias realmente tratam de elementos ônticos. A prova disso é que várias teorias antes pensadas como verdadeiras, como a teoria do flogisto, foram abandonadas no decorrer da história. Diante disso, parece impossível falar do mundo ôntico de modo definitivo⁵⁴. Não obstante, é possível admitir uma ontologia geral (cf. Pessoa, 2005, p. 182), uma visão de realidade que pode não ser confirmada como uma descrição do aspecto ôntico, mas é consistente com ontologias específicas admitidas pela ciência. Assim, poderíamos substituir a dimensão ôntica por uma dimensão ontológica compatível com o conhecimento científico que possuímos até agora. Espaço e tempo, por exemplo, são tomados como princípios básicos na física, ainda que não possamos demonstrar se são elementos que existem objetivamente ou, aos moldes de Kant, são apenas intuições básicas de sensibilidade presentes apenas no nível subjetivo.

Mas nem tudo em uma ontologia geral é fácil de definir – as teorias científicas discordam em diversos pontos, como quais tipos de entidades básicas existem fisicamente. Na teoria quântica, por exemplo, é defensável que ondas sejam as entidades mais básicas, outros acreditam que sejam partículas e dualistas podem argumentar que as entidades mais fundamentais sejam tanto ondas quanto partículas. Fora isso, também seria possível defender que existem elementos ainda mais básicos do que nossas partículas elementares e uma ontologia geral não deveria considerar apenas as partículas catalogadas pelo Modelo Padrão. Desse modo, a decisão mais sensata seria procurar uma ontologia que deixasse de lado essas questões controversas⁵⁵ (Pessoa, 2005, pp. 182-183).

Para explorar os conceitos deste capítulo, tomaremos como ontologia geral apenas a existência de entidades distribuídas no espaço e no tempo, sem especificar a natureza de tais entidades⁵⁶. Além disso, essa ontologia será pensada em termos de escalas, consideradas como

54 Ainda que nossas teorias correspondam à realidade objetiva, tudo que falarmos sobre essa realidade ainda estará no nível ontológico, uma vez que não conseguimos exprimir afirmações a respeito desse domínio objetivo sem uma linguagem.

55 No próximo capítulo, porém, tomaremos uma ontologia geral bastante controversa.

56 Essa ontologia parece o caso mais geral, mas ainda assim pode ser controversa, já que ideias como o realismo estrutural ôntico consideram que nem mesmo entidades existem, apenas estruturas. Ver Ladyman et. al.

dimensões adicionais da física (Pessoa, 2005, pp. 183-184). Dimensões podem ser entendidas como especificações necessárias para fixar a localização de um objeto físico com respeito a um determinado ponto de referência. Considerando a teoria da relatividade e desconsiderando teses mais controversas como a teoria das cordas, há quatro dimensões necessárias para isso: três dimensões espaciais e uma dimensão temporal. Para nossos fins, contudo, consideraremos também que há dimensões de escala. Com respeito ao espaço, a escala especifica os níveis trabalhados com base no tamanho, dividindo o mundo em níveis microscópicos e macroscópicos. Com respeito ao tempo, a escala determina os níveis de mudança: rápida ou lenta. Essa preocupação com escalas é importante, pois mesmo que algo seja bem definido no espaço e no tempo, resta ainda a dúvida com respeito à informação sobre qual escala deve ser considerada – afinal, há uma diferença considerável entre o tamanho de um país (centenas de quilômetros), o tamanho de uma casa (dezenas de metros) e o tamanho de uma bactéria (frações de milímetros), bem como uma diferença apreciável entre os picossegundos necessários para uma proteína se dobrar, os segundos que uma pessoa demora para se deslocar de um cômodo da casa para outro e as centenas de anos que o plástico demora para se decompor⁵⁷ (Pessoa, 2011, p. 1).

Feitas essas considerações, podemos definir como sondas epistemológicas podem ser aplicadas em uma ontologia geral como definida acima. Grosso modo, uma sonda epistemológica é um elemento abstrato capaz de conhecer a ontologia em questão e determinar descrições verdadeiras a respeito dela. Obviamente, isso exigirá um maior esclarecimento do que se pretende dizer com “determinar”.

2.3.2. Determinação e Separabilidade com Sondas Epistemológicas

Para compreender a noção de determinação seria conveniente apresentar um exemplo de sonda epistemológica. Talvez a mais conhecida delas seja o “demônio de Laplace”⁵⁸, um ser dotado dos seguintes poderes (cf. Pessoa, 2005, p. 184; 2011, p. 3): (i) onisciência instantânea, significando que pode conhecer o estado de todo o universo em um único instante de tempo, com resolução e precisão absoluta em todas as escalas espaciais; (ii) erudição

(2007) para mais detalhes.

57 Uma questão em aberto que vale comentar é se a escala temporal ou a escala espacial podem ser derivadas a partir de outras.

58 O nome “demônio” é interpretado aqui a partir do termo grego *daimon*, como um ser dotado de poderes sobrehumanos. Caso a palavra ainda soe negativa, podemos imaginar um anjo ao invés de um demônio, como foi feito em Floridi (2009).

nomológica, ou seja, é capaz de conhecer com exatidão leis de evolução temporal do universo;(iii) supercomputação, sendo capaz de executar qualquer cálculo, por mais complexo que seja, em um tempo desprezível e (iv) não-perturbação, ou seja, nada que o demônio fizer causará qualquer influência no universo em análise. Com esses atributos, o demônio de Laplace seria uma sonda capaz de informar qual será o estado do universo em um instante futuro a partir dos dados de um instante presente. No entanto, para isso acontecer é necessário que o universo seja estritamente determinista. Isso significa que, dados dois instantes t_1 e t_2 , tal que t_2 seja posterior a t_1 e cada instante envolva todas as escalas espaciais do universo⁵⁹, o estado do universo em t_2 seria completamente fixado pelo estado do universo em t_1 . Se o universo atender a tais condições, então o demônio de Laplace é capaz de fornecer o estado em t_2 e há uma determinação temporal máxima. Caso o universo não atenda essas condições, o instante anterior não determina o posterior e o universo é indeterminista ou tiquista (cf. Pessoa, 2011, p. 2). Isso exemplifica um caso de determinação ocorrendo no tempo, mas também podemos pensar em casos de determinação no espaço, mais precisamente, uma determinação escalar espacial (Pessoa, 2011, p. 4).

Uma determinação escalar no espaço pode ser entendida quando tomamos dois cortes S_1 e S_2 de nosso universo em escalas diferentes, sendo que ambos capturam todos os aspectos no espaço e no tempo. Considere, por exemplo, que S_1 abarca todo o espaço-tempo na escala microscópica (digamos, o nível das partículas elementares), enquanto S_2 acolhe o espaço-tempo em escala macroscópica. Nesse caso, haverá uma determinação escalar espacial se S_1 determinar univocamente S_2 . Em outras palavras, todo o estado global macroscópico está fixado pelo estado global microscópico.

Para ilustrar a ideia de determinação escalar espacial é possível definir uma outra sonda epistemológica chamada demônio escalar (cf. Pessoa, 2005, p. 185). Seus poderes são semelhantes ao demônio de Laplace, com a diferença de que ele restringe seu olhar apenas para uma escala espacial determinada (o nível das partículas elementares, por exemplo), mas é capaz de abarcar o universo como um todo na dimensão temporal. Em linhas gerais, um demônio escalar é dotado dos seguintes poderes: (i) onisciência temporal instantânea, significando que pode conhecer toda a distribuição temporal do universo em uma dada escala espacial⁶⁰; (ii) erudição nomológica com respeito às leis de escala, ou seja, é capaz de conhecer com exatidão como elementos de escalas maiores são definidos em termos de

59 Isso não é tão simples considerando a teoria da relatividade, mas ainda seria possível (cf. Pessoa, 2011, p. 2).

60 Equivalentemente, conhece o estado do universo em todos os instantes de tempo de uma única vez.

elementos de escalas menores; (iii) supercomputação, sendo capaz de executar qualquer cálculo, por mais complexo que seja, em um tempo desprezível e (iv) não-perturbação, ou seja, nada que o demônio fizer causará qualquer influência no universo em análise. Em suma, essa sonda seria capaz de capturar toda a informação contida na escala microscópica e, conhecendo as leis que permitem transitar de uma escala para outra (como as leis de ponte nagelianas), poderia descrever tudo que ocorre na escala macroscópica. Se isso ocorrer, dizemos que há determinação escalar espacial. Havendo esse tipo de determinação em nosso mundo, a informação e as leis dos níveis mais fundamentais seriam suficientes para o demônio escalar descrever o nível menos fundamental.

Uma outra aplicação interessante de sondas epistemológicas em contexto espacial envolve relações de composição. Considere duas ou mais regiões espaciais onde alguma propriedade é bem definida. Dizemos que há separabilidade espacial no universo se sempre for o caso de que a consideração em separado de cada uma dessas regiões é suficiente para definir completamente o conjunto dela (cf. Pessoa, 2011, p. 5). Isso ocorre, por exemplo, em um universo que segue a teoria gravitacional newtoniana: é possível, em princípio, analisar separadamente as propriedades de cada corpo celeste (como posição, massa e velocidade) para depois compô-las e obter uma descrição gravitacional de todo o sistema. Ainda que isso seja muito difícil na prática, é teoricamente possível se recorreremos a uma sonda epistemológica como o demônio de Laplace (Pessoa, 2011, p. 6).

Vale ressaltar que o uso de sondas epistemológicas não é o único recurso possível para avaliar as determinações temporais e espaciais que ocorrem em uma dada ontologia. Também podemos utilizar o chamado “método das cópias” (Pessoa, 2005, p. 187; 2011, pp. 3-4). Seguindo esse método, ao invés de imaginarmos um ser abstrato, fora do espaço-tempo e com poderes supercomputacionais, imaginamos uma réplica exata de nosso universo em suas menores escalas espaciais e temporais.

Dados esses dois universos, haverá determinação temporal se os dois universos apresentarem uma evolução temporal independente, mas apresentarem estados indiscerníveis em sequência durante a evolução temporal – se o universo A iniciar no estado S_1 e passar para o estado S_2 em seguida, o universo B iniciará em um estado indiscernível de S_1 e entrará em um estado indiscernível de S_2 no instante seguinte.

Para que haja determinação escalar espacial, o universo A deve apresentar um estado S_1 em uma escala espacial microscópica que abarque toda a dimensão temporal enquanto sua

cópia, o universo B, deve apresentar um estado S_2 indiscernível de S_1 , de tal modo que se A apresentar um estado S_3 em escala espacial macroscópica, B necessariamente deverá apresentar um estado S_4 indistinguível de S_3 . De modo análogo, a separabilidade espacial também pode ser avaliada por meio de cópias: se um universo A for separado em duas ou mais partes no espaço e apenas essas partes forem copiadas em um universo B, então B deve necessariamente apresentar o mesmo todo de A.

Vale notar que isso se aproxima bastante da superveniência global discutida na Seção 2.2.3, onde dois mundos possíveis indiscerníveis nas propriedades de base também serão indiscerníveis em suas propriedades supervenientes. Aplicando o conceito à determinação espacial escalar, se a escala espacial macroscópica é superveniente à escala espacial microscópica, então a cópia da escala microscópica de um dado universo implica na cópia de sua escala macroscópica. Comparando com a noção de reducionismo trivial trabalhada anteriormente, as propriedades do todo são definidas totalmente pelas propriedades de base. Assim, não haveria razão para dois mundos indistinguíveis na base apresentarem uma diferença no todo e, portanto, qualquer mundo possível indistinto do nosso em termos microscópicos também o será em termos macroscópicos, de modo que o reducionismo é compatível com a superveniência global. Por outro lado, se o todo não for totalmente determinado pelas partes, talvez existam outros fatores que tornem dois mundos microscopicamente indistinguíveis macroscopicamente diferentes.

Tanto sondas epistemológicas como métodos de cópias são úteis para discutir a questão do reducionismo na realidade natural. Enquanto o método de cópias segue um caminho mais ontológico, as sondas epistemológicas (como o próprio nome sugere) apresenta uma justificativa epistemológica para argumentar que relações entre níveis em um mundo natural concebido sob a perspectiva reducionista podem ser conhecidas em princípio, embora, vale ressaltar mais uma vez, dificilmente seres humanos teriam capacidade de encontrar todas essas interações e ciências específicas sejam necessárias do ponto de vista prático. Voltaremos a mencionar tanto sondas epistemológicas quanto cópias do universo no próximo capítulo.

2.4. Objeções ao Reduacionismo

Assim como qualquer posição filosófica, o reducionismo também pode ser confrontado com visões contrárias. Sendo assim, antes de prosseguir seria interessante discutir

brevemente alguns argumentos contrários à posição reducionista, a saber, a realização múltipla, o tiquismo, o emergentismo e o holismo.

2.4.1. Realização Múltipla

Uma primeira objeção possível ao reducionismo é a ideia de realização múltipla, ou seja, o conceito de que um único aspecto definido em alto nível pode ser determinado por diferentes tipos de realizadores em nível mais baixo. Música, por exemplo, seria multiplamente realizável, já que uma mesma melodia pode ser tocada em diferentes instrumentos musicais.

Na filosofia, discussões sobre realização múltipla são recorrentes na filosofia da mente (e.g. Putnam, 1967): uma propriedade mental considerada na psicologia (como dor ou prazer) pode ser exercida por diferentes tipos de sistema nervoso, já que humanos, polvos, cães e talvez marcianos possuem cérebros muito diferentes entre si.

Já Fodor (1974) considera a realização múltipla no contexto da ciência. Eventos em áreas científicas de alto nível podem ser definidos pelas mais diversas situações físicas em baixo nível – uma transação financeira pode ser feita por dinheiro ou por cartão de crédito, ambas as situações indicam um mesmo evento em contexto econômico, mas envolvem aspectos físicos nada parecidos.

Em linhas gerais, a realização múltipla é utilizada como um argumento contrário à ideia de que um dado tipo de elemento é totalmente definido em termos de outro tipo de elemento. Aqui é importante traçar uma distinção entre tipo (*type*) e instância (*token*). Considere a sequência de números (1, 1, 2). Essa sequência apresenta dois tipos de número (1 e 2), mas um mesmo número é instanciado duas vezes, logo temos duas instâncias do número 1. Um outro exemplo seriam peças em um jogo de xadrez. A coleção de peças brancas no xadrez contém apenas um tipo de peça denominada peão, mas esse tipo de peça aparece em oito instâncias.

Se um único tipo definido em alguma descrição de alto nível pode ser realizado de diferentes maneiras por tipos definidos em uma descrição de baixo nível, então não podemos dizer que um tipo de alto nível é totalmente definido por um outro tipo de baixo nível. Considere novamente o exemplo do jogo de xadrez. Uma peça do tipo peão pode ser construída por diferentes materiais: madeira, marfim, etc. Se esses materiais são tipos

diferentes e cada tipo pode realizar o tipo “peão do jogo de xadrez”, então não podemos dizer que o tipo peão é completamente determinado por algum tipo específico de material e, portanto, não podemos dizer que o tipo da peça de xadrez se reduza ao tipo de material.

Semelhantemente, no caso da relação entre aspectos mentais e materiais, não podemos dizer que um mesmo tipo mental (como dor) é totalmente realizado por um único tipo material, já que cérebros diferentes entre si podem realizar um mesmo tipo mental. Se um evento definido em uma ciência de alto nível (como uma transação econômica) pode ser realizado por diferentes eventos em ciências de baixo nível (como diferentes eventos físicos), então um evento descrito pela ciência de alto nível não pode ser definido por um único tipo de evento em nível inferior. Diante disso, uma visão reducionista que entenda um tipo de alto nível como totalmente determinado por um tipo de baixo nível não estaria correta.

No entanto, embora a realização múltipla ataque uma redução geral entre tipos, ainda é possível defender uma relação redutiva entre instâncias. Nesse sentido, cada instância de alto nível poderia ser completamente definida por uma instância de baixo nível. Assim, cada instanciação do tipo dor seria totalmente determinada por uma instância específica de tipo material (uma instância de dor para cérebros humanos, outra instância de dor para cérebros caninos, outra para cérebros marcianos e assim por diante). Semelhantemente, cada instanciação de uma transação econômica pode ser definida completamente por um evento físico (uma transação com dinheiro é totalmente definida por um determinado evento físico, uma transação com cartão é definida por completo por meio de outro evento de baixo nível e assim por diante).

Seguindo esse raciocínio é possível pensar que relações de redução entre ciências de diferentes níveis ainda se dão, mas elas devem ser definidas localmente e não de maneira generalizada. Como diz Jaegwon Kim:

Se a mesma teoria psicológica é verdadeira para humanos, répteis e marcianos, espécies psicológicas postuladas por essa teoria devem ter realizações em fisiologias humanas, reptilianas e marcianas. Isso implica que a teoria é localmente redutível de três maneiras, para humanos, répteis, e marcianos. Se a dependência do mental no físico significa alguma coisa, deve significar que as regularidades postuladas por essa psicologia comum deve ter diferentes explicações físicas para as três espécies. Toda a ideia de realização física envolve a possibilidade de explicar fisicamente propriedades e regularidades psicológicas, e a suposição de multiplicar essas realizações, chamada múltipla realização, envolve um compromisso com a possibilidade de múltiplas reduções explicatórias da psicologia. (Kim, 1992, p. 20)

Em resumo, o fato de uma única ciência de alto nível ser aplicável a diferentes casos

ainda não elimina uma concepção reducionista, pois esses diferentes casos de aplicação da ciência de alto nível podem ser interpretados como diferentes casos de redução – ocorre apenas que as relações de redução são locais e não gerais. Nesse sentido, dizer que o nível dos eventos físicos define completamente o nível dos eventos psicológicos ou eventos econômicos seria apenas dizer que cada evento de alto nível pode ser definido inteiramente por algum evento de baixo nível e ainda assim seria razoável afirmar que há redução de um nível maior para um nível menor em cada um desses casos.

2.4.2. Tiquismo

A palavra tiquismo se origina de *tiché*, vocábulo grego para acaso. É um termo utilizado por Charles Sanders Peirce (1839-1914) para designar uma ideia contrária à noção de determinismo (Pessoa, 2011, p. 2)⁶¹. Seguindo essa ideia, caso nosso universo seja tiquista não seria possível haver uma determinação temporal exata (ver seção 2.3). Desse modo, mesmo que uma sonda epistemológica como o demônio de Laplace pudesse analisar o universo físico em um instante determinado seria impossível para o demônio definir com exatidão o estado do universo em um instante posterior ao estado atual.

Com o advento da mecânica quântica no século XX e as dificuldades em justificar a imprevisibilidade de estados quânticos, a tese de que o universo é fundamentalmente indeterminista começou a florescer. No entanto, por volta dos anos 50, David Bohm (1917-1992) propôs uma interpretação da mecânica quântica determinística baseada em variáveis ocultas (Bohm, 1952). Isso significa que nossas dificuldades em prever certos estados físicos não implicam que o universo seja indeterminista. Por outro lado, ainda não estamos autorizados a afirmar com certeza o caráter determinista do universo. Sendo assim, essa questão ainda está aberta. Podemos dizer, no entanto, que se o universo for determinista, então o estado do universo em um dado instante é suficiente para fixar o estado do universo em um instante posterior.

2.4.3. Emergentismo

Um conceito que quase sempre aparece em oposição à ideia de reducionismo é

61 Poderíamos também utilizar simplesmente o termo “indeterminismo”.

emergentismo. Grosso modo, emergentismo remete à noção de emergência, o surgimento de entidades ou propriedades de alto nível que não podem ser definidas em termos de níveis inferiores. O pensamento emergentista começa a ser discutido com mais profundidade a partir do século XIX com autores britânicos como John Stuart Mill (1806-1873), Alexander Bain (1818-1903), George Henry Lewes (1817-1878), Samuel Alexander (1859-1938), Lloyd Morgan (1852-1936) e Charlie Broad (1887-1971).

Podemos entender a ideia de emergentismo como uma oposição à determinação escalar discutida na seção 2.3. Se uma sonda epistemológica analisando escalas microscópicas não consegue determinar as escalas macroscópicas, mesmo de posse das leis de escalas (as regras que direcionam como elementos de nível superior são obtidos a partir do nível inferior), então aspectos macroscópicos emergem do nível microscópico e não podem ser determinados apenas pelo conhecimento do nível microscópico e as leis de escala. Mas essa ideia de emergentismo é mais radical. Há outras noções emergentistas mais flexíveis.

Jaegwon Kim, ao estudar a noção de emergência, identifica cinco princípios que podem estar envolvidos nessa ideia (cf. Kim, 1999, pp. 20-23): (i) o arranjo de entidades de níveis inferiores definem novas entidades em níveis superiores; (ii) propriedades em níveis superiores surgem das propriedades e relações definidas em um nível inferior; (iii) propriedades emergentes em níveis superiores não podem ser previstas a partir da informação sobre as propriedades definidas em nível inferior; (iv) propriedades emergentes em um nível inferior não poderiam ser explicadas em termos de propriedades de nível inferior; e (v) propriedades emergentes teriam poderes causais independentes do nível de base, ou seja, não seriam meros epifenômenos.

Os dois primeiros princípios também são compartilhados por concepções reducionistas, afinal, a ideia básica que adotamos para o reducionismo neste texto é que o nível inferior defina completamente o nível superior. Sendo assim, tanto reducionistas quanto emergentistas concordariam com o fato de entidades e propriedades em nível superior dependerem de entidades e propriedades em nível inferior. No caso do reducionismo, contudo, o nível inferior define inteiramente o nível superior, enquanto no emergentismo o nível inferior não seria suficiente para obter todos os aspectos de nível superior.

Enquanto os dois primeiros princípios são mais ontológicos, os princípios (iii) e (iv) envolvem questões epistêmicas: propriedades emergentes não podem ser previstas (terceiro princípio) e também não podem ser explicadas (quarto princípio). Isso pode ser compatível

com uma visão reducionista em âmbito puramente ôptico – a determinação do nível superior em termos do nível inferior pode ser uma realidade independente da possibilidade de conseguirmos prever ou explicar essa determinação.

O físico Philip Anderson, por exemplo, é um defensor dessa ideia. Anderson concorda que existem leis em níveis fundamentais que definem as leis em níveis menos fundamentais, mas discorda que seja possível simplesmente partir dessas leis básicas e reconstruir todo o restante do universo em maior escala – níveis superiores podem depender do que ocorre em níveis inferiores, mas apresentam propriedades autônomas em relação aos níveis mais básicos (Anderson, 1972, p. 393). Em suma, para Anderson há um reducionismo em dimensão ôptica, mas não epistêmica. O argumento dado pelo autor envolve quebras espontâneas de simetria. Um exemplo conhecido de quebra de simetria é a ilustração de uma esfera colocada bem no centro do pequeno topo de uma montanha em formato de chapéu mexicano, mas com condições tão delicadas de equilíbrio estático que é praticamente certo que a esfera rolará por um dos lados da montanha em algum momento. Enquanto a esfera não iniciar seu movimento há uma simetria perfeita em todas as direções, mas, a partir do momento que a esfera se mover e cair montanha abaixo, a simetria será quebrada. Há diversos exemplos de simetrias na natureza (como nas relações entre partículas e anti-partículas) e algumas podem ser quebradas espontaneamente, como a paridade da molécula de sacarose. Segundo Anderson, não é possível derivar o comportamento coletivo dos átomos de uma molécula de sacarose partindo apenas de propriedades individuais de átomos de hidrogênio, carbono e oxigênio, já que uma quebra espontânea de simetria tornaria impossível prever se a molécula assumiria uma posição dextrógira ou levógira apenas com essa informação (Anderson, 1972, p. 394).

Lembramos, contudo, que o fato de não conseguirmos prever ou explicar algo a partir dos níveis inferiores não significa que seja impossível prever o que ocorre nas escalas mais elevadas em princípio. Um demônio escalar, como discutido na Seção 2.3, poderia derivar as propriedades superiores conhecendo as propriedades fundamentais e as leis que ligam um nível a outro. A imprevisibilidade humana não implica uma imprevisibilidade em princípio. Enquanto os níveis inferiores definem completamente os níveis superiores, de modo que essas determinações possam ser conhecidas em princípio, ainda há compatibilidade com uma visão reducionista mesmo no cenário epistêmico – ainda que não possamos prever o arranjo da molécula de sacarose, um demônio escalar poderia fazê-lo em princípio.

Conciliar reducionismo e emergentismo parece mais difícil quando consideramos o

quinto princípio apresentado por Kim, a saber, poderes causais atribuídos às propriedades emergentes. Supondo que a mente seja emergente ao cérebro, então aspectos mentais seriam capazes de exercer uma influência causal na base neural. Nesse sentido, minha decisão de ir à cozinha comer um pedaço de chocolate seria causado totalmente por um elemento mental que influenciou tal decisão. Na literatura filosófica, poderes causais exclusivos dos níveis superiores são incluídos em uma noção chamada *causação descendente* (*downward causation*).

Diante desse problema, Bedau (2003) faz uma distinção entre emergência nominal, emergência fraca e emergência forte. Emergência nominal seria simplesmente a ideia de uma propriedade macroscópica que não pode ser uma propriedade microscópica - a liquidez da água não é uma propriedade das moléculas de H₂O, por exemplo. Emergência fraca seria a noção de que propriedades de alto nível podem ser derivadas a partir de um conhecimento das interações de baixo nível, mas isso não ocorre de modo trivial, de tal modo que provavelmente apenas um demônio escalar poderia derivá-las. Já emergência forte considera que propriedades de alto nível incluem poderes causais irreduzíveis, permitindo assim *causação descendente*. Uma posição reducionista é compatível com emergência nominal (o nível superior pode ter propriedades exclusivas, mas elas seriam totalmente definidas em termos do nível inferior) e emergência fraca (em princípio, seria possível compreender como determinar as propriedades em maiores escalas conhecendo as propriedades em menores escalas, ainda que seja uma tarefa complexa), mas não seria compatível com emergência forte (se o nível superior tem poderes causais irreduzíveis ao nível inferior, então o nível inferior não pode determinar completamente tais poderes).

Se nosso mundo possui propriedades emergentes em um sentido forte ainda é uma questão em aberto. Podemos dizer apenas que se o que ocorre em um nível inferior determina totalmente o que ocorre em um nível superior, sondas epistemológicas como o demônio escalar ou um método de cópias seriam capazes de fazer previsões exatas dos níveis mais altos, o que não pode ocorrer no caso da emergência forte.

2.4.4. Holismo e Não-Separabilidade

A seção 2.3 tratou da aplicação de sondas epistemológicas na definição de propriedades de um sistema inteiro com base na análise de partes espacialmente separadas.

Um mundo em que todas as propriedades podem ser reduzidas desse modo apresenta separabilidade espacial, enquanto um mundo com propriedades em que isso não ocorre seria um mundo espacialmente holista (cf. Pessoa, 2011, p. 5). Grosso modo, holismo seria a visão de que nem todos os aspectos do todo de um sistema podem ser obtidos a partir de uma unificação de aspectos de suas partes tomados separadamente.

Um argumento a favor do holismo espacial de nosso universo é apresentado com base no que é dado pela teoria quântica, área responsável por estudar o comportamento de sistemas físicos muito pequenos, como elétrons e fótons. Não temos condições de apresentar todas as peculiaridades da teoria quântica no presente texto⁶², mas podemos expor o básico da intuição a partir de um experimento mental simples.

Considere como sistema quântico uma partícula capaz de assumir dois valores de *spin*⁶³ (*up* e *down*). De acordo com a teoria quântica, antes desse valor de *spin* ser medido, a partícula pode se encontrar em uma sobreposição de estados com respeito ao *spin* – grosso modo, há tanto uma probabilidade não-nula de a partícula ser medida com *spin up* quanto uma probabilidade da mesma partícula ser medida com *spin down*.

Tome agora um par A e B desse tipo de partícula. É possível preparar esse par de partículas em um estado emaranhado (*entangled*), ou seja, um único estado que determina a probabilidade de se obter um resultado referente às duas partículas. No caso das partículas de nosso exemplo ocorrerá o seguinte: se Alice efetuar uma medida do *spin* da partícula A com um aparelho na condição x ⁶⁴ e obter *up* como resultado, então Bob, ao efetuar uma medida do *spin* da partícula B com um aparelho na mesma condição x obterá *down* como resultado. Caso Alice efetue uma medida do *spin* da partícula A com um aparelho na condição x e obter *down* como resultado, então Bob, ao efetuar uma medida do *spin* da partícula B com um aparelho na mesma condição x , deverá obter *up* como resultado. A mesma situação ocorreria se Bob efetuasse sua medida antes: Alice encontraria sempre um valor de *spin* oposto ao de Bob se efetuar uma medida com um aparelho na mesma condição de Bob⁶⁵.

O conhecido teorema de Bell mostrou que para esse tipo de experimento, quando se adota uma interpretação realista da mecânica quântica, a medida do *spin* de uma partícula

62 Ver Pessoa (2006) para um estudo detalhado dos conceitos fundamentais da teoria quântica.

63 Um momento angular intrínseco a partículas como elétrons, tomado aqui com valor $+1/2$ (para *up*) e $-1/2$ (para *down*).

64 Por exemplo, um aparelho de Stern-Gerlach orientado em uma determinada direção (Pessoa, 2006, p. 39-40).

65 Essa é uma propriedade conhecida como anti-correlação perfeita. Considerando medições realizadas em aparelhos de Stern-Gerlach, tal propriedade é válida para qualquer direção dos ímãs dos aparelhos, desde que sejam as mesmas para as duas partículas (Pessoa, 2006, pp. 215-16).

afeta o resultado da medida de *spin* da outra partícula – as medições não são independentes. Essa peculiaridade quântica é chamada de não-separabilidade: com pares emaranhados, é impossível analisar apenas uma das partículas separadamente sem influenciar a outra partícula e, conseqüentemente, afetar o estado do par como um todo. O estado do par emaranhado, antes de qualquer medição, seria uma sobreposição não-fatorável de dois estados: (i) a probabilidade de medir *up* para A e *down* para B e (ii) a probabilidade de medir *down* para A e *up* para B. Ao medirmos qualquer uma das partículas individualmente, o estado do par se torna fatorado, ou A com *up* e B com *down*, ou A com *down* e B com *up*. Sendo assim, se tomarmos as partes separadamente, não obtemos o estado inicial do todo com a sobreposição de (i) e (ii), mas sim ou (i) ou (ii). Em outras palavras, se o estado de cada partícula em separado for considerado, o estado de cada uma será uma mistura estatística, e a reunião dessa informação de cada partícula não recupera todas as informações de correlação do par emaranhado. Nesse sentido, haveria não-separabilidade espacial em nosso universo e, portanto, o holismo se aplicaria.

Esse argumento é aceitável se considerarmos comportamentos quânticos como princípios fundamentais de nossa realidade física. Como já visto no primeiro capítulo, autores como 't Hooft e Fredkin sugerem uma realidade ainda mais básica da qual efeitos quânticos são deriváveis e as dificuldades encontradas no estudo de pares emaranhados como apresentado acima seriam apenas ilusórias. Nesse sentido, uma sonda epistemológica capaz de analisar essa realidade mais básica poderia, em princípio, inferir a totalidade das escalas maiores tomando partes separadas no nível mais fundamental. Mais uma vez, tudo dependerá da metafísica de nosso universo físico⁶⁶.

2.5. Considerações Gerais sobre Reduccionismo

Neste capítulo procuramos traçar uma ideia geral do conceito de reduccionismo. Seguindo um caminho mais realista como proposto por Weinberg, reduccionismo pode ser avaliado de modo significativo, onde leis mais fundamentais determinam completamente leis menos fundamentais, ou trivial, onde propriedades mais fundamentais determinam completamente propriedades menos fundamentais. Propriedades em níveis mais básicos podem definir propriedades em níveis menos básicos por meio de uma relação de

⁶⁶ Uma interpretação da mecânica quântica que mais se aproxima de uma descrição separabilista é a interpretação causal de David Bohm (1952), mas nesse caso há ação à distância (não-localidade).

superveniência, mas, uma vez que superveniência não é uma condição suficiente para garantir redução, também é necessário especificar que níveis superiores explicam níveis inferiores, no sentido de teorias sobre níveis inferiores permitirem compreender as razões de teorias em níveis superiores funcionarem tão bem – compreender o nível básico necessariamente levaria a compreender o nível menos básico. Para ilustrar essa ideia podemos recorrer a modelos abstratos como sondas epistemológicas e cópias de universo. A partir de uma ontologia geral com elementos distribuídos no espaço e no tempo, além de dimensões de escala adicionais, seres como o demônio de Laplace e o demônio escalar podem ser definidos para mostrar que se o reducionismo é verdadeiro, então um conhecimento completo do estado presente do universo juntamente com suas leis de evolução temporal seria suficiente para definir um estado futuro e um conhecimento completo de escalas menores no espaço juntamente com leis de transição de escala seria suficiente para conhecer escalas maiores. Semelhantemente, em um universo reducionista, uma cópia exata de nosso universo físico em um determinado instante garante que o instante posterior será como o original e uma cópia exata das menores escalas de nosso universo garante que as maiores escalas da cópia serão como as maiores escalas do universo original. Objeções contra o reducionismo incluem a realização múltipla e a possibilidade de vivermos em uma realidade indeterminista, emergentista ou holista. A realização múltipla não impede que casos de redução sejam pensados caso a caso e as outras objeções invocam aspectos metafísicos do mundo natural que ainda permanecem em aberto e são tão questionáveis quanto o reducionismo.

No capítulo a seguir, aplicaremos as noções apresentadas aqui em um universo aos moldes de Fredkin.

Capítulo 3: Reduccionismo em um Universo Digital

O primeiro capítulo deste trabalho apresentou o modo como Edward Fredkin compreende os fundamentos da realidade física a partir de uma perspectiva computacional e o segundo capítulo procurou traçar uma ideia geral do conceito de reduccionismo. O objetivo deste último capítulo é conectar esses dois assuntos tratados até agora de modo separado, procurando mostrar em que sentido a visão de Fredkin é reducionista. Para tanto, iniciaremos com a definição de uma ontologia geral do universo digital proposto pelo autor. Feito isso, mostraremos como essa ontologia pode ser dimensionada em diferentes níveis, de modo que seja possível examinar como sondas epistemológicas seriam compatíveis nesse universo. Por fim, será discutido se um mundo baseado em computação digital aos moldes de Fredkin também não poderia ser investigado em uma linha não-reducionista.

3.1. Ontologia Geral Digital

Foi afirmado no capítulo anterior que diante da impossibilidade de definir o ôntico, poderíamos pelo menos definir uma ontologia que fosse coerente com as teorias mais aceitas pela ciência. Para fins de explicação conceitual de uma visão de mundo reducionista, costumamos adotar uma ontologia com entidades definidas no espaço e no tempo sem afirmar nada a respeito da natureza dessas entidades. Com o mundo digital defendido por Fredkin, porém, temos uma ontologia mais específica. Lembramos que Fredkin defende um princípio de Natureza Finita, onde espaço e tempo são tomados como discretos. Além disso, cada unidade de espaço apresentaria um estado bem definido em cada unidade de tempo, de modo que as transições de um estado para outro podem ser representadas por algum modelo computacional. Nas palavras de Fredkin:

Há espaço, há tempo, há bits, há um processo digital simples – e não há nada mais na Filosofia Digital. (Fredkin, 2003, p. 195)

Em suma, o próprio Fredkin já define uma ontologia geral constituída de espaço, tempo, *bits* (ou informação) e um processo digital que direcione as transições de estados correspondentes às unidades espaciais e temporais. Além desses elementos básicos, não haveria mais nada na Filosofia Digital. Tudo que há para se falar em um mundo digital

fredkiniano seria uma distribuição de informação no espaço que pode se alterar de acordo com o processo digital responsável por realizar as transições. Definimos assim a ontologia geral do mundo de Fredkin como informação contida em unidades discretas de espaço e tempo que pode ser representada digitalmente e se modifica de acordo com regras de transição específicas.

Vale notar, porém, que a natureza definitiva dessa informação é desconhecida. Fredkin diz que vivemos em um universo que se comporta como um sistema computacional, mas ele considera tal universo como parte de um domínio maior chamado apenas de Outro (ver Seção 1.8.4). Fredkin não apresenta aspectos específicos do Outro – esse domínio exterior pode ser muito diferente com relação à física de nosso universo, apresentando um número diferente de dimensões ou até mesmo uma estrutura espaço-temporal diferenciada (Fredkin, 2003, p. 193). Não obstante, o Outro deve ser capaz de apresentar estados que a informação necessária para a computação correspondente aos processos físicos existentes em nosso universo. É sabido que muitas coisas podem servir como portadores de informação. A rotação intrínseca de um elétron (seu *spin*), por exemplo, pode exprimir um *bit* de informação (no caso de um elétron, dois estados possíveis: *spin* +1/2 e *spin* -1/2). Mas a natureza da informação poderia ser qualquer outra coisa. Para Fredkin, a natureza da informação está relacionada com o modo como um arranjo de coisas se apresenta, independente do que essas coisas sejam (cf. Fredkin, 2004, p. 35). Como já vimos (Seção 1.3), haverá informação em um determinado arranjo de coisas havendo interpretação de significado da configuração do arranjo em algum processo. A condição necessária é que esse processo apresente transições de estados correspondentes à evolução que identificamos em nosso mundo natural. Pelo modelo de MD sugerido pelo autor, a informação de nosso mundo poderia ser dada em termos de momento angular associados a algum elemento material elementar, mas não parece haver impedimentos lógicos para imaginar que qualquer outra coisa poderia servir como portador de informação, desde que o processamento dessa informação permita derivar as leis de nosso mundo físico⁶⁷. Se Fredkin não define especificamente as propriedades do Outro, então não podemos dizer muito sobre ele além do fato de ser uma realidade que admite computação universal, permitindo assim a realização de alguma computação que resulta no que encaramos como nosso universo físico.

67 Tegmark (2007), por exemplo, sugere que nosso universo é fundamentalmente uma estrutura matemática. Aproximando essa ideia do pensamento de Fredkin, um modelo de MD poderia partir do princípio de que as unidades básicas de informação são simplesmente *bits* puros e o universo físico seria simplesmente uma estrutura computacional. Um realismo estrutural ôntico que admite apenas a existência de estruturas oferece uma argumentação metafísica que pode reforçar uma visão como essa (e.g. Ladyman et. al., 2007).

Diante disso, nossa ontologia geral será restrita ao nosso mundo: vivemos em um universo com espaço, tempo, informação e regras de transição, ainda que não seja possível definir precisamente uma ontologia para o local onde esse universo computador se encontra. Não obstante, independente do que exatamente origina a informação, os *bits* contidos no espaço e no tempo são processados de acordo com uma regra específica.

Como discutido no primeiro capítulo, Fredkin não apresenta um modelo definitivo do programa de nosso universo, mas parte do princípio de que se a realidade física é resultante de processos computacionais, então deve ser possível derivar as leis conhecidas da física a partir de tais processos. Nas palavras do autor:

O que precisamos exigir da Filosofia Digital é a eventual habilidade de derivar, partindo de nossos modelos de processos fundamentais na Filosofia Digital, as mesmas equações matemáticas que constituem a base da ciência hoje. (Fredkin, 2003, p. 190)

Assim sendo, partindo apenas de uma programação básica seria possível definir as leis que conhecemos hoje. Se a natureza é completamente definida pela computação, isso está bem próximo da visão de mundo reducionista discutida no capítulo anterior. Mas, para refinar essa ideia, além da ontologia geral definida acima, também seria importante aplicar os métodos de sondas epistemológicas e cópias discutidos no capítulo anterior para mostrar em que sentido o reducionismo se aplica ao universo digital. Para tanto, será esclarecido como funcionam as escalas no universo de Fredkin na seção seguinte.

3.2. Autômatos Celulares e Níveis

Fredkin defende que o modelo mais adequado para descrever o processamento de dados distribuídos em unidades discretas de espaço seria um autômato celular. Relembramos que um autômato celular é um modelo computacional constituído de unidades mínimas (células) capazes de assumir, cada uma, um número finito de estados. Essas células estão relacionadas entre si por uma relação de vizinhança, de modo que a regra de transição de um estado para outro dependa do estado da própria célula e do estado das células vizinhas. Assim, Fredkin acredita que a computação fundamental de nosso mundo se dê conforme as regras de um autômato celular: cada unidade espacial modificaria seu estado de acordo com as unidades mais próximas no espaço.

O nível mais básico de um autômato celular é chamado por Daniel Dennett de nível físico (Dennett, 1991, p. 39). Mais precisamente, seria o nível constituído pelas células, seus estados e a regra de transição aplicada. O programa atuará apenas nesse nível, agindo simultaneamente em todas as células. Tudo que um programador precisa fazer é definir quais estados ocorrem e qual será a regra de transição responsável por alterar os estados. O Jogo da Vida, discutido na Seção 1.4.2, por exemplo, envolve apenas dois estados (Viva e Morta) e as regras de transição para cada célula seria: manter o estado atual caso tenha exatamente duas vizinhas vivas ou se tornar viva caso tenha exatamente três vizinhas vivas ou se tornar morta para qualquer número de vizinhas diferente de dois ou três. Com esses dados, um programador pode desenvolver seu próprio mundo digital apenas indicando como as transições devem ocorrer. Há também o fator tempo. Podemos dizer que cada transição leva para um próximo instante ou uma próxima unidade de tempo. No entanto, o intervalo de tempo da transição de uma unidade para outra não necessariamente é o tempo mínimo. Fredkin, ao sugerir um modelo de autômato celular para o mundo físico, considerou que cada transição do autômato também deve levar em conta um tempo separado para o programa detectar os estados de cada célula e de sua vizinhança (ver Seção 1.7). Assim, cada unidade de tempo de transição seria dividida em outras subunidades necessárias para que a informação presente em cada célula seja devidamente processada. Seja como for, ainda seriam unidades discretas de tempo. Em suma, o nível físico teria variáveis relacionadas às coordenadas espaciais das células, ao tempo (o tempo do programa examinar os estados nas células e executar a transição), à regra de transição e aos estados possíveis para cada uma das células.

Mas além do nível físico, onde o programa atua diretamente, o maior charme dos autômatos celulares, como já vimos, está no fato de permitirem a formação de certos padrões quando tomados em maiores escalas. A ideia de escala, como discutido no segundo capítulo, pode ser pensada como dimensões adicionais em nossa ontologia geral. Assim, se pensarmos no nível físico apenas como aquilo que ocorre em cada célula individualmente, podemos pensar também em algum nível que considere o que ocorre em certos aglomerados de células durante certos intervalos de tempo. Esse nível pode ser incorporada em uma classe distinta chamada nível de *design* (Dennett, 1991, p. 39). Basicamente, não consideramos esse nível em termos de células em cada instante de tempo, mas em termos de fatias de espaço e tempo que permitem tomar, de uma única vez, o arranjo de várias células durante várias gerações. Assim, é possível avaliar o comportamento coletivo das células em uma perspectiva mais

ampla.

Tomemos mais uma vez o Jogo da Vida como exemplo. Se focarmos nossa atenção para apenas uma célula, tudo que podemos dizer sobre ela é seu estado individual (Viva ou Morta) em cada geração. Se tomarmos uma escala um pouco maior, digamos, uma fileira de três células durante três gerações, será possível verificar mais padrões, mas ainda não seriam muitos. Um *blink* (Figura 10), por exemplo, só poderia ser identificado quando consideramos uma área de três células por três células em três gerações. Já quando consideramos uma fatia maior do tabuleiro constituída, por exemplo, de uma área por volta de sete células horizontais por sete células verticais e uma fatia de tempo constituída de cinco gerações, podemos perceber configurações mais interessantes além do *blink*, como o *glider*, já discutido na seção 1.4.2 e repetido na figura 20 abaixo:

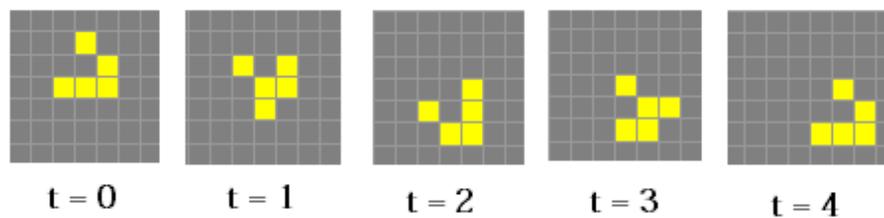


Figura 20: Padrão para cinco primeiras gerações de um glider.

Assim, diferentes escalas podem ser tomadas e, portanto, diferentes níveis de *design* podem ser definidos. Em cada um desses níveis de *design*, podemos identificar mais comportamentos de grupos de células que podem ser definidos como elementos próprios daquele nível.

Explicando melhor com um exemplo, o grupo de cinco células vivas constituindo o *glider* parece formar um único corpo que se locomove rotacionalmente na diagonal quando observamos seu comportamento de uma perspectiva de maior escala. No entanto, esse tipo de comportamento não é uma regra definida pelo programador no nível físico. O movimento do *glider* é algo que só pode ser contemplado em uma escala maior: só podemos incluir uma ontologia de *gliders* para áreas maiores de células e maiores intervalos de tempo. A própria ideia de movimento já aparece como uma novidade: note que as células do Jogo da Vida não se movem em nenhum momento, tudo que ocorre são transições de estado, mudanças no conteúdo informacional de cada célula. Porém, nessa escala mais ampla, se definimos *gliders*

como objetos próprios do nível de *design*, então também devemos definir propriedades específicas dos *gliders* – e movimento com relação a outras células e objetos seria uma delas.

Se tomarmos fatias maiores, a quantidade de entidades presentes na ontologia em níveis de *design* só tende a aumentar. A Figura 14 no capítulo 1, por exemplo, é um caso de configuração estável que gera *gliders* automaticamente após um certo número de gerações. Em nenhum momento o programador incluiu uma nova regra, tudo se mantém inalterado no nível físico. Semelhantemente, também é possível encontrar configurações capazes de destruir *gliders*. Considere a Figura 21 a seguir:

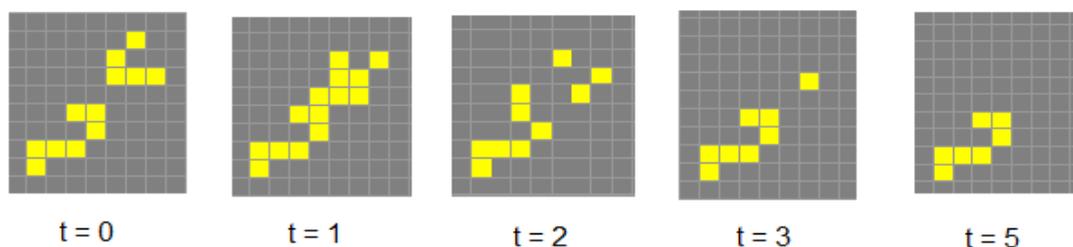


Figura 21: Padrão de um eater “devorando” um glider.

Vemos no canto inferior esquerdo um padrão conhecido como *eater*⁶⁸ (Poundstone, 1985, p. 40) com um *glider* logo acima. O *glider* tende a se movimentar diagonalmente para baixo e para a esquerda, colidindo com o *eater* na próxima geração. O resultado é uma modificação tanto na configuração do *glider* quanto na configuração do *eater* na geração seguinte. Na quarta geração, o *eater* retorna a seu estado inicial acompanhado de uma única célula viva logo acima. Na quinta geração, por fim, essa célula viva restante é destruída, restando apenas o *eater*. Em resumo, o *glider* é “devorado” pelo *eater*: começamos com um *glider* na primeira geração e terminamos sem ele na quinta geração, enquanto permanecemos com o *eater*. Cabe ressaltar que o *eater* é um devorador de *gliders* seguindo a configuração da Figura 21 – outros padrões não são devorados e podem até destruir o padrão *eater*. Parece haver uma clara interação entre elementos de alto nível e parece perfeitamente possível definir leis para níveis além da menor escala (no nível físico), ainda que apenas esse nível mais básico envolva a atuação do programa⁶⁹.

68 Uma tradução literal seria “devorador”.

69 Mostramos regras de alto nível que claramente diferem das regras do autômato atuando em baixo nível. Porém, uma questão interessante envolve a possibilidade de desenvolver um autômato celular de alto nível

Nesses poucos exemplos encontrados no Jogo da Vida já podemos reconhecer claramente dois níveis de autômato celular: o nível físico, onde consideramos cada célula individualmente alterando seu estado de acordo com os estados da vizinhança, e níveis de *design* em escalas mais ampla, onde aglomerados de células permitem definir uma ontologia composta de entidades como *gliders* e *eaters*. Além disso, parece possível reconhecer regularidades próprias no nível superior, como o fato de *gliders* se movimentarem e *eaters* conseguirem devorar *gliders*. Esse tipo de descrição só pode ocorrer em níveis mais altos e não tem lugar no nível físico, definido apenas pelas regras de transição. Comparando com a terminologia utilizada no segundo capítulo, o nível físico dos autômatos celulares corresponde à dimensão ôptica do mundo, enquanto os níveis de *design* correspondem à dimensão ontológica.

O exemplo do Jogo da Vida estimula uma reflexão mais profunda sobre a realidade imaginada por Fredkin. Segundo o autor, nosso mundo físico se comporta como um autômato celular em que regras de transição atuam em cada unidade básica de espaço. Assim como um único programa atuando em nível físico parece suficiente para definir objetos complexos como *gliders* e *eaters* no nível de *design* do Jogo da Vida, um único programa poderia bastar para definir entidades como elétrons e quarks no nível superior do mundo real com propriedades especiais que não apareciam explicitamente no nível mais básico. Dada uma configuração inicial, o programa rodaria apenas nas células básicas que culminariam em padrões estáveis associados a nossas partículas elementares em alguma escala que, por sua vez, definiriam átomos, moléculas e as demais entidades macroscópicas com seus respectivos comportamentos descritos por nossas atuais teorias científicas.

Uma vez que níveis e escalas podem ser considerados em autômatos celulares, resta ainda entender de que maneira esses níveis podem ser tomados em uma visão reducionista, ou seja, esclarecer se os níveis superiores de *design* dos autômatos celulares são completamente definidos pelo nível físico. Podemos utilizar os métodos de sondas epistemológicas ou de cópias para esse fim.

com “macrocélulas” utilizando as regras de baixo nível. Isso deve ser difícil em níveis de *design* muito próximos do nível físico – o movimento de um *glider* é bem diferente da dinâmica de objetos básicos que apenas mudam seu estado de acordo com a vizinhança. Mas como nosso mundo permite a execução do Jogo da Vida, um autômato celular diferente do que Fredkin propõe para o universo, em um certo sentido podemos dizer que nosso mundo seria um autômato celular capaz de simular outros autômatos celulares – ainda que as regras referentes à dinâmica desses outros autômatos só se torne aparente em uma escala muito alta. Isso está ligado ao princípio de computabilidade universal dos autômatos celulares (ver seção 1.5).

3.3. Sondando um Mundo Digital

No segundo capítulo foram apresentadas algumas sondas epistemológicas, recursos abstratos que permitem compreender como ocorre determinação em um mundo específico. Uma delas é o demônio de Laplace, o conceito de um ser fora do espaço-tempo, mas capaz de descrever com precisão o estado do universo em algum instante partindo dos dados analisados em um instante anterior por conta de sua onisciência de todos os estados do universo em um dado instante, sua onisciência de todas as leis de evolução da natureza e seu grande poder de computação.

Não parece muito difícil definir a tarefa do demônio de Laplace em um mundo fredkiniano. Em um universo desse tipo o espaço é discreto e todas as unidades básicas de espaço apresentam um estado bem definido. O poder de onisciência do demônio de Laplace seria suficiente para abarcar todos os estados distribuídos espacialmente. Feito isso, o conhecimento das regras de transição do programa permitiriam ao demônio determinar como seriam todos os estados espaciais no próximo instante de tempo discreto.

Um problema da aplicação do demônio de Laplace no universo de Fredkin diz respeito ao seu poder computacional. Como apresentado na Seção 1.8.3, Fredkin declara que qualquer predição exata de estados futuros dentro do universo computacional é impossível. Uma vez que estamos inseridos no programa, nossos poderes cognitivos também são resultantes da ação do autômato, nosso cérebro faz parte do sistema e, portanto, nossa memória nunca poderia abarcar todos os estados espaciais, tornando logicamente impossível calcular um estado futuro mais rapidamente do que o próprio autômato celular levaria para calcular (Fredkin, 2003, p. 210).

Por outro lado, como o demônio de Laplace, por definição, é externo ao espaço-tempo (na terminologia de Fredkin, o demônio estaria no “Outro”), ainda seria possível dizer que os poderes do demônio não estão limitados ao processamento do universo digital, já que nada impede que as leis naturais do Outro sejam diferentes das leis definidas no programa – talvez o processamento do nosso universo seja mais lento do que o processamento no Outro e, portanto, um demônio de Laplace seja capaz de prever com precisão estados futuros de nosso universo⁷⁰. Em suma, o universo de Fredkin é determinista, ainda que seus estados futuros não

⁷⁰ Se o Outro também for um programa rodando em outra realidade, então o demônio de Laplace não poderia prever com exatidão os estados futuros do Outro – mas, se a realidade em que se encontra o Outro apresentar

sejam previsíveis para habitantes do universo.

Tomemos mais uma vez o Jogo da Vida como ilustração. Suponha uma versão do Jogo da Vida programada em algum computador de nosso mundo de modo que cada geração demore um minuto inteiro para se alterar. Assim, um *glider* como mostrado na Figura 20 levaria quatro minutos para executar um ciclo completo. Nesse caso, poderíamos facilmente cumprir o papel do demônio de Laplace com respeito ao Jogo da Vida e determinar o estado da próxima geração antes que o computador faça isso, pelo menos para um caso com poucas células, onde nossa memória fosse capaz de armazenar todos os estados (um tabuleiro de Jogo da Vida com uma área de cinco células por cinco células já seria suficiente para prever o comportamento de um *glider*, por exemplo). Uma vez que os estados de todas as células e as regras de transição sejam conhecidos, a tarefa seria factível, já que podemos calcular os estados futuros sem interferir no sistema.

A partir disso, podemos adaptar essa ideia para o universo de Fredkin: a sonda teria um poder de computação maior do que o processamento do universo, apresentaria uma memória grande o suficiente para capturar todos os estados definidos no espaço, conheceria as leis de transição e não afetaria o funcionamento do sistema por tentar prever os próximos estados. Assim, estados futuros poderiam ser previstos por um demônio de Laplace em um universo digital como proposto por Fredkin. Como o autor propõe um autômato celular reversível no tempo, também podemos dizer que o demônio de Laplace seria capaz de definir com precisão um estado anterior do universo. Conhecendo todos os dados, o demônio poderia adaptar as regras de transição para verificar com precisão em que estado o universo se encontrava no passado⁷¹.

Além do demônio de Laplace, responsável por ilustrar a determinação temporal do mundo em análise, também podemos pensar em como aplicar uma sonda epistemológica envolvendo determinação em escalas, ou seja, algum modo de verificar se é possível conhecer em princípio o comportamento do universo em escala macroscópica a partir da escala microscópica. Para tanto, como discutido na seção 2.3.2, podemos definir um demônio escalar que, a partir dos estados e leis das escalas microscópicas seria capaz de determinar completamente os estados e leis em escalas superiores.

um processamento maior, poderíamos definir um outro demônio de Laplace e assim por diante. Mas como desconhecemos a natureza exata do Outro, isso é apenas especulação.

⁷¹ Nesse caso, vale lembrar que a reversibilidade do universo de Fredkin é baseada em um princípio de conservação de informação (cf. Seção 1.6), de modo que o demônio de Laplace não teria dúvidas quanto ao estado anterior do universo.

O universo de Fredkin é fundamentalmente um autômato celular. Vimos que autômatos celulares podem ser analisados em vários níveis que mudam de acordo com a escala. No nível físico, o nível mais elementar onde o programa atua, apenas o comportamento das células individuais é considerado, enquanto nos níveis superiores é possível definir objetos com comportamentos diferenciados, como *gliders* e *eaters* no caso do Jogo da Vida. Sendo assim, um demônio escalar trabalhando no universo fredkiniano deve ser capaz de definir o que ocorre em níveis de *design* a partir do nível físico.

Uma maneira de compreender os níveis maiores em termos dos níveis menores é entender como o nível mais básico explica os níveis menos básicos. No caso de autômatos celulares, devemos lembrar que o algoritmo atua apenas no nível físico, alterando o estado de cada célula individualmente com base em sua vizinhança – não há nenhuma atuação do programa nos níveis de *design*. Não obstante, se apenas as regras de transição das células determinam o sistema, então parece razoável imaginar que há algum modo de vincular o que ocorre em níveis superiores em termos de níveis inferiores.

Nesse ponto, poderíamos procurar por pontes que ligam diferentes níveis de organização (ver Seção 2.2.1). Lembramos do raciocínio de redução entre teorias desenvolvido por Nagel: se teorias de nível fundamental explicam teorias de nível menos fundamental, mas as teorias menos básicas apresentam um vocabulário diferenciado em relação às teorias mais básicas, então devem existir regras de ponte que permitam traçar uma equivalência entre as definições de alto nível e as definições de baixo nível. A temperatura de um corpo, por exemplo, é um conceito de alto nível que pode ser definido em termos de conceitos trabalhados em níveis menores, como a média dos valores de energia cinética das moléculas do corpo. Semelhantemente, é perfeitamente possível que exista um modo de definir os comportamentos de alto nível utilizando apenas as descrições válidas no nível físico.

Mais uma vez podemos tomar o Jogo da Vida como exemplo. Considere um padrão simples no Jogo da Vida, como o *block* (Figura 22):

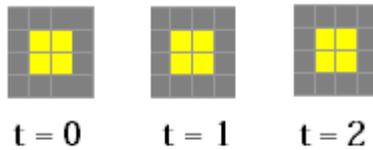


Figura 22: Padrão block para três gerações.

Pelas regras definidas no nível físico, cada célula com exatamente três vizinhas deve assumir o estado viva. Atentando para a Figura 22, vemos que na primeira geração temos quatro células vivas e cada uma delas tem exatamente três vizinhas vivas. No nível mais elementar, portanto, podemos dizer que todas essas células permanecerão no estado viva na próxima geração. Já em um nível mais elevado na hierarquia, considerando uma escala um pouco maior, podemos definir um *block* como um objeto formado por quatro células vivas juntas que não se separam no decorrer das gerações caso não apareçam outras células vivas na vizinhança. O comportamento do *block* pode ser explicado em termos das regras físicas: como todas as células do *block* possuem exatamente três vizinhas vivas, então elas nunca sairão do estado viva sem a presença de novas vizinhas. Desse modo, o conceito de *block*, significativo apenas em um nível superior, teria uma descrição equivalente utilizando apenas o vocabulário do nível mais básico e o nível físico permitiria compreender as razões do comportamento do *block*.

Considere como segundo exemplo autômatos celulares unidimensionais elementares trabalhados na Seção 1.4.1. O nível físico desse tipo de autômato considera, mais uma vez, as regras de transição em cada célula com base no estado da célula e os estados de suas vizinhas. Aspectos de níveis superiores se tornam mais evidentes quando tomamos um intervalo consideravelmente alto de transições. Vimos na seção 1.4.1 que a regra 90 (adotando a classificação de Wolfram) gera um padrão fractal em um fator $\log_2 3$ (cf. Wolfram, 1983, p. 607) a partir de uma determinada condição inicial. O surgimento desse padrão pode ser explicado com base nas regras de transição atuando no nível físico.

Supondo um demônio escalar avaliando o Jogo da Vida, ele conseguiria inferir um *block* atentando apenas para o nível físico. Conhecendo as regras de transição das células e os estados assumidos pelas células no tabuleiro, o poder de processamento do demônio permitiria a ele determinar o padrão *block* em escala maior conhecendo apenas os princípios da escala menor e a regra de vínculo entre a escala menor e a escala maior: se é possível definir um *block* em termos do nível físico, então um demônio escalar poderia definir o *block*

em um nível de *design*. Semelhantemente para um autômato celular unidimensional, o demônio escalar conseguiria inferir o padrão fractal sabendo apenas que a regra 90 atua em nível físico e sabendo as relações de ponte que permitem descrever esse padrão de alto nível em termos das regras do autômato.

Para um universo fredkiniano, a atividade do demônio seguiria esse mesmo raciocínio. Conhecendo a programação aplicada aos estados no nível mais básico e as pontes entre as descrições dos níveis, o demônio poderia determinar como seria o comportamento dos níveis superiores, definindo interações entre elementos de escalas maiores (como partículas elementares, átomos, moléculas, células e assim por diante) a partir das interações entre os estados assumidos pelas unidades espaciais.

Sondas epistemológicas também podem ser utilizadas na questão da separabilidade espacial, ou seja, conhecendo separadamente aspectos de um sistema em diferentes regiões espaciais permitir definir o aspecto do todo quando essas partes separadas forem reunidas. No caso de autômatos celulares, se sempre for o caso do conhecimento dos estados de cada célula em separado permitir definir o estado de um padrão coletivo, então há separabilidade espacial nos autômatos. No Jogo da Vida, saber a localização de cada célula viva permitiria definir um objeto inteiro em nível superior – verificar separadamente o estado e vizinhança de cada célula na figura 22 seria suficiente para definir um *block*, por exemplo. Semelhantemente, o conhecimento do estado de cada unidade espacial do autômato celular de nosso mundo permitiria o conhecimento do estado de um objeto macroscópico que essas unidades em conjunto definiriam em uma unidade de tempo específica⁷².

3.4. Copiando um Mundo Digital

Um segundo método para verificar um aspecto reducionista é o chamado método das cópias (ver seção 2.3.2). Enquanto o método por sondas epistemológicas apresenta um meio de mostrar como é possível conhecer estados de alto nível em princípio, o método de cópias segue uma linha mais ontológica: ao invés de invocarmos um observador fora do espaço-tempo, consideramos um mundo possível exatamente como o nosso em seu nível mais básico.

72 Muitas questões podem surgir dessa discussão: autômatos celulares não-determinísticos, caracterizados por mudanças de estado aleatórias, não poderiam ter o estado futuro determinado pelo demônio de Laplace, mas talvez princípios de alto nível pudessem ser definidos por um demônio escalar. Semelhantemente, talvez fosse possível pensar em um autômato celular onde não ocorresse separabilidade espacial, mas as relações entre níveis pudessem se manter. Essas questões não serão discutidas aqui, mas vale notar que elas podem aparecer.

Se esse mundo também se comportar como o nosso mundo em níveis superiores, então podemos dizer que as interações básicas determinam o que acontece nos patamares menos fundamentais. Em suma, admitimos superveniência do nível superior com respeito ao nível inferior: não há como mudar aspectos superiores sem modificar aspectos inferiores (ver seção 2.2.3).

Isso não parece muito difícil de realizar em autômatos celulares. Modelos computacionais podem ser executados em diferentes tipos de substratos físicos, portanto é extremamente simples pensarmos na cópia de um autômato celular. Considere mais uma vez o Jogo da Vida como exemplo. Tanto um computador A quanto um computador B podem rodar um programa que implemente as regras do Jogo da Vida. Suponha que o computador B inicie o Jogo da Vida exatamente no mesmo estado que um computador A inicie esse programa. Além disso, as regras de transição funcionarão do mesmo modo para as duas instâncias do programa, ou seja, a cada geração os computadores exibirão exatamente os mesmos estados. Mas se os objetos de alto nível dependem exclusivamente dos objetos de baixo nível, então o computador B exibirá os mesmos comportamentos de nível superior que aparecem no computador A. Se encontrarmos *gliders* em A, também esperamos encontrar *gliders* em B se comportando exatamente da mesma maneira. Se o comportamento de alto nível em B for diferente do comportamento de alto nível em A, isso mostraria que A e B se diferenciam de algum modo no nível físico: talvez os estados das células sejam diferentes nos dois programas ou as regras de transição não sejam as mesmas. Porém, se de fato as mesmas regras do Jogo da Vida são aplicáveis nos dois programas com a mesma configuração de células, o esperado é que os algoritmos sejam executados do mesmo modo.

Fredkin defende que nosso universo é um programa implementado em uma realidade exterior (Outro). Mais uma vez vale repetir que não sabemos nada sobre o Outro, mas se esse cenário externo admitir várias implementações de um mesmo programa, então podemos pensar no método das cópias aplicado ao nosso universo de modo semelhante aos programas rodando em diferentes computadores⁷³. Considere que nosso universo, de fato, seja um programa executado no Outro. Considere também que uma cópia de nosso programa também possa ser executada nesse mesmo cenário, com as mesmas condições iniciais e as mesmas regras de transição. Sendo assim, novamente podemos esperar que cada unidade espacial do

⁷³ Essa metodologia é utilizada apenas a título de ilustração. Não é de modo algum necessário que existam cópias reais de um dado mundo para aplicar o método das cópias. Basta que haja possibilidade de pensar em uma cópia que apresente todos os aspectos microscópicos do mundo em questão.

universo original tenha um correspondente no espaço do universo cópia em cada instante de tempo. Se as entidades e propriedades de alto nível dependerem do que ocorre em níveis mais básicos, então podemos esperar que elas também ocorram na cópia. Conseqüentemente, se nosso universo for mesmo digital do modo que está sendo proposto aqui e fizermos uma cópia exata de todos os estados de suas células, desde seu instante inicial até o presente instante, seguindo exatamente as mesmas regras de transição, então todo o universo copiado seguiria a mesma evolução do nosso universo e podemos concluir que uma pessoa idêntica a mim estará escrevendo essas mesmas palavras em um *laptop* nesse universo vizinho. Desse modo, valeria superveniência, isto é, qualquer diferença mínima nas ocorrências de alto nível poderá ser entendida como alguma diferença de baixo nível - ou não há uma correspondência exata entre os estados das células espaciais de cada universo em cada instante ou as regras de transição para cada universo são diferentes. Além disso, essa superveniência seria forte, pois considerando a superveniência global discutida na seção 2.2.3 e o que foi discutido acima, um mundo exatamente idêntico ao nosso no nível físico, tanto em seus estados individuais e em suas regras, não poderia resultar em um mundo diferente em nível de *design*⁷⁴.

3.5. Autômatos Celulares e Anti-Reduccionismo

As duas seções anteriores mostraram em que sentido autômatos celulares podem ser encarados sob uma perspectiva reducionista. Basicamente, se é possível explicar completamente comportamentos de alto nível (como *eaters* devorando *gliders* no Jogo da Vida) apenas em termos das regras de transição atuando no nível físico, então parece razoável aplicar reduccionismo a um mundo físico computacional baseado em autômatos celulares, como é o caso da hipótese de Fredkin. Porém, também é esperado que argumentos anti-reducionistas apareçam nessa discussão. Diante disso, serão comentadas nesta seção algumas objeções ao reduccionismo em autômatos celulares (e conseqüentemente em universos

74 Talvez seja interessante mencionar que existe uma noção de superveniência chamada superveniência humeana, definida por David Lewis (1986, p. x) que segue a negação de David Hume com respeito à necessidade das leis naturais. De acordo com a superveniência humeana, nosso mundo macroscópico seria superveniente a um nível microscópico e toda mudança em nível macroscópico deveria ser acompanhada de uma mudança em nível microscópico. Porém, as mudanças em nível microscópico não estariam submetidas a nenhuma regularidade necessária – não haveria conexão causal entre a configuração de estados no espaço em um instante inicial e as configurações espaciais de estado um instante posterior. A perspectiva de Fredkin, portanto, seria incompatível com uma superveniência humeana, pois as mudanças microscópicas – no nível físico – obedecem princípios mais básicos definidos pela regra do autômato. Partindo do princípio de que a realidade microscópica apresenta uma regularidade mais fundamental do que a distribuição espaço-temporal de estados, a superveniência no mundo de Fredkin não pode ser humeana.

baseados nesses modelos).

3.5.1: Mundos Digitais Multiplamente Realizáveis

Um dos argumentos contra o reducionismo mencionado no capítulo anterior é o argumento da múltipla realização. Em linhas gerais, o fato de aspectos de alto nível admitirem vários modos de realização contraria a ideia de que elementos de um determinado tipo sejam reduzidos a um único tipo (Seção 2.4.1).

Embora a realização múltipla possa ser problemática para uma redução de tipos, não é uma preocupação quando consideramos uma redução de instâncias. Tome mais uma vez o exemplo da mente: mesmo que um tipo de estado mental como a sensação de dor não possa se reduzir a um único tipo de cérebro, ainda é possível argumentar que ocorre redução para cada instância de dor: a dor humana pode se reduzir ao cérebro humano, a dor canina pode se reduzir ao cérebro de um cachorro, a dor marciana pode se reduzir ao cérebro de um marciano e assim por diante.

No caso da perspectiva digital de Fredkin, a realização múltipla também não é problemática. Considere a questão da informação, por exemplo. Fredkin pressupõe que informação não depende do tipo de material que a carrega, ou seja, os *bits* processados pelo programa da natureza podem se originar de qualquer elemento capaz de fornecer estados com significado passível de interpretação por um processo (Seção 1.3). De fato, a informação no mundo de Fredkin é multiplamente realizável e não se reduz a um único tipo de portador, porém, esse universo em especial se reduz a algum elemento em especial capaz de manifestar informação.

Também é interessante notar que a própria evolução do universo é multiplamente realizável por diferentes regras de autômato celular, ou seja, duas ou mais regras podem gerar os mesmos estados no decorrer do tempo quando partem de uma mesma condição inicial. Considere, por exemplo, a regra 4 do autômato celular unidimensional elementar tratado na seção 1.4.1. Essa regra faz com que uma célula assumo estado 1 (preenchido) apenas se ambas as suas vizinhas estiverem no estado 0 (vazio). Para um estado inicial com apenas uma célula no estado 1 teremos um padrão de acordo com a figura 23:



Figura 23: Padrão de evolução do autômato regido pela Regra 4 de Wolfram com uma célula inicial preenchida.

Ocorre que exatamente esse mesmo padrão pode ser obtido pela regra 12. A regra 12, assim como a regra 4, torna uma célula preenchida se uma célula preenchida apresentar as duas vizinhas vazias, mas também faz isso para o caso da célula preenchida apresentar a vizinha da esquerda vazia e a vizinha da direita preenchida. Iniciando o autômato com apenas uma célula preenchida, o padrão no decorrer das transições se torna idêntico ao mostrado na figura 23. Aplicando isso a um universo físico definido por autômatos celulares, seria possível argumentar que o desenvolvimento de nosso universo no decorrer do tempo é consistente com mais de uma regra, já que uma outra regra poderia resultar na mesma evolução partindo de uma mesma condição inicial⁷⁵. Diante disso, universos digitais baseados em autômatos celulares são multiplamente realizáveis por diferentes regras subdeterminadas pela evolução do universo.

Mais uma vez, a realização múltipla não é um problema na perspectiva reducionista quando consideramos um reducionismo de instâncias. Ainda que diferentes regras possam determinar dois universos indiscerníveis no espaço e no tempo, não se segue que não possamos reduzir cada um dos universos como instâncias à sua respectiva ontologia de espaço, tempo, informação e regra. No exemplo do autômato celular unidimensional elementar, a instância do padrão de evolução baseado na regra 4 é completamente definido por um programa que segue a regra 4 e a instância do padrão de evolução baseado na regra 12 é completamente definido por um programa que segue a regra 12. Ainda que os padrões sejam indiscerníveis, cada instância foi feita em uma regra diferente. Note que se os autômatos iniciassem em uma outra condição inicial, a evolução de cada um deles poderia ocorrer de modo diferente. Suponha que ao invés de uma única célula preenchida, ambos iniciassem com duas células adjacentes preenchidas. Para o autômato que segue a regra 4, onde apenas casos

⁷⁵ Vale mencionar que esse problema é mais preocupante para interessados em conhecer a regra dentro do universo. Um demônio de Laplace conhece as regras de transição por definição e diferenciaria os dois universos de mesma evolução, mas com regras diferentes desde o início. Para quem está “dentro da máquina”, porém, resta essa subdeterminação de regras.

em que uma célula preenchida não tenha vizinhas preenchidas garante que o próximo estado seja preenchido, o padrão resultante será totalmente vazio, sem nenhuma célula preenchida. Já para o autômato que segue a regra 12, o padrão resultante será algo semelhante ao que encontramos na figura 23 (com a diferença de que há uma célula preenchida adicional na primeira linha). Se a semelhança entre as regras não resiste a condições iniciais contrafactuais, então não podemos dizer que esses dois casos de autômatos são a mesma coisa.

3.5.2. Mundos Digitais Não-Deterministas

Uma outra objeção ao reducionismo seria insinuar que se autômatos celulares não forem deterministas, então o presente não pode determinar completamente o futuro. Essa objeção ataca o determinismo, mas isso afetaria apenas um reducionismo do futuro em relação ao passado – ainda que estados futuros não possam ser conhecidos por um demônio de Laplace, isso não afetaria um reducionismo de escalas, onde um demônio escalar pudesse prever comportamento macroscópico partindo do comportamento microscópico e de leis de conexão entre escalas. De modo análogo, também é possível discutir a questão da reversibilidade em autômatos celulares – se não há reversibilidade, então o passado não pode ser totalmente definido a partir do presente. Como já comentado na seção 1.6, o Jogo da Vida é um exemplo de autômato celular não-reversível de modo que nem sempre é possível determinar o estado de uma geração anterior a partir de um estado da geração atual. No caso do universo digital aqui tratado, Fredkin toma o cuidado de exigir um autômato celular determinístico e reversível para o seu mundo por meio da conservação de informação. Desse modo, ainda seria possível determinar completamente tanto estados futuros quanto passados a partir do estado atual, pelo menos em princípio (Fredkin, 2003, p. 210).

3.5.3. Computação Emergente

Uma visão mais desafiadora ao reducionismo é o emergentismo, a aceitação da possibilidade de emergência (Seção 2.4.3). Na presente discussão envolvendo autômatos celulares, isso significaria que há aspectos no nível de *design* que não poderiam ser definidos em termos do que ocorre em nível físico. Um desses aspectos emergentes seria computação.

A computação no nível físico de autômatos celulares depende de interações locais –

uma célula altera seu estado em função de sua vizinhança. Diante disso, uma questão que surge é se autômatos celulares seriam capazes de determinar sua configuração global. Esse é o princípio do chamado problema da classificação de densidade.

Em linhas gerais, esse problema de classificação consiste no seguinte: é possível construir um autômato celular unidimensional finito, com dois estados para cada célula, capaz de determinar se há mais células em um estado do que em outro na geração inicial? Mais precisamente, dado um autômato celular unidimensional finito com células que podem estar preenchidas ou vazias, queremos que esse autômato pare⁷⁶ com todas as células preenchidas caso existam mais células preenchidas do que vazias no instante inicial ou que pare com todas as células vazias caso não existam mais células vazias do que preenchidas no instante inicial, isto é, o número de células vazias é maior ou igual ao de preenchidas. Solucionar esse problema envolve uma perspectiva global, já que a densidade de células preenchidas (ou vazias) em relação ao total de células do autômato é um aspecto do todo e não das partes separadas⁷⁷. Trata-se, portanto, de uma tarefa não-trivial para autômatos celulares, sistemas que lidam com interações locais – para apresentar a solução desejada, o autômato deve encontrar um modo de unir informações de todas as suas partes.

É provado que não existe um autômato celular capaz de solucionar esse problema precisamente (cf. Land & Belew, 1995). Não obstante, ao aplicar técnicas de algoritmo genético⁷⁸ com autômatos celulares unidimensionais definidos para raios de vizinhança maiores do que 1 ($r > 1$) é possível encontrar algumas evoluções de autômato associadas a determinadas regras e determinadas condições iniciais em que a solução desejada aparece. Um exemplo de evolução obtida por esse método seria a regra Φ_{17083} descoberta por James Crutchfield e Melanie Mitchell (1995) que permite gerar um padrão como mostrado na Figura 24 a seguir.

76 “Parar” aqui significa não realizar mais transições de estado.

77 Isso também torna o problema da classificação de densidade um caso interessante para avaliar a questão do holismo em autômatos celulares.

78 Muito grosso modo, um algoritmo genético seria a versão computacional da evolução genética. O computador analisa vários candidatos à solução de um determinado problema computacional, mantendo aqueles que mais se aproximam da solução e combinando aspectos desses candidatos para gerar variações com potencial de alcançar a solução do problema. Ver Mitchell (1996) para uma introdução sobre o assunto.

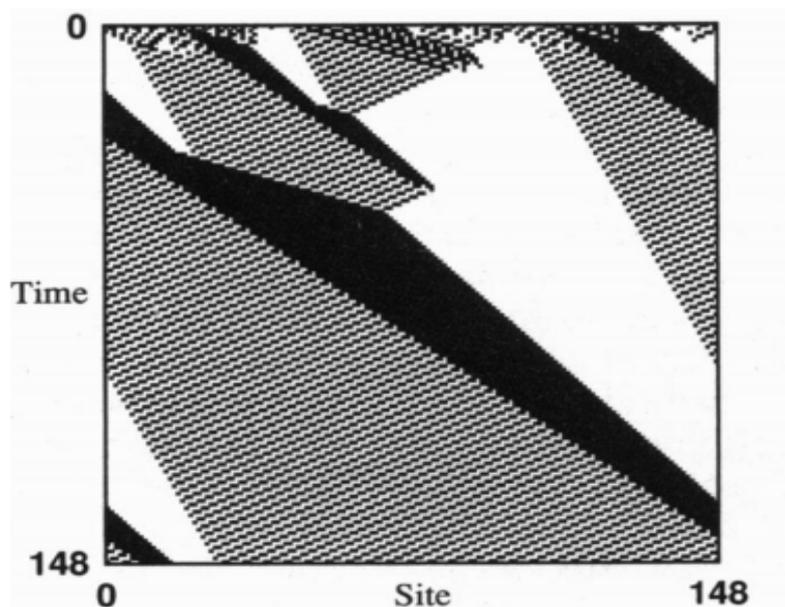


Figura 24: Padrão de evolução da regra Φ_{17083} para um estado inicial com mais células vazias do que preenchidas. O eixo horizontal representa a distribuição espacial e o eixo vertical a distribuição temporal. Na geração 250 (não mostrada na imagem), o autômato terá todas as suas células vazias, correspondendo a solução desejada para o problema de classificação de densidade (cf. Crutchfield & Mitchell, 1995, p. 10744).

Uma descrição de alto nível dessa distribuição espaço-temporal poderia ser feita em termos de regiões pretas (apenas com células preenchidas), brancas (apenas com células vazias) e cinzas (tanto com células vazias como preenchidas) que se “expandem”, se “contraem” e se “movimentam”. Crutchfield e Mitchell associaram esses comportamentos com trocas de informação no decorrer da distribuição espaço-temporal que poderiam representar algum tipo de computação de alto nível. Os autores identificaram cinco padrões de evolução dessas regiões que parecem seguir uma lógica específica e, por meio da aplicação da teoria da computação em autômatos celulares, filtraram o diagrama da figura 24 para apresentar um novo diagrama (figura 25) em que as regiões são substituídas por domínios equivalentes mais compactos chamados de “partículas”⁷⁹.

⁷⁹ As peculiaridades técnicas não serão tratadas aqui, mas será suficiente aceitar que é possível definir “partículas” equivalentes ao comportamento das regiões consideradas na figura 24. Ver Hanson & Crutchfield (1992) para mais detalhes.

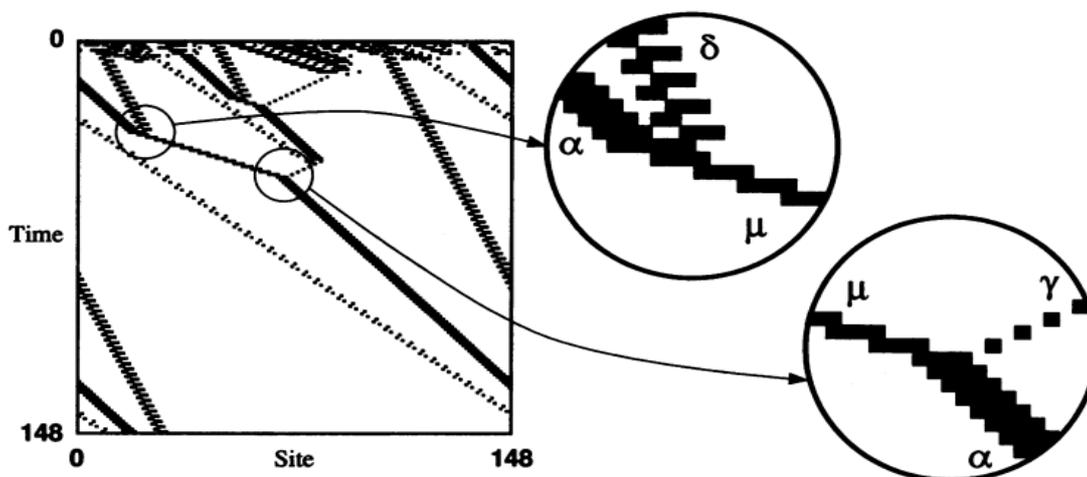


Figura 25: Análise do comportamento da regra Φ_{17083} apresentada acima. O diagrama é o mesmo da figura 24, mas filtrado por meio da aplicação de ferramentas teóricas computacionais especiais, permitindo que comportamentos encontrados nas regiões da figura anterior sejam representados por trajetórias na distribuição espaço-temporal encaradas como “partículas” que interagem entre si por meio de uma lógica própria que pode ser associada a processos computacionais de alto nível. Na figura, os padrões de partículas são identificados pelas letras gregas. Os dois círculos ampliam dois casos de interações entre partículas.

Assim, essas “partículas” constituirão elementos portadores de informação que se deslocam pela distribuição espaço-temporal do sistema e interações entre elas serão associadas a operações lógicas envolvidas em uma computação (Crutchfield & Mitchell, 1995, p. 10744).

A computação resultante das interações entre essas “partículas” ocorre apenas em um nível de *design* e não no nível básico das células, o chamado nível físico do autômato. Não obstante, essa computação de alto nível parece influenciar o baixo nível, pois leva a cada célula individualmente a terminar vazia no nível mais básico. Nessas condições, aspectos de alto nível afetariam causalmente aspectos de baixo nível, possibilitando uma emergência forte no sentido de Bedau (2003) comentado na seção 2.4.3.

O mundo de Fredkin certamente não é emergente no sentido forte, uma vez que, no universo digital definido pelo autor, o alto nível nunca modifica o baixo nível, antes é completamente definido por ele. Por outro lado, Fredkin usa o termo “emergente” em seus textos. Nas palavras do autor:

O campo da Mecânica Digital é totalmente definido por seis constantes fundamentais. Todas as outras constantes numéricas do modelo, todas as leis, todas as equações diferenciais e o conjunto de partículas e suas características são todos propriedades emergentes desse campo. (Fredkin, 2003, p. 209)

Se isso não pode ser encarado como emergência forte (afinal, o “totalmente definido” parece garantir que os aspectos fundamentais regulam tudo), podemos questionar se envolve uma emergência fraca. Mais uma vez, segundo Bedau (2003), propriedades emergem de modo fraco se as propriedades de baixo nível permitem derivar as propriedades de alto nível de modo não-trivial. Mais precisamente, isso significa que não há modo mais simples de definir os aspectos superiores além de uma simulação computacional dos aspectos inferiores (Bedau, 2003, p. 15). Ocorre que isso é justamente a proposta de Fredkin: todas as leis e propriedades de nível superior são definidas a partir de uma computação básica, de tal modo que não seria possível expressar totalmente o que acontece nas escalas macroscópicas sem um conhecimento completo das interações no nível do programa elementar, algo que só seres fora do programa (como um demônio de Laplace ou um demônio escalar) conseguiriam. Nesse sentido, o pensamento de Fredkin parece admitir pelo menos uma noção fraca de emergência.

É verdade que Crutchfield e Mitchell conseguiram identificar padrões de alto nível (assim como conseguimos identificar *gliders* no Jogo da Vida) e associar operações computacionais a tais padrões que culminam em todas as células esvaziarem individualmente. Não obstante, se essas “partículas” de alto nível podem ser completamente definidas pelo comportamento de baixo nível e não há uma evidência clara de que elas não sejam meros epifenômenos resultantes da evolução das células, então poderíamos dizer que, ao fim e ao cabo, tudo que ocorre é simplesmente a aplicação da regra de nível inferior, ainda que os autores tenham percebido padrões peculiares em uma perspectiva mais ampla. Poderíamos dizer que o resultado final com todas as células vazias foi alcançado simplesmente pela aplicação da regra Φ_{17083} no nível físico e nada mais do que isso – simplesmente havia mais células vazias do que preenchidas no início, o autômato alcança um estado global estável com todas as células vazias quando essa regra é aplicada e esse resultado final pode ser perfeitamente justificado pela ação da regra (um demônio de Laplace poderia prever esse resultado apenas analisando o nível físico). Ainda que seja possível apresentar uma descrição de alto nível que defina um algoritmo mais elaborado para solucionar o problema da classificação, tal descrição não implica a existência de aspectos emergentes no sentido forte.

3.5.4. Limite de Informação

Um outro argumento ligado a problemas de reducionismo em um mundo digital seria a

limitação de informação em nosso mundo. De acordo com uma estimativa feita por Lloyd (2002), levando em conta todas as leis conhecidas da física, nosso universo suportaria cerca de 10^{120} operações elementares de computação quântica em cerca de 10^{90} *bits* registrados em campos quânticos (ou 10^{120} *bits*, se graus de liberdade gravitacionais forem levados em conta). Se considerarmos escalas ainda menores do que as conhecidas atualmente, talvez esse número aumente, mas ainda haveria uma quantidade de informação finita no nível físico. Mas, se o mundo não obedecer um reducionismo escalar, além da informação do nível físico, também seria preciso especificar a informação dos objetos macroscópicos em níveis superiores.

Considere, por exemplo, a eficácia enzimática de uma proteína. Há 20 variedades de aminoácidos, portanto uma cadeia peptídica de n aminoácidos pode ser arranjada em 20^n formas diferentes, de modo que cada sequência poderia assumir uma grande quantidade de conformações. Supondo que cada aminoácido possa assumir, digamos, 5 orientações diferentes, então o número total de conformações será de 5^n . Combinando 20^n com 5^n teremos cerca de 10^{2n} estruturas moleculares diferentes. Se n for da ordem de 100 para uma dada proteína, teríamos cerca de 10^{200} estruturas moleculares. Se a computação de cada estrutura exigir pelo menos um *bit* de informação, então tratar computacionalmente todos os aspectos dessa proteína exigiria uma quantidade informacional maior do que os 10^{120} *bits* disponíveis em nível físico. Se o programa de nosso universo exigir informações adicionais para tratar de objetos macroscópicos, então apenas a informação de nível mais básico não seria suficiente para descrever toda a realidade natural⁸⁰.

Essa questão é tratada em Pexton (2015). O autor apresenta a ilustração de um grande arquiteto de *software* que pretende desenvolver uma cópia exata de nosso universo (Pexton, 2015, p. 310). Contudo, ele possui uma disponibilidade limitada de informação para desenvolver essa simulação, mais precisamente, a mesma quantidade de informação presente em nosso mundo natural. Se a informação do nível fundamental não for suficiente para definir a informação de entidades macroscópicas, então utilizar todos os *bits* disponíveis apenas no nível básico não garantirá que o programa contemple essas entidades de nível superior. Nas palavras do autor:

80 Recentemente, Ringel e Kovrizhin (2017) descobriram que nem mesmo computação quântica seria capaz de tratar adequadamente todos os efeitos quânticos conhecidos com a quantidade de informação que temos disponível em nosso universo. Esse estudo confronta os adeptos da visão computacional de natureza, mas, como é comentado na próxima nota, o modelo de Fredkin pressupõe uma realidade externa que não necessariamente apresenta os mesmos limites de informação do nosso mundo físico.

Qualquer processo suficientemente complexo, como a vida, não aparecerá na simulação, uma vez que não há informação suficiente no computador para calcular os resultados corretos para esses sistemas dadas as restrições informacionais na simulação. Mas, se o Arquiteto permitir ao computador utilizar os termos gerais e dependências estruturais de entidades de alto nível como unidades básicas para computação, então a simulação pode recriar fenômenos como a vida. O Arquiteto poderia conseguir isso não apenas programando as regras microfísicas, mas também regras para sistemas agregados em níveis mais altos. Ou o Arquiteto poderia programar a simulação com a habilidade heurística de reconhecer termos dos tipos de nível mais alto quanto novos sistemas agregados são formados (a simulação poderia então descobrir e aplicar novas regras probabilísticas, novos padrões contrafactuais, que se aplicassem a esses termos de tipos). De qualquer modo, se o Arquiteto quer simular processos complicados como a vida usando apenas a quantidade de informação que nosso universo possui, então ele deve permitir que o programa use entidades de alto nível como termos básicos nos cálculos para comprimir os algoritmos da simulação. O Arquiteto não pode evitar ter esses termos à mão, ou colocá-los em um conjunto de regras nas quais o programa possa reconhecer quando deve mudar para o uso desses termos. Apenas se a quantidade de informação microfísica fosse infinita o Arquiteto poderia evitar usar esses termos de alto-nível. (Pexton, 2015, p. 310)

Em suma, para um arquiteto interessado em simular processos de alto nível utilizando a informação disponível em nosso universo, algoritmos baseados nas entidades de nível superior teriam que ser implementados.

Pexton apresentou essa ilustração para um cenário que admite entidades emergentes, no sentido de que são ontologicamente dependentes das entidades de nível básico (isto é, não podem existir se as entidades de nível inferior não existirem), mas apresentam algum aspecto irreduzível que não pode ser definido nem mesmo em princípio em termos do que ocorre no nível inferior (Pexton, 2015, p. 309). Isso significa que se todo o nível microscópico de um mundo for copiado, não será suficiente para garantir a existência de entidades em nível macroscópico ou, considerando uma análise por sonda epistemológica, mesmo um demônio escalar não conseguiria definir como seriam as entidades macroscópicas partindo do nível microscópico e das leis de ponte entre níveis. Assim sendo, um programa que processa apenas a informação do nível básico não conseguiria obter todas as informações necessárias para processar o que ocorre em níveis menos básicos.

Note, porém, que essa situação não cabe no mundo fredkiniano. No universo definido pela MD, apenas o autômato celular básico é suficiente para gerar toda a diversidade da natureza. A emergência na MD, como já mencionado na seção anterior, é do tipo mais fraco. Embora padrões e comportamentos novos possam ser definidos em alto nível em princípio, eles são completamente redutíveis à dinâmica do nível das células. Um computador rodando o Jogo da Vida não precisa de memória adicional para que um *glider* seja formado – nesse caso, células vivas na configuração correspondente e as regras de transição bastam para que toda a

dinâmica de nível superior apareça e não haja necessidade de informação adicional⁸¹.

3.5.5. O Problema da Consciência

Um último desafio ao reducionismo no mundo digital de Fredkin que vale a pena mencionar seria a dificuldade em explicar certos aspectos relacionados à mente. Embora o autor não chegue a tratar de detalhes sobre o assunto na maior parte de sua obra, há um artigo em específico (Fredkin, 1982) onde esse tema é trabalhado como tópico central. Nesse trabalho, Fredkin examina o conceito de “alma” (o que podemos entender como mente) e a define como um elemento informacional, associado a um arranjo de elementos materiais (Fredkin, 1982, p. 5). Se lembrarmos que o conceito de informação para Fredkin é definido como um estado que tenha significado para algum processo, então mente seria uma determinada configuração de um corpo que pode ser reconhecida por um processo. Em outras palavras, mente poderia ser encarada como um conjunto de dados passível de identificação por um programa. Segundo o autor:

A alma é uma entidade informacional, que é construída a partir de estados e arranjo de coisas materiais. (Fredkin, 1982, p. 5)

O autor diferencia entre a noção de “alma estática” e “alma dinâmica”. A primeira seria simplesmente informação digital associada a programas e dados armazenados na memória de um computador, enquanto a segunda deve ser entendida como a consequência da informação digital sendo processada por algum programa em funcionamento – a alma dinâmica estaria no controle do corpo assim como um computador está no controle de um robô (Fredkin, 1982, pp. 5-6). Esse modo de pensar possibilita uma ideia de “sobrevivência da alma” – se os dados de sua mente forem salvos em um sistema de armazenamento, eles poderiam ser reproduzidos em qualquer computador capaz de rodar um programa que reconheça aqueles dados. Também permitiria a ideia de que uma alma existe antes do nascimento – se os dados existiam gravados em um *pendrive* antes de serem iniciados pelo

81 Vale lembrar que a definição de informação de Fredkin envolve algo que possua significado no contexto de algum processo. No caso do processo que regula as transições da natureza, apenas o nível físico possui significado e, por isso, apenas o que ocorre nas células contará como informação para o programa. Isso não implica que elementos definidos em alto nível não possam ter significado para algum outro processo – talvez para as operações de um demônio escalar localizado no Outro. Como as propriedades do Outro não são completamente definidas, seria possível que a quantidade de *bits* exigida para tratar da computação dessas entidades de alto nível esteja disponível sem limitações no Outro.

programa que roda o seu nascimento, então sua alma tem alguma existência antes disso (Fredkin, 1982, pp. 19-21).

Em suma, se o universo é um autômato celular que evolui de modo a gerar padrões que podem ser definidos em alto nível, então muitos desses padrões poderiam ser identificados por algum processo externo (talvez um *software* do Outro) como entidades dotadas de alma (ou, em nossos termos, mente) por exibirem comportamentos que podem ser identificados com atividades como pensar, sentir, aprender, etc. Assim, uma pessoa expressando alguma emoção ao ouvir uma música seria um ser mental no contexto de um processo que identifique essa situação de alto nível como uma expressão de atividade mental, ainda que tudo isso seja apenas resultado de uma configuração específica dos estados das células em baixo nível.

Utilizando a terminologia da filosofia da mente, a noção fredkiniana de mente parece se aproximar de uma visão conhecida como funcionalismo. Grosso modo, o funcionalismo considera que sistemas possuem mente com base na função que seus componentes desempenham e não com base no que esses componentes são feitos. Sendo assim, computadores poderiam apresentar atividade mental por exercerem algumas funções relacionadas a atividades cognitivas – como jogar xadrez ou realizar operações matemáticas. De acordo com Fredkin, a mente é associada à informação digital que pode ser processada por um programa, portanto, estados mentais dependeriam apenas do modo como a estrutura dos componentes é examinada por um processo em especial, independentemente do tipo de elemento arranjado. Diante disso, Fredkin parece adotar um tipo de funcionalismo.

O problema é que a visão funcionalista pode encontrar obstáculos quando se tenta explicar consciência. Mais precisamente, explicar o “problema difícil” da consciência, isto é, buscar um entendimento de certos aspectos qualitativos e subjetivos da experiência consciente. Embora métodos das ciências naturais atuais consigam oferecer explicações envolvendo a dinâmica e a estrutura da experiência consciente, não é uma tarefa tão simples revelar o que exatamente seria o conteúdo da consciência – com os recursos da ciência contemporânea não é muito difícil mostrar quais regiões do cérebro estão ativas quando nossa retina é atingida por ondas eletromagnéticas com frequência correspondente à cor vermelha, mas isso é bem diferente de entender o que exatamente seria a sensação de observar um objeto vermelho. Segundo o filósofo da mente David Chalmers:

O que torna o problema difícil difícil e quase único é que ele vai além de problemas sobre o desempenho de funções. Para ver isso, note que mesmo quando explicamos o desempenho de todas as funções cognitivas e comportamentais na vizinhança da experiência – discriminação perceptual, categorização, acesso interno, relato verbal – ainda pode haver uma outra questão sem resposta: *Por que o desempenho dessas funções é acompanhado de experiência?* (Chalmers, 1995, p. 202. Grifo no original)

Compreender o papel das experiências conscientes na atividade mental é um problema enfrentado por muitos trabalhos em filosofia da mente. Não temos condições de explorar todos os desdobramentos dessa discussão no presente texto, mas devemos observar que se a consciência envolver um aspecto novo, algo que não pode ser redutível a explicações físicas, então uma visão completamente reducionista encontrará problemas.

Vimos que Fredkin entende a mente em termos de informação digital – alguém apresenta estados mentais com base no arranjo de componentes básicos. Mas isso não parece responder o desafio do problema difícil da consciência. Não fica muito claro como a sensação de “vermelhidão” que temos ao olhar um tomate pode ser totalmente definida a partir de dados digitais. Por outro lado, isso não significa que necessariamente elementos relativos à consciência (incluindo propriedades qualitativas como cores e sabores) sejam irredutíveis a propriedades mais básicas – afinal, se padrões no nível de *design* apresentam aspectos que não estão contemplados na programação básica, talvez aspectos de consciência possam ser explicados pelo nível mais básico de alguma maneira que ainda não conhecemos⁸². Sendo o caso, um demônio escalar poderia utilizar as leis de ponte que ligam esses níveis para visualizar como a consciência surge. Não obstante, também não podemos descartar completamente a possibilidade de a consciência, de fato, ser algo especial e irredutível a um nível físico. Nessa segunda opção, o reducionismo falharia e seria necessário entender o mundo digital sob outra perspectiva. Por enquanto, a questão permanece em aberto.

82 Talvez consciência seja um aspecto elementar presente nos elementos do Outro que definem o programa de nossa realidade, uma visão conhecida como panpsiquismo, ou, no mínimo, os elementos básicos de nossa realidade computacional apresentariam propriedades especiais precursoras de consciência quando tomadas em conjunto, visão conhecida como pamprotopsiquismo (Chalmers, 2013).

Considerações Finais

Este trabalho examinou a perspectiva de natureza tomada por Edward Fredkin sob o ponto de vista do reducionismo, isto é, a admissão da possibilidade da diversidade do mundo natural ser totalmente definida em termos de uma ontologia mínima constituída de espaço discreto, tempo discreto, estados diferenciáveis que exprimem informação digital em cada unidade de espaço e uma regra de transição que define como esses estados se modificam no decorrer do tempo de acordo com os estados presentes na vizinhança espacial. Basicamente, o universo se comportaria como algum tipo de autômato celular determinístico que admite computação universal, conservação de informação e reversibilidade.

Em linhas gerais, a visão reducionista considerada neste texto, apoiada na visão de Weinberg, alega que as disciplinas científicas trabalham com elementos que são completamente determinados por princípios mais básicos. A biologia, por exemplo, investiga seres vivos que são totalmente determinados por princípios químicos que, por sua vez, são totalmente definidos por princípios físicos. Uma visão reducionista mais trivial alega que certas propriedades são completamente definidas por propriedades mais básicas: a dureza do diamante seria uma propriedade inteiramente determinada pelo arranjo dos átomos de carbono, por exemplo. Seja como for, o reducionista concordará com a afirmação de que um elemento X é totalmente definido por algum outro elemento Y.

Seguindo essa linha de raciocínio, a visão reducionista de natureza mais radical considerará que toda a diversidade do mundo natural se resumirá a uma dinâmica subjacente em algum nível fundamental. Podemos assim definir uma ontologia geral para representar essa estrutura básica. No caso da visão defendida por Fredkin, essa ontologia geral, como mencionado acima, seria constituída de um autômato celular básico. Assim como ocorre em autômatos celulares que conhecemos, como o Jogo da Vida e o autômato celular unidimensional elementar, padrões com comportamentos interessantes podem ser definidos quando consideramos uma escala maior. Tomando o exemplo do Jogo da Vida, entidades como *gliders* podem ser definidas como objetos que se deslocam pelo tabuleiro do autômato, ainda que a programação básica atuando nas células não envolva movimento – apenas alteração de cada célula individualmente. Semelhantemente, os padrões que identificamos como o comportamento de objetos macroscópicos seria resultado da ação coletiva de todas as células do autômato celular básico. Ainda que se possa argumentar uma emergência de

entidades novas, se conseguirmos traçar uma correspondência entre o comportamento básico com esse comportamento de alto nível, o reducionismo se manterá – se *gliders* não são nada além de um comportamento coletivo de células e pode ser totalmente explicado em termos das células, então uma visão reducionista ainda pode ser mantida, já que a redução pode ser justificada pelo menos em princípio. Podemos, porém, argumentar que não temos capacidade cognitiva o suficiente para prever todos os padrões que se formarão em alto nível e apenas uma sonda epistemológica (como um demônio escalar) que atua no universo inteiro seria capaz de definir o comportamento de alto nível em termos do nível básico de modo satisfatório. Mas se essa definição pode ser feita pelo menos em princípio, então o reducionismo pode ser mantido em dimensão ontológica, embora seja compatível com uma emergência em sentido fraco.

Há, contudo, pelo menos dois pontos que merecem destaque como pontos problemáticos no pensamento de Fredkin.

O primeiro deles envolve a localização de nosso “universo computador”. Fredkin assume que, devido à contingência das regras do autômato que define nossa realidade, o computador em que vivemos pode estar localizado em uma realidade mais ampla chamada apenas de “Outro”, um local que pode conter leis físicas muito diferentes das nossas. A invocação dessa realidade é bastante conveniente, especialmente por sugerir a ideia de que nosso universo pode ter um propósito (talvez a natureza tenha sido programada para alcançar um determinado resultado) e também por permitir espaço, tempo e informação como elementos já dados em nosso nível mais básico (a origem última desses elementos teria que ser investigada no “Outro”, uma realidade fora de nosso alcance). Ocorre que essa realidade externa parece cair muito mais em um domínio metafísico do que em um domínio da experiência, carecendo de maiores evidências empíricas. Isso torna difícil estudar o Outro em termos mais próximos das práticas científicas, embora seja uma ideia bastante rica filosoficamente.

Um segundo ponto seria entender como o mundo fredkiniano pode contornar o problema difícil da consciência, já que acaba caindo em um funcionalismo. Talvez aspectos qualitativos como a sensação de ver a cor vermelha possam ser reduzidos a princípios básicos com desenvolvimentos científicos futuros, mas se não for o caso, será mais difícil sustentar que a sensação de “vermelhidão” que se tem ao ver um tomate seja simplesmente o resultado de um programa rodando em nível básico. Talvez uma saída para essa última situação seja

admitir que um dos estados básicos do autômato seja justamente algo relacionado à consciência – temos consciência pelo fato das células do programa elementar apresentarem um aspecto consciente, nos forçando a admitir que consciência é um dos princípios básicos da realidade. Teríamos assim espaço, tempo, informação, regra de transição e consciência. Mas isso é apenas uma especulação que exigiria uma formulação bem mais complexa do que nosso espaço permite.

Lembramos que a ideia de Fredkin é apenas uma entre várias correntes recentes que encaram a realidade natural em termos de computação. Fredkin, em especial, é um dos autores que levam essa hipótese às últimas consequências – nosso mundo não apenas se comporta como uma máquina computacional, mas é literalmente um computador funcionando em uma realidade mais ampla. Filosoficamente, não se pode negar que é uma ideia interessante, mas suas implicações científicas também devem ser levadas em conta – sendo verdadeira, a perspectiva de Fredkin implicará que vivemos em uma realidade física finita e determinística com espaço e tempo discretos incluída em um ambiente maior que, em tese, admite uma pluralidade de leis físicas. Talvez muito da visão sugerida não seja tão objetiva, já que reflete muito da personalidade do autor que tende a valorizar descrições computacionais como mais precisas e, portanto, estratégias de atividade científica que encaram o mundo como computador seriam preferíveis a outras. Entretanto, também é inegável que muitos fenômenos naturais são computáveis e podem ser simulados, mostrando que a natureza, de fato, possui alguma relação com computação. A tendência das pesquisas futuras será um desenvolvimento cada vez maior de computadores e, conseqüentemente, simulações de sistemas naturais serão cada vez mais frequentes. Diante disso, perspectivas como a defendida por Fredkin também aparecerão com mais frequência e reflexões filosóficas de ideias desse tipo serão cada vez mais exigidas. Esperamos que este modesto trabalho possa contribuir nesse sentido.

Referências

- Anderson, P. (1972). More is Different. *Science*, 177: 393-396.
- Aristóteles (1991). Physics (W. D. Ross, Trad.). In J. Barnes (ed.), *The Complete Works of Aristotle – Volume I*. Princeton: Princeton University Press.
- Banks, E. (1971). *Information Processing and Transmission in Cellular Automata* (Tese de doutorado). MIT Project MAC. <http://hdl.handle.net/1721.1/6891>.
- Bedau, M. (2003). Downward Causation and Autonomy in Weak Emergence. *Principia Revista Internacional de Epistemologica*, 6: 5-50.
- Bell, J. S. (1966). On the Problem of Hidden Variables in Quantum Mechanics. *Reviews of Modern Physics*, 38: 447-475.
- Bennett, C. H. (1973). Logical reversibility of computation. *IBM Journal of Research and Development*, 17 (6), 525-532.
- Beraldo-de-Araújo, A. & Baravalle, L. (2016). The Ontology of Digital Physics. *Erkenntnis*: 1-21. DOI 10.1007/s10670-016-9866-y.
- Berkelamp, E.; Conway, J. & Guy, R. (2004). *Winning Ways for Your Mathematical Plays: Volume 4*. Wellesley: A K Peters. (Obra original publicada em 1982).
- Berto, F. & Tagliabue, J. (2014). The World is Either Digital or Analogue. *Synthese*, 191 (3), 481–497.
- Bickle, J. (1998). *Psychoneural Reduction: The New Wave*. Cambridge, MA: MIT Press.
- Bohm, D. (1952). A Suggested Interpretation of the Quantum Theory in Terms of 'Hidden' Variables, I and II. *Physical Review*, 85 (2), 166-193.
- Bostrom, N. (2003). Are We Living In a Computer Simulation?. *Philosophical Quaterly*, 53 (211): 243-255.
- Carnap, R. (1934). *The Unity of Science*. London: Kegan Paul, Trench, Trubner, and Co.
- Chalmers, D. (1995). Facing up to the Problem of Consciousness. *Journal of Consciousness Studies*, 2: 200-219.
- _____ (2013). Panpsychism and Panprotopsychism. *The Amherst Lecture in Philosophy*, 8: 1-35.
- Churchland, P. (1986). *Neurophilosophy*. Cambridge, MA: MIT Press.
- Crutchfield, J. & Mitchell, M. The Evolution of Emergent Computation. *Proceedings of the National Academy of Sciences*, 92 (23): 10742-10746.
- Dennett, D. (1991). Real Patterns. *The Journal of Philosophy*, 88 (1): 27-51.

- Deutsch, D. (1985). Quantum Theory, the Church-Turing Principle and the Universal Quantum Computer. *Proceedings of the Royal Society of London A*, 400: 97–117.
- Einstein, A.; Podolsky, B & Rosen, N. (1935), Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?. *Physical Review*, 47: 777-780.
- Esfeld, M. & Sachse, C. (2011). *Conservative Reductionism*. New York: Taylor & Francis Group.
- Feynman, R. (1965). *The Character of Physical Law*. Massachusetts: MIT Press.
- _____ (1982), Simulating Physics with Computers. *International Journal of Theoretical Physics*, 21 (6/7), 467-488.
- Feyerabend, P. (1962). Explanation, Reduction and Empiricism. In: *Minnesota Studies in the Philosophy of Science III*, H. Feigl & G. Maxwell (eds.). Minneapolis: University of Minnesota Press, pp. 28-97.
- Floridi, L. (2009). Against Digital Ontology. *Synthese*, 168 (1), 151–178.
- Fodor, J. (1974). Special Sciences (or The Desunity of Science as a Working Hypothesis). *Synthese*, 28: 97-115.
- Fredkin, E. (1982). *On the Soul*. Esboço do autor: 1-31.
https://www.cs.auckland.ac.nz/~jaru003/vitae/on_the_soul.pdf
- _____ (1990). Digital Mechanics: An Information Process Based on Reversible Universal Cellular Automata. *Physica D*, 45: 254-270.
- _____ (1993). A New Cosmogony. In D. J. Matzcke (coord.), *PhysComp '92: Proceedings of the Workshop on Physics and Computation*, Los Alamitos: IEEE Computer Society Press, pp. 116–121.
- _____ (2003). An Introduction to Digital Philosophy. *International Journal of Theoretical Physics*, 42 (2), 189-247.
- _____ (2004). Five Big Questions with Pretty Simple Answers. *IBM Journal of Research and Development*, 48 (1), 31-45.
- _____ (2013). Discrete Theoretical Processes. In H. Zenil (ed.), *A Computable Universe: Understanding and Exploring Nature as Computation*. Singapore: World Scientific Publishing Co., pp. 365-380.
- Fredkin, E. & Toffoli, T. (1982). Conservative Logic. *International Journal of Theoretical Physics*, 21 (3), 219-253.
- Hacking, I. (1983). *Representing and Intervening: Introductory Topics in the Philosophy of Natural Science*. Cambridge: Cambridge University Press.
- Hagar, A. (2014). *Discrete or Continuous? The Quest for Fundamental Length in Modern Physics*. Cambridge: Cambridge University Press.

- _____ (2016). Ed Fredkin and the Physics of Information – An Inside Story of an Outsider Scientist. *Information and Culture*, 52 (3), 419-443.
- Hanson, J. & Crutchfield, J. (1992). The Attractor-Basin Portrait of a Cellular Automaton. *Journal of Statistical Physics*, 66 (5-6): 1415-1462.
- Hare, R. M. (1952). *The Language of Morals*. London: Oxford Clarendon Press.
- Heidegger, M. (1971) [1929]. Sobre a Essência do Fundamento. In: *Sobre a Essência do Fundamento, A Determinação do Ser do Ente segundo Leibniz, Hegel e os Gregos*. E. Stein (trad.). J.G.N. Moutinho (rev.). São Paulo: Livraria Duas Cidades, pp. 27-80.
- Hogan, C. (2012). Interferometers as Probes of Planckian Quantum Geometry. *Physical Review D*, 85 (6), 064007.
- Hooker, C. (1981). Towards a general theory of reduction. Part I: Historical and scientific setting. Part II: Identity in reduction. Part III: Cross-categorical reduction. *Dialogue*, 20: 38–59, 201–236, 496–529.
- Hossenfelder, S. (2011). Testing Super-deterministic Hidden Variable Theories. *Foundations of Physics*, 41 (9), 1521-1531.
- Illachinski, A. (2001). *Cellular Automata: A Discrete Universe*. Singapore: World Scientific Publishing.
- Jones, R. (2000). *Reductionism: Analysis and Fullness of Reality*. London: Buckwell University Press.
- Kim, J. (1984). Concepts of Supervenience. *Philosophy and Phenomenological Research*, 45 (2): 153-176.
- _____ (1990). Supervenience as a Philosophical Concept. *Metaphilosophy*, 21 (1-2): 1-27.
- _____ (1992). Multiple Realization and The Metaphysics of Reduction. *Philosophy and Phenomenological Research*, 52 (1): 1-26.
- _____ (1999). Making Sense of Emergence. *Philosophical Studies*, 95: 3-36.
- _____ (2008). Reduction and Reductive Explanation: Is One Possible without the Other?. In J. Hohwy & J. Kallestrup (eds.), *Being Reduced: New Essays on Reduction, Explanation, and Causation*. New York: Oxford University Press, pp. 93-114.
- Kleene, S. (1967). *Mathematical Logic*. New York: Wiley.
- Kuhn, T. (2011) [1962]. *A Estrutura das Revoluções Científicas*. São Paulo: Perspectiva.
- Ladyman, J.; Ross, D.; Spurrett, D. & Collier, J. (2007). *Every Thing Must Go*. Oxford: Oxford University Press.
- Land, M. & Belew, R. (1995). No Perfect Two-State Cellular Automata for Density Classification Exists. *Physical Review Letters*, 74 (25): 1548-1550.
- Landauer, R. (1961). Irreversibility and Heat Generation in the Computing Process. *IBM Journal of Research and Development*, 5: 183-191.

- Lewis, D. (1986). *Philosophical Papers, vol. II*. Oxford: Oxford University Press.
- Longo, G. O. & Vaccaro, A. (2013). *Bit Bang: La Nascita della Filosofia Digitale*. Santarcangelo di Romagna: Maggioli.
- Lloyd, S. (2002). Computational Capacity of the Universe. *Physical Review Letters*, 88 (23): 237901 1-4.
- _____ (2006). *Programming the Universe: A Quantum Computer Scientist Takes on the Cosmos*. New York: Knopf.
- Margolus, N. (1984). Physics-like Models of Computation. *Physica D*, 10: 81-95.
- McLaughlin, B. & Bennett, K. (2014). Supervenience. In: Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*. <http://plato.stanford.edu/archives/spr2014/entries/supervenience/>.
- Miller, D. B. & Fredkin, E. (2005). *Two State, Reversible, Universal Cellular Automata in Three Dimensions*. <http://arxiv.org/abs/nlin/0501022>.
- _____ (2012). *Circular Motion of Strings in Cellular Automata, and Other Surprises*. <http://arxiv.org/abs/1206.2060>.
- Mitchell, M. (1996). *An Introduction to Genetic Algorithms*. Cambridge: MIT Press.
- Nagel, E. (2006) [1961]. *La Estructura de la Ciencia: Problemas de la Lógica de la Investigación Científica*. Néstor Míguez (Trad.). Barcelona: Paidós.
- Noether, E. (1918). Invariante Variationsprobleme. *Nachr. d. Konig. Gesellsch. d. Wiss. zu Gottingen, Math-phys. Klasse*, 235-257.
- Pessoa Jr., O. (2005). Fisicalismo Redutivo e Sondas Epistemológicas. In: *Anais do III encontro da rede paranaense de pesquisa em história e filosofia da ciência*. E.S.O.; Calazans, A.; Calazans, V.F.B. (orgs.). Curitiba: SCHLA/UFPR, pp. 179-190.
- _____ (2006). *Conceitos de Física Quântica: Volumes I & II*. São Paulo: Livraria da Física.
- _____ (2011). *Definição de Propriedades "Superempíricas" Como Relações Entre Fatias do Universo*. Recuperado de: <http://www.fflch.usp.br/df/opessoa/Pessoa-AFHIC-4.pdf>.
- Pexton, M. (2015). Emergence and Fundamentality in a Pancomputationalist Universe. *Minds and Machines*, 25(4): 301-320.
- Piccinini, G. & Scarantino, A. (2011). Information Processing, Computation, and Cognition. *Journal of Biological Physics*, 37 (1), 1-38.
- Putnam, H. (1967). Psychological Predicates. In: W.H. Capitan and D.D. Merrill (eds.), *Art, Mind, and Religion*. Pittsburgh: University of Pittsburgh Press, 37-48.
- Poundstone, W. (1985). *The Recursive Universe: Cosmic Complexity and the Limits of Scientific Knowledge*. Oxford: Oxford University Press.

- Rendell, P. (2016). *Turing Machine Universality of the Game of Life*. Cham: Springer.
- Ringel, Z. & Kovrizhin, D. (2017). Quantized gravitational responses, the sign problem, and quantum complexity. *Science Advances*, 3 (9): e1701758.
- Ruse, M. (2005). Reductionism. In: T. Honderich (Ed.), *The Oxford Companion to Philosophy* (p. 793). New York: Oxford University Press.
- Salmon, W. (1989) *Four Decades of Scientific Explanation*. Pittsburgh: University of Pittsburgh Press.
- Schaffer, J. (2003). Is There a Fundamental Level?. *Noûs*, 37 (3): 498-517.
- Schaffner, K. (1967). Approaches to Reduction. *Philosophy of Science*, 34: 137-147.
- Schmidhuber, J. (1997). A Computer Scientist's View of Life, Universe, and Everything. In C. Freska (ed.), *Foundations of Computer Science: Potential – Theory – Cognition*, Lecture Notes in Computer Science. Springer, pp. 201-208.
- Searle, J. (1992). *The Rediscovery of Mind*. Cambridge: The MIT Press.
- Shannon, C. (1948). A Mathematical Theory of Communication. *Bell System Technical Journal*, 27: 379–423, 623–656.
- Silberstein, M. (2002). Reduction, Emergence and Explanation. In P. Machamer & M. Silberstein (orgs.), *The Blackwell Guide of the Philosophy of Science*. Oxford: Blackwell Publishers, pp. 80-107.
- Sipser, M. (2013). *Introduction to the Theory of Computation*. Third Edition. Boston: Cengage Learning.
- Sklar, L. (1967). Types of Inter-theoretic Reduction. *The British Journal of the Philosophy of Science*, 18: 109-124.
- Smolin, L. (2001). *Three Roads to Quantum Gravity*. New York: Basic Books.
- Tegmark, M. (2007). The Mathematical Universe. *Foundations of Physics*, 38: 101-150.
- 't Hooft, G. (2001). How Does God Play Dice? (Pre-)Determinism at the Planck Scale. In R.A. Bertlmann & A. Zeilinger (eds.), *Quantum [Un]speakables, From Bell to Quantum Information*. Berlin: Springer-Verlag, pp.307-316.
- _____ (2015). *The Cellular Automaton Interpretation of Quantum Mechanics: A View on the Quantum Nature of our Universe, Compulsory or Impossible?*. <https://arxiv.org/abs/1405.1548v3>.
- Turing, A. M. (1936). On Computable Numbers, with an Application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society*, Series 2, 42: 230-265.
- van Riel (2014). *The Concept of Reduction*. Dordrecht: Springer.

Weinberg, S. (1996) [1992]. *Sonhos de uma Teoria Final: A Busca das Leis Fundamentais da Natureza*. Rio de Janeiro: Rocco.

_____ (2001). *Facing Up: Science and its Cultural Adversaries*. Cambridge: Harvard University Press.

Wolfram, S. (1983). Statistical Mechanics of Cellular Automata. *Reviews of Modern Physics*, 55: 601-644.

_____ (1984). Universality and Complexity in Cellular Automata. *Physica 10D*: 1-35.

_____ (2002). *A New Kind of Science*. Champaign: Wolfram Media.

Wright, R. (1988). Did the Universe Just Happen?. *The Atlantic*, 261 (4).
<http://www.theatlantic.com/past/docs/issues/88apr/wright.htm>.

Zuse, K. (1970) [1969]. *Calculating Space*. Cambridge: MIT Press.