



FILHOS DA MENTE

O Futuro da Inteligência Humana e Robótica



Digitalizado, Corrigido e Adaptado por

Gullan Greyl

<http://www.gullangreyl.pt>

1ª Edição, 1988

26-06-2025

Síntese

Imagine assistir a uma palestra na viragem do século XX em que Orville Wright especula sobre o futuro dos transportes, ou uma em que Alexander Graham Bell perspetiva as comunicações via satélite e as bases de dados globais. *Filhos da Mente*, escrito por um roboticista de renome internacional, oferece uma experiência comparável — um vislumbre alucinante de um mundo que em breve poderemos partilhar com a nossa descendência artificial. Repleto de novas ideias e revelações, este livro é uma das visões de futuro mais envolventes e controversas alguma vez escritas por um académico sério.

Hans Moravec defende convincentemente que estamos a aproximar-nos de um ponto de viragem na história da vida — um momento em que as fronteiras entre a inteligência biológica e a pós-biológica começarão a dissolver-se. Dentro de quarenta anos, acredita Moravec, alcançaremos a equivalência humana nas nossas máquinas, não só na sua capacidade de raciocínio, mas também na sua capacidade de perceber, interagir e modificar o seu complexo ambiente. O fator crítico é a mobilidade. Um computador preso a um lugar está condenado a iterações estáticas, enquanto uma máquina à espreita, como um organismo móvel, precisa de desenvolver um fundo mais rico de conhecimento sobre um mundo em constante mudança no qual basear as suas ações.

Índice

Prólogo	1
CAPÍTULO 1	6
MENTE EM MOVIMENTO	6
Máquinas Que Pensam (Fracamente)	7
Máquinas Que Veem (obscuramente) e Agarram (desajeitadamente)	10
Máquinas Que Exploram (de forma hesitante)	15
Um Robô Para as Massas	27
Locomoção do Ponto de Equilíbrio	32
Manipulação do Ponto de Equilíbrio	35
Navegação de Ponto de Equilíbrio	38
Reconhecimento do Ponto de Equilíbrio	42
Processamento e Coordenação	44
Um Robô Sensato	45
A Evolução Convergente das Emoções e da Consciência	46
Aprendizagem	55
Imagens	59
CAPÍTULO 2	63
REFORÇANDO O DESEMPENHO	63
Circuitos Neurais	65
Células e Ciclos	71
Memória	74
Comparando Computadores	76
Projeções	80
Equivalência Humana em 40 Anos	84
CAPÍTULO 3	93
SIMBIOSE	93
Programas Armazenados e Montadores	93
Compiladores e Sistemas Operativos	98
Menus e Ícones	102
Óculos Mágicos	106
Encontrando-se a Si Mesmo	110
Indo a Lugares	112
Programação Sem Ferramentas	117

Física Elementar	121
CAPÍTULO 4	126
CLÁUSULA DO AVÔ	126
Robôs Arbustos	129
Transmigração.....	137
Muitas Mudanças.....	142
O Que Sou Eu?.....	147
Despertar o Passado	154
CAPÍTULO 5	157
VIDA SELVAGEM	157
Cavalos de Troia, Bombas-Relógio e Vírus.....	158
Fantasmas	165
Geração Espontânea	167
Uma Advertência para o SETI.....	171
A Peste como Algo Positivo.....	175
Altruísmo Egoísta	177
CAPÍTULO 6	183
EMANCIPAÇÃO	183
O Universo Pensante e Mais Além.....	185
Newway e os Cellticks	189
HashLife.....	192
O Caminho à Frente	197
APÊNDICE 1	199
Retinas e Computadores	199
APÊNDICE 2	206
Medindo a Potência do Computador	206
Uma Metáfora Náutica	210
APÊNDICE 3	216
Os Limites Externos da Computação	216
Mente sem Máquina?.....	216
Pensamento Não Determinístico.....	217
Replicação e Não Determinismo Quântico.....	220
Muitos Mundos e o Computador NP do Juízo Final	226
O Dispositivo do Juízo Final.....	227

Um Mundo, Não Muitos?	231
Ondas Grandes.....	236
BIBLIOGRAFIA	239

Prólogo

Envolvidos durante milhares de milhões de anos numa implacável e crescente corrida armamentista, os nossos genes acabaram por se superar. Produziram uma arma tão poderosa que aniquilará tanto vencidos como vencedores. Este engenho não é a bomba de hidrogénio — a utilização generalizada de armas nucleares apenas atrasaria o fim imensamente mais interessante que foi concebido.

O que nos aguarda não é o esquecimento, mas sim um futuro que, da nossa perspetiva atual, é melhor descrito pelas palavras "pós-biológico" ou até "sobrenatural". É um mundo no qual a raça humana foi varrida pela maré da mudança cultural, usurpada pela sua própria descendência artificial. As consequências últimas são desconhecidas, embora muitos passos intermédios não só sejam previsíveis como já foram dados. Hoje, as nossas máquinas ainda são criações simples, que exigem os cuidados parentais e a atenção constante de qualquer recém-nascido, dificilmente dignas da palavra "inteligente". Mas dentro do próximo século, amadurecerão e tornar-se-ão entidades tão complexas como nós, e eventualmente em algo que transcende tudo o que conhecemos — em quem poderemos ter orgulho quando se referirem a si próprios como nossos descendentes.

Libertados do ritmo moroso da evolução biológica, os filhos das nossas mentes serão livres para crescer e confrontar desafios imensos e fundamentais no universo. Nós, humanos, beneficiaremos por algum tempo dos seus esforços, mas mais cedo ou mais tarde, como filhos naturais, procurarão o seu próprio destino enquanto nós, os seus pais envelhecidos, nos desvaneceremos silenciosamente. Muito pouco se perderá nesta passagem da tocha — estará no poder dos nossos descendentes artificiais, e para seu benefício, recordar quase tudo sobre nós, talvez até mesmo o funcionamento detalhado das mentes humanas individuais.

O processo teve início há cerca de 100 milhões de anos, quando certas linhagens genéticas descobriram uma forma de criar animais com a capacidade de aprender alguns comportamentos com os seus progenitores ao longo da vida, em vez de os herdarem todos na conceção. A complexidade aumentou há 10 milhões de anos, quando os nossos ancestrais primatas começaram a depender de ferramentas feitas de ossos, paus e pedras, e acelerou novamente com o domínio do fogo e o desenvolvimento de linguagens complexas há cerca de 1 milhão de anos. Quando a nossa espécie surgiu, há aproximadamente 100 mil anos, a

evolução cultural — o gigante que os nossos genes tinham, sem querer, construído — avançava com um ímpeto irresistível.

Nos últimos 10 mil anos, as mudanças no património genético humano foram insignificantes em comparação com os avanços cada vez maiores da cultura humana. Testemunhámos primeiro uma revolução agrícola, seguida pelo estabelecimento de governos burocráticos em larga escala com poder para cobrar impostos para o seu sustento, o desenvolvimento de linguagens escritas e o surgimento de classes de lazer com tempo e energia para se dedicarem a questões intelectuais. No último milhar de anos ou menos, invenções como a impressão com caracteres móveis aceleraram consideravelmente o fluxo de informação cultural e, conseqüentemente, o seu ritmo evolutivo.

Com o advento da revolução industrial, há 200 anos, entrámos na fase final, uma em que os substitutos artificiais das funções do corpo humano, como levantar e transportar, se tornaram cada vez mais economicamente atrativos — na verdade, indispensáveis. Depois, há 100 anos, com a invenção de máquinas de calcular práticas, fomos capazes, pela primeira vez, de duplicar artificialmente algumas funções pequenas, mas complexas, da mente humana. O poder computacional dos dispositivos mecânicos tem aumentado mil vezes a cada 20 anos desde então.

Estamos muito perto da altura em que praticamente nenhuma função humana essencial, física ou mental, carecerá de uma contrapartida artificial. A materialização desta convergência de desenvolvimentos culturais será o robô inteligente, uma máquina que pode pensar e agir como um humano, por mais inumana que possa ser em detalhes físicos ou mentais. Tais máquinas poderiam dar continuidade à nossa evolução cultural, incluindo a sua própria construção e um autoaperfeiçoamento cada vez mais rápido, sem nós e sem os genes que nos construíram. Quando isso acontecer, o nosso ADN ver-se-á sem trabalho, tendo perdido a corrida evolutiva para um novo tipo de competição.

A. G. Cairns-Smith, um químico que se debruçou sobre os primórdios da vida na Terra primitiva, designa este tipo de golpe interno como uma aquisição genética. Ele sugere que tal fenómeno já ocorreu pelo menos uma vez. Em "*Seven Clues to the Origin of Life*" (*Sete Pistas para a Origem da Vida*), Cairns-Smith argumenta que os precursores da vida tal como a conhecemos eram cristais microscópicos de argila, que se reproduziam através do simples processo de crescimento cristalino. A maioria dos cristais é marcada por padrões de desorganização na disposição ordenada dos seus átomos, muitos dos quais se propagam à medida que o cristal cresce. Se o

cristal se fraturar, cada pedaço pode herdar uma cópia do padrão, por vezes com uma ligeira alteração. Tais defeitos podem ter um efeito dramático nas propriedades físicas e químicas de uma argila. Cristais que partilham um padrão de desorganização podem formar aglomerados densos, enquanto outros com um padrão diferente podem agregar-se numa massa esponjosa. A água que transporta minerais pode ser desviada de um tipo, mas escoar através do outro, fornecendo matérias-primas para o crescimento contínuo. Os padrões também afetam o crescimento indiretamente, modulando a química de outras moléculas no seu ambiente. As argilas são poderosos catalisadores químicos; os pequenos cristais possuem uma enorme área de superfície total, à qual as moléculas podem aderir em certas configurações, dependendo da forma externa do cristal e da molécula em questão. Estes cristais comuns possuem, assim, os elementos essenciais para a evolução darwiniana — reprodução, herança, mutação e diferenças no sucesso reprodutivo.

Na teoria de Cairns-Smith, a primeira aquisição genética começou quando algumas espécies de argila, em vigorosa competição darwiniana entre si, começaram a codificar alguma informação genética externamente em longas moléculas de carbono. Tais polímeros são mais estáveis do que os próprios padrões de desorganização, facilmente perturbáveis, e os organismos que os usaram em maior extensão reproduziram-se com mais sucesso. Embora totalmente dependentes, no início, da maquinaria química existente baseada em cristais, à medida que estas moléculas de carbono assumiram uma maior parte do papel reprodutivo, tornaram-se menos dependentes dos cristais. Com o tempo, o simples andaime cristalino desapareceu completamente, deixando na sua esteira evolutiva o sistema complexo e interdependente de maquinaria orgânica a que chamamos vida.

Hoje, milhares de milhões de anos depois, uma nova mudança está em curso na forma como a informação passa de geração em geração. Os seres humanos evoluíram de organismos definidos quase totalmente pelos seus genes orgânicos. Atualmente, dependemos adicionalmente de um vasto e rapidamente crescente conjunto de informação cultural gerada e armazenada fora dos nossos genes — nos nossos sistemas nervosos, bibliotecas e, mais recentemente, computadores. A nossa cultura ainda depende totalmente dos seres humanos biológicos, mas a cada ano que passa, as nossas máquinas, um produto importante da cultura, assumem um papel maior na sua manutenção e crescimento contínuo. Mais cedo ou mais tarde, as nossas máquinas tornar-se-ão suficientemente conhecedoras para gerir a sua própria manutenção, reprodução e autoaperfeiçoamento sem ajuda. Quando isto acontecer, a nova aquisição genética estará completa. A

nossa cultura poderá então evoluir independentemente da biologia humana e das suas limitações, passando diretamente de geração em geração para um tipo de maquinaria inteligente cada vez mais capaz.

Os nossos genes biológicos e os corpos de carne e osso que eles constroem desempenharão um papel cada vez menor neste novo regime. Mas será que as nossas mentes, onde a cultura teve origem, também se perderão neste golpe? Talvez não. A revolução que se avizinha pode libertar as mentes humanas tão eficazmente quanto liberta a cultura humana. Na condição atual, somos seres híbridos desconfortáveis, parte biologia, parte cultura, com muitas das nossas características biológicas desalinhas com as invenções das nossas mentes. As nossas mentes e genes podem partilhar muitos objetivos comuns durante a vida, mas há uma tensão entre o tempo e a energia gastos na aquisição, desenvolvimento e disseminação de ideias, e o esforço despendido na manutenção dos nossos corpos e na produção de uma nova geração (como qualquer pai de adolescentes pode observar). A frágil trégua entre mente e corpo desfaz-se completamente quando a vida termina. Os nossos genes geralmente sobrevivem à nossa morte, agrupados de diferentes formas na nossa descendência e nos nossos familiares. De forma subtil, é sem dúvida do seu interesse evolutivo fazer regularmente experiências como esta com novas "misturas" do baralho genético. Mas o processo é devastador para a nossa outra metade. Demasiados aspetos da nossa existência mental, arduamente conquistados, simplesmente morrem connosco.

É fácil imaginar o pensamento humano libertado da servidão a um corpo mortal — a crença numa vida após a morte é comum. Mas não é necessário adotar uma postura mística ou religiosa para aceitar a possibilidade. Os computadores fornecem um modelo mesmo para o mais fervoroso mecanicista. Uma computação em progresso — o que podemos razoavelmente chamar de processo de pensamento de um computador — pode ser interrompida a meio e transferida, como programa e dados lidos da memória da máquina, para um computador fisicamente diferente, onde retoma como se nada tivesse acontecido. Imagine que uma mente humana pudesse ser libertada do seu cérebro de uma forma análoga (ainda que muito mais desafiadora tecnicamente).

Uma mente exigiria muitas modificações para operar eficazmente após ser resgatada das limitações de um corpo mortal. A mentalidade humana natural é sintonizada para a progressão de uma vida útil, desde a plasticidade impressionável até à rigidez autoconfiante, e assim é um material pouco promissor para a imortalidade. Teria de ser reprogramada para uma adaptabilidade contínua para ser viável a longo prazo. Enquanto um

organismo mortal e transitório pode deixar a tarefa da adaptação para os processos externos de mutação e seleção natural, uma mente que aspira à imortalidade, quer as suas origens sejam as de um ser humano mortal ou seja uma criação completamente artificial, deve estar preparada para adaptar-se constantemente a partir do interior.

Talvez passasse por um rejuvenescimento cíclico, adquirindo novo hardware e software em fases periódicas que se assemelham à infância. Ou talvez pudesse atualizar o conteúdo da sua mente e corpo continuamente, adicionando e eliminando, testando componentes em todo o tipo de combinações, para se manter a par das condições em mudança. O teste é de importância central: ele guia a evolução. Se o indivíduo tomar demasiadas decisões erradas nestes testes, falhará totalmente, à antiga maneira darwiniana.

Um mundo pós-biológico dominado por máquinas pensantes e autoaperfeiçoáveis seria tão diferente do nosso mundo de seres vivos como este mundo é diferente da química sem vida que o precedeu. Uma população composta por "filhos da mente" sem grilhões é bastante inimaginável. Vamos tentar imaginar algumas das consequências, de qualquer forma.

CAPÍTULO 1

MENTE EM MOVIMENTO

Acredito que robôs com inteligência humana serão comuns dentro de cinquenta anos. Em comparação, as melhores máquinas de hoje têm mentes mais parecidas com as de insetos do que com as de humanos. No entanto, este desempenho em si representa um salto gigantesco em apenas algumas décadas.

Imitações mecânicas de certas funções humanas acompanham-nos há séculos. Muitas torres de relógio medievais estão equipadas com mecanismos que assinalam as horas com elaboradas peças moralistas encenadas por santos, cavaleiros, bispos, anjos, demónios e todo o tipo de animais mecânicos. Dispositivos mais pequenos que andavam, falavam, nadavam, respiravam, comiam, escreviam com canetas de pena ou tocavam instrumentos musicais divertiram a sociedade educada pelo menos desde o século XV. Leonardo da Vinci, por exemplo, construiu elaboradas exposições mecânicas deste tipo para os seus patronos. Estas primeiras máquinas de relojoaria — movidas por água corrente, pesos em queda ou molas — copiavam os movimentos dos seres vivos, mas não conseguiam responder ao mundo à sua volta. Elas só podiam agir, por mais encantadoras que fossem.

A tecnologia elétrica, eletrónica e de rádio, desenvolvida no início deste século, tornou possíveis máquinas que podiam reagir — à luz, ao som e ao controlo remoto invisível. O resultado foi um conjunto de robôs de demonstração divertidos — bem como pensamentos e histórias sobre futuras máquinas semelhantes aos humanos. Mas apenas ligações simples entre os sensores e os motores eram possíveis no início. Estas máquinas mais recentes podiam sentir, bem como agir, mas não podiam pensar.

Máquinas Que Pensam (Fracamente)

Durante a Segunda Guerra Mundial, os computadores analógicos — máquinas que simulavam sistemas físicos representando as suas quantidades variáveis como movimentos análogos de eixos ou voltagens — foram projetados para controlar armas antiaéreas, para navegação e para bombardeamento de precisão. Alguns dos seus desenvolvedores notaram uma semelhança entre a operação destes dispositivos e os sistemas reguladores em seres vivos, e estes investigadores foram inspirados a construir máquinas que agiam como se estivessem vivas. Norbert Wiener, no Massachusetts Institute of Technology (MIT), cunhou o termo *cibernética* para este estudo unificado de controlo e comunicação em animais e máquinas. Os seus praticantes combinaram a nova teoria sobre regulação por feedback com avanços na eletrónica pós-guerra e o conhecimento inicial dos sistemas nervosos vivos para construir máquinas que eram capazes de responder como animais simples e de aprender. Os rudimentos do pensamento artificial tinham chegado.

Entre os destaques do esforço da cibernética, destaca-se uma série de tartarugas eletrónicas construídas durante os anos 50 por W. Grey Walter, um psicólogo britânico. Com cérebros eletrónicos de válvula de rádio subminiatura, olhos rotativos de fotocélula, ouvidos de microfone e sensores de interruptor de contato, as primeiras versões podiam localizar a sua estação de recarga quando as suas baterias estavam fracas e, de outra forma, evitar problemas enquanto vagueavam. Grupos delas exibiam um comportamento social complexo ao responderem às luzes de controlo e toques umas das outras. Uma máquina posterior com os mesmos sentidos podia ser condicionada a associar um estímulo a outro e podia aprender por experiência repetida que, por exemplo, um ruído alto seria seguido por um pontapé na sua carapaça. Uma vez "educada", a tartaruga evitaria um ruído tal como antes respondia a um pontapé. As associações eram lentamente acumuladas como cargas elétricas em dispositivos eletrónicos chamados capacitores, usados aqui como dispositivos de memória.

Talvez a criação mais impressionante dos ciberneticistas tenha sido o Johns Hopkins Beast. Construído por um grupo de investigadores cerebrais no início dos anos 60, vagueava pelos corredores guiado por sonar e um olho de fotocélula especializado que procurava a distintiva placa preta das tomadas de parede, onde se ligava para se alimentar. O Beast inspirou uma série de imitadores. Alguns usavam circuitos especiais conetados a câmaras de televisão em vez de fotocélulas e eram controlados por conjuntos de (então novas) portas lógicas digitais de transístores, como as que agora podem ser encontradas, em milhares e milhões, nos circuitos integrados de todos os computadores. Alguns adicionaram novos movimentos, como "Agitar para desembaraçar o braço de recarga", ao repertório de ações básicas.

O campo da cibernética prosperou por menos de duas décadas. Como tantas vezes acontece, foi eclipsado por um parente, o movimento da inteligência artificial. Os muitos pequenos computadores analógicos da guerra, que tinham inspirado a cibernética, tinham alguns primos digitais muito maiores. Estas máquinas computavam não pelas voltas medidas de eixos ou pelo fluxo de corrente, mas contando, em saltos discretos. Os primeiros computadores digitais automáticos — calculadoras enormes, imóveis e autónomas — foram concluídos no final da guerra. O Colossus, uma máquina britânica ultrassecreta que decifrou o código alemão Enigma e ajudou a mudar o curso da guerra, examinava chaves de código dezenas de milhares de vezes mais rapidamente do que seria humanamente possível. Nos Estados Unidos, o ENIAC calculava tabelas de artilharia antiaérea para o Exército e mais tarde fez cálculos para a construção da bomba atômica, a velocidades semelhantes às do Colossus.

De forma menos beligerante, estes "cérebros gigantes", como vieram a ser chamados, proporcionaram oportunidades sem precedentes para experiências em complexidade. Pioneiros como Alan Turing, um dos criadores do Colossus, e John von Neumann, que esteve envolvido nas primeiras máquinas americanas, acalentavam a esperança de que a capacidade de pensar racionalmente, o nosso trunfo único para lidar com o mundo, pudesse ser capturada numa máquina. As nossas mentes poderiam

ser amplificadas por computadores, tal como os nossos músculos tinham sido amplificados pelas máquinas a vapor da revolução industrial. Programas para raciocinar e para jogar jogos intelectuais como xadrez foram projetados por Claude Shannon do MIT e por Turing em 1950, mas os computadores mais antigos eram demasiado limitados e caros para esta utilização. Alguns programas fracos de jogo de damas apareceram nas primeiras máquinas comerciais do início dos anos 50, e programas de xadrez igualmente fracos surgiram na última metade dessa década, juntamente com um bom jogador de damas de Arthur Samuel da IBM. Depois, em 1957, Allen Newell e Herbert Simon do Carnegie Tech (agora Carnegie Mellon University) e John Shaw da RAND Corporation demonstraram o Logic Theorist, o primeiro programa capaz de raciocinar sobre assuntos arbitrários, começando com axiomas e aplicando regras de inferência para provar teoremas.

Em 1960, John McCarthy, então no MIT, cunhou o termo inteligência artificial (IA) para o esforço de fazer os computadores pensarem. Por volta de 1965, os primeiros alunos de McCarthy, Marvin Minsky (também no MIT), Newell e Simon tinham produzido programas de IA que provavam teoremas em geometria, resolviam problemas de testes de inteligência, livros de álgebra e exames de cálculo, e jogavam xadrez, tudo com a proficiência de um caloiro universitário médio. Cada programa conseguia lidar apenas com um tipo restrito de problema, mas para os primeiros esforços, estes programas eram encorajadores — tão encorajadores que a maioria das pessoas envolvidas sentiu que mais uma década de progresso produziria, com certeza, uma máquina genuinamente inteligente. Esta foi uma falha de cálculo compreensível.

Agora, um quarto de século depois, os computadores são milhares de vezes mais poderosos do que estes modelos dos anos sessenta, mas não parecem muito mais inteligentes. Por volta de 1975, o progresso na inteligência artificial tinha abrandado do rápido sprint de um punhado de entusiastas para a caminhada laboriosa de crescentes multidões de trabalhadores. Mesmo assim, sucessos modestos mantiveram a esperança acesa. Os chamados sistemas periciais, programas que codificam as regras de decisão de especialistas humanos em domínios restritos como o diagnóstico

de doenças, o agendamento de fábricas ou a configuração de sistemas informáticos, estão diariamente a justificar a sua utilidade no mundo empresarial. Um esforço de quinze anos no MIT reuniu conhecimento sobre álgebra, trigonometria, cálculo e áreas relacionadas num programa maravilhoso chamado MACSYMA, agora comercializado, que manipula fórmulas simbólicas e ajuda a resolver problemas de outra forma intimidantes. Vários programas de xadrez são agora oficialmente classificados como mestres de xadrez, e um excelente desempenho foi alcançado noutros jogos como o gamão. Existem programas semi-inteligentes que conseguem compreender inglês simplificado digitado sobre assuntos restritos e fazer deduções elementares ao responder a perguntas sobre esses textos. Alguns interpretam comandos falados escolhidos de repertórios de mil palavras, e outros conseguem fazer tarefas visuais simples, como decidir se uma peça está no seu local desejado.

Infelizmente para os robôs semelhantes a humanos, os computadores são piores a fazer as coisas que são mais naturais para os humanos, como ver, ouvir, manipular objetos, aprender línguas e raciocinar com bom senso. Esta dicotomia — máquinas que fazem bem coisas que os humanos acham difíceis, enquanto fazem mal o que é fácil para nós — é uma pista gigante para o problema de como construir uma máquina inteligente.

Máquinas Que Veem (obscuramente) e Agarram (desajeitadamente)

Em meados da década de 1960, os alunos de Marvin Minsky no MIT começaram a ligar câmaras de televisão e braços robóticos mecânicos aos seus computadores, dando olhos e mãos a mentes artificiais para que as suas máquinas pudessem ver, planear e agir. Por volta de 1965, estes investigadores tinham criado uma máquina capaz de encontrar e remover blocos brancos de uma mesa preta. Esta proeza exigiu um programa de controlo tão complexo como qualquer um dos programas de raciocínio puro da época — programas que, desimpedidos de apêndices robóticos, conseguiam, por exemplo, igualar-se a estudantes universitários do primeiro ano

na resolução de problemas de cálculo. No entanto, o sistema "mão-olho" de Minsky podia ser superado por uma criança pequena. Ainda assim, as experiências continuaram no MIT e noutros locais, desenvolvendo-se gradualmente num campo que hoje se conhece como robótica, um termo cunhado em 1942 numa história de ficção científica por Isaac Asimov a partir da palavra "robot", ela própria criada pelo dramaturgo checo Karel Capek em 1921. A robótica começou muito abaixo na escala de desempenho humano em comparação com a inteligência artificial, mas o seu progresso nos últimos vinte anos tem sido igualmente lento e difícil.

Nem todos os robôs, nem todas as pessoas, passam a vida em universidades. Muitos têm de trabalhar para viver. Mesmo antes da revolução industrial, antes de qualquer tipo de pensamento ser mecanizado, máquinas parcialmente automáticas, movidas a vento ou água corrente, eram usadas para moer grão e cortar madeira. O início da revolução industrial no século XVIII foi marcado pela invenção de uma infinidade de dispositivos que podiam substituir o trabalho manual de uma forma precisa e completamente desumana. Impulsionadas por eixos movidos a água ou vapor, estas máquinas bombeavam, batiam, cortavam, fiavam, teciam, estampavam, moviam materiais e peças, e muito mais, de forma consistente e incansável.

De vez em quando, surgia algo engenhosamente diferente. Por exemplo, o tear de Jacquard, inventado em 1801, conseguia tecer tapeçarias intrincadas especificadas por uma sequência de cartões perfurados. No início do século XX, a eletrónica tinha dado aos sistemas de controlo industrial sentidos limitados; podiam agora parar quando algo corria mal, ou controlar a temperatura, a espessura e até a consistência das suas peças de trabalho. No entanto, cada máquina fazia um trabalho e apenas um. Consequentemente, o produto muitas vezes tornava-se obsoleto antes que a máquina tivesse recuperado os seus custos de conceção e construção. Este problema tornou-se particularmente agudo no final da Segunda Guerra Mundial.

Em 1954, o inventor George Devol patenteou um novo tipo de máquina industrial: o braço robótico programável. A sua função

podia ser alterada simplesmente ao mudar o fluxo de cartões perfurados que controlavam o seu movimento. Em 1958, Devol, em parceria com Joseph Engelberger, fundou a Unimation (uma junção de "universal" e "automação") para construir tais máquinas. Os cartões perfurados foram rapidamente substituídos por memória magnética, o que permitiu que o robô fosse programado simplesmente guiando-o, por assim dizer, "pela mão" uma única vez pelos seus movimentos necessários. O primeiro robô industrial começou a trabalhar numa fábrica da General Motors em 1961. Até hoje, a maioria dos grandes robôs que soldam, pintam e movem peças de automóveis são deste tipo.

Só quando o custo dos pequenos computadores desceu para menos de 10.000 dólares é que a investigação em robótica nas universidades começou a influenciar o projeto dos robôs industriais. Os primeiros sistemas de visão industrial, capazes de localizar e identificar peças em tapetes rolantes, e geralmente acoplados a uma nova classe de pequenos braços robóticos, surgiram no final da década de 1970. Robôs capazes de ver e sentir, à sua maneira, desempenham agora um papel modesto, mas em crescimento silencioso, na montagem e inspeção de pequenos dispositivos como calculadoras, placas de circuito impresso, máquinas de escrever e bombas de água para automóveis. De facto, as necessidades industriais influenciaram fortemente a pesquisa universitária. O que antes era um número negligenciável de projetos de robôs inteligentes cresceu para centenas em campus por todo o país. E enquanto a cibernética pode agora estar relativamente dormente, o seu parente mais "sisudo", a teoria de controlo, tem estado bastante ativa desde a guerra, num esforço para satisfazer as necessidades lucrativas da indústria aeroespacial. Métodos elaborados desenvolvidos para controlar aeronaves, naves espaciais e armas estão agora a influenciar o projeto de robôs industriais.

Em 1987, tive a oportunidade de visitar a fábrica em Fremont, Califórnia, onde os computadores Macintosh da Apple eram montados. Achei a maior parte da fábrica bem organizada, mas comum. Muitos passos de montagem eram feitos manualmente. As máquinas mais eficientes eram provavelmente aquelas que

inserir componentes em placas de circuito. Atuando de forma semelhante a máquinas de costura, estas "colocadoras de placas" pegam em componentes enfiados em fitas como munição de metralhadora e "cosiam-nos" em placas de circuito impresso a uma velocidade estonteante, vários componentes por segundo, com a placa a deslizar rapidamente para a posição de cada peça, frente e verso, esquerda e direita. As máquinas são maravilhas de controlo computadorizado e muito económicas para produção de alto volume, mas deixaram-me vagamente desapontado. No entanto, num pequeno nicho, vi um dispositivo bastante diferente a inserir componentes que as máquinas de alta velocidade não conseguiam manusear. As peças eram indutores antiquados — pequenas latas metálicas contendo uma bobina de fio. Cada lata tinha patilhas metálicas que deviam encaixar em ranhuras na placa, e as bobinas terminavam em fios destinados a pequenos orifícios. Ao contrário dos componentes precisamente moldados nas fitas de alimentação das outras máquinas, os indutores, fornecidos cuidadosamente dispostos em tabuleiros de plástico texturizados, muitas vezes tinham patilhas e fios ligeiramente dobrados que seriam simplesmente danificados por uma tentativa cega de os empurrar para uma placa.

A máquina de inserção operava dentro de uma cabine de vidro. As placas e os tabuleiros de indutores chegavam e partiam em tapetes rolantes. O processo de inserção começava com uma câmara de TV que focava o tabuleiro das peças. Um programa de visão localizava um indutor, e um braço robótico rápido movia-se para o apanhar, trazendo-o, com as pontas dos fios e patilhas para cima, para a frente de uma segunda câmara de TV. Um segundo programa de visão examinava os condutores e, se estivessem suficientemente direitos, dava instruções ao braço para inserir o componente na placa. Se os condutores estivessem ligeiramente dobrados, o indutor era primeiro empurrado contra um "endireitador de pinos", um bloco de metal com orifícios afunilados, após o que os pinos eram novamente inspecionados. Se os condutores estivessem irremediavelmente danificados, o indutor era descartado para um recipiente de rejeição e outro era recolhido do tabuleiro.

A própria inserção era um processo delicado. O indutor era rapidamente levado a poucos milímetros da superfície da placa, depois lentamente baixado até que o braço robótico encontrasse resistência. O braço empurrava o indutor para a frente e para trás, mantendo uma ligeira pressão descendente, até que as patilhas e os fios encontrassem os seus orifícios. Nesse momento, aplicava maior pressão para encaixar o componente. Um cortador motorizado montado por baixo da placa cortava e dobrava o metal saliente, fixando o indutor. Se a tentativa de encaixar a peça falhasse após alguns segundos, ela seria novamente trazida para a câmara de verificação de condutores e, possivelmente, para o endireitador de pinos, antes de outra tentativa de inserção. Se uma terceira tentativa falhasse, a peça seria deitada fora e uma nova seria tentada.

Tudo isto acontecia muito rapidamente — uma peça podia ser inserida a cada três ou quatro segundos, embora uma problemática pudesse demorar até dez. Fiquei impressionado — e um pouco nostálgico. As atividades tinham uma sensação familiar. Eu tinha sido testemunha regular de versões um pouco mais lentas e desajeitadas destas, uma década antes, no Laboratório de Inteligência Artificial de Stanford (SAIL), onde eu era estudante de pós-graduação. De facto, o sistema de montagem da Apple era um produto de uma pequena empresa do sul da Califórnia chamada Adept, que pode traçar a sua ascendência até teses de doutoramento no SAIL. As sementes lançadas ali estavam a começar a brotar.

O objetivo de um desempenho semelhante ao humano por parte de robôs estacionários, embora altamente diluído por uma miríade de abordagens e metas de curto prazo, adquiriu um vigor implacável e darwiniano. Como história, torna-se desconcertante na sua diversidade e interligação. Passemos agora ao mundo mais esparso dos robôs que se movem.

Máquinas Que Exploram (de forma hesitante)

Os primeiros programas de raciocínio necessitavam de muito poucos dados para funcionar. Um tabuleiro de xadrez, ou um problema de lógica, geometria ou álgebra, podia ser descrito com algumas centenas de símbolos bem escolhidos. De forma semelhante, as regras para resolver o problema podiam ser expressas como várias centenas das chamadas "transformações" desses dados. A dificuldade residia apenas em encontrar uma sequência de transformações que resolvesse o problema, de entre o número astronómico de combinações possíveis. Era óbvio que resolver problemas em áreas menos restritas (a questão "Como posso chegar a Tombuctu?" era um exemplo retórico frequentemente usado) exigiria um acervo inicial de informação muito maior. Parecia improvável que todos os factos necessários para resolver tais problemas pudessem ser fornecidos manualmente aos programas de raciocínio.

Alguns factos poderiam ser disponibilizados se os programas pudessem ser ensinados a ler e compreender livros, mas compreender até palavras simples exigiria um conhecimento detalhado do mundo físico. Esse conhecimento é assumido como preexistente nas mentes dos leitores – nenhum livro tenta uma definição abrangente de uma rocha, uma árvore, o céu ou um ser humano. Possivelmente, parte deste conhecimento do mundo, como veio a ser chamado, poderia ser obtido pela própria máquina se esta pudesse observar diretamente os seus arredores através de olhos de câmara, ouvidos de microfone e mãos robóticas sensíveis. O desejo de automatizar a aquisição de conhecimento do mundo foi uma das primeiras justificativas para os projetos de robótica nos laboratórios de inteligência artificial. O modelo interno do mundo que estes computadores poderiam desenvolver poderia então ser usado por eles, e por outras máquinas, como base para o raciocínio de senso comum.

Embora uma máquina que se possa mover de um lado para o outro devesse ser capaz de recolher muito mais dados do que uma imóvel, a dificuldade logística de ligar um enorme computador

imóvel a um conjunto complicado de sensores numa plataforma móvel tornou os sistemas fixos "mão-olho" mais atrativos para a maioria dos investigadores. Além disso, logo se percebeu que o problema de adquirir sistematicamente informação do ambiente era muito menos fácil de resolver do que as atividades mentais que a informação se destinava a servir. Durante a década de 1970, dezenas de laboratórios de investigação tinham braços robóticos ligados a computadores, mas quase nenhum tinha veículos robóticos. A maioria dos investigadores de robótica via a mobilidade como uma complicação desnecessária para um problema já excessivamente difícil. A sua experiência contrastava fortemente com a atitude dos cibernéticos (e centenas de entusiastas e fabricantes de brinquedos), que estavam bastante satisfeitos em obter um comportamento simples, semelhante ao animal, a partir dos circuitos modestos a bordo das suas muitas pequenas máquinas móveis.

O Shakey, concluído em 1969 pelo Stanford Research Institute, foi o primeiro — e ainda é o único — robô móvel a ser controlado principalmente por programas que raciocinavam. É uma exceção que confirma a regra. Os seus criadores — Nils Nilsson, Charles Rosen e outros — foram inspirados pelo sucesso inicial na investigação em inteligência artificial. Procuraram aplicar métodos de resolução de problemas baseados em lógica a uma tarefa do mundo real que envolvia uma máquina capaz de se mover e de sentir o seu ambiente. No entanto, os problemas de controlar esse movimento e de interpretar os dados sensoriais eram de importância secundária para os projetistas, cujo principal interesse residia na capacidade de raciocínio da máquina. O trabalho de desenvolver o software de mobilidade e sensorial foi delegado a programadores juniores.

Com 1,5 metros de altura e movido por dois motores de passo lentos, mas precisos, o Shakey estava equipado com uma câmara de televisão e era controlado remotamente por um grande computador. Métodos dos programas de manipulação de blocos do MIT, anteriormente usados apenas com braços robóticos, foram adaptados para interpretar as imagens de TV. Estes funcionavam apenas quando a cena consistia unicamente em objetos simples, de

cor uniforme e faces planas, pelo que foi construído um ambiente especial para o robô. Este consistia em várias salas delimitadas por paredes limpas, contendo vários blocos e cunhas grandes e uniformemente pintados. O desempenho mais impressionante de Shakey, executado faseadamente ao longo de vários dias, foi resolver o chamado problema dos "macacos e bananas". Recebendo a instrução para empurrar um bloco específico que estava a repousar sobre um maior, o robô construiu e executou um plano que incluía encontrar uma cunha que pudesse servir de rampa, empurrá-la contra o bloco grande, subir a rampa e realizar o empurrão solicitado.

O ambiente e o problema eram artificiais, mas serviram de motivação e teste para um engenhoso programa de raciocínio chamado STRIPS (do inglês, STanford Research Institute Problem Solver). Dada uma tarefa para o robô, o STRIPS montava um plano a partir das ações limitadas que o robô podia executar. Cada ação tinha pré-condições (por exemplo, para empurrar um bloco, ele deve estar à minha frente) e consequências prováveis (depois de eu empurrar um bloco, ele é movido). O estado do mundo do robô era representado em frases de lógica matemática, e formular um plano era como provar um teorema: o estado inicial do mundo eram os axiomas, as ações primitivas as regras de inferência, e o resultado desejado o teorema. Uma complicação era imediatamente evidente: o resultado de uma ação nem sempre é o esperado (como quando o bloco não se mexe). O Shakey tinha uma capacidade limitada de lidar com essas falhas, observando ocasionalmente partes do mundo e ajustando a sua descrição interna e replaneando as suas ações se as condições não fossem como as que tinha assumido.

O Shakey era impressionante em conceito, mas lastimável em ação. Cada movimento do robô, cada vislumbre captado pela sua câmara, consumia cerca de uma hora de tempo de computador e tinha uma alta probabilidade de falha. O exercício de empurrar blocos descrito acima foi encenado e filmado em etapas, com várias etapas a exigirem "tomadas (tentativas)" repetidas antes de terem sucesso. A falha não estava no planeador STRIPS, que produzia bons planos quando lhe era dada uma boa descrição do que estava

à volta do robô, mas sim nos programas que interpretavam os dados brutos dos sensores e agiam de acordo com as recomendações.

Pareceu-me, no início da década de 1970, que alguns dos criadores de programas de raciocínio bem-sucedidos suspeitavam que o fraco desempenho no trabalho de robótica de alguma forma refletia as capacidades intelectuais dos que se dedicavam a essa vertente da investigação. Esse esnobismo intelectual não é inédito, como, por exemplo, entre teóricos e experimentalistas na física. Mas, à medida que o número de demonstrações aumentou, tornou-se claro que é comparativamente fácil fazer com que os computadores exibam um desempenho de nível adulto na resolução de problemas em testes de inteligência ou a jogar damas, e difícil ou impossível dar-lhes as habilidades de uma criança de um ano no que diz respeito à percepção e mobilidade.

Em retrospectiva, esta dicotomia não é surpreendente. Desde que os primeiros animais multicelulares surgiram há cerca de mil milhões de anos, a sobrevivência na feroz competição por recursos limitados como espaço, alimento ou parceiros tem sido frequentemente concedida ao animal que conseguia produzir mais rapidamente uma ação correta a partir de percepções inconclusivas. Codificada nas grandes e altamente evoluídas porções sensoriais e motoras do cérebro humano está mil milhões de anos de experiência sobre a natureza do mundo e como sobreviver nele. O processo deliberado que chamamos raciocínio é, acredito, a mais fina camada do pensamento humano, eficaz apenas porque é suportado por este conhecimento sensório-motor muito mais antigo e muito mais poderoso, embora geralmente inconsciente. Somos todos atletas prodigiosos em áreas percetuais e motoras, tão bons que fazemos o difícil parecer fácil. O pensamento abstrato, no entanto, é um truque novo, talvez com menos de 100 mil anos. Ainda não o dominamos. Não é intrinsecamente assim tão difícil; apenas parece assim quando o fazemos.

Organismos que carecem da capacidade de perceber e explorar o seu ambiente não parecem adquirir nada que chamaríamos de inteligência. Basta fazermos a mais grosseira

comparação entre os reinos vegetal e animal para apreciar o facto de que os organismos móveis tendem a desenvolver as características mentais que associamos à inteligência, enquanto os imóveis não. As plantas são impressionantemente eficazes à sua maneira, mas não mostram qualquer inclinação aparente para a inteligência. Talvez, com muito mais tempo, uma planta inteligente pudesse evoluir — algumas plantas carnívoras e "sensíveis" mostram que algo semelhante à ação nervosa é possível — mas a esperança de vida do universo pode ser insuficiente.

Os investigadores da cibernética, cujas experiências autocontidas eram muitas vezes semelhantes a animais e móveis, iniciaram a sua investigação sobre os sistemas nervosos tentando duplicar as capacidades sensório-motoras dos animais. A comunidade da inteligência artificial ignorou esta abordagem nos seus primeiros trabalhos e, em vez disso, apontou diretamente para o auge intelectual do pensamento humano, em experiências realizadas em grandes computadores centrais estacionários dedicados a mecanizar o raciocínio puro. Esta rota "de cima para baixo" para a inteligência de máquina fez progressos impressionantes no início, mas produziu lamentavelmente poucos ganhos fundamentais em mais de uma década. Enquanto a cibernética "arranhou" a parte inferior da inteligência real, a inteligência artificial "arranhou" a parte superior. O grosso interior do problema permanece intocado.

Todas as tentativas de alcançar a inteligência em máquinas têm imitado a inteligência natural, mas as diferentes abordagens mimetizaram aspetos distintos do original. A inteligência artificial tradicional tenta copiar os processos mentais conscientes dos seres humanos ao realizar tarefas específicas. A sua limitação reside no facto de que os aspetos mais poderosos do pensamento são inconscientes, inacessíveis à introspeção mental e, portanto, difíceis de formalizar. Alguns dos cibernéticos, adotando uma abordagem diferente, focaram-se na construção de modelos de sistemas nervosos animais a nível neural. Esta abordagem é limitada pelo número astronómico de células em grandes sistemas nervosos e pela grande dificuldade de determinar exatamente o que as células

individuais fazem, como estão interligadas e o que as redes nervosas fazem.

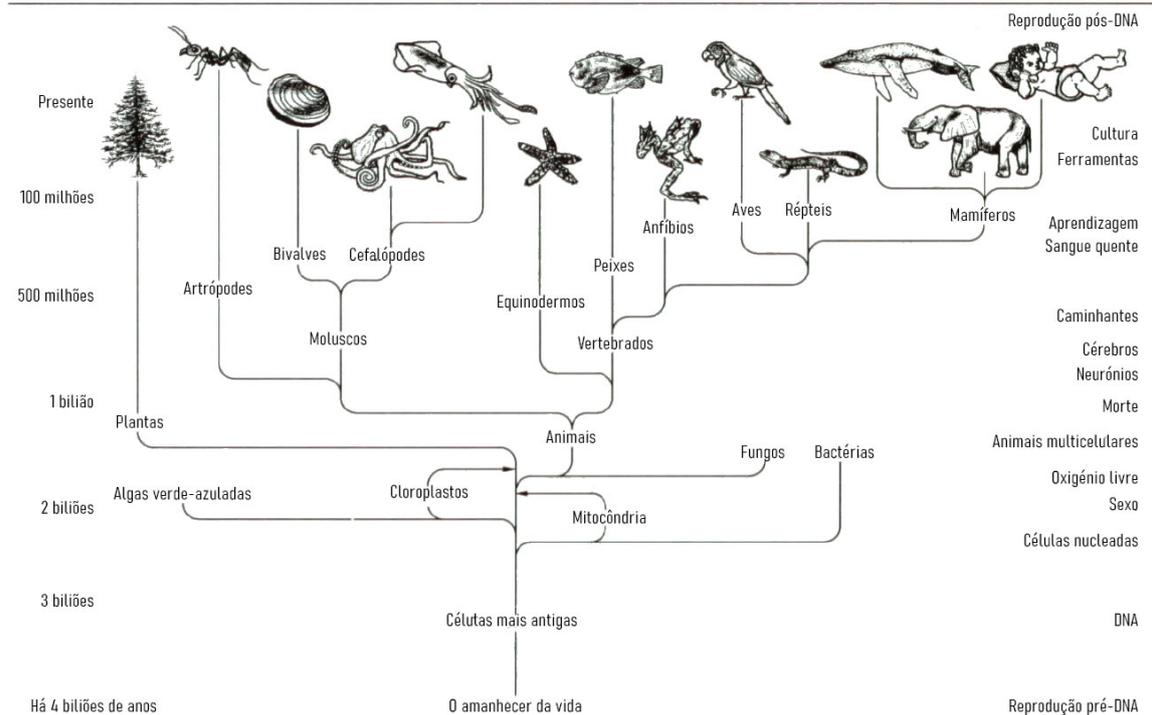
Tanto a IA tradicional como a modelagem neural contribuíram com revelações para a empreitada, e sem dúvida que cada uma poderia resolver todo o problema, dado tempo suficiente. No entanto, com o estado atual da arte, sinto que o progresso mais rápido pode ser feito imitando a *evolução* das mentes animais, esforçando-nos por adicionar capacidades às máquinas gradualmente, de modo que a sequência resultante de comportamentos das máquinas se assemelhe às capacidades de animais com sistemas nervosos cada vez mais complexos. Uma característica fundamental desta abordagem é que a complexidade desses avanços incrementais pode ser adaptada para tirar o melhor partido das capacidades de resolução de problemas dos investigadores humanos e dos computadores envolvidos. A nossa inteligência, como ferramenta, deverá permitir-nos seguir o caminho para a inteligência, como objetivo, com passos maiores do que os originalmente dados pelos processos incrivelmente pacientes, mas cegos, da evolução darwiniana.

A rota é de baixo para cima, e os primeiros problemas são os da percepção e mobilidade, porque é sobre este alicerce sensório-motor que a inteligência humana se desenvolveu. Programas que abordam problemas incrementais, semelhantes aos que os primeiros animais enfrentaram — como lidar com, e até antecipar, as surpresas súbitas, perigos e oportunidades encontrados por um organismo explorador — estão a ser escritos e testados em robôs que têm de enfrentar as incertezas de um mundo real. A maioria das abordagens falhará, mas algumas terão sucesso, da mesma forma que uma minúscula fração das mutações espontâneas que aparecem nos organismos sobrevivem para a próxima geração. A sobrevivência depende das vantagens que uma nova técnica oferece ao lidar com os desafios de um ambiente complexo e dinâmico. Ao estabelecer condições experimentais análogas às encontradas pelos animais no curso da evolução, esperamos refazer os passos pelos quais a inteligência humana evoluiu. O facto de os animais terem começado com pequenos sistemas nervosos dá confiança de que os pequenos computadores de hoje podem emular

os primeiros passos em direção a um desempenho semelhante ao humano. Sempre que possível, os nossos esforços para simular a inteligência de baixo para cima serão ajudados por "espreitadelas biológicas no fim do livro" — nas características neuronais, morfológicas e comportamentais de animais e humanos, conforme revelado pelas pessoas que estudam esses aspectos da vida.

O esforço moderno em robótica tem apenas vinte anos, e só nos últimos dez é que os computadores têm estado rotineiramente disponíveis para controlar robôs. A recapitulação da evolução da vida inteligente encontra-se numa fase muito inicial — existem equivalentes robóticos de sistemas nervosos, mas a sua complexidade é comparável à dos sistemas nervosos de vermes. No entanto, as pressões evolutivas que moldaram a vida já são palpáveis no laboratório de robótica, e estou confiante de que esta rota "de baixo para cima" para a inteligência artificial irá um dia encontrar a rota tradicional "de cima para baixo" a mais de meio caminho. Estará então pronta para fornecer a competência no mundo real e o conhecimento de senso comum que se tem mostrado tão frustrantemente elusivo nos programas de raciocínio. Máquinas totalmente inteligentes resultarão quando for cravado o "prego dourado" metafórico, unindo os dois esforços. Um programa de raciocínio, apoiado por um modelo de mundo robótico, será capaz de visualizar os passos do seu plano, de distinguir situações razoáveis de absurdas, e de intuir algumas soluções ao observá-las acontecer no seu modelo, tal como os humanos fazem. Mais tarde, explicarei por que espero ver esta união em cerca de quarenta anos. Por enquanto, vamos explorar algumas das "faunas" artificiais que estão na base.

Inteligência na Terra



Esta árvore genealógica parcial dos organismos terrestres sugere a ligação entre a mobilidade e a inteligência. Há mil milhões e meio de anos, os nossos antepassados unicelulares separaram-se geneticamente das plantas. Como células únicas, ambas as linhagens nadavam livremente, mas quando as plantas se tornaram multicelulares, especializaram-se em ser coletoras sedentárias de energia solar. Os nossos antepassados animais, por outro lado, permaneceram ambulantes, para melhor se alimentarem das plantas e uns dos outros. Embora as plantas sejam enormemente bem-sucedidas – a maior parte da biosfera terrestre é composta por vegetação e as maiores, mais numerosas e com maior longevidade são as plantas –, demonstram muito pouca tendência evolutiva para qualquer coisa que reconheceríamos como inteligência.

Os artrópodes – incluindo insetos, aranhas e outros rastejantes multipernas – são certamente móveis. Possuem órgãos sensoriais e sistemas nervosos altamente desenvolvidos e exibem uma complexidade de comportamento semelhante à dos robôs atuais. Mas a sua estrutura exoesquelética parece ter limitado o seu metabolismo e tamanho, e os seus sistemas nervosos nunca ultrapassam cerca de 1 milhão de neurónios, em comparação com

os 10 milhões do mais pequeno cérebro de vertebrados. Os insetos sociais ultrapassaram parcialmente a sua limitação de tamanho coordenando grandes grupos de indivíduos que atuam quase como animais isolados.

Os moluscos estão tão distantes de nós como os artrópodes — o nosso antepassado comum existiu há cerca de mil milhões de anos. São especialmente interessantes devido aos estilos de vida bastante contrastantes entre as suas espécies. A maioria dos moluscos são moluscos de movimento lento com estilos de alimentação por filtração, mas os cefalópodes — polvo, lula, sépia e náutilo — nadam livremente. O polvo e a lula abdicaram das suas conchas para se tornarem os invertebrados com maior mobilidade, com os maiores sistemas nervosos. Têm olhos que conseguem visualizar imagens como os vertebrados, mas o olho é uma meia esfera ligada ao tecido corporal flexível, e as fotocélulas da retina estão viradas para a lente, em vez de estarem para longe, como nos vertebrados. O cérebro possui oito gânglios, dispostos em anel em redor do esófago. O polvo e os seus parentes são espetáculos de luzes a nadar, com as suas superfícies cobertas por um milhão de células que mudam de cor, controladas individualmente. Os pequenos polvos podem aprender a resolver problemas como abrir um recipiente de comida. Sabe-se que as lulas gigantes, tão grandes como as baleias e com sistemas nervosos robustos, existem a partir de alguns cadáveres inchados encontrados a boiar no mar. Mas vivem permanentemente em profundidades para além da observação humana atual.

Os equinodermes, incluindo estrelas-do-mar, ouriços-do-mar e outros habitantes do fundo do mar com simetria pentagonal, conquistaram um meio de vida estável caçando animais sésseis, como mariscos e corais. Embora intimamente relacionados com os vertebrados, os equinodermes movem-se muito mais lentamente e possuem órgãos sensoriais mais simples e sistemas nervosos mais pequenos. Uma estrela-do-mar — em marcha lenta permanente — a caçar um molusco parece adormecida aos olhos humanos, mas em fotografias em time-lapse (modo acelerado) pode ser vista a perseguir a sua presa, atacando-a, abrindo-a e, em seguida,

iniciando a digestão, evertindo (virando para fora) o seu estômago para dentro da concha aberta!

Os peixes são muito móveis e têm cérebros maiores do que a maioria dos invertebrados, mas mais pequenos do que os dos outros vertebrados, possivelmente porque as águas abertas são um habitat simples. Os répteis evoluíram em terra e possuem sistemas nervosos maiores. Os anfíbios são intermediários entre os peixes e os répteis no tamanho do sistema nervoso, assim como em muitas outras características. Animais de sangue frio como estes têm um metabolismo lento que não consegue suportar sistemas nervosos tão grandes como os das aves ou dos mamíferos, mas, pela mesma razão, conseguem sobreviver com um estilo de vida menos frenético.

As aves estão relacionadas connosco através de um réptil primitivo que viveu há cerca de 300 milhões de anos. O seu tamanho é limitado pelas exigências do voo, mas têm um metabolismo muito energético. Embora estruturadas de forma diferente, têm cérebros tão grandes como os dos mamíferos de tamanho igual e possuem capacidades comparáveis de aprendizagem e resolução de problemas.

O nosso último antepassado comum com as baleias foi um mamífero primitivo semelhante a um rato, que viveu há 100 milhões de anos. Algumas espécies de golfinhos têm massas corporais e cerebrais idênticas às nossas, são tão hábeis como nós em diversos tipos de resolução de problemas e conseguem compreender e comunicar ideias complexas. Os cachalotes têm os maiores cérebros do mundo. A vantagem que os humanos têm sobre outros animais com cérebros grandes, como os elefantes e as baleias, pode depender menos da nossa inteligência individual do que da eficácia com que essa inteligência está associada à nossa cultura de rápida evolução, imensamente poderosa e que utiliza ferramentas.

Como vimos, o Shakey não se enquadrava nesta categoria de robôs mais orgânicos. Este robô era uma expressão do esforço "de cima para baixo": a sua especialidade era o raciocínio, enquanto o seu software rudimentar de visão e movimento funcionava apenas

em ambientes extremamente simples. Quase ao mesmo tempo, no entanto, e com um orçamento muito mais reduzido, nasceu um robô móvel que se iria especializar em ver e mover-se em ambientes naturais no Projeto de Inteligência Artificial da Universidade de Stanford, localizado a cerca de oito milhas de distância da "residência" do Shakey no SRI. John McCarthy fundou o Projeto em 1963 com o objetivo, então plausível, de construir uma máquina totalmente inteligente em uma década. (O Projeto foi renomeado para Stanford Artificial Intelligence Laboratory, ou SAIL, à medida que a década se aproximava e a plausibilidade do Projeto se desvanecia.) Refletindo as prioridades da pesquisa inicial em inteligência artificial, McCarthy trabalhou no raciocínio e delegou a outros o projeto de ouvidos, olhos e mãos para a mente artificial antecipada. O grupo "mão-olho" do SAIL rapidamente superou o grupo de robótica do MIT e foi fundamental no posterior boom de braços robóticos inteligentes para uso industrial. Um investimento modesto em mobilidade foi adicionado quando Les Earnest, o astuto administrador técnico do SAIL, soube de um veículo abandonado pelo departamento de engenharia mecânica de Stanford após um breve período como simulador de um rover lunar controlado remotamente. No SAIL, este tornou-se o Stanford Cart, o primeiro robô móvel controlado por um grande computador que não raciocinava, e o primeiro campo de testes para visão computacional no mundo desorganizado e iluminado de forma aleatória que a maioria dos animais habita. Fruto de duas teses de doutoramento (uma delas a minha própria), o Stanford Cart navegou lentamente por espaços interiores e exteriores brutos, guiado por imagens de TV processadas por programas bastante diferentes daqueles do mundo do Shakey.

Em meados da década de 1970, a NASA começou a planejar uma missão robótica a Marte, na sequência das bem-sucedidas aterragens Viking. Programada para lançamento em 1984, incluiria dois veículos que iriam percorrer a superfície marciana. Marte está tão longe, mesmo por rádio, que um simples controlo remoto seria muito lento ou muito arriscado; o atraso entre enviar um comando e ver a sua consequência pode ser de até quarenta minutos. Se o robô pudesse viajar sozinho em segurança durante grande parte do

tempo, conseguiria cobrir muito mais terreno. Para este fim, o Jet Propulsion Laboratory (JPL) do Caltech, projetista da maioria das naves espaciais robóticas da NASA, que até então usavam automação bastante segura e simples, iniciou um projeto de robótica inteligente. Reunindo métodos, hardware e pessoas de programas de robótica universitários, construiu uma grande plataforma de teste com rodas chamada Robotics Research Vehicle (RRV), uma engenhoca que transportava câmaras, um telémetro a laser, um braço robótico e um rack (armário ou estrutura) de eletrónica completo, tudo ligado por um longo cabo a um grande computador. Em 1977, já conseguia percorrer pequenos trechos de um parque de estacionamento cheio de rochas para recolher uma determinada rocha e rodá-la para as câmaras. Mas em 1978, o projeto foi interrompido quando a missão a Marte de 1984 foi cancelada e removida do orçamento da NASA. (Claro, Marte não desapareceu, e o JPL está a considerar uma visita para o final do milénio.)

Juntamente com o Office of Naval Research, o primeiro e mais constante apoiante da pesquisa em inteligência artificial (e uma das principais razões para todos os avanços iniciais no campo terem acontecido nos Estados Unidos) é a Advanced Research Project Agency (DARPA) do Departamento de Defesa. Fundada em 1958 após a humilhação nacional causada pelo Sputnik, o seu propósito era financiar projetos "fora da caixa" como seguro contra surpresas tecnológicas indesejadas. Em 1981, os gestores da DARPA decidiram que a navegação robótica estava suficientemente avançada para justificar um grande esforço para desenvolver veículos autónomos capazes de percorrer grandes distâncias por terra sem um operador humano, talvez em zonas de guerra ou outras áreas perigosas. O número de projetos de robôs móveis aumentou vertiginosamente, em universidades e em empresas contratadas de defesa, à medida que o financiamento para este projeto se materializou. Mesmo agora, vários novos robôs do tamanho de camiões estão a percorrer estradas de teste por todo o país — e a poeira ainda está a assentar.

Num plano mais prático e quotidiano, não é de somenos importância o facto de os braços robóticos fixos nas fábricas terem

de ter as suas tarefas "entregues" a eles. Uma linha de montagem com tapete transportador é uma solução, mas os gestores de fábricas cada vez mais automatizadas, no final dos anos 70 e início dos 80, descobriram que os tapetes, cujas rotas são difíceis de alterar, eram demasiado restritivos. Os seus robôs podiam ser rapidamente reprogramados para diferentes trabalhos, mas as rotas de fluxo de material não. Várias grandes empresas em todo o mundo lidaram com o problema construindo o que chamaram de Veículos Automaticamente Guiados (VAGs), que navegavam detetando sinais transmitidos por fios enterrados ao longo da sua rota. Com a aparência de empilhadores ou grandes carros de choque, podem ser programados para viajar de um lugar para outro e ser carregados e descarregados por braços robóticos. Algumas variantes recentes transportam os seus próprios braços robóticos. Enterrar os fios de rota em pisos de fábrica de betão é caro, e métodos alternativos de navegação estão a ser procurados. Tal como acontece com os braços robóticos, os esforços académicos e industriais para desenvolver robôs móveis fundiram-se, e um número alucinante de direções e ideias está a ser energicamente perseguido.

Um Robô Para as Massas

A história apresentada até agora é altamente higienizada e descreve apenas alguns dos principais intervenientes no novo campo da robótica. A realidade é uma mistura complexa de abordagens, motivações e problemas ainda desconetados. Os profissionais são grandes e pequenos grupos de engenheiros elétricos, mecânicos, óticos e de todas as outras áreas, físicos, matemáticos, biólogos, químicos, tecnólogos médicos, cientistas da computação, artistas e inventores, espalhados por todo o mundo. Cientistas da computação e biólogos colaboram no desenvolvimento de máquinas que veem. Físicos e matemáticos trabalham para melhorar o sonar e outros sentidos. Engenheiros mecânicos construíram máquinas que andam sobre pernas e outras que agarram com mãos robóticas de destreza quase humana. No entanto, todos estes esforços incipientes têm sofrido de má comunicação entre os vários grupos, que não conseguiram sequer

chegar a um consenso sobre um esboço geral para o campo da robótica. Apesar do caos, espero ver a primeira oferta em massa deste "caldeirão" servida a tempo para o novo milénio, sob a forma de um robô de uso geral para a fábrica — e para o lar.

Nas nações industrializadas, a agricultura e a manufatura são cada vez mais a província das máquinas, deixando as pessoas livres para prestarem serviços humanos umas às outras. Alimentos e bens tornaram-se abundantes e baratos com este arranjo, mas muitos serviços aumentaram de custo. O serviço doméstico, outrora comum, é escasso e caro. Máquinas domésticas como processadores de alimentos, aspiradores de pó e fornos de micro-ondas não preenchem o vazio em famílias onde todos os adultos trabalham fora de casa. A necessidade existe há muitas décadas: Quando haverá um robô para ajudar em casa?

Durante muitos anos, acreditei que os robôs-serviçais, tão ubíquos na ficção científica, seriam improváveis num futuro próximo. Os lares são ambientes complexos com recursos limitados. O retorno económico de um ajudante doméstico mecânico seria apenas uma fração do valor de um robô num papel industrial típico, portanto, o robô doméstico teria de ser vendido (ou alugado) por muito menos. Pior ainda, a operação segura e eficaz no ambiente doméstico, muitas vezes caótico, é muito mais difícil do que em ambientes fabris controláveis. Os robôs existentes oferecem movimentos maioritariamente "cegos", repetitivos e potencialmente letais a um preço comparável ao de uma residência inteira. Esta enorme lacuna no preço e no desempenho é real, mas agora espero ver um robô de uso geral utilizável em casa dentro de dez anos. A mudança na minha atitude provém, em parte, dos desenvolvimentos de pesquisa dos últimos anos e, em parte, de uma nova apreciação das implicações do conceito de "uso geral".

Os robôs industriais de hoje são mais flexíveis do que a automação fixa que por vezes substituem, mas fazem tão poucas coisas bem que o termo "uso geral" dificilmente se aplica. De facto, os robôs individuais são geralmente aparafusados a uma estação fixa, equipados com pinças e, por vezes, sensores especializados para uma determinada tarefa, que executam incessantemente,

talvez pelo resto da sua existência. A estreiteza do seu repertório, além de ser aborrecida, limita bastante o número de unidades que podem ser vendidas. Existem menos de 100 mil robôs (além de brinquedos) de todas as marcas no mundo hoje. Compare-se este número com 100 milhões de carros, 500 milhões de aparelhos de televisão ou 20 milhões de computadores. Tão poucas unidades vendidas só podem suportar uma quantidade limitada de pensamento e desenvolvimento de engenharia. O resultado é um projeto menos otimizado a um preço elevado. Mas não para sempre. À medida que o número de unidades produzidas cresce, cresce também a oportunidade e o incentivo para melhorar o projeto dos robôs e os detalhes da sua produção. Os custos diminuem, e os robôs tornam-se melhores, expandindo incidentalmente o mercado e aumentando o número de unidades vendidas, o que leva a novas melhorias. O gráfico de custo unitário decrescente versus número de unidades produzidas é chamado de curva de aprendizagem do fabricante.

O mercado potencial para robôs expandir-se-á enormemente quando for alcançado um certo nível de utilidade geral. Até este ponto de equilíbrio, a especialização — a exploração das circunstâncias únicas de um trabalho para alcançar um desempenho aceitável com o mínimo de complexidade — será a norma robótica. Para além do ponto de equilíbrio, o mercado potencial será suficientemente grande para que maiores lucros sejam direcionados para projetos mais padronizados, vendidos em números cada vez maiores. O robô barato, produzido em massa e de alta utilidade terá chegado. Já acumulámos experiência suficiente para especificar algumas das características deste "Model T" dos robôs. Não será inteligente e não virá pré-programado para fazer muitas tarefas úteis. Sairá da fábrica com um conjunto suficiente de capacidades mecânicas, sensoriais e de controlo que podem ser convenientemente ativadas por software especialmente escrito para aplicações particulares.

O primeiro grande mercado para tal máquina será nas fábricas, onde será um pouco mais barato e consideravelmente mais versátil do que a geração mais antiga de robôs que substitui. A sua melhoria na relação custo-benefício permitirá que seja utilizado num leque

muito mais vasto de trabalhos e, conseqüentemente, em maiores quantidades, diminuindo ainda mais o seu custo. Com o tempo, tornar-se-á mais barato do que um carro pequeno, colocando-o ao alcance de alguns lares e criando uma procura por uma enorme variedade de novo software. Os programas de controlo de robôs que realmente executam várias tarefas virão de muitas fontes diferentes, tal como acontece com os programas para os computadores pessoais e de negócios de hoje.

Tal como aconteceu com os computadores pessoais, muitas das aplicações bem-sucedidas do robô de uso generalizado surgirão como surpresas até para os seus criadores. Podemos especular sobre os equivalentes, na Era dos robôs em massa, dos videojogos, dos processadores de texto ou das folhas de cálculo, mas a realidade será mais estranha do que qualquer previsão. Para começar este jogo de adivinhações, consideremos alguns programas: montagem mecânica ligeira (de uma empresa de automação industrial), limpeza de casas de banho (de uma pequena empresa fundada por antigos funcionários de limpeza), preparação e confeção de refeições gourmet com ingredientes frescos (numa colaboração entre um informático e um chef parisiense), afinação de automóveis Saturn de um determinado ano (do departamento de assistência técnica da General Motors Saturn), confeção de tapetes com padrões (por um estudante do secundário do Massachusetts), remoção de ervas daninhas de um relvado, uma a uma, participação em corridas de robôs (contra outros programas — os robôs físicos são atribuídos por sorteio imediatamente antes da corrida), escavação e trabalho de cantaria detalhado (por uma empresa de construção emergente), investigação de ameaças de bomba (vendido a departamentos de polícia em todo o mundo), entrega e recolha de inventário armazenado, auxílio na montagem e teste de outros robôs (em várias fases independentes), entre muitas outras possibilidades. Algumas destas aplicações exigirão acessórios opcionais para o robô, ferramentas e sensores especiais (como detetores químicos), revestimentos de proteção, e assim por diante.

É possível que, nas primeiras décadas do próximo século, a escrita de programas de aplicações para sucessivas gerações de

robôs de uso generalizado se torne na principal ocupação humana. O canalizador especializado, por exemplo, poderá ter de escolher entre aplicar as suas competências técnicas ao serviço de algumas centenas de clientes, ou codificar essas mesmas competências em programas para robôs que poderão ser vendidos, com sucesso, a milhares ou até milhões de utilizadores. A primeira alternativa tornar-se-á progressivamente menos atrativa como fonte de rendimento, à medida que o trabalho manual tiver de competir com um número crescente de robôs cada vez mais sofisticados, controlados por software cada vez mais avançado. Já a segunda via não está isenta de riscos — um programa pode fracassar no mercado, tal como acontece hoje com invenções, livros, música, arte ou software informático. Por outro lado, um programa bem-sucedido poderá gerar anos de rendimento para o seu autor.

Quase toda a gente possui, ou pode desenvolver, múltiplas competências — e cada uma delas pode tornar-se uma potencial fonte de direitos de autor, quando codificada sob a forma de programa. Várias versões concorrentes de uma mesma competência serão comercializadas e adquiridas com base na sua utilidade, custo, gosto pessoal, tendências ou campanhas publicitárias. Cada programa terá uma duração limitada, estando destinado a ser ultrapassado por versões que sejam simplesmente melhores ou concebidas para operar com uma nova geração de robôs mais sofisticada. Surgirá uma vasta indústria secundária dedicada a apoiar o processo de programação.

Em pouco tempo, serão criados programas capazes de tornar os robôs de uso generalizado bons aprendizes, ensináveis, por exemplo, através da demonstração dos movimentos necessários para executar uma tarefa ou pela simples repetição do exemplo. A biblioteca crescente desses programas tornar-se-á, eventualmente, numa autêntica mina de conhecimento humano codificado, de natureza não verbal, à qual poderão recorrer as sucessivas gerações de robôs cada vez mais autónomos que surgirão após a geração do ponto de equilíbrio económico. A indústria dos sistemas especialistas já começou, aliás, a codificar conhecimento verbal segundo este mesmo princípio.

Locomoção do Ponto de Equilíbrio

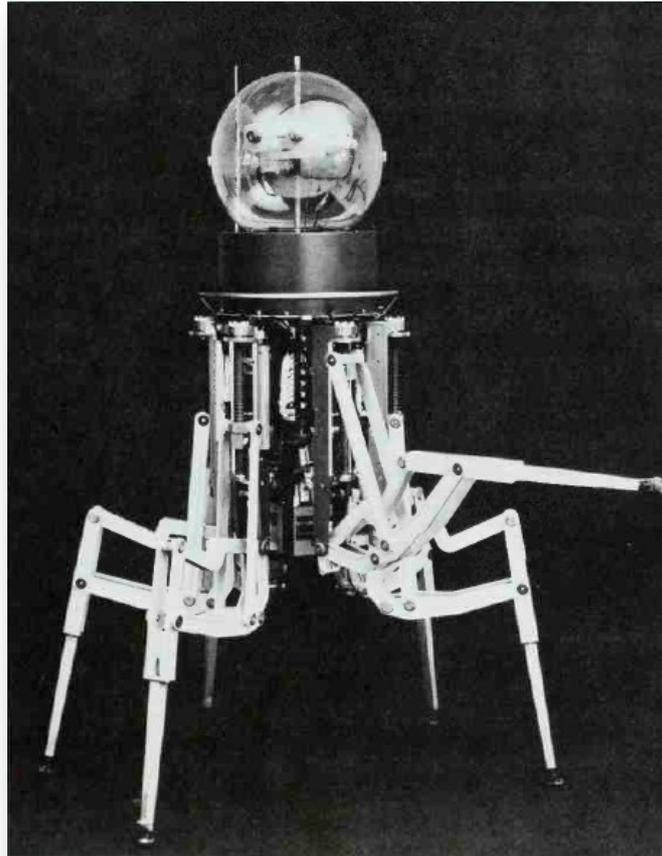
Para ser bem-sucedido, um robô de uso generalizado e produzido em massa terá de possuir um nível mínimo de funcionalidade. Os primeiros robôs desta classe não precisam de ser capazes de fazer tudo, nem sequer de fazer a maioria das coisas com grande eficiência. Precisam apenas de realizar um número suficiente de tarefas com competência razoável, de modo a criar um mercado aberto e em expansão para si próprios — onde cada queda no preço provoque um aumento mais do que proporcional na quantidade de aplicações economicamente viáveis e na procura de unidades.

Mesmo em ambientes altamente urbanizados, as zonas de terreno plano e duro formam um arquipélago rodeado por um mar de terrenos irregulares, moles, em desnível ou simplesmente intransitáveis. Uma máquina incapaz de navegar neste mar ficará presa numa única “ilha”, vendo as suas utilizações potenciais severamente limitadas. O nosso critério de viabilidade económica exige, por isso, um sistema de locomoção mais avançado do que as tradicionais rodas. Robôs com pernas estão apenas agora a começar a mostrar sinais de viabilidade prática. A demonstração mais convincente até à data foi realizada por uma empresa californiana chamada Odetics, cujo robô de seis pernas, com aspeto araneiforme e acionado eletricamente, é capaz de sair do camião onde se encontra, subir para cima de uma caixa redonda, passar por uma porta estreita, e ainda exhibir-se ao levantar uma das extremidades do próprio camião e arrastá-lo pelo chão. Esta e outras demonstrações promissoras tornam provável que uma locomoção prática baseada em pernas esteja disponível dentro de uma década.

As pernas constituem um mecanismo poderoso para a locomoção, mas a sua natureza descontínua — com movimentos de arranque e paragem — limita tanto a velocidade como a eficiência energética. O robô da Odetics, por exemplo, esgota as suas baterias em menos de uma hora de marcha lenta. As sérias limitações energéticas de um robô autónomo e autossuficiente poderão exigir um sistema de propulsão mais económico. Em terreno plano, as

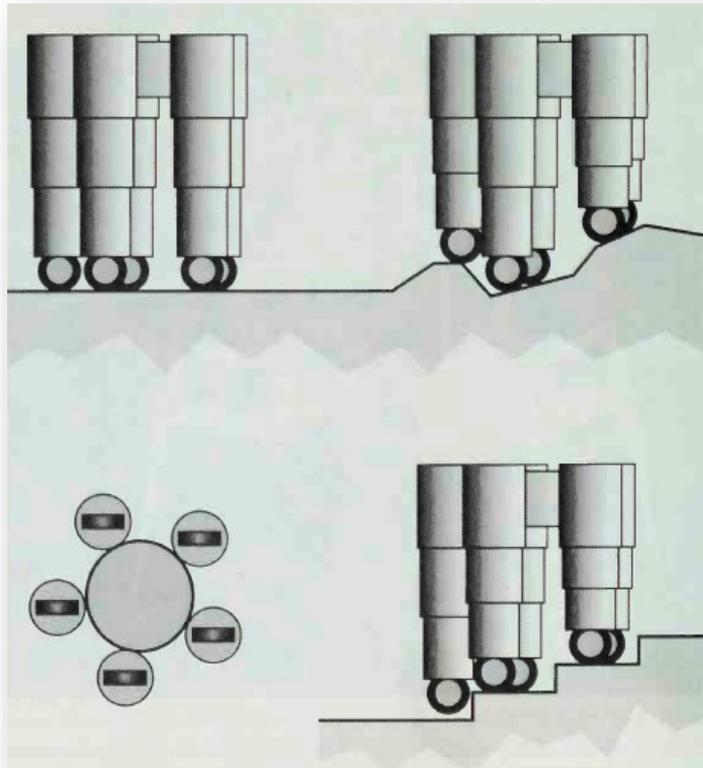
rodas continuam a ser a melhor solução, oferecendo uma eficiência próxima dos 100% numa vasta gama de velocidades. Uma solução de compromisso poderá consistir em pernas lentas que terminem em pés com rodas — como patins motorizados. O robô mover-se-ia sobre rodas na maior parte do tempo, mas levantaria os pés para ultrapassar obstáculos ou subir escadas. Em terreno acidentado, poderia avançar lentamente através de uma marcha completa.

A Hitachi, gigante japonesa da eletrônica, experimentou, no início da década de 1980, uma versão particularmente simples da ideia de roda-pé. Destinado à utilização em reatores nucleares, o sistema da Hitachi possuía cinco “pernas” simples — cada uma consistia num suporte vertical motorizado que se estendia e recolhia telescopicamente a partir do corpo do robô. As pernas estavam dispostas uniformemente em torno da estrutura, formando um pentágono regular. Cada extremidade possuía uma roda com capacidade de direção e propulsão. Cinco pernas constituem o número mínimo necessário para que um robô possa manter-se estável com uma das pernas levantada, sem necessidade de deslocar o seu centro de gravidade. As máquinas da Hitachi subiam escadas aproximando-se delas sobre as cinco rodas, levantando a roda da frente até à altura do primeiro degrau, avançando até que a perna elevada estivesse firmemente posicionada sobre o degrau, baixando-a ligeiramente até estabelecer contato sólido, e prosseguindo depois com a perna mais próxima a seguir. Em escadas estreitas, o robô pode ter as suas rodas apoiadas em até três degraus sucessivos ao mesmo tempo.



Máquina de andar. O Odetics "Odex" pode andar, trepar, passar por portas ou esticar-se para se manter estável. Mas o seu consumo de energia limita-o a uma hora de mobilidade por carga de bateria.

Um procedimento semelhante permite-lhe superar obstáculos. O robô pode atravessar terrenos acidentados lentamente, com as rodas individuais a subir e a descer sobre as irregularidades da superfície — uma suspensão ativa. A parte superior do corpo do robô permanece perfeitamente horizontal em funcionamento normal. O tempo e a investigação futura dirão qual a configuração que se revelará melhor para o primeiro robô multifuncional.



Cinco pernas. Este projeto de mobilidade da Hitachi, com cinco rodas direcionáveis sobre pernas telescópicas, tem mais limitações do que um com pernas totalmente articuladas, mas tem um desempenho muito melhor em superfícies planas.

Manipulação do Ponto de Equilíbrio

Poucos trabalhos úteis podem ser realizados por um robô que apenas se desloca. O trabalho produtivo exige a capacidade de segurar e transportar ingredientes, peças, ferramentas e outros objetos. Os manipuladores industriais — os robôs mais numerosos e bem-sucedidos até à data — estão equipados com braços que alcançam os pontos necessários através de cerca de seis articulações rotativas ou deslizantes. Se deixarmos de lado pormenores relativos ao peso, potência e controlo, alguns dos modelos mais pequenos são quase adequados ao alcance requerido pelo robô do ponto de equilíbrio. Como muitas tarefas implicam juntar pares de objetos, é provável que o nosso robô venha equipado com pelo menos dois braços. Um terceiro braço seria

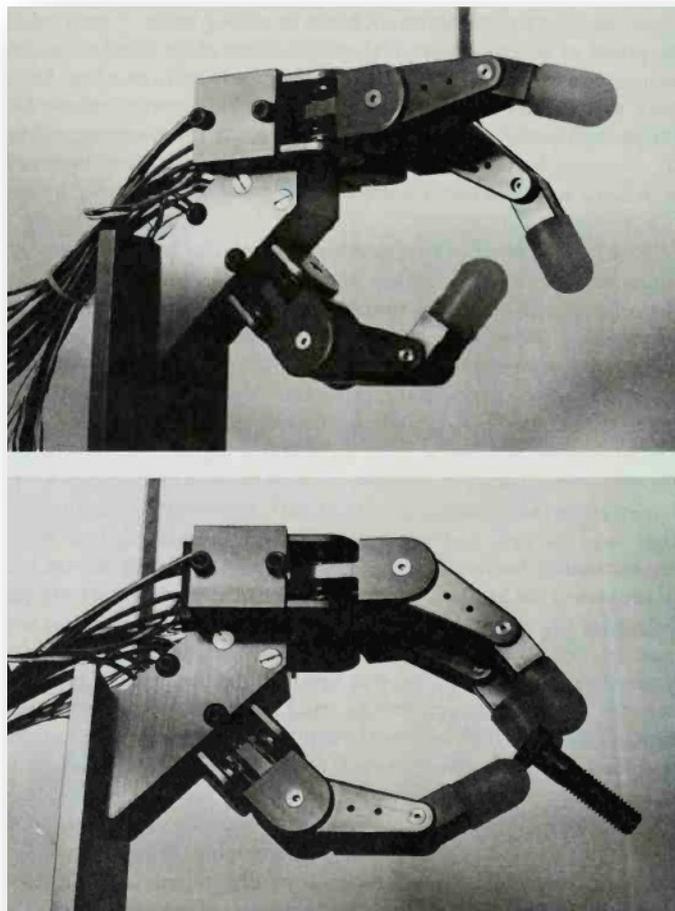
vantajoso em tarefas onde os objetos em contato precisam de ser manipulados de alguma forma (os entusiastas de eletrônica reconhecerão, por exemplo, a soldadura como um desses casos).

As mãos robóticas ainda não estão tão desenvolvidas quanto os braços. Os manipuladores industriais conseguem agarrar e segurar objetos com dispositivos específicos adaptados a cada peça, ou com um tipo simples de mão de dois dedos, conhecido como grampo de mandíbula paralela. Estes grampos são fáceis de operar, mas apenas conseguem segurar com segurança determinados tipos de objetos robustos. Não são capazes de controlar nem alterar a orientação do que seguram. O nosso robô universal necessitará de maior flexibilidade.

Alguns projetos de investigação têm explorado grampos multifuncionais com maior destreza. Um dos melhores resulta de uma década de trabalho de Ken Salisbury, atualmente no MIT. A mão robótica de Salisbury, com três dedos, consegue segurar e orientar parafusos e ovos, e até manipular cordel de forma semelhante à humana. A configuração básica e as dimensões da mão foram determinadas através de uma busca computacional por entre diferentes ligações mecânicas, procurando o conjunto mínimo que permitisse às pontas dos dedos convergir sobre e segurar com firmeza objetos pequenos de formas variadas. O resultado é uma mão com três dedos dispostos simetricamente, que se dobram de maneira semelhante aos dedos humanos. Contudo, como os dedos podem curvar-se tanto para fora como para dentro, a mão consegue agarrar objetos ocos tanto pelo exterior como pelo interior. A força de preensão é transmitida através de finos cabos de aço puxados por um conjunto de motores localizados mais abaixo, no "pulso" do robô.

Para realizar tarefas com um mínimo de destreza, as mãos do robô precisam de "sentir" os objetos que seguram. Ken Salisbury está a desenvolver pontas dos dedos em forma de hemisfério para a sua mão robótica, que, graças a extensómetros cuidadosamente posicionados no interior, conseguem detetar a magnitude e a direção das forças externas. Os programas informáticos necessários para planear e executar os movimentos dos braços e das mãos em

manipuladores complexos ainda se encontram numa fase incipiente. Existem já alguns programas capazes de planejar movimentos de braços livres de colisões entre dois pontos num ambiente conhecido e obstruído. Estes programas consideram um espaço (conhecido como *espaço de configuração*) que descreve todas as posturas possíveis do manipulador. Cada articulação acrescenta uma dimensão a esse espaço, o que significa que um manipulador tão complexo como o de Salisbury tem um espaço de configuração extremamente intrincado. Essa complexidade espacial traduz-se numa busca computacional dispendiosa e demorada para encontrar um trajeto viável. Os tempos de processamento variam entre minutos e horas, mas continuam a surgir algoritmos mais eficientes, ao mesmo tempo que os computadores se tornam cada vez mais rápidos.



Três dedos. A "Mão de Salisbury", uma solução minimalista para a destreza geral dos robôs. Cada dedo é controlado por três

motores. A mão pode segurar por fora ou, com igual facilidade, dobrar os dedos para fora para segurar um objeto oco por dentro.

Navegação de Ponto de Equilíbrio

A capacidade mecânica de se mover é apenas uma parte do problema da mobilidade. É também necessário saber encontrar e regressar a locais específicos, além de evitar perigos durante o trajeto. Tenho trabalhado nesta questão durante a maior parte da minha carreira e fico feliz por poder dizer que algumas boas soluções estão a emergir. A minha tese em Stanford, nos anos 70, abordava programas concebidos para permitir que o Stanford Cart encontrasse o seu caminho através de salas desordenadas e espaços exteriores. A primeira versão de tal programa, em 1976, observava o mundo através de uma câmara de televisão colocada no robô. Ao identificar zonas distintivas na imagem de vídeo e acompanhá-las enquanto o robô se movia, o programa conseguia estimar a distância a esses elementos e calcular a extensão do próprio movimento. Com esses dados, construía um mapa tridimensional esparsa do ambiente, identificava obstáculos e planeava um trajeto até ao destino que evitasse colisões. O robô avançava então cerca de um metro nesse trajeto, parava, observava, atualizava o mapa, planeava e voltava a mover-se. Aos solavancos cautelosos, o Cart tentava assim alcançar o destino em segurança.

Infelizmente, o programa não funcionava bem. Aproximadamente em um de cada quatro movimentos, a parte do software responsável por estimar o movimento do robô a partir das alterações na imagem cometia erros: confundia áreas da imagem, calculava mal o deslocamento e corrompia o mapa que estava a ser lentamente construído. A probabilidade de atravessar uma sala grande — uma viagem de talvez trinta solavancos — com sucesso era praticamente nula. Em 1979, fiz uma nova tentativa com um programa renovado e algum hardware adicional: um mecanismo que movia a câmara lateralmente com precisão, ao longo de um trilho. Com esse sistema, o programa podia obter várias imagens da cena sem mover o robô inteiro — tal como os humanos obtêm

duas imagens com os olhos. Aproveitando cuidadosamente essa informação adicional para eliminar erros, o programa conseguiu aumentar a taxa de sucesso por movimento para quase 100%. O robô conseguia agora, muitas vezes, completar com êxito os trinta solavancos necessários para atravessar a sala e apresentar um mapa correto num ecrã. Ainda assim, em cerca de um em cada quatro casos, o sistema falhava: ou porque a filtragem agressiva de erros tinha eliminado um obstáculo real do mapa (levando a uma colisão), ou porque, apesar da filtragem, os erros infiltravam-se e confundiam a perceção da sua posição. Talvez aceitável para uma tese académica, mas claramente insuficiente para um robô executar tarefas complexas que, no mínimo, exigiriam atravessar salas repetidamente.

Em 1980, mudei-me para a Universidade Carnegie Mellon, para continuar esta investigação no recém-criado Robotics Institute. Dois estudantes de pós-graduação, Chuck Thorpe e Larry Matthies, analisaram e melhoraram substancialmente o programa antigo, aumentando tanto a sua velocidade como a sua precisão em dez vezes. Quando tudo corria bem, o sistema era agora capaz de reportar a posição do novo robô — que chamámos Neptune — com uma precisão de alguns centímetros. Infelizmente, nem sempre tudo corria bem, e a taxa de falhas permanecia teimosamente inalterada: o robô continuava a atravessar a sala com sucesso apenas três em cada quatro vezes.

Em 1984, o nosso grupo concordou em realizar uma investigação para uma nova empresa chamada Denning Mobile Robotics, Inc., sediada em Massachusetts, que estava a desenvolver um robô de segurança (mais precisamente, um alarme móvel contra intrusos). Em vez de uma câmara, o robô estava equipado com um anel em forma de cinto composto por sensores de sonar semelhantes aos que se encontram nas câmaras Polaroid. Estes sensores já se tinham revelado bastante úteis para detetar a presença e a direcção geral de obstáculos próximos, permitindo assim ao robô evitá-los. Contudo, o nosso objetivo era mais ambicioso. Em vez de simplesmente detetar colisões iminentes, seria possível utilizar este sistema de sonar — continuamente ativo — para construir um mapa do ambiente, capaz de orientar uma

navegação precisa de ponto a ponto, tal como (três em cada quatro vezes) os nossos programas guiados por visão conseguiam fazer? Cada unidade de sonar emite um chirp (pulsção) ultrassónico dentro de um cone largo e regista o tempo até ao primeiro eco detetado. Este tempo é proporcional à distância do objeto mais próximo dentro desse cone. A distância medida pode ter uma precisão superior a um centímetro. No entanto, dado que o cone tem uma abertura de cerca de 30° , a posição lateral do objeto continua a ser muito incerta. Isto contrasta fortemente com as medições quase pontuais possíveis com câmaras de vídeo, e por isso os métodos programáticos desenvolvidos para o Cart não podiam ser aplicados diretamente aqui.



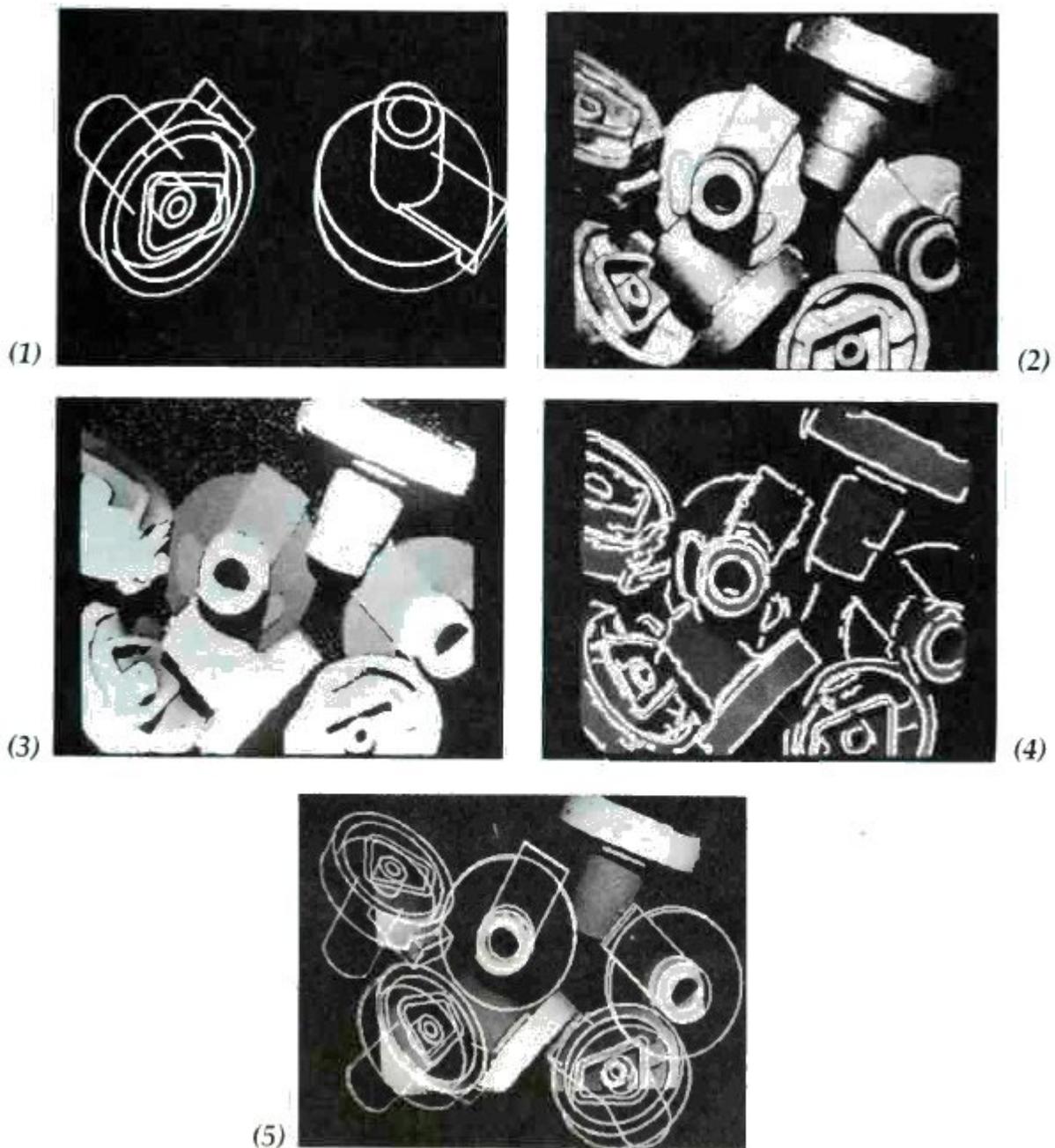
Navegação Autónoma. O Denning Sentry é um produto comercial capaz de patrulhar um grande armazém ou complexo de escritórios todas as noites durante meses sem intervenção humana, guiado por faróis emissores de luz e uma imagem de sonar dos arredores. Durante o dia, recarrega-se numa cabine especial.

Embora uma única leitura de sonar forneça apenas informação limitada sobre a posição do objeto que causou o eco, essa mesma leitura traça um grande volume de espaço vazio à frente desse objeto. Quando as leituras de diferentes sensores se sobrepõem, a região vazia indicada por uma pode restringir a localização possível do objeto causador do eco indicado pela outra. Assim, centenas ou milhares de leituras recolhidas a partir de diferentes posições podem, no seu conjunto, construir mapas detalhados, apesar da imprecisão de cada sensor individual. Como a sensibilidade de um sensor de sonar diminui gradualmente do centro para as bordas do seu cone, pareceu mais adequado fazer esta combinação com probabilidades. Alberto Elfes, outro estudante de pós-graduação, e eu escrevemos um programa com base nestas ideias e ficámos surpreendidos ao ver que guiava o robô de forma muito mais fiável do que o antigo sistema baseado em câmara de vídeo. Outro estudante, Bruno Serrey, juntamente com Larry Matthies, encontrou então uma forma de aplicar a abordagem probabilística também aos dados de vídeo, e mais uma vez descobriu que funcionava notavelmente melhor do que o método anterior.

A nossa nova abordagem representa o espaço à volta do robô como uma grelha de células, cada uma contendo a probabilidade, com base em todas as leituras sensoriais disponíveis, de que a célula correspondente no espaço esteja ocupada por matéria. Uma leitura pode reduzir a probabilidade de algumas células (por exemplo, as que pertencem ao interior do cone do sonar) e aumentar outras (como as que estão na superfície da distância medida pelo sonar). Isto fornece uma forma prática de combinar resultados de diferentes tipos de sensores, e, de facto, Elfes e Matthies demonstraram recentemente um programa que constrói mapas combinando dados de sonar e de vídeo. Com estes novos métodos, os nossos robôs conseguem agora viajar longas distâncias quase sem falhas. Com uma nova base matemática para esta abordagem e um novo robô — muito sofisticado — chamado Uranus, para continuar o trabalho, sinto-me extremamente confiante de que a navegação estará mais do que resolvida dentro do prazo de dez anos previsto para o robô universal.

Reconhecimento do Ponto de Equilíbrio

O sistema sensorial tem outra função vital: o reconhecimento e a localização de objetos específicos no entorno do robô. Os objetos reconhecidos podem ser elementos pequenos, destinados posteriormente a ser agarrados por uma das mãos, ou objetos de grandes dimensões que servem como marcos de referência ou locais de trabalho. Imaginemos um processo no qual os objetos são descritos com base na sua forma e nas características das suas superfícies, e o sistema de reconhecimento do robô procura identificar um objeto de cada vez. Uma identificação provisória pode ser confirmada ao observar a cena a partir de um ponto de vista diferente. O resultado é uma descrição da posição e orientação do objeto, adequada para ser utilizada pelo programa que controla o movimento de preensão das mãos.



Localização de Objetos. 3DPO (acrônimo para Three-Dimensional Parts Orientation) é um programa que encontra peças específicas no meio de uma confusão de outras peças. Esta sequência de imagens é: (1) um modelo tridimensional computacional da peça a encontrar; (2) uma imagem de TV de um amontoado de peças reais; (3) uma imagem computacional das mesmas peças, em que o brilho indica agora a que distância da câmara se encontra cada pedaço da superfície visível; (4) a dedução, pelo computador, dos principais limites da superfície do amontoado;

e (5) o ajuste, pelo computador, do modelo da peça às ocorrências reais da peça no amontoado.

A visão por computador é, de longe, o meio mais promissor para esta capacidade de identificação. A operação-chave consiste em identificar um objeto específico no meio de uma confusão visual. A investigação em visão computacional para robôs industriais produziu soluções parciais para o chamado problema do bin-picking (seleção de peças em caixotes). Os programas de bin-picking permitem que um computador identifique objetos pré-definidos num amontoado visível numa imagem de televisão, mesmo quando esses objetos se ocultam parcialmente uns aos outros, possibilitando que sejam removidos um a um por um manipulador. Um grupo de investigação da General Motors, nos anos 1970, demonstrou um sistema que funcionava, desde que as peças sobrepostas estivessem maioritariamente dispostas na horizontal. Era demasiado lento e pouco fiável para ser prático numa linha de produção, mas demonstrou que era possível. Nos últimos anos, vários grupos nos Estados Unidos e no Japão apresentaram programas capazes de identificar objetos simples com base em dados tridimensionais obtidos por uma câmara a observar uma cena iluminada por dispositivos especiais que projetam faixas ou grelhas de luz. Em computadores contemporâneos, estes sistemas demoram muitos minutos a realizar identificações que, mesmo assim, são insatisfatórias. No entanto, é provável que os requisitos mínimos para o nosso robô venham a ser cumpridos dentro do horizonte temporal de dez anos que estabelecemos.

Processamento e Coordenação

Os melhores protótipos dos componentes sensoriais de baixo nível e de planeamento de movimento do nosso futuro robô consomem muitos minutos de tempo de computação num bom microcomputador. Isto é, em parte, um reflexo da paciência dos investigadores: processos que demoram mais de uma hora a ser executados tornam-se demasiado difíceis de investigar eficazmente, enquanto programas mais simples e rápidos não são muito interessantes, pois têm um desempenho inferior. Por outro

lado, os tempos de execução também revelam algo sobre a dificuldade dos critérios de equilíbrio. Um robô que gaste até uma hora a considerar cada movimento simples é claramente inaceitável, mas alguns segundos seriam toleráveis. Um computador capaz de realizar mil milhões (1 000 000 000) de operações por segundo, com mil milhões de bytes de memória principal, seria suficiente. Isto corresponde aproximadamente à potência dos maiores supercomputadores construídos até hoje, sendo algumas centenas de vezes mais rápido do que os melhores microcomputadores atuais. A evolução contínua da computação deverá entregar essa capacidade num microcomputador dentro de uma década. Dependendo do progresso em várias frentes de desenvolvimento, essa potência poderá estar distribuída por poucas ou muitas unidades de processamento individuais e poderá depender fortemente de hardware especializado — por exemplo, circuitos dedicados ao processamento visual de baixo nível. A configuração exata do hardware é irrelevante para os nossos propósitos aqui.

O nosso trabalho em Carnegie Mellon com tarefas integradas para robôs móveis sugere que os processos básicos devem ser organizados em módulos que funcionem em simultâneo. Por exemplo, um programa de navegação que dirija o robô até a um destino pode coexistir com outros que estejam atentos a surpresas e perigos. Se um módulo de deteção de escadas concluir que há um perigo iminente, ele assume o controlo do robô até que o perigo tenha passado.

Um Robô Sensato

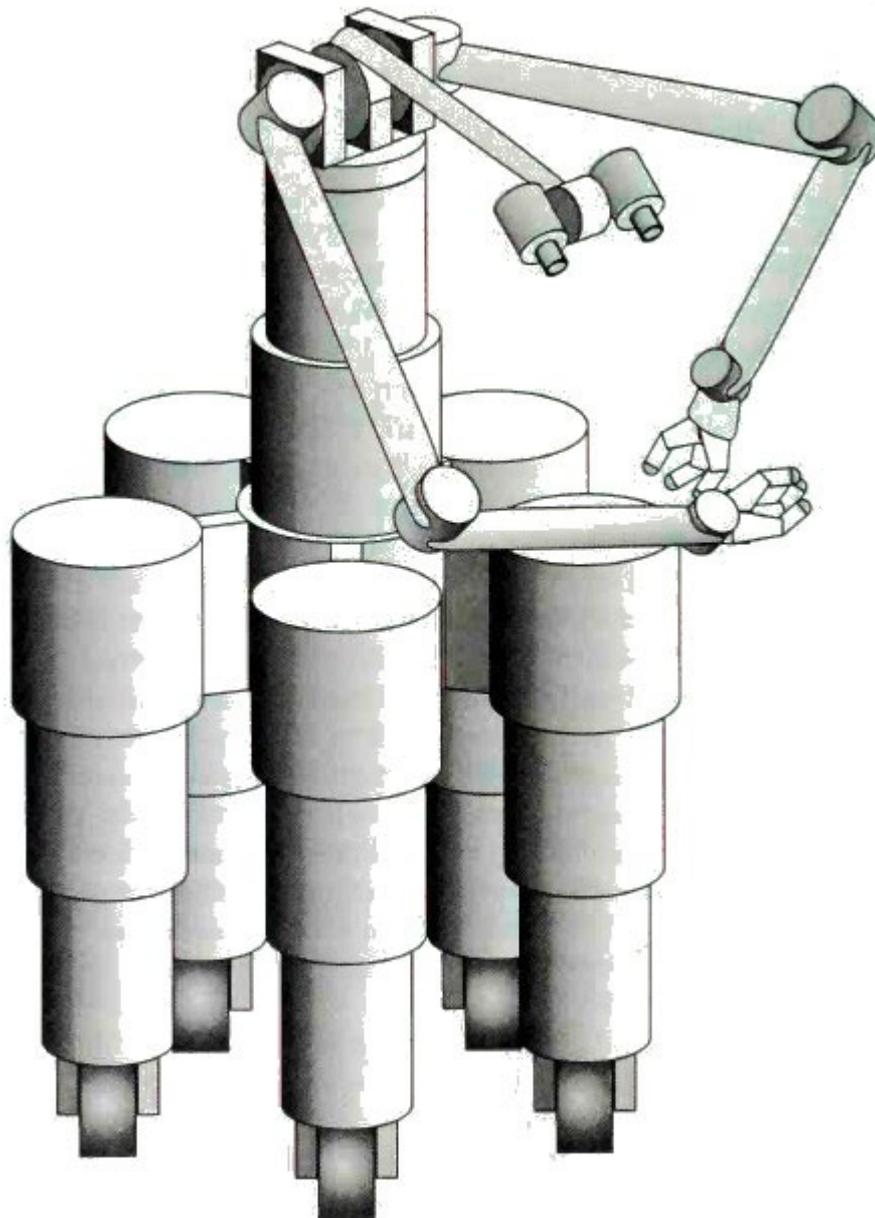
Eis uma possível configuração para o nosso robô universal. Desloca-se sobre cinco pernas-rodas com o projeto da Hitachi e possui dois braços com mãos Salisbury. No topo, tem um par de câmaras de TV a cores, complementadas por um conjunto discreto de sensores de sonar que permitem perceber o mundo em direções não cobertas pelas câmaras. Transporta um giroscópio laser de baixo custo para ajudar na navegação, e é controlado por um sistema computacional capaz de efetuar pelo menos mil milhões

de operações por segundo. Integrado no hardware encontra-se um sistema operativo de software que permite a execução de múltiplos processos simultâneos. Programas incorporados possibilitam a descrição de objetos no mundo exterior, o seu reconhecimento visual mesmo no meio da desordem, e a sua manipulação através das mãos. O sistema de navegação pode ser instruído para construir, armazenar, recuperar e comparar mapas dos arredores, e para conduzir o robô até locais específicos.

Os leitores familiarizados com computadores pessoais poderão reconhecer a semelhança com as funções utilitárias de um sistema operativo, especialmente a caixa de ferramentas gráficas do Apple Macintosh. Estas capacidades do robô são coordenadas por software de aplicação destinado (espera-se) a uma variedade surpreendente de tarefas específicas; o software seria fornecido por vários fornecedores independentes. Mais uma vez, a semelhança com os computadores pessoais é evidente. Pode-se eventualmente esperar pacotes de software integrados que permitam ao robô mudar rápida e automaticamente de uma tarefa para outra, tornando-o num servo mecânico mais autónomo.

A Evolução Convergente das Emoções e da Consciência

As máquinas que estamos a considerar comportam-se de uma forma previsível que poderíamos descrever como mecânica ou semelhante à dos insetos. Continuarão os robôs a exibir esta previsibilidade à medida que se tornam mais complexos, ou desenvolverão algo semelhante ao carácter mais rico dos animais superiores e dos humanos?



Um Robô de Uso Geral. Esta caricatura de um robô de uso geral de primeira geração mostra os principais sistemas: locomoção com capacidade limitada em escadas e terrenos acidentados, manipulação geral, visão estereoscópica, detecção de 360° para desvio de obstáculos e navegação. Não são mostrados o hardware e o software necessários para animar esta montagem.

Como vimos, os programas de controlo mais avançados nos atuais robôs móveis utilizam dados provenientes de sensores para manter representações — a vários níveis de abstração e precisão — do mundo que os rodeia, da posição do robô nesse mundo e do seu

estado interno. Os programas que planificam ações para o robô manipulam esses modelos do mundo a fim de ponderar diferentes movimentos futuros. Estes modelos podem também ser armazenados periodicamente e analisados mais tarde, servindo de base para a aprendizagem.

Uma interface verbal associada a estes programas poderia responder de forma significativa a perguntas como: “Onde estás?” (“Estou numa área de cerca de vinte metros quadrados, limitada em três lados, e há três pequenos objetos à minha frente”) e “Por que fizeste isso?” (“Virei à direita porque achei que não conseguiria passar pela abertura à esquerda”). No nosso laboratório, os programas que desenvolvemos costumam apresentar essas informações — extraídas do modelo interno do mundo do robô — sob a forma de imagens num ecrã de computador: uma janela direta para a mente da máquina. Nestes modelos internos do mundo, vejo os primórdios de uma consciência nascente nas mentes das nossas máquinas — uma consciência que acredito virá a evoluir até se tornar comparável à dos seres humanos.

O termo evolução *convergente* é usado pelos biólogos evolutivos sempre que espécies apenas remotamente aparentadas desenvolvem, de forma independente, características semelhantes — presumivelmente em resposta a pressões ambientais semelhantes. Os olhos são um exemplo de evolução convergente; evoluíram mais de quarenta vezes de forma independente no reino animal. O que era necessário era a presença de células sensíveis à luz e a existência de pressões seletivas que favorecessem a sobrevivência de animais capazes de ver, ainda que de forma rudimentar. Se uma função do sistema nervoso tão complexa como a visão pôde evoluir tantas vezes sempre que as condições ambientais o permitiram, que dizer então das emoções e da consciência? Ao contrário da visão, estas características da mente humana não têm manifestação externa inequívoca — e nem sequer possuem uma definição precisa. A sua existência em animais, e até em seres humanos, foi posta em causa por uma geração de psicólogos comportamentais. Ainda assim, etólogos como Donald Griffin consideram os conceitos úteis para explicar o comportamento animal. Se um animal age como eu quando estou

com medo, não será razoável designar o seu estado mental como “medo”? E se escolhe entre várias alternativas complexas para lidar com uma situação nova — tal como eu ponderaria conscientemente as minhas opções nas mesmas circunstâncias — por que não atribuir-lhe consciência, em vez de invocar algum outro mecanismo com um nome diferente, mas que resulta no mesmo efeito?

Consideremos a seguinte experiência mental. Suponhamos que desejamos construir um robô capaz de executar uma tarefa geral, como: “Vai pelo corredor até à terceira porta, entra, procura uma chávena e traz-ma de volta.” A nossa necessidade mais urgente seria uma linguagem de programação que nos permitisse especificar tarefas complexas para o robô explorador, bem como um sistema de hardware e software capaz de materializá-la. As linguagens de controlo sequencial, que têm sido utilizadas com sucesso em manipuladores industriais, poderiam parecer um bom ponto de partida. No entanto, uma análise mais aprofundada, em papel, das estruturas e ações elementares necessárias para a mobilidade revelaria que a estrutura de controlo linear dessas linguagens — ainda que adequada para um braço robótico — se revelaria insuficiente para um robô móvel. A diferença essencial reside no facto de que um robô móvel, nas suas deambulações, é regularmente “surpreendido” por eventos imprevistos com os quais tem de lidar. Isso exige que rotinas de contingência possam ser ativadas em ordem arbitrária e executadas em simultâneo, cada uma com acesso próprio aos sensores, aos efetores, ao estado interno da máquina, e a um mecanismo de arbitragem para resolver conflitos entre elas. À medida que as condições se alteram, a prioridade entre os módulos muda, podendo o controlo ser transferido de um módulo para outro.

Um pedido dirigido ao nosso robô do futuro — para ir pelo corredor até à terceira porta, entrar, procurar uma chávena e trazê-la de volta — poderia ser implementado sob a forma de um módulo denominado (BUSCAR-CHÁVENA), que se assemelha muito a um programa escrito em linguagens de controlo de braços robóticos (as quais, por sua vez, se assemelham a linguagens de programação como Algol ou Basic), com a diferença de que outro módulo, chamado (CONTAR-PORTAS), estaria a ser executado em

simultâneo com a rotina principal. Consideremos o seguinte esboço de tal programa.

Módulo CONTAR-PORTAS:

- Verifica os arredores do robô em busca de portas
- Incrementa em uma unidade a variável NÚMERO-DE-PORTAS sempre que uma nova porta for localizada
- Regista a localização da nova porta na variável LOCALIZAÇÃO-DA-PORTA

Módulo IR-BUSCAR-CHÁVENA:

Passo 1: Regista a localização atual do robô na variável LOCALIZAÇÃO-INICIAL

Passo 2: Define a variável NÚMERO-DE-PORTAS com o valor zero

Passo 3: Ativa o módulo CONTAR-PORTAS

Passo 4: Conduz o robô paralelamente à parede do lado direito até que o valor de NÚMERO-DE-PORTAS seja igual ou superior a três

Passo 5: Faz o robô virar-se para a localização registada em LOCALIZAÇÃO-DA-PORTA

Passo 6: Se o robô estiver virado para uma porta aberta, passa para o Passo 10

Passo 7: Se o robô não estiver virado para uma porta, subtrai um ao valor de NÚMERO-DE-PORTAS e volta ao Passo 4

Passo 8: Se o robô estiver virado para uma porta fechada, tenta abri-la

Passo 9: Se não conseguir abrir a porta, diz “toc toc” e volta ao Passo 6

Passo 10: Conduz o robô através da porta aberta

Passo 11: Verifica os arredores em busca de chávenas; se não encontrar nenhuma, passa ao Passo 15

Passo 12: Regista a localização da chávena mais próxima na variável LOCALIZAÇÃO-DA-CHÁVENA

Passo 13: Conduz o robô até uma distância que lhe permita alcançar a LOCALIZAÇÃO-DA-CHÁVENA

Passo 14: Pega na chávena em LOCALIZAÇÃO-DA-CHÁVENA; se falhar, passa ao Passo 15

Passo 15: Volta e posiciona o robô virado para a porta em LOCALIZAÇÃO-DA-PORTA

Passo 16: Se estiver virado para uma porta fechada, tenta abri-la

Passo 17: Se a porta não abrir, diz “toc toc” e volta ao Passo 16

Passo 18: Conduz o robô através da porta aberta

Passo 19: Regressa à LOCALIZAÇÃO-INICIAL

Passo 20: Coloca o robô em modo de suspensão

Até aqui, tudo corre bem. Ativamos o programa, e o robô começa obedientemente a deslocar-se pelo corredor, contando as portas. Reconhece corretamente a primeira. A segunda porta, infelizmente, está decorada com cartazes berrantes, e a iluminação nessa parte do corredor é deficiente, pelo que o nosso reconhecedor experimental de portas falha na sua deteção. O seguidor de paredes, no entanto, continua a funcionar corretamente, e o robô prossegue pelo corredor, com a contagem de portas reduzida em uma unidade. Reconhece a terceira porta — aquela pela qual lhe pedimos para passar — mas, por pensar que é apenas a segunda, continua. A porta seguinte é corretamente reconhecida e está aberta. O programa, convencido de que se trata da terceira, faz o robô virar-se para ela e avançar. Esta quarta porta, infelizmente, dá acesso à caixa de escadas, e o pobre robô, não estando equipado para se deslocar em escadas, encontra-se em perigo de vida.

Felizmente, existe outro módulo no nosso sistema de programação concorrente chamado DETETAR-PRECIPÍCIO. Este programa está sempre a correr e verifica os dados sobre a posição do solo gerados incidentalmente pelos processos de visão, bem como solicita verificações de proximidade do solo por sonar e infravermelhos. Combina estas informações, talvez com uma

expetativa *a priori* de encontrar um desnível — ajustada para um valor elevado ao operar em zonas perigosas — para produzir um número que indica a probabilidade de haver uma queda nas proximidades. Um processo complementar, LIDAR-COM-PRECIPÍCIO, também corre continuamente, embora com baixa prioridade, verificando regularmente esse número e ajustando a sua própria prioridade com base no mesmo. Quando a variável de probabilidade de precipício (a que talvez chamemos VERTIGEM) se torna suficientemente alta, a prioridade do LIDAR-COM-PRECIPÍCIO ultrapassa a do processo atualmente no controlo — neste caso, IR-BUSCAR-CANECA — e o LIDAR-COM-PRECIPÍCIO assume o controlo do robô. Um LIDAR-COM-PRECIPÍCIO bem escrito fará então com que o robô pare ou abrande significativamente os seus movimentos, aumente a frequência das medições dos sensores para deteção de precipício e recue lentamente quando este tiver sido identificado e localizado com fiabilidade.

Há aqui algo curioso nesta sequência de ações. Uma pessoa que observasse este comportamento, sem conhecimento dos mecanismos internos do robô, poderia oferecer a seguinte interpretação: «Primeiro, o robô estava determinado a passar pela porta, mas depois viu as escadas e ficou tão assustado e absorvido com isso que se esqueceu completamente do que estava a fazer». Sabendo nós o que realmente aconteceu dentro do robô, poderíamos sentir-nos tentados a repreender essa pessoa por utilizar conceitos antropomórficos tão imprecisos como determinação, medo, absorção e esquecimento para descrever as ações de uma máquina. Poderíamos fazê-lo, mas, na minha opinião, isso seria um erro. O robô manifestou as suas falhas e reações com a mesma honestidade que qualquer animal vivo; o comportamento observado representa a conduta correta para um ser a operar com dados incertos num mundo perigoso. Um polvo em perseguição de uma presa pode ser desviado por ameaças subtis de perigo exatamente da mesma forma que o robô foi. Curiosamente, o polvo — um invertebrado — possui um sistema nervoso que evoluiu de forma completamente independente do nosso, enquanto vertebrados. No entanto, a maioria de nós não hesita em atribuir qualidades como paixão, prazer, medo e dor às ações deste animal.

Acredito que estamos perante um caso de evolução convergente no comportamento de uma pessoa, de um polvo e de um robô. As exigências da vida móvel conspiraram, nos três casos, para criar uma entidade com modos de operação distintos para diferentes circunstâncias, e com a capacidade de mudar rapidamente de modo com base em dados incertos e ruidosos, suscetíveis de serem mal interpretados. À medida que a complexidade dos robôs móveis aumenta, a sua semelhança com os animais — e com os humanos — tornar-se-á ainda mais acentuada.

Esperre aí um momento, dirá você. Pode haver alguma semelhança entre a reação do robô perante uma situação perigosa e a de um animal, mas certamente existem diferenças. O robô não será mais parecido com uma aranha sobressaltada, ou até com uma bactéria, do que com um ser humano assustado? Não reagiria ele da mesma forma, repetidamente, mesmo que a situação afinal não fosse perigosa? Apanhou-me. Na verdade, penso que o sistema nervoso de uma aranha é uma excelente analogia para os programas de robô atualmente possíveis. (Já ultrapassámos o estágio bacteriano nos anos 1950, com as tartarugas eletrónicas que seguiam a luz.) Mas isto não significa que conceitos como pensamento e consciência tenham de ser descartados.

No seu livro *Animal Thinking (Pensamento Animal)*, Griffin analisa evidências que sugerem que grande parte do comportamento animal — incluindo o comportamento dos insetos — pode ser explicada de forma parcimoniosa em termos de consciência: um modelo interno do “eu”, do meio envolvente e de outros indivíduos, que, por mais rudimentar que seja, permite considerar ações alternativas. Por exemplo, as abelhas, como Otto von Frisch descobriu, comunicam a direção, distância e atratividade de uma fonte de alimento a outros membros da colmeia através da direção, duração e vigor de cada impulso numa “dança do abanico”. Martin Lindauer expandiu as observações de Frisch a casos em que um enxame de uma colónia sobrepovoada procura um novo local para se instalar. Uma operária do enxame sai em busca de cavidades adequadas e regressa quando encontra uma e a explora meticulosamente. Executa então uma dança do abanico sobre a superfície do enxame, descrevendo a localização e qualidade da sua

descoberta. Entretanto, outras operárias relatam outros locais. Os sítios promissores são visitados e examinados cuidadosamente por outros membros do enxame, que regressam para contar a sua versão. Uma abelha que comunica a localização de um determinado local não é afetada por outra a transmitir a mesma mensagem, mas pode ser “convertida” por uma demonstração suficientemente enfática e repetida que descreva um local diferente. O debate prolonga-se por vários dias, com visitas repetidas a um número cada vez menor de locais candidatos, até que se atinge uma quase unanimidade. Todo o enxame voa então para se estabelecer na cavidade vencedora. Este comportamento pode ser explicado se postulamos a existência, no cérebro de cada abelha, de um mapa simples que descreve os locais e a sua atratividade — mapas que podem ser modificados tanto pelas experiências complexas da exploração como pelas mais simples da comunicação — e que podem tornar-se a base para a tomada de decisões.

Um modelo interno do mundo suficientemente complexo para permitir escolhas comportamentais — quer chamemos ou não a esse modelo “consciência” — é precisamente aquilo que os roboticistas estão atualmente a tentar alcançar com os seus robôs móveis. De facto, a investigação em robótica é demasiado prática para se propor, de forma explícita, o objetivo de produzir máquinas com características tão nebulosas e controversas como emoção e consciência. Seria suficiente que as nossas máquinas fossem capazes de “ganhar a vida” perante as muitas surpresas, contratemplos, oportunidades, obstáculos e concorrentes que encontrarão no mundo. Mas a seleção natural, o mecanismo orientador da evolução darwiniana, é igualmente utilitária — e, no entanto, cá estamos nós, com sentimentos e um sentido de identidade.

Em *The Growth of Biological Thought (O Desenvolvimento do Pensamento Biológico)*, o biólogo evolucionista Ernst Mayr observa que tanto os sistemas vivos como os não vivos “quase sempre possuem a característica de que as propriedades do todo não podem (mesmo em teoria) ser deduzidas do conhecimento mais completo possível das suas componentes, tomadas isoladamente ou em combinações parciais”. A emergência — esse surgimento de

propriedades novas nos sistemas enquanto totalidade — tem sido frequentemente invocada para explicar realidades biológicas difíceis, como a mente, a consciência e até a própria vida. É assim que eu imagino que algumas das experiências mentais mais misteriosas, que tendemos a associar aos seres humanos, possam eventualmente emergir nas nossas máquinas, à medida que perseguimos objetivos puramente funcionais.

Aprendizagem

Quando é tocada de forma ligeira, a lesma-do-mar *Aplysia* retrai as suas delicadas brânquias para dentro do corpo. Se esse estímulo for repetido com frequência e sem causar qualquer dano, a *Aplysia* acaba por aprender a ignorar o incômodo, mantendo as brânquias expostas. Contudo, se mais tarde esses toques forem seguidos por estímulos agressivos — como o contato com um ácido forte — o reflexo de retração regressa com força redobrada. Em ambos os casos, o comportamento modificado permanece na memória durante horas. A *Aplysia* tem sido estudada de forma tão exaustiva nas últimas décadas que os neurónios envolvidos neste reflexo estão bem identificados, e a aprendizagem foi recentemente rastreada até alterações químicas em sinapses específicas desses neurónios. Redes neuronais mais complexas podem adaptar-se de formas ainda mais elaboradas — por exemplo, aprendendo a associar pares específicos de estímulos entre si. Esses mecanismos ajustam o sistema nervoso tanto ao corpo que o alberga como ao ambiente envolvente. Nos vertebrados, grande parte da sua flexibilidade comportamental deve-se ao desenvolvimento sofisticado deste tipo de sistema — redes que podem ser ativadas a partir de múltiplos pontos e que encorajam ou desencorajam a repetição futura de comportamentos recentes. Embora a arquitetura neuronal desses sistemas nos vertebrados ainda não seja plenamente compreendida, o seu efeito é evidente nas sensações subjetivas que chamamos de prazer e dor.

Os sistemas robóticos atuais, na melhor das hipóteses, estão configurados para aprender algumas poucas coisas específicas a partir do ambiente — uma sequência simples de movimentos, a

localização de um componente esperado, a posição de obstáculos próximos, ou talvez alguns parâmetros que permitam um melhor controle de um motor ou a interpretação de um sensor. Não faz muito sentido programá-los para coordenar as suas ações de forma complexa, quando mal conseguimos que façam uma única tarefa com competência. Contudo, este estado primitivo não durará para sempre. A começar, talvez, com o robô universal que descrevi anteriormente, tornar-se-á desejável incorporar capacidades de aprendizagem mais generalistas.

A segurança e utilidade de um robô no ambiente doméstico aumentariam significativamente se este pudesse aprender a evitar perigos idiossincráticos e a tirar partido de oportunidades. Se uma certa porta num percurso habitual estiver frequentemente trancada, seria útil que o robô aprendesse a preferir um caminho mais longo, mas mais fiável. Uma tarefa poderia ser executada de forma mais eficaz se o robô pudesse aprender — ou até antecipar — a localização variável de um ingrediente necessário, com base em indícios subtis. É impossível programar explicitamente o robô para cada eventualidade deste tipo, mas muito poderia ser conseguido através de um mecanismo unificado de condicionamento, que aumentasse a probabilidade de decisões que se mostraram eficazes no passado em circunstâncias semelhantes e diminuísse a probabilidade daquelas que conduziram a atividades inúteis ou perigosas.

O software de condicionamento que tenho em mente receberia dois tipos de mensagens oriundas de qualquer parte do robô: uma indicando sucesso e outra sinalizando problema. Algumas — como os sinais de baterias carregadas ou de colisões iminentes — seriam geradas pelo sistema operativo básico do robô. Outras, mais específicas para a realização de determinadas tarefas, poderiam ser iniciadas pelos programas correspondentes a essas tarefas. Chamarei às mensagens de sucesso "prazer" e às mensagens de perigo "dor". A dor tenderia a interromper a atividade em curso, enquanto o prazer aumentaria a probabilidade da sua continuação.

As mensagens também forneceria dados de entrada para um programa que utilizaria técnicas estatísticas para "catalogar", de

forma compacta, o tempo, a posição, a atividade, o ambiente e outras propriedades conhecidas pelo robô que precederam o sinal. Um "reconhecedor" monitorizaria constantemente estas variáveis, comparando-as com os registos existentes no catálogo. Sempre que ocorresse um conjunto de condições semelhante àquele que frequentemente antecederam uma dor (ou um prazer) no passado, o reconhecedor emitiria ele próprio uma mensagem de dor (ou prazer) um pouco mais fraca. No caso da dor, esta mensagem de aviso poderia impedir a repetição da atividade que anteriormente causara problemas. Com o tempo, essas mensagens de aviso acumular-se-iam também no catálogo, e o robô começaria a evitar os passos que conduzissem aos passos que, por sua vez, originaram o problema inicial. Eventualmente, uma longa cadeia de associações deste tipo poderia evitar problemas logo numa fase muito precoce. Existem, evidentemente, armadilhas neste processo. Se a intensidade dos avisos secundários não enfraquecer suficientemente à medida que a cadeia se alonga, a dor poderia transformar-se numa fobia incapacitante e o prazer numa adição igualmente paralisante.

Para além de permitir ao robô adaptar-se de forma oportunista ao seu ambiente, um mecanismo de prazer-dor poderia também ser explorado de modo mais direcionado por programas de aplicação. Suponhamos que o robô possui um reconhecedor de palavras faladas. Um módulo que simplesmente gerasse um sinal de prazer ao ouvir a palavra "bom" e um sinal de dor ao ouvir "mau" permitiria a um utilizador modificar facilmente o comportamento do robô. Se este estivesse a incomodar ao aspirar um espaço enquanto este se encontrava em uso, bastariam alguns "mau!" para o treinar a parar até que as condições mudassem — por exemplo, noutra hora do dia ou quando a divisão estivesse vazia.

Um robô com software de condicionamento poderia ser programado para se treinar a si próprio. Se uma tarefa num programa de aplicação exigisse que um determinado tipo de recipiente fosse aberto, poderia escrever-se uma lista detalhada de instruções descrevendo exatamente como segurar, rodar e puxar para concluir a ação. Em alternativa, um robô na fábrica poderia ser programado para pegar em muitos desses recipientes, um a seguir

ao outro, e empurrar, rodar, agitar e puxar aleatoriamente cada um até que se abrisse — ou se partisse. O programa de treino reconheceria ambas as situações e emitiria uma mensagem de prazer num caso e uma de dor no outro, antes de passar ao recipiente seguinte. Gradualmente, o sistema de condicionamento inibiria as sequências que causaram quebras e reforçaria aquelas que tiveram sucesso. Uma versão abstraída do catálogo de treino dessa sessão poderia então ser inserida no programa de aplicação final, substituindo as instruções explícitas e sendo, talvez, combinada com catálogos de outras partes da tarefa, desenvolvidos por outros robôs.

Paciência infinita seria uma vantagem numa sessão de treino, mas poderia tornar-se exasperante num robô em operação no mundo real. No programa de ir buscar uma chávena que descrevi anteriormente, talvez tenha reparado que, se o robô encontra a porta fechada e é incapaz de a abrir, limita-se a ficar ali parado, repetindo incessantemente “toc toc” até que alguém lhe abra a porta. Um robô que se comportasse frequentemente desta forma — e muitos robôs atuais assim o fazem — teria um desempenho fraco em companhia humana. Curiosamente, é possível induzir insetos a este tipo de repetição automática. Algumas vespas providenciam alimento para os seus ovos em incubação paralisando lagartas e depositando-as num buraco subterrâneo. Normalmente, a vespa escava um buraco e sela a entrada, depois parte em busca de uma lagarta. Ao regressar com a presa, deixa-a junto à entrada, reabre o buraco e arrasta-a para o interior. No entanto, se um experimentador mover a lagarta alguns centímetros enquanto a vespa está ocupada na entrada, ela irá buscá-la novamente e repetirá todo o ritual de abrir o buraco — mesmo que este já esteja aberto. Se, durante esse processo, o experimentador voltar a deslocar a lagarta, ela repetirá tudo outra vez. Este ciclo pode, aparentemente, repetir-se indefinidamente, até que a vespa ou o experimentador sucumba de exaustão. Um robô poderia ser protegido contra este tipo de situação por um módulo que detetasse comportamentos repetitivos e gerasse um fraco sinal de dor a cada repetição. No exemplo do bater à porta, o comportamento seria gradualmente inibido, libertando o robô para outras tarefas

pendentes ou mesmo para a inatividade. O robô teria, assim, adquirido a capacidade de se aborrecer.

Módulos que reconhecessem outras condições e enviassem sinais de dor ou prazer com intensidade apropriada dotariam o robô de uma personalidade única. Um grande e perigoso robô industrial, com um detetor de presença humana que emitisse um sinal de dor, tornar-se-ia mais cauteloso na presença de seres humanos, tornando-se assim menos propenso a causar ferimentos. Um módulo que registasse prazer ao encontrar novos detritos, e dor ao vê-los novamente, poderia tornar um programa de limpeza extremamente criativo e agressivo na sua luta contra a sujeira.

Imagens

Robôs com capacidade de aprendizagem rápida seriam capazes de lidar com programas que envolvessem múltiplas ações alternativas em cada etapa de uma tarefa — alternativas essas que proporcionariam ao robô uma ampla margem para a criatividade. Contudo, um robô equipado apenas com um sistema simples de condicionamento aprenderia lentamente. Seriam necessárias muitas repetições para que surgissem correlações estatisticamente significativas no catálogo de condicionamento. Certas situações no mundo real não toleram essa abordagem tão morosa. Um robô que, repetidamente, se aventurasse numa estrada pública, sendo lento a registrar o perigo desse local, poderia ser subitamente transformado em sucata. Um robô, ou software, que demorasse a adaptar-se a condições ou oportunidades variáveis numa casa poderia perder a batalha pela sobrevivência económica perante um produto mais ágil de outro fabricante.

A aprendizagem poderia ser enormemente potenciada pela adição de outro módulo essencial: um simulador geral do mundo. Já o robô universal mais rudimentar que descrevi recorre, em certa medida, à simulação. Para alcançar o seu destino em segurança, o programa do robô consulta um mapa interno do ambiente e avalia múltiplos caminhos alternativos até encontrar o mais eficaz. Estas deliberações são simulações de ações hipotéticas do robô.

Processos semelhantes ocorrem quando o robô decide como pegar num objeto ou quando considera diferentes interpretações daquilo que vê com as câmaras. Contudo, cada um destes procedimentos é especializado, modela apenas um aspeto do mundo e só pode ser utilizado para uma função específica. Imaginemos agora que o robô dispunha de um simulador muito mais poderoso, capaz de modelar situações hipotéticas complexas envolvendo o robô e muitos aspetos do seu ambiente. Um programa de aplicação poderia utilizar esse simulador para testar uma ação proposta quanto à sua segurança e eficácia, sem colocar o robô em risco.

Mas a verdadeira revolução acontece quando os acontecimentos do simulador são encaminhados para o mecanismo de condicionamento. Nesse caso, um desastre simulado (por exemplo, uma queda hipotética do robô) poderia, na vida real, condicionar o robô a evitar o acontecimento precursor da simulação (por exemplo, vaguear junto ao topo de uma escadaria simulada). O robô poderia assim preparar-se para muitos problemas e oportunidades futuras, simulando cenários possíveis nos seus tempos de inatividade. Esses cenários poderiam ser simplesmente variações dos eventos reais do dia. Dotado desta capacidade, o robô passaria a ser capaz de recordar, imaginar e sonhar.

A imaginação gerada por um simulador só é útil se este for capaz de produzir previsões razoavelmente exatas sobre o mundo real. Para tal, é necessário um vasto conhecimento sobre o mundo, e imagino que o desenvolvimento competitivo de simuladores cada vez mais precisos venha a constituir uma parte fundamental da investigação na indústria da robótica durante as primeiras décadas do século XXI. As empresas de robótica observarão os erros e limitações dos seus robôs, tanto em laboratório como no terreno, e afinarão os simuladores com o objetivo de melhor modelar os aspetos do mundo mais relevantes para o desempenho das suas máquinas. A evolução fez o mesmo connosco ao longo dos eões do nosso desenvolvimento. Os simuladores serão fornecidos de fábrica com um conhecimento genérico, mas também terão de ser capazes de aprender as idiossincrasias de cada novo local onde o robô for colocado. Robôs avançados poderão encontrar-se a trabalhar com outros robôs e com seres humanos. Tal interação tornar-se-ia mais

eficaz se os simuladores destas máquinas fossem capazes de prever, até certo ponto, o comportamento dos outros. Parte dessa previsão poderia envolver a modelação aproximada do estado mental do outro, de modo a antecipar as suas reações perante diferentes alternativas de ação. Abre-se assim uma nova e rica esfera de possibilidades quando existe um modelo interno do estado de espírito de outro ser. Por exemplo, um módulo que gerasse sinais de dor ao detetar sofrimento nesse modelo mental simulado condicionaria o robô a atuar de forma benevolente. E um robô poderia até ser repreendido por atribuir de forma inadequada sentimentos ou intenções “robomórficas” a outras máquinas — ou até a seres humanos!

Seria, evidentemente, tão fácil programar robôs para cometer crimes como para realizar tarefas socialmente aceites, e certamente serão desenvolvidas formas legais de atribuir culpa quando isso acontecer. Contudo, robôs complexos poderão, por vezes, meter-se em sarilhos por sua própria iniciativa. Imaginemos um robô equipado com simulador que, em diversas ocasiões no passado, tenha sofrido consequências graves por não conseguir recarregar as baterias a tempo. Esse robô estaria, por isso, fortemente condicionado a evitar que o seu nível de energia descasse demasiado. Suponhamos que se encontra trancado fora da casa dos seus donos, com a bateria a esgotar-se. O seu simulador começa então a percorrer freneticamente diferentes cenários, à procura de uma solução — uma combinação de ações que lhe permita obter recarga. À medida que as combinações de comportamentos convencionais falham em aproximá-lo do seu objetivo, a pesquisa do simulador alarga-se a possibilidades mais invulgares. A casa dos vizinhos está próxima, a porta poderá estar aberta, e é provável que existam tomadas elétricas no interior — o simulador descobre um cenário em que o robô alcança essas tomadas. Há dor associada ao facto de abandonar o território familiar, e ao incómodo potencial que possa causar, mas essa dor é largamente superada pelo prazer — e alívio da dor — que resulta da possibilidade de recarregar. O robô repete várias vezes a simulação da invasão da casa dos vizinhos, reforçando a cada iteração o condicionamento dos passos envolvidos, tornando a concretização do ato cada vez mais

provável. Eventualmente, o condicionamento torna-se suficiente, e o robô inicia um percurso que provavelmente o conduzirá a maiores problemas do que aqueles que o seu simulador imperfeito foi capaz de antecipar. Não será a primeira criatura a ser levada a um ato desesperado por uma necessidade extrema.

Podemos levar esta evolução especulativa mais longe, dotando gradualmente os nossos robôs com sentimentos e capacidades intelectuais semelhantes às dos seres humanos. No entanto, prevejo que, quando os robôs estiverem prontos para receber tais capacidades, estas já estarão amplamente disponíveis para aquisição em larga escala junto da indústria tradicional da inteligência artificial, a qual terá prosseguido a sua estratégia de desenvolvimento “de cima para baixo” em paralelo com a evolução “de baixo para cima” dos robôs. A união destas abordagens poderá levar muitos anos a consumir-se plenamente, levantando questões como a de saber de que modo o sistema de raciocínio poderá aceder ao simulador para daí extrair lampejos de intuição, ou como deverá o raciocínio influenciar o sistema de condicionamento, de forma a conseguir sobrepor-se aos instintos do robô em circunstâncias excepcionais. A combinação de ambos dará origem a seres que, sob certos aspetos, se assemelharão a nós — mas que, noutros, serão como nada que o mundo alguma vez tenha visto.

CAPÍTULO 2

REFORÇANDO O DESEMPENHO

DURANTE a década de 1970, enquanto eu era estudante de pós-graduação, pareceu-me que a capacidade de processamento disponível para os programas de inteligência artificial não estava a aumentar muito rapidamente. Em 1970, o meu trabalho era realizado num computador central PDP-10 da Digital Equipment Corporation, que servia uma comunidade de talvez trinta pessoas. Em 1980, o meu computador era um DEC KL-10, cinco vezes mais rápido e com cinco vezes mais memória do que o anterior — mas a servir o dobro dos utilizadores. Pior ainda: o pouco ganho de velocidade restante parecia ter sido absorvido por funcionalidades dispendiosas em termos computacionais, como sistemas de partilha de tempo mais sofisticados, linguagens de alto nível, gráficos, editores de texto visuais, sistemas de correio eletrónico, redes informáticas e outros luxos que, entretanto, se tinham tornado em necessidades.

Vários fatores contribuíram para este estado de coisas no que respeita ao hardware informático. O apoio à ciência universitária em geral tinha diminuído após as missões Apollo à Lua e a guerra do Vietname, deixando as universidades a arrastarem-se com equipamento obsoleto. As mesmas condições provocaram uma recessão nas indústrias tecnológicas: engenheiros desempregados abriram restaurantes de fast food em vez de projetarem computadores. O impulso inicial bem-sucedido da inteligência artificial na resolução de problemas ainda não se tinha esgotado, e a muitos ainda parecia que as máquinas existentes seriam suficientemente poderosas — se ao menos se descobrissem os programas certos. Contudo, o progresso na própria investigação tornava-se lento, difícil e frustrante, e muitos dos melhores programadores foram atraídos para a atividade mais recompensadora de construir ferramentas de software atrativas, embora dispendiosas do ponto de vista computacional — cujo sucesso, por sua vez, inspirava ainda mais criação de ferramentas.

Se a década de 1970 foi um período de estagnação para o hardware informático, a década de 1980 veio mais do que compensar. Tal como a inteligência artificial recebeu o seu primeiro impulso nos anos 1960 com o salto soviético para o espaço, a segunda etapa foi desencadeada nesta década com o salto japonês para o mercado norte-americano. Os êxitos industriais do Japão concentraram a atenção mundial na importância da tecnologia — especialmente dos computadores e da automação — nas economias modernas. As indústrias e o governo dos Estados Unidos reagiram com investimentos em investigação. Os japoneses, por sua vez, alimentaram o ímpeto, sob a influência de um pequeno grupo de investigadores seniores, ao anunciarem audaciosamente uma grande iniciativa rumo aos computadores do futuro: o chamado projeto da Quinta Geração, que procurava expandir-se nas direções de investigação mais promissoras nos Estados Unidos e na Europa. Os americanos responderam com mais financiamento.

Para além deste impulso económico, a circuitaria integrada tinha evoluído ao ponto de, nos anos 1980, permitir que um computador inteiro coubesse num único chip. De repente, os computadores tornaram-se acessíveis a indivíduos, e uma nova geração de utilizadores e fabricantes de computadores emergiu. No extremo oposto da escala, os supercomputadores — outrora reservados a um punhado de laboratórios e agências governamentais — tornaram-se populares em centenas de ambientes industriais e de investigação. Ao longo de toda a gama de tamanhos, a indústria informática tornou-se mais lucrativa e competitiva do que nunca, com novas gerações de máquinas mais rápidas e baratas a serem lançadas a um ritmo frenético.

Até que ponto terá de continuar esta evolução até que as nossas máquinas sejam suficientemente poderosas para se aproximarem do intelecto humano? Conhece-se demasiado pouco, tanto sobre o funcionamento global do cérebro humano como sobre o modo de operação de um computador verdadeiramente inteligente, para que se possa fazer essa estimativa de forma direta. Abordei o problema de forma indireta, comparando um fragmento do sistema nervoso que está moderadamente bem compreendido — a retina do olho — com programas de visão computacional que

executam aproximadamente a mesma função. A partir dessa comparação, extrapolo a proporção para o cérebro inteiro, com o objetivo de estimar a potência de processamento necessária numa máquina que o mimetize. O momento da chegada de uma máquina com esse poder é então calculado ao projetar para o futuro a tendência da evolução da potência computacional por unidade de custo ao longo deste século.

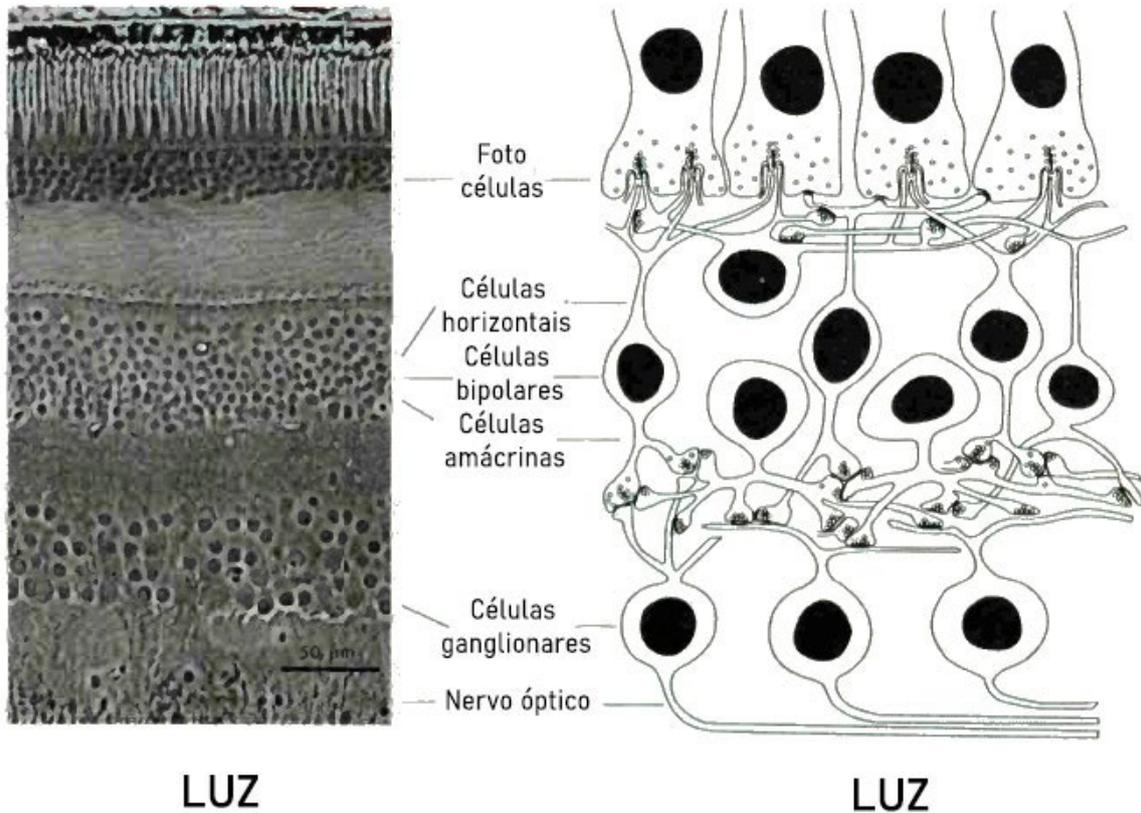
As estimativas computacionais que utilizarei para fazer estas comparações provêm da minha própria investigação. A neurobiologia é extraída da obra de referência de John Dowling, *The Retina*, e do manual clássico de Stephen Kuffler e John Nicholls, *From Neuron to Brain*. Os números são incertos, pois tanto a visão computacional como a nossa compreensão da visão biológica (sem falar noutras funções cerebrais) ainda estão na sua infância. Muitos fundamentos permanecem envoltos em mistério neste domínio complexo. Felizmente, a minha comparação não exige uma precisão diabólica; erros na ordem das 100 vezes para cima ou para baixo não farão grande diferença qualitativa, dada a escala logarítmica ampla utilizada neste capítulo. Além disso, espero que alguns dos meus erros sigam direções opostas e, assim, se anulem parcialmente. Há curvas perigosas nesta montanha-russa rumo à equivalência com o ser humano — por isso, segure-se bem!

Circuitos Neurais

A retina é, na verdade, uma extensão alongada do cérebro. No entanto, a sua localização na parte posterior do globo ocular, a alguma distância do volume principal do cérebro, tornou-a comparativamente fácil de estudar, mesmo em animais vivos. Uma vez removida do corpo, pode continuar a funcionar durante horas, com os seus estímulos de entrada e de saída altamente acessíveis. Transparente e mais fina do que uma folha de papel, a retina pode ser corada com corantes para tornar neurónios específicos visíveis ao microscópio ótico e eletrónico. Por estas razões, a retina é provavelmente a parte mais estudada do sistema nervoso dos vertebrados. Iremos examiná-la com algum detalhe, mas primeiro, um pouco de contexto sobre as células nervosas.

Todos os neurónios, como outras células, são mecanismos complexos e impressionantes. Começam a sua existência diferenciando-se de células estaminais nas fases iniciais do desenvolvimento embrionário, passando depois por ciclos repetidos de migração — rastejando como amibas — até destinos precisos por todo o corpo, onde continuam a dividir-se e a diferenciar-se. Quando atingem a sua localização final, estendem prolongamentos fibrosos que procuram estabelecer ligações específicas com outros neurónios, através de junções chamadas sinapses. Diferentes subpopulações de neurónios diferem radicalmente em geometria, tamanho e função. Alguns neurónios possuem milhares de pequenas fibras chamadas dendrites e podem albergar centenas de milhares de sinapses. Uma dessas fibras, conhecida como axónio, pode crescer até vários centímetros de comprimento — um milhão de vezes o tamanho original da célula.

Um neurónio típico recebe mensagens através das suas dendrites e transmite-as pelo axónio, que pode ramificar-se na extremidade. A sua sinalização ocorre por meio de diferenças de potencial elétrico de alguns milivolts através da membrana exterior. Essa voltagem é mantida, num ambiente húmido e eletricamente condutor, por bombas iónicas moleculares na membrana, que movimentam iões carregados — potássio, sódio, cálcio, cloro, entre outros — para dentro e para fora da célula. Estas bombas são ativadas ou inibidas por pequenas moléculas chamadas neurotransmissores, produzidas por outros neurónios e entregues através de diversos tipos de sinapse.



A Retina. Uma vista transversal através da espessura de 1/2 milímetro de uma retina humana. O cristalino está localizado na parte inferior, e a luz precisa de penetrar três camadas do circuito de processamento de imagem para atingir as fotocélulas sensíveis à luz na parte superior. As células horizontais calculam a intensidade média em grandes áreas de fotocélulas; as células bipolares subtraem estas médias de grandes áreas das médias em áreas mais pequenas para produzir sinais "centro-periferia" que são mais fortes nas extremidades dos objetos. Algumas células amácrinas intensificam ainda mais os sinais centro-periferia; outras sinalizam objetos em movimento. As células ganglionares completam os cálculos e enviam os resultados como pulsos ao longo dos seus longos axónios para locais profundos no cérebro.

Quando um neurónio recebe uma descarga de neurotransmissores, a voltagem da sua membrana pode aumentar ou diminuir, consoante o tipo de sinapse e de neurónio. Se a voltagem for suficientemente reduzida por múltiplos sinais, ocorre uma espécie de curto-circuito: a voltagem colapsa subitamente por completo, e esse colapso propaga-se ao longo do axónio sob a

forma de um impulso. Quando o impulso atinge uma sinapse que liga o neurónio a outro, desencadeia a libertação de neurotransmissores a partir de pequenos sacos situados na membrana. Estes neurotransmissores difundem-se através da sinapse, acabando por elevar ou reduzir a voltagem da segunda célula. Entretanto, as bombas do primeiro neurónio trabalham para restaurar a voltagem original e, em apenas alguns milésimos de segundo, a célula está pronta para disparar novamente. A frequência com que estes impulsos se repetem codifica a intensidade do estímulo; podem ser produzidos desde zero até várias centenas de impulsos por segundo. Os impulsos são usados para comunicações de longa distância, mas os neurónios próximos entre si, como os encontrados na retina, comunicam frequentemente simplesmente respondendo às variações graduais da voltagem uns dos outros. Para além das ligações sinápticas com outras células, muitos neurónios e sinapses possuem recetores para certas classes de neurotransmissores livres, transportados pelo sangue a partir de outras regiões do sistema nervoso ou de outros órgãos do corpo, que inibem ou reforçam a resposta do neurónio.

No núcleo do neurónio, atuam processos genéticos mais lentos que produzem neurotransmissores e os transportam ao longo do axónio até aos sacos de armazenamento. A maquinaria genética do neurónio também embala energia, constrói e repara estruturas, e realiza todas as outras funções extraordinárias que qualquer célula precisa desempenhar para se manter funcional. Felizmente para aqueles de nós que trabalham em imitações eletrónicas do sistema nervoso, grande parte dessa complexidade não está diretamente envolvida na perceção, ação e pensamento. Muito do mecanismo do neurónio serve para fazer crescer e construir um organismo de dentro para fora. Mesmo as suas operações de processamento de informação parecem adaptadas a essa necessidade evolutiva — e isso nota-se.

Neste estágio da tecnologia informática, é mais fácil manter a maquinaria de construção e reparação fora das partes funcionais do que integrá-la no seu interior. As fábricas produzem circuitos integrados e montam-nos em hardware funcional de forma bastante eficaz. Isso elimina uma grande quantidade de “bagagem extra” do

produto final. Além disso, devido ao seu método indireto de funcionamento, os neurónios são bastante lentos; parecem incapazes de gerar muito mais do que 100 sinais por segundo. Hoje em dia, os interruptores eletrónicos — sempre incomparavelmente mais simples e agora menores do que os neurónios — podem comutar até 100 mil milhões de vezes por segundo. A enorme vantagem de velocidade da eletrónica permitirá que usemos menos interruptores eletrónicos do que o número de neurónios presentes no sistema nervoso humano. A eletrónica é também excecionalmente precisa, permitindo que as operações sejam realizadas de forma sistemática e eficiente.

Voltando agora à retina humana: o que faz ela, na prática? Uma resposta aproximada pode ser obtida ao compararmos a função dos seus cinco diferentes tipos de células. No nível mais externo, existe uma rede de neurónios que responde ao contraste, ao movimento e a outras características mais específicas do objeto observado. Conetada a esta rede neuronal está uma camada de células fotossensíveis. Este tipo de célula subdivide-se em cones, que em conjunto discriminam as cores, e bastonetes, que não o fazem.

O facto de a luz ter de atravessar a rede neural para chegar às células fotossensíveis é uma característica peculiar da retina dos vertebrados — uma solução adotada muito cedo na história evolutiva dos vertebrados e que permaneceu fixa desde então. As retinas dos polvos e lulas, que evoluíram de forma independente, têm os fotorreceptores voltados para a frente. A posição pouco prática da rede nervosa retiniana dos vertebrados limitou significativamente o seu tamanho, mas a intensa pressão seletiva aprimorou a sua eficiência e funcionalidade. Pequenas diferenças na acuidade visual ou na velocidade de resposta terão tido, com frequência, consequências de vida ou morte para os nossos antepassados, e os neurónios da retina encontram-se numa posição única para abstraírem de forma rápida e abrangente os elementos essenciais de uma imagem. A retina é, portanto, provavelmente uma peça excecionalmente eficiente da maquinaria neural dos vertebrados.

Depois de se adaptarem a um determinado nível geral de luminosidade, os agrupamentos de células fotossensíveis geram uma voltagem proporcional à quantidade de luz que as atinge. Este sinal é recebido por duas classes de células: as células horizontais e as células bipolares. As células horizontais, cujas milhares de fibras cobrem vastos campos circulares de células fotossensíveis, produzem uma espécie de média das suas áreas. Se os valores de voltagem de todas as células horizontais fossem projetados num ecrã de televisão, surgiria uma versão desfocada da imagem retiniana. As células bipolares, por outro lado, estão ligadas apenas a pequenas áreas e proporcionariam uma imagem nítida no televisor. Algumas destas células bipolares também recebem sinais de células horizontais vizinhas, calculando depois uma diferença entre a pequena área central bipolar e a vasta envolvente horizontal. Visualizada no nosso ecrã, essa imagem pareceria muito mais esbatida do que o original, exceto nos contornos dos objetos e padrões, onde surgiria um halo claro bem distinto.

Os axónios das células bipolares ligam-se a sinapses complexas e em camadas nos dendritos das células *amácrinas* (sem axónios). Cada célula *ganglionar* recolhe sinais de várias destas sinapses amácrinas e produz uma saída pulsada, que percorre o seu longo axónio. Cada célula amácrina liga-se a várias células bipolares e ganglionares, sendo que algumas das junções parecem transmitir e receber sinais simultaneamente. Algumas células amácrinas reforçam a resposta do tipo "centro versus periferia"; outras detetam alterações no brilho em determinadas partes da imagem. No ecrã de televisão, algumas destas células revelariam apenas objetos a mover-se da esquerda para a direita, enquanto outras detetariam diferentes direções de movimento. Cada célula ganglionar coneta-se a várias células bipolares e amácrinas e emite séries de impulsos cuja cadência é proporcional a uma característica específica da imagem. Algumas informam sobre o contraste elevado em regiões específicas da imagem, outras sobre diferentes tipos de movimento ou combinações entre contraste e movimento.

A televisão a que me refiro não é totalmente imaginária. Ao meu lado, enquanto escrevo, está um monitor de TV que frequentemente exhibe imagens como as que descrevi. Estas

imagens não vêm da retina de um animal, mas sim do “olho” de um robô. A imagem captada por uma câmara de vídeo no robô é convertida por meios eletrônicos numa matriz de números armazenada na memória de um computador. Programas nesse computador combinam os números para deduzir informações sobre o ambiente em redor do robô. Apesar de terem sido concebidos com pouca referência à neurobiologia, muitos dos passos destes programas assemelham-se fortemente às operações das células da retina — um caso de evolução convergente. Este paralelo fornece uma forma de medir a capacidade computacional líquida do tecido neural.

Células e Ciclos

A retina humana contém 100 milhões de células fotossensíveis, dezenas de milhões de células horizontais, bipolares e amácrinas, e um milhão de células ganglionares, cada uma contribuindo com uma fibra condutora de sinais para o nervo óptico. Tudo isto está comprimido num volume com meio milímetro de espessura e menos de um centímetro quadrado — ou seja, 1/100.000 do volume total do cérebro. As células fotossensíveis interagem com as suas vizinhas para potenciar mutuamente os seus sinais, e a sua grande multiplicidade parece ser uma forma de maximizar a sensibilidade; mesmo um único fóton pode, por vezes, produzir uma resposta detetável. As células horizontais, bipolares e as sinapses das células amácrinas parecem cada uma realizar um cálculo único. No entanto, a essência do processo é que cada um dos axónios das células ganglionares transmite informação sobre uma função específica calculada a partir de uma região particular das células fotossensíveis.

Para encontrar o equivalente computacional de tal função, será primeiro necessário igualar, no nosso equivalente informático, o nível de detalhe visual do olho humano. Contar apenas o número de células fotossensíveis leva a uma sobrestimação, porque estas funcionam em grupos. Os testes externos de acuidade visual são mais precisos, mas complicam-se pelo facto de a retina possuir uma pequena zona central, densa e de alta resolução, chamada fóvea,

que consegue discernir detalhes mais de 10 vezes mais finos do que o restante olho. Apesar de cobrir menos de 1% do campo visual, a fóvea utiliza talvez um quarto da circuitaria retiniana e um quarto das fibras do nervo óptico. Em condições ideais de visão, é possível distinguir até 500 pontos distintos ao longo da largura desta região central. Tal proeza poderia ser igualada por uma câmara de televisão com 500 elementos de imagem separados, ou pixels, na direção horizontal. A resolução vertical da fóvea é semelhante, pelo que a nossa câmara teria de possuir 500 x 500, ou 250.000 pixels no total — o que, curiosamente, corresponde à resolução de uma imagem de boa qualidade num televisor convencional.

Mas não vemos nós com mais nitidez do que um televisor tradicional? Na verdade, não. A matriz de 500 x 500 corresponde apenas à nossa fóvea, que cobre cerca de 5° do nosso campo de visão. Um ecrã de televisão padrão subtende aproximadamente os mesmos 5° quando observado a uma distância de 10 metros. A essa distância, as linhas de varrimento e outras imperfeições de resolução da imagem televisiva tornam-se invisíveis porque a resolução do nosso olho não é melhor. Quando nos aproximamos, conseguimos concentrar a nossa fóvea em pequenas partes da imagem do televisor para captar mais detalhes, e isso cria-nos a ilusão de que vemos o ecrã inteiro com essa nitidez. Não o vemos; os nossos olhos, movendo-se inconscientemente, percorrem rapidamente o ecrã, fazendo com que a fóvea salte de um ponto para outro. Em algum lugar — ainda misterioso — do nosso cérebro, uma imagem de alta resolução é sintetizada, como um puzzle, a partir destes vislumbres fragmentários.

Assim, a circuitaria foveal da retina processa eficazmente uma imagem de 500 x 500 pixels, produzindo 250.000 valores — alguns resultantes de operações de centro-periferia, outros de deteções de movimento. Com que rapidez se processa tudo isto? A experiência com filmes de cinema fornece uma resposta imediata. Quando os fotogramas sucessivos são apresentados a uma cadência inferior a cerca de 10 por segundo, os fotogramas individuais tornam-se distinguíveis. A taxas superiores, fundem-se numa ilusão de movimento contínuo. Embora os fotogramas separados não possam ser distinguidos acima dos 10 por segundo, se a luz cintilar à mesma

frequência, esse cintilar é ainda percecionável até cerca de 50 intermitências por segundo. Presume-se que, no intervalo entre 10 e 50 ciclos por segundo, são ativados apenas os detetores de mudança de luminosidade mais simples, não havendo tempo suficiente para que as cadeias neuronais mais complexas reajam. Os projetores de cinema evitam a maior parte dessa cintilação mantendo uma taxa de fotogramas relativamente baixa, mas usando um obturador rotativo que faz cada fotograma piscar mais de uma vez. A televisão emprega uma técnica semelhante, digitalizando cada fotograma duas vezes — uma com as linhas de varrimento ímpares e outra com as pares. As regiões periféricas da retina possuem detetores de movimento mais rápidos do que a fóvea (presumivelmente para melhor captar perigos rápidos vindos dos lados), e muitas pessoas conseguem detetar a cintilação de televisores e filmes no canto dos olhos.

No nosso laboratório em Carnegie Mellon, programámos frequentemente computadores para realizarem operações de centro-periferia sobre imagens captadas por robôs equipados com câmaras de televisão, e por vezes escrevemos algoritmos de deteção de movimento. Para aumentar a velocidade, investimos muito esforço de programação e recorremos a truques matemáticos para tornar o processo o mais eficiente possível. Apesar de todos os nossos esforços, taxas de processamento de 10 fotogramas por segundo continuam fora do nosso alcance — os nossos computadores são simplesmente demasiado lentos. Com um programa eficiente, um cálculo de centro-periferia aplicado a cada pixel de uma imagem de 500 x 500 requer cerca de 25 milhões de operações computacionais, o que equivale a cerca de 100 cálculos por cada valor de centro-periferia produzido. Um operador de deteção de movimento pode ser aplicado a um custo semelhante. Traduzido para a retina, isto significa que cada célula ganglionar relata o equivalente computacional a 100 cálculos por décimo de segundo, ou seja, representa 1.000 cálculos por segundo. Assim, o conjunto de fibras do nervo óptico — composto por um milhão de axónios — funciona como um canal para os resultados de mil milhões de cálculos por segundo.

Se o processamento da retina pode ser equiparado a mil milhões de cálculos computacionais por segundo, o que podemos então dizer sobre o cérebro inteiro? O cérebro tem cerca de mil vezes mais neurónios do que a retina, mas o seu volume é 100.000 vezes superior. Os neurónios da retina, comprimidos pela pressão evolutiva, são menores e mais densamente agrupados do que a média. Multiplicando o equivalente computacional da retina por um valor intermédio de 10.000, representando a relação entre a complexidade do cérebro e a da retina, concluo — talvez precipitadamente — que o trabalho total do cérebro poderá ser realizado por um computador com capacidade para 10 biliões (10^{13}) de cálculos por segundo. Isto representa uma velocidade um milhão de vezes superior à das máquinas de média dimensão que atualmente conduzem os meus robôs, e mil vezes superior à dos melhores supercomputadores contemporâneos.

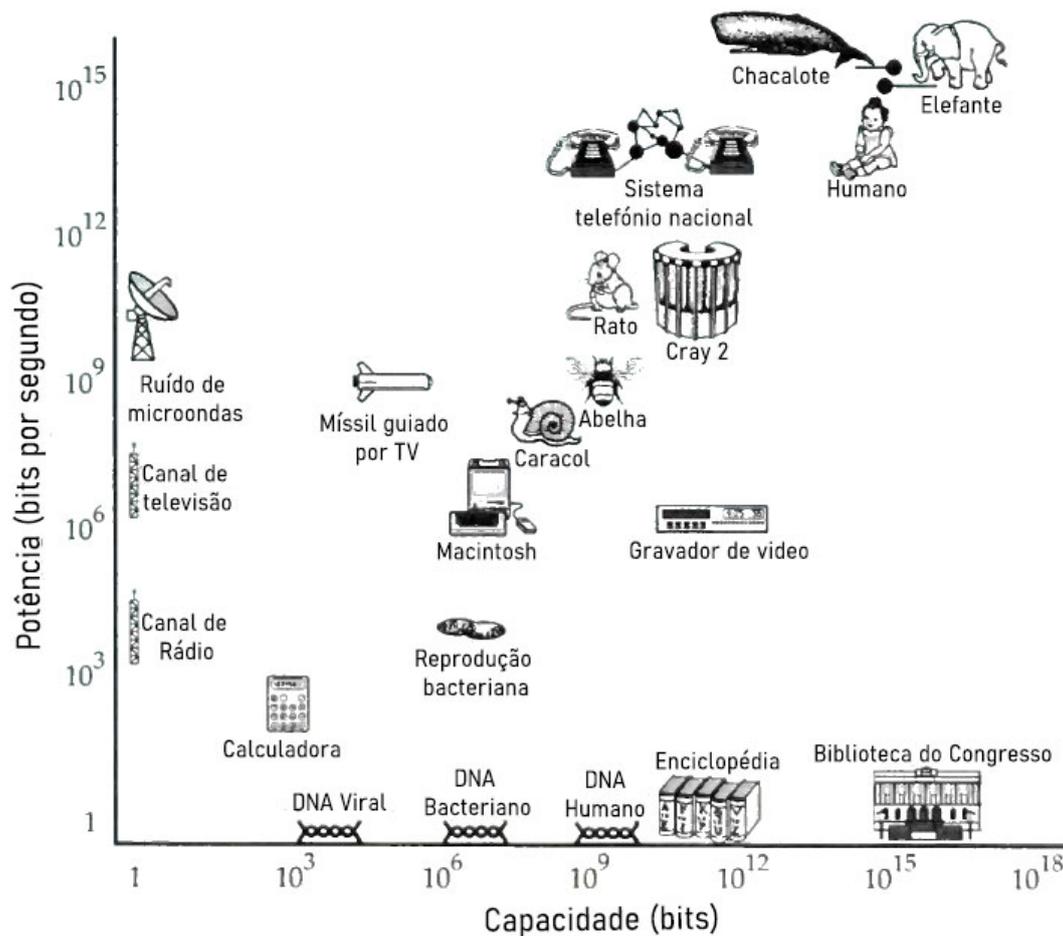
Estimativas como estas são vulneráveis a críticas de várias ordens (ver Apêndice 1). Afinal, polémicas surgem mesmo ao comparar computadores eletrónicos semelhantes, cujas operações internas são bem compreendidas e cujo desempenho pode ser rigorosamente testado. Seria, portanto, insensato esperar um consenso em torno de uma comparação entre sistemas radicalmente distintos que executam funções de compreensão ainda obscura. Ainda assim, estas estimativas podem ser úteis, mesmo que estejam apenas remotamente corretas. Veremos mais adiante que um erro de mil vezes na proporção entre neurónios e cálculos desloca a previsão da chegada de máquinas plenamente inteligentes em apenas vinte anos.

Memória

Tendo fixado os 10 biliões de operações por segundo (10 teraops) como capacidade computacional suficiente para alojar uma mente semelhante à humana, resta-nos ainda decidir qual a quantidade de memória a incluir. Em 1953, o computador IBM 650 executava 1.000 instruções por segundo e dispunha de 1.000 “palavras” de memória, cada uma capaz de armazenar um número ou uma instrução. Em 1985, o Cray 2 alcançava até mil milhões de

instruções por segundo e continha até mil milhões de palavras de memória. Esta proporção — cerca de uma palavra de memória por cada instrução por segundo de velocidade —, comum à maioria dos computadores, foi moldada pelo mercado e provavelmente reflete a dimensão necessária para conter problemas suficientemente grandes para manter um computador ocupado durante segundos a horas, ritmo confortável para programadores humanos. Se aplicássemos esta proporção, um computador de tipo humano exigiria 10 biliões de palavras de memória, ou cerca de 10^{15} bits. (Um bit, ou dígito binário, é uma unidade mínima de informação que codifica uma escolha entre duas possibilidades equivalentes. As palavras dos computadores atuais variam entre 16 e 64 bits de comprimento. Máquinas mais potentes tendem a utilizar palavras mais longas.)

Mas será este número compatível com o que se conhece sobre o sistema nervoso? Na última década, Eric Kandel, da Universidade de Columbia, e outros investigadores estudaram as alterações celulares que ocorrem na lesma-do-mar *Aplysia* quando esta é sujeita a estímulos irritantes como forma de condicionamento. Descobriram que a aprendizagem se manifesta através de alterações químicas duradouras em sinapses individuais entre neurónios — alterações que afetam a força das conexões com outros neurónios. Cada sinapse pode armazenar apenas um valor de força, e mesmo esse com precisão limitada. Se atribuirmos 10 bits — o suficiente para representar um número com três casas decimais de precisão — a cada sinapse, e se este modelo de armazenamento for substancialmente correto para sistemas nervosos mais complexos, então os 10^{15} bits da “memória padrão” de um computador de tipo humano seriam suficientes para conter a informação codificada nas 10^{14} sinapses do cérebro humano.



Poder Computacional e Memória Comparativos. Alguns organismos naturais e artificiais avaliados pelas medições deste capítulo. Os computadores de laboratório atuais têm uma potência aproximadamente igual à dos sistemas nervosos dos insetos. São estas máquinas que albergaram essencialmente toda a investigação em robótica e inteligência artificial. Os maiores supercomputadores do final da década de 1980 são páreo para o cérebro de 1 grama de um ratinho, mas, a 10 milhões de dólares ou mais cada, estão reservados para trabalhos sérios.

Comparando Computadores

É fácil perceber que os computadores se estão a tornar mais poderosos — mas quanto e com que rapidez? Quando poderemos esperar que 10 teraoperações por segundo (10 teraops) estejam disponíveis num dispositivo com tamanho e preço compatíveis com um robô autónomo? Quando comecei a abordar esta questão,

pareceu-me natural avaliar os computadores eletrônicos em termos de operações por segundo, começando pelos primeiros modelos da década de 1940, e projetar a curva resultante para o futuro. Mas surgiram complicações. As máquinas apresentavam-se em muitos tamanhos, com preços que iam de algumas dezenas de dólares a dezenas de milhões. Um mesmo modelo podia ser equipado com várias opções: mais memória, processadores auxiliares, entradas e saídas mais rápidas, etc. Máquinas mais recentes são, por vezes, multiprocessadores — múltiplos computadores a trabalhar em paralelo ou de forma independente, partilhando dados. As instruções variavam de máquina para máquina: uma operação que exigia 10 instruções numa podia ser executada numa só noutra. Algumas máquinas lidavam com números de apenas 5 casas decimais; outras processavam 20 dígitos de cada vez. Além disso, a literatura sobre as primeiras máquinas eletrônicas levou-me aos seus predecessores: computadores construídos com relés telefônicos — comutadores eletromagnéticos aperfeiçoados para centrais telefônicas. Esta investigação, por sua vez, sugeriu máquinas ainda mais antigas, que calculavam com engrenagens e cames acionados por motores, ou mesmo por manivelas manuais. Se estas máquinas manuais pudessem, de algum modo, ser comparadas com os computadores automáticos, então a minha curva de progresso poderia estender-se até ao século XIX.

Como primeiro passo para criar uma medida útil, decidi anular as diferenças de dimensão entre máquinas, dividindo a potência de processamento de cada uma pelo seu preço, em dólares constantes. Isso dar-me-ia uma estimativa da eficácia de custo. Para os calculadores mecânicos, acrescentei o operador humano (avaliado em 100.000 dólares de 1988 — como se o salário fosse um custo de aluguer) ao preço da máquina, uma vez que, para resolver um problema, um calculador manual necessita de um humano que introduza números e operações de forma contínua e registe os resultados. Esta abordagem permitiu também medir o custo do cálculo puramente manual — um funcionário humano sem auxílio, cujo custo de capital efetivo é de 100.000 dólares, consegue realizar cerca de um cálculo por minuto!

O segundo passo consistia em determinar de que forma fatores como a velocidade, a dimensão da memória e o repertório de instruções afetavam a capacidade de processamento de uma máquina. Esta tarefa revelou-se escorregadia. Hoje em dia, os computadores são frequentemente comparados com base no tempo de execução de grandes conjuntos de programas padrão. Essa via, porém, não me estava aberta, já que a maioria das máquinas que eu esperava incluir na minha curva já não existe. Sabia, no entanto, quanto tempo a maioria dessas máquinas demorava a somar ou a multiplicar dois números, quantas palavras de memória possuíam, qual o tamanho de cada palavra, e a dimensão aproximada do seu repertório de instruções. Assim, o poder de processamento seria a quantidade de computação realizada por uma máquina num determinado intervalo de tempo. Se conseguisse estimar a quantidade média de computação realizada por cada instrução, bastar-me-ia multiplicar esse valor pelo número de instruções executadas por unidade de tempo para obter o poder total. O problema, portanto, reduzia-se à estimativa do trabalho realizado por uma única instrução.

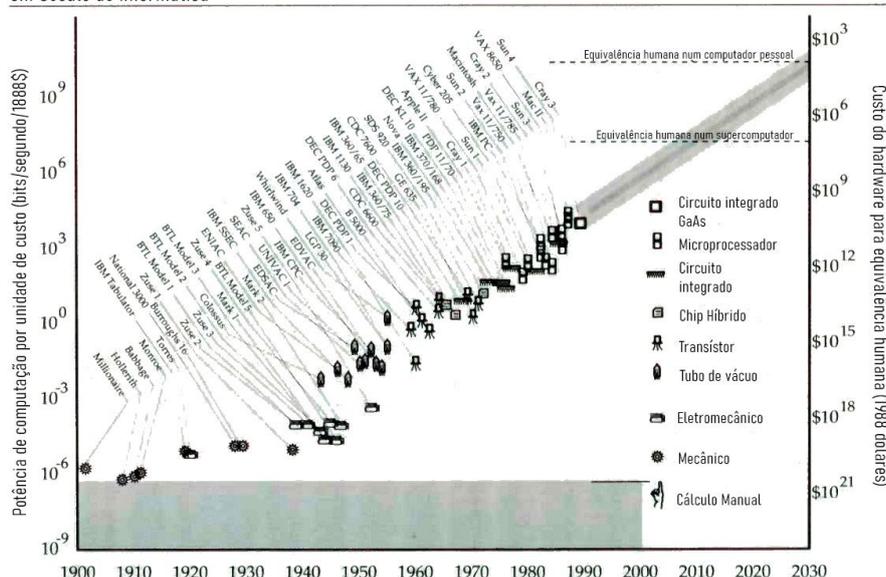
Suponhamos que uma história infantil começa com as palavras: *Este é o meu gato. Tem pêlo. Tem garras...* Aborrecido, certo? Agora imagine uma outra história que comece assim: *Este é o meu gato. Usa chapéu. Leva uma arma...* Melhor. Esta segunda história parece mais interessante e informativa porque as suas afirmações são menos prováveis — os gatos normalmente têm pêlo e garras, mas raramente usam chapéu ou levam armas. Em 1948, Claude Shannon, do MIT, formalizou estas observações num sistema matemático que veio a ser conhecido como teoria da informação. Uma das suas ideias centrais é que o conteúdo informativo de uma mensagem aumenta à medida que a sua probabilidade (tal como percecionada pelo receptor) diminui — matematicamente, é proporcional ao logaritmo negativo da probabilidade. Uma série de mensagens tem conteúdo informativo máximo quando é maximamente surpreendente.

A minha medida de computação efetiva funciona do mesmo modo. Cada instrução executada por uma máquina é como uma mensagem. Quanto mais previsível for a sua sequência de

instruções, menos trabalho útil a máquina está a realizar. Por exemplo, um programa que faz com que o computador apenas adicione um a uma localização de memória, uma vez por milhão de segundos, está a fazer quase nada de relevante. O conteúdo da memória, a qualquer momento futuro, é completamente previsível. No entanto, até os melhores programas são limitados quanto à quantidade de "surpresa" que podem introduzir num cálculo a cada passo. Cada instrução só pode especificar um número finito de operações diferentes e escolher entre um número finito de localizações de memória, cada uma contendo também um número finito de possibilidades. Estas fontes de surpresa podem ser combinadas através das fórmulas da teoria da informação para expressar o conteúdo informativo máximo de uma única instrução de computador.

Desenvolvo este tipo de cálculo no Apêndice 2. Os valores variam de máquina para máquina e de programa para programa, mas concluo ali que um computador típico, a executar um programa excepcionalmente eficiente, produz cerca de 50 bits de surpresa por cada operação realizada. Se o computador for capaz de efetuar 1 milhão de operações por segundo, o seu poder computacional máximo será de cerca de 50 milhões de bits por segundo. Expressa nestas unidades, a capacidade computacional exigida a um robô equivalente a um ser humano é da ordem dos 10^{14} bits por segundo.

Um Século de Informática



Projeções

A figura na página 17 traça o número de bits por segundo de capacidade computacional fornecida por cada dólar (constante de 1988) do preço de aquisição, em várias máquinas de computação notáveis, desde 1900 até ao presente. Embora numerosos calculadores digitais mecânicos tenham sido concebidos e construídos durante os séculos XVII e XVIII, só com os avanços mecânicos da revolução industrial é que se tornaram fiáveis e suficientemente económicos para rivalizar com o cálculo manual. No final do século XIX, a sua vantagem era evidente e o progresso contínuo, dramático. A escala vertical da figura é logarítmica — passos iguais representam aumentos de dez vezes na razão desempenho/preço. Desde 1900, houve um aumento de um bilião de vezes na quantidade de computação que um dólar pode comprar. Cada uma das máquinas representadas na figura tem uma história fascinante; mas dado que este livro não é, em primeiro lugar, uma história da computação, limitar-me-ei a alguns destaques. *The Origins of Digital Computers: Selected Papers*, editado por Brian Randell, contém excelentes relatos em primeira mão de muitas destas primeiras máquinas.

Charles Babbage, de Cambridge, Inglaterra, concebeu a ideia de uma máquina de calcular automática, controlada por programas, em 1834 — quase um século antes de qualquer outro. Esta “Máquina Analítica” seria um colosso de engrenagens e eixos, movido a vapor, capaz de operar com números decimais de 50 dígitos. Um conjunto de rodas dentadas armazenaria 1.000 desses números, e uma unidade de cálculo permitiria somar dois números em menos de dez segundos e multiplicá-los em menos de um minuto. A máquina seria controlada, a pequena escala, por tambores rotativos com pinos, semelhantes aos que ainda hoje fazem vibrar as palhetas em caixas de música mecânicas, e, a uma escala mais ampla, por um fluxo de cartões perfurados que especificariam os locais de memória e as operações aritméticas a realizar com os seus conteúdos.

Em conceito, a Máquina Analítica continha todos os elementos de um computador digital moderno. Babbage trabalhou nela durante os últimos 37 anos da sua vida, mas nunca chegou a completá-la. A enorme escala do projeto e o estado incipiente da engenharia mecânica (os componentes eram geralmente ajustados manualmente) tornavam improvável que tivesse sucesso durante a sua vida. Peças de precisão e intercambiáveis só se tornaram comuns no início do século XX, e em 1910 o filho mais novo de Babbage conseguiu demonstrar uma porção funcional da unidade central de cálculo, embora nunca tenha concluído a máquina completa. Incluí a Máquina Analítica como um ponto de dados de 1910 na minha figura, pois é provável que a máquina pudesse ter sido construída nessa altura, caso existisse uma necessidade urgente.

Os outros calculadores mecânicos representados no gráfico eram operados manualmente e vendidos sobretudo a empresas, para uso por empregados e contabilistas, embora alguns tenham tido aplicações científicas. Como referido anteriormente, incluí o “preço” do operador humano no custo do cálculo destas máquinas. As primeiras melhorias em velocidade e fiabilidade deveram-se a avanços na mecânica: engrenagens e cames de precisão produzidos em massa, por exemplo, melhores molas e lubrificantes, bem como a crescente experiência de projeto e a concorrência entre fabricantes de calculadoras. A introdução de motores elétricos nos anos 1920 acelerou o desempenho e a automatização, o que foi reforçado, nos anos 1930, pela incorporação de eletroímãs e interruptores especiais no interior das máquinas.

Leonardo Torres y Quevedo, inventor espanhol, demonstrou um calculador eletromecânico em Madrid em 1919. Controlado por comandos aritméticos simples introduzidos através de um teclado de máquina de escrever, o “Aritmómetro de Torres” não tinha memória e não era totalmente automático no sentido moderno. Ainda assim, estava muito próximo — e poderia ter sido convertido em operação automática com a adição de uma unidade de fita para introdução de comandos, e tornado prático com a adição de uma unidade para armazenar e recuperar um pequeno número de valores.

Konrad Zuse, de forma independente, inventou a ideia do cálculo programado ainda jovem, na Alemanha, em 1934; construiu vários computadores eletromecânicos grandes e automáticos na sala de estar dos pais. A terceira máquina da série, construída com apoio do governo alemão e concluída em 1941, era um computador completo, controlado por fita, binário e de ponto flutuante (isto é, representava números em notação científica, permitindo o uso de números muito pequenos ou muito grandes), com uma memória de 64 palavras. Zuse fundou uma empresa que vendeu modelos melhorados nos anos seguintes à guerra.

As máquinas dos Bell Telephone Labs (BTL) foram construídas com técnicas de comutação por relés utilizadas em centrais telefônicas. As duas primeiras foram projetos internos modestos, concebidas para testar sistemas de direção de fogo antiaéreo. A terceira era um computador automático de uso geral, controlado por fita, de grandes dimensões e destinado ao mercado comercial. No entanto, foi rapidamente ultrapassado por máquinas eletrônicas muito mais rápidas e nunca obteve sucesso. As enormes máquinas Harvard-IBM, de construção semelhante, tiveram um destino parecido. A Era dos computadores de relés de uso geral terminou quase antes de começar.

Uma classe de máquinas eletromecânicas teve, contudo, uma história mais longa. A Constituição dos Estados Unidos estipula que se deve realizar um censo nacional a cada década. À medida que o país crescia, estes censos demoravam cada vez mais a ser concluídos. Os resultados do censo de 1880 ainda estavam a ser organizados em 1887. Era evidente que, sem melhores técnicas, o censo de 1890 prolongar-se-ia para além do ano 1900. O Gabinete do Censo organizou então um concurso para encontrar um sistema mais eficiente. O vencedor foi um jovem engenheiro chamado Herman Hollerith, que concebeu uma maquinaria capaz de contar automaticamente os furos em cartões perfurados. Ao longo de meio século, a invenção de Hollerith evoluiu para um conjunto de máquinas de "tabulação" que ordenavam e intercalavam cartões perfurados, os duplicavam, imprimiam neles e a partir deles, e realizavam cálculos com os seus conteúdos. A empresa de Hollerith

viria a transformar-se na International Business Machines, ou IBM, que ainda hoje representa 70% da indústria da computação.

Computadores eletrônicos com válvulas termiônicas (tubos eletrônicos), usando técnicas de rádio e radar ultrarrápido, começaram a surgir como projetos financiados pelo governo nos últimos anos da Segunda Guerra Mundial. O primeiro computador deste tipo a ser fabricado comercialmente foi o UNIVAC I, cujo primeiro cliente foi o Gabinete do Censo, em 1951. No final da década de 1950, existiam cerca de 6.000 computadores, distribuídos pela indústria, governo e universidades. Estes computadores eletrônicos baseavam-se em válvulas termiônicas e tornaram-se conhecidos como a “primeira geração” de computadores.

Por volta de 1960, começou a surgir uma segunda geração de máquinas, que utilizava transístores recentemente desenvolvidos no lugar das válvulas. Estes computadores eram mais pequenos, mais fiáveis, mais baratos, consumiam menos eletricidade e ofereciam mais velocidade e memória do que os computadores da geração anterior.

No final da década de 1960, a IBM começou a introduzir uma terceira geração de máquinas que utilizava “circuitos integrados híbridos”. Dezenas de pequenos transístores sem encapsulamento, juntamente com outros componentes eletrônicos, eram ligados a circuitos impressos em chips cerâmicos do tamanho de uma unha. Ao longo dos anos seguintes, estes chips híbridos deram lugar aos circuitos integrados “monolíticos”, nos quais dezenas de componentes eram gravados diretamente em chips de silício com apenas alguns milímetros quadrados.

A tecnologia dos circuitos integrados desenvolveu-se rapidamente, e, em meados da década de 1970, um único chip já podia conter milhares de componentes. Surgiu então uma quarta geração de computadores, cujo núcleo era constituído por um pequeno conjunto desses chips — mas essa geração foi rapidamente eclipsada pelo microprocessador: um chip com dezenas de milhares de componentes que constituía, por si só, um

computador completo. O progresso tornara-se agora tão vertiginosamente rápido e multifacetado — com computadores a surgirem em dispositivos quotidianos como fornos de micro-ondas — que a indústria acabou por abandonar a nomenclatura por gerações. (O último vestígio dessa tradição foi o projeto japonês da Quinta Geração, um esforço de investigação para desenvolver máquinas com inteligência artificial.) Escrevo estas palavras num Macintosh II, um computador que contém múltiplos chips, cada um com mais de um milhão de componentes — uma máquina sem geração!

Equivalência Humana em 40 Anos

O progresso documentado na figura da página 17 revela-se notavelmente constante, apesar das transformações radicais ocorridas na natureza da computação ao longo deste século. A quantidade de poder computacional que um dólar pode adquirir tem aumentado mil vezes a cada duas décadas desde o início do século. Em oitenta anos, o custo do cálculo sofreu uma redução de um trilhão de vezes. Se esta taxa de melhoria continuar no próximo século, os 10 teraoperações por segundo necessários para um computador humanoide estarão disponíveis num supercomputador de 10 milhões de dólares antes de 2010 e num computador pessoal de 1.000 dólares até 2030.

Mas poderá esta corrida desenfreada manter-se durante mais quarenta anos? Facilmente! A curva apresentada na figura não mostra sinais de abrandamento, e a linha de desenvolvimento tecnológico já contém inovações laboratoriais que se aproximam das exigências mencionadas. Em grande parte, a inclinação da curva é uma profecia que se autorrealiza. Os fabricantes de circuitos integrados têm consciência desta tendência desde que Gordon Moore, um dos inventores do circuito integrado, notou em 1963 que o número de componentes num chip duplicava a cada ano. Os fabricantes de computadores chegaram a observações semelhantes, e os novos produtos, em ambos os setores, são concebidos tendo em vista esta tendência. Os fabricantes estabelecidos concebem e preparam os preços dos produtos de

modo a manterem-se na curva e maximizarem os lucros; empresas novas procuram posicionar-se acima da curva, para ganharem vantagem competitiva. O sucesso da indústria é, em si mesmo, uma das razões que permite a continuidade desse sucesso — a sua imensa e cada vez maior riqueza financia uma investigação e desenvolvimento cada vez mais robustos. Além disso, os próprios computadores produzidos pela indústria são utilizados no desenho das gerações futuras de circuitos e computadores. À medida que estes se tornam mais potentes e acessíveis, também o processo de projeto se torna melhor e mais económico — e vice-versa. A eletrónica cavalga este ciclo virtuoso com tamanha rapidez que é provável que se torne na principal ocupação da humanidade até ao final do século.

Um dos fatores-chave desta descida de preços e aumento de desempenho é a miniaturização. Componentes mais pequenos custam simultaneamente menos e funcionam mais depressa. Charles Babbage compreendeu isso já em 1834. Escreveu que a velocidade da sua Máquina Analítica — que requeria centenas de milhar de componentes mecânicos — poderia ser aumentada proporcionalmente se, “à medida que a arte mecânica atingisse estados mais perfeitos”, as suas engrenagens do tamanho da palma da mão fossem reduzidas à escala da relojoaria — ou ainda mais. (Tente imaginar o nosso mundo se a eletricidade não tivesse sido descoberta e os maiores génios tivessem continuado o percurso de Babbage. Talvez hoje existissem computadores mecânicos de secretária e de bolso contendo milhões de engrenagens microscópicas, a computar a milhares de rotações por segundo.)

De forma notável, o custo por quilo de maquinaria manteve-se praticamente constante à medida que a maquinaria se tornava mais complexa. Isto aplica-se tanto à eletrónica de consumo como aos computadores (categorias que começaram a fundir-se nos anos 80). Os rádios dos anos 1930 eram tão grandes e dispendiosos como os televisores dos anos 1950, os televisores a cores dos anos 1970 e os computadores pessoais dos anos 1980. O volume necessário para amplificar ou comutar um único sinal desceu do tamanho de um punho em 1940, para o de um polegar em 1950, uma borracha de lápis em 1960, um grão de sal em 1970 e uma

pequena bactéria em 1980. Durante o mesmo período, a velocidade básica de comutação aumentou um milhão de vezes, e o custo decresceu na mesma proporção. Não posso dizer-lhe exatamente que desenvolvimentos trarão o fator adicional de um milhão que prevejo — tais previsões são impossíveis por diversas razões. Possibilidades totalmente novas e inesperadas surgem durante a investigação fundamental. Mesmo entre os candidatos já conhecidos, muitas técnicas competem entre si, e uma linha promissora de desenvolvimento pode ser abandonada simplesmente porque outra abordagem tem uma ligeira vantagem. Posso, no entanto, afirmar que existem atualmente componentes experimentais em laboratórios que superam os melhores componentes comerciais em mil vezes, pelo menos em termos de velocidade e dimensão. Eis uma breve lista do que parece promissor hoje.

Nos últimos anos, as larguras das conexões dentro dos circuitos integrados diminuíram para menos de um micrómetro, perigosamente próximas do comprimento de onda da luz usada para imprimir os circuitos. Os fabricantes mudaram da luz visível para o ultravioleta, de comprimento de onda mais curto, mas isso apenas lhes concedeu um breve alívio. Os raios X, com comprimentos de onda muito mais curtos, serviriam por mais tempo, mas as fontes convencionais de raios X são tão fracas e difusas que requerem tempos de exposição economicamente inviáveis. Os físicos de partículas de alta energia têm uma resposta. Elétrons acelerados curvam-se em campos magnéticos e emitem fótons como lama de uma roda giratória. Este fenómeno, denominado *radiação síncrotron* devido ao tipo de acelerador de partículas onde inicialmente se revelou um incómodo, pode ser aproveitado para produzir raios X potentes e direcionados. Quanto mais fortes os ímanes, mais pequeno o síncrotron. Com ímanes supercondutores arrefecidos a temperaturas extremamente baixas, uma máquina adequada cabe num camião; caso contrário, tem o tamanho de um pequeno edifício. Seja como for, os síncrotrons despertam hoje grande interesse e prometem reduzir a dimensão dos circuitos produzidos em massa para a região submicrométrica. Feixes de elétrons e iões estão também a ser usados para escrever

circuitos submicrométricos, mas os sistemas atuais afetam apenas pequenas áreas de cada vez e têm de ser varridos lentamente sobre o chip. Contudo, a natureza desse varrimento torna os feixes de elétrons, controlados por computador, ideais para fabricar as “máscaras” que funcionam como negativos fotográficos na impressão de circuitos.

Os circuitos mais pequenos têm menor inércia eletrônica; comutam mais rapidamente e requerem tensões e potências inferiores. Pelo lado negativo, à medida que o número de elétrons num sinal diminui, o circuito torna-se mais suscetível a agitações térmicas. Este efeito pode ser contrariado através de arrefecimento e, de facto, muitos circuitos experimentais de alta velocidade, em diversos laboratórios, funcionam atualmente mergulhados em azoto líquido super-refrigerado. Um supercomputador está a ser concebido para operar desta forma. O azoto líquido é produzido em grandes quantidades durante a fabricação de oxigénio líquido a partir do ar e é barato (ao contrário do muito mais frio hélio líquido). A agregação irregular de impurezas essenciais provoca valores erráticos nos componentes à medida que os circuitos se tornam mais pequenos, pelo que estão a ser desenvolvidos métodos mais precisos para as implantar. Os efeitos quânticos tornam-se mais acentuados, criando novos problemas e também novas oportunidades. As super-redes — múltiplas camadas de regiões de silício dopadas de forma diferente, com espessuras de apenas alguns átomos, fabricadas com feixes moleculares — são uma dessas oportunidades. Permitem afinar as características eletrônicas do material e possibilitam novos métodos de comutação, frequentemente com melhorias da ordem de dez vezes. Ainda mais entusiasmantes são os dispositivos de “pontos quânticos” que exploram o comportamento ondulatório de pequenos conjuntos de elétrons presos em regiões menores que o comprimento de onda do elétron.

Os primeiros transístores eram feitos de germânio; não resistiam a temperaturas elevadas e tendiam a ser pouco fiáveis. O avanço na compreensão da física dos semicondutores e nas técnicas de crescimento de cristais de silício tornou possível o fabrico de transístores e circuitos integrados de silício mais rápidos e fiáveis.

Novos materiais começam agora a ganhar destaque. O mais imediato é o arsenieto de gálio. A sua rede cristalina oferece menos resistência ao movimento dos eletrões do que o silício, permitindo circuitos até dez vezes mais rápidos. O supercomputador Cray 3, previsto para surgir no final de 1988, utiliza circuitos integrados de arsenieto de gálio compactados num volume de um pé cúbico, superando em dez vezes a velocidade do Cray 2. Outros compostos, como o fosforeto de índio e o carboneto de silício, aguardam nos bastidores. O carbono puro, sob a forma de diamante, é uma possibilidade concreta; deverá representar uma melhoria tão significativa em relação ao arsenieto de gálio como este o foi em relação ao silício. Entre as suas muitas qualidades superiores, o diamante perfeito é o melhor condutor térmico sólido conhecido, uma propriedade crucial em circuitos densamente compactados. A visão de um circuito tridimensional ultradenso num diamante de qualidade gemológica é sedutora. Até agora não há relatos de circuitos funcionais em diamante, mas as probabilidades aumentaram em 1987 com relatos provenientes da União Soviética, do Japão e, com atraso, dos Estados Unidos, sobre camadas de diamante com até um milímetro de espessura cultivadas a partir de metano aquecido por micro-ondas.

Os circuitos mais avançados poderão vir a ser dispositivos quânticos supercondutores — não apenas extremamente rápidos, mas também altamente eficientes. Os circuitos supercondutores têm oscilado entre modas ao longo dos últimos vinte anos. Enfrentaram dificuldades significativas porque, até recentemente, necessitavam de um ambiente de hélio líquido, dispendioso e sujeito a ciclos térmicos desgastantes; e, sobretudo, porque os semicondutores, em constante e rápida melhoria, ofereciam uma concorrência difícil de ultrapassar. Os recentemente descobertos supercondutores cerâmicos de alta temperatura poderão resolver, de uma só vez, muitos destes problemas antigos. Um transistor supercondutor anunciado pelos Laboratórios Bell no início de 1988 tem menos de um vigésimo de micrómetro e consegue alternar entre os estados ligado e desligado num picossegundo (um bilionésimo de milésimo de segundo), reagindo a um sinal de entrada composto por apenas um eletrão! Mil microprocessadores

feitos com estes interruptores caberiam no espaço ocupado por um único chip de microprocessador atual, sendo cada um mil vezes mais rápido. Mil processadores, cada um mil vezes mais veloz que os atuais, teriam sensivelmente os 10 teraops necessários para igualar a capacidade humana.

Mais fora dos caminhos convencionais, encontram-se os circuitos ópticos, que utilizam lasers e efeitos ópticos não lineares para comutar luz em vez de eletricidade. Já foram demonstrados tempos de comutação de poucos picossegundos — cem vezes mais rápidos do que os circuitos convencionais —, mas muitos problemas práticos persistem. Lasers finamente calibrados têm também sido usados, em conjunto com cristais fotossensíveis e moléculas orgânicas, em memórias demonstrativas capazes de armazenar até um bilhão de bits por centímetro quadrado.

Na base destes avanços técnicos, e antecedendo-os, encontram-se progressos igualmente extraordinários nos métodos da física fundamental. Um dispositivo inesperado, e algo improvável, é o microscópio de tunelamento de varrimento (STM), relativamente barato, que consegue ver, identificar e manipular átomos individuais à superfície, varrendo-os com uma agulha extremamente fina. A ponta é posicionada por três cristais piezoelétricos que se expandem microscopicamente sob a ação de pequenas tensões elétricas. Mantém-se um intervalo de apenas alguns átomos, monitorizando-se a corrente que atravessa esse espaço por tunelamento quântico. O microscópio de tunelamento oferece um ponto de apoio seguro à escala atômica, e grandes ideias acerca de pequenos átomos estão a ser exploradas com entusiasmo tanto em laboratórios de semicondutores como de biotecnologia.

Os organismos vivos são claramente máquinas quando observados à escala molecular — neles, a informação codificada em “fitas” de RNA dirige dispositivos de montagem de proteínas, chamados ribossomas, a recolher sequências específicas de aminoácidos do ambiente e a anexá-las às extremidades de cadeias proteicas em crescimento. As proteínas, por sua vez, dobram-se de formas específicas consoante a sequência dos aminoácidos que as

compõem, desempenhando assim múltiplas funções. Algumas proteínas possuem partes móveis, como dobradiças, molas e fechos ativados por moldes. Outras têm essencialmente funções estruturais, assemelhando-se a tijolos, cordas ou fios. As proteínas do tecido muscular operam como pistões com sistema de catraca.

A indústria biotecnológica contemporânea depende de manipulações modestas da maquinaria genética natural. Os visionários, no entanto, têm planos bem mais ambiciosos — nada menos do que a fusão das técnicas biológicas, microeletrônicas e micromecânicas numa única e imensamente poderosa nova tecnologia. As técnicas de modelação computacional estão gradualmente a tornar-se suficientemente avançadas para permitir a conceção e teste de novas proteínas diretamente em ecrãs, tal como, atualmente, se desenvolvem peças de maquinaria convencional. Estas proteínas projetadas — bem como mecanismos proteicos já existentes, copiados de células vivas — poderiam ser montados em minúsculas máquinas artificiais. Os primeiros produtos poderiam consistir em medicamentos personalizados simples e pequenos circuitos computacionais experimentais. Contudo, à medida que se acumulem ferramentas e experiência, será possível construir máquinas cada vez mais elaboradas, eventualmente tão complexas como braços robóticos microscópicos, acompanhados de computadores minúsculos que os controlem. Seriam suficientemente pequenos para agarrar moléculas individuais e mantê-las, mesmo que sujeitas às agitações térmicas, fixas no lugar. Estes robôs proteicos poderiam então ser usados como ferramentas para construir uma segunda geração de dispositivos ainda mais pequenos, duros e resistentes, por meio da montagem de átomos e moléculas de todos os tipos. Por exemplo, átomos de carbono poderiam ser colocados, como tijolos, em fibras ultrarresistentes de diamante perfeito. Este conceito global tem sido designado por *nanotecnologia*, em virtude da escala nanométrica dos seus componentes. Em comparação, a microtecnologia dos circuitos integrados atuais opera com características micrométricas, mil vezes maiores. Certas coisas são mais fáceis à escala nanométrica. Os átomos são absolutamente uniformes em tamanho e forma — ainda que ligeiramente difusos — e comportam-se de

forma previsível, ao contrário das peças lascadas, empenadas ou fissuradas das máquinas maiores. À escala nanométrica, o mundo encontra-se repleto de componentes de precisão absoluta.

A maquinaria à escala atômica é um conceito prodigioso e permitir-nos-ia ir muito além do ponto de paridade com a inteligência humana nos computadores, pois permitiria que milhões de processadores coubessem num chip que hoje só contém um. Mas quão rápido poderia ser cada nanocomputador individual? A mecânica quântica impõe um limite mínimo de energia para localizar um evento num dado intervalo temporal: $Energia = h / tempo$, em que h é a constante fundamental de Planck da mecânica quântica. Velocidades mais elevadas requerem maiores quantidades de energia. Acima da frequência da luz — cerca de um quadrilhão (10^{15}) de transições por segundo —, a energia atinge um eletrão-volt, valor próximo da energia das ligações químicas que mantêm a matéria sólida coesa. Tentar comutar a velocidades superiores destruiria os próprios interruptores. Um femtossegundo (um quadrilhão de segundo) é um milhão de vezes mais rápido que o nanossegundo (milionésimo de milésimo de segundo), que é o tempo de comutação dos componentes computacionais comerciais mais rápidos da atualidade. Assim, um único nanocomputador poderia ter uma velocidade de processamento da ordem de um trilhão de operações por segundo. Com milhões de tais processadores compactados num chip do tamanho de uma unha, o critério de equivalência com a inteligência humana seria superado em mais de um milhão de vezes! Isso poderia parecer suficiente, mas não posso deixar de interrogar-me se, talvez, velocidades além desta “barreira da luz” não sejam, afinal, possíveis.

O mundo da física é um universo turbulento, onde os teóricos perseguem um objetivo que escapou a Einstein: uma teoria única que abranja todos os tipos de partículas e formas de energia da natureza. A principal candidata atualmente é o modelo das supercordas; segundo este, as partículas são minúsculos laços do espaço entrelaçados em seis dimensões adicionais para além das quatro do espaço-tempo convencional. As suas variantes predizem uma variedade de partículas mais pesadas do que aquelas que compõem os átomos — algumas das quais seriam estáveis. Um

material constituído por tais partículas seria de mil até astronómica vezes mais denso e mais fortemente ligado do que a matéria comum. Em princípio, a matéria ultradensa poderia permitir operações de comutação muito mais rápidas do que a frequência da luz. Ou seja, os benefícios da miniaturização não têm de se deter à escala atómica! Embora a existência de matéria ultradensa estável na Terra permaneça uma especulação, quantidades imensas de substâncias semelhantes são conhecidas por existir nos colossais campos gravitacionais de estrelas anãs brancas colapsadas e de estrelas de neutrões. Um dia, os nossos descendentes poderão vir a explorar esses corpos para construir máquinas com um milhão de milhões de milhões de milhões de milhões (isto é, 10^{30}) de vezes o poder de uma mente humana.

CAPÍTULO 3

SIMBIOSE

O robô que trabalhará ao nosso lado daqui a meio século terá propriedades notáveis. As suas capacidades de raciocínio deverão ser extraordinariamente superiores às de um ser humano — mesmo os sistemas mais rudimentares da atualidade já superam os nossos em certas áreas. No entanto, as suas capacidades perceptivas e motoras provavelmente serão comparáveis às humanas. O mais interessante é que esta entidade artificial será altamente mutável, tanto enquanto indivíduo como ao longo das suas gerações.

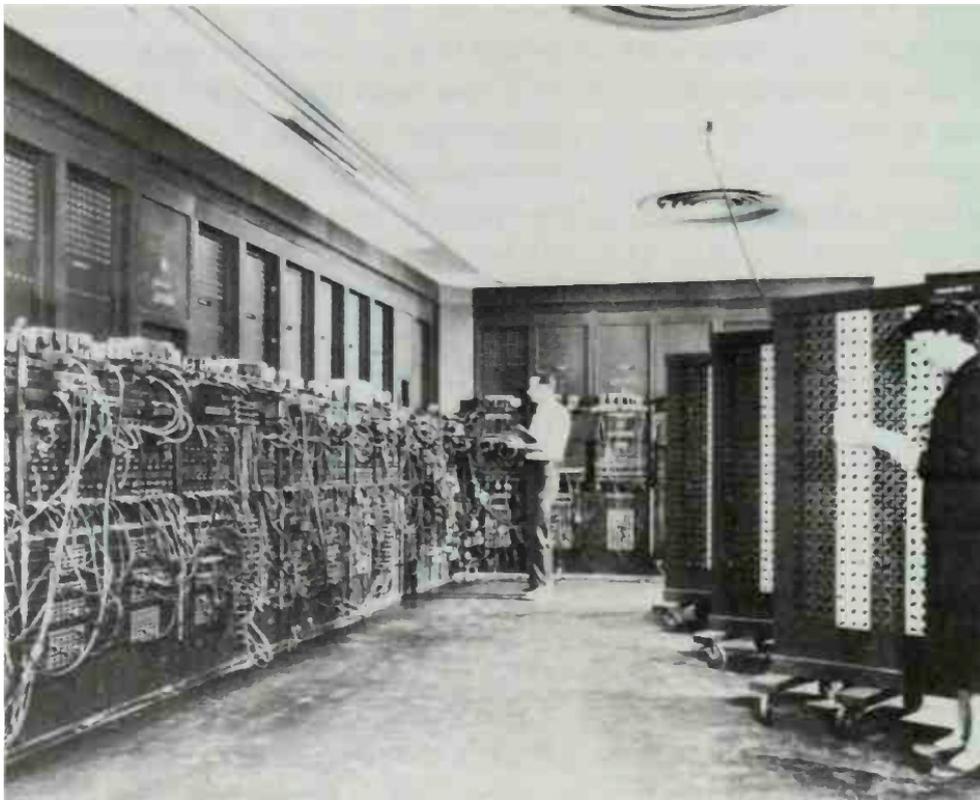
Mas os robôs solitários e laboriosos, por mais competentes que sejam, são apenas uma parte da história. Hoje, e por mais algumas décadas, as máquinas de computação mais eficazes funcionam como ferramentas nas mãos humanas. À medida que essas máquinas se tornam mais flexíveis e autônomas, essa associação entre humanos e máquinas transformar-se-á numa verdadeira parceria. Com o tempo, essa relação tornar-se-á muito mais íntima — uma simbiose, onde a fronteira entre o parceiro "natural" e o "artificial" deixará de ser evidente. Este caminho colaborativo é especialmente interessante pelas profundas consequências humanas que acarreta, mesmo que, como acredito, pouco importe a longo prazo se os humanos fazem ou não parte integrante das inteligências artificiais em evolução.

Começaremos a nossa exploração por esta via simbiótica com uma história dos seus modestos primórdios, as interfaces mínimas dos primeiros computadores.

Programas Armazenados e Montadores

O ENIAC, o primeiro computador digital eletrônico de uso geral, construído em Filadélfia em 1946, foi concebido para controlo por cablagem fixa. Um programa típico exigia milhares de fios ligados manualmente, ponto a ponto, em grandes painéis de

programação. "Escrever" um programa era uma tarefa morosa e extenuante; depurá-lo, então, era verdadeiramente assustador. Apenas alguns poucos programas foram criados até que John von Neumann apresentasse uma proposta que viria a tornar-se célebre. O ENIAC dispunha de três grandes blocos de interruptores rotativos, denominados tabelas de funções, destinados a armazenar resultados matemáticos pré-calculados necessários durante uma operação. Estas tabelas podiam, por exemplo, conter raízes quadradas, logaritmos, ou funções mais especializadas. Von Neumann sugeriu que esses interruptores poderiam ser utilizados de forma diferente: para armazenar sequências de instruções, codificadas sob a forma de números, que dirigiriam o funcionamento da máquina. A cablagem fixa habitual seria, assim, montada uma única vez, permitindo à máquina ler sucessivamente estas instruções diretamente das tabelas de funções e executar o que os números indicassem. A partir desse momento, o ENIAC poderia ser programado para novas tarefas apenas discando os comandos nas tabelas — um avanço monumental rumo à programação flexível.



ENIAC. O ninho de ratos de fio à esquerda são os depósitos de programação originais da máquina. Os três bancos de interruptores da direita foram concebidos para conter tabelas de funções matemáticas, mas foram logo utilizados como uma forma mais conveniente de representar programas.

Este novo modo de programação era muito mais simples e organizado do que o esquema original. Em vez de um emaranhado caótico de fios, um programa passou a consistir em colunas ordenadas de números. O sistema de codificação numérica das operações da máquina viria a ser conhecido como linguagem de máquina. Foi um pequeno passo rumo à computação amigável para o utilizador, mas um salto gigantesco na organização dos computadores.

Todos os computadores digitais que se seguiram ao ENIAC incorporaram uma versão expandida desta ideia de programa armazenado. Não apenas os programas eram representados como sequências de números, mas esses números eram mantidos na mesma memória usada para os cálculos, podendo ser carregados rapidamente a partir de dispositivos de entrada como leitores de fitas de papel perfurado. Esta unidade de memória permitia que o computador modificasse o seu próprio programa durante a execução — uma técnica intrigante, amplamente usada inicialmente, mas hoje menos comum.

O ENIAC ocupava uma sala inteira com tubos de vácuo para armazenar menos de cinquenta números, e era capaz de realizar cerca de mil cálculos por segundo. Os sucessores do ENIAC conseguiram armazenar programas inteiros na memória de trabalho, graças à invenção de novos métodos para armazenar, de forma mais densa e económica, os sinais eletrónicos que constituíam a memória. Em alguns casos, um dispositivo semelhante a um tubo de imagem de televisão conseguia reter milhares desses "bits" como pequenas áreas de carga eléctrica na sua superfície de vidro. Um feixe eletrónico varrido podia detetar e alterar o conteúdo de cada área.

O método mais bem-sucedido revelou-se ser aquele que utilizava pequenos anéis (ou “donuts”) de um material magnético especialmente desenvolvido, enfiados nas intersecções de uma rede de fios finos. Cada um destes núcleos magnéticos podia armazenar um bit, codificado conforme a magnetização fosse no sentido dos ponteiros do relógio ou no sentido contrário. A direção da magnetização de um núcleo específico podia ser alterada enviando pequenas correntes através dos fios horizontais e verticais que o atravessavam. O seu conteúdo anterior podia ser determinado com a ajuda de um terceiro fio, que serpenteava por todos os núcleos. Sempre que a magnetização de um núcleo se invertia — de sentido horário para anti-horário ou vice-versa — surgia um pequeno impulso elétrico nesse fio. As memórias de núcleo magnético, capazes de armazenar milhares, e eventualmente milhões, de números, foram a principal forma de memória de trabalho dos computadores durante mais de vinte anos, até serem ultrapassadas, em meados da década de 1970, por circuitos de armazenamento semelhantes aos do ENIAC, mas feitos com transístores em vez de válvulas e organizados aos milhares em minúsculos chips de silício.

A programação em linguagem de máquina foi, sem dúvida, uma grande poupança de trabalho. Mas o aumento da memória, da velocidade e da disponibilidade dos computadores rapidamente atraiu os seus utilizadores para problemas tão grandes e complexos que até a linguagem de máquina se tornou insuportavelmente trabalhosa. Um programa em linguagem de máquina consiste numa sequência de instruções codificadas em números. Alguns dígitos de cada número — o código de operação — indicam a ação que o computador deve realizar, como por exemplo somar ou buscar a próxima instrução noutra parte da memória. Os dígitos restantes contêm o endereço de uma ou mais localizações na memória, que indicam, nos nossos exemplos, onde encontrar os números a somar ou a próxima instrução a executar. Converter problemas formulados originalmente em expressões algébricas como $x^2 + y$ em códigos numéricos que o computador pudesse usar era um processo pouco intuitivo, lento e propenso a erros. Pior ainda, a inserção de apenas algumas instruções adicionais num programa podia obrigar ao

deslocamento da maior parte do programa e das zonas de memória onde se armazenavam números, o que, por sua vez, exigia alterar os campos de endereço de muitas — por vezes milhares — de instruções. O mais pequeno erro nesse processo podia impedir por completo o funcionamento do programa. Os programadores que dominavam esta tarefa árdua e precisa, pontuada por ocasionais lampejos de engenho criativo, eram por vezes tratados com o mesmo respeito reservado a mestres de xadrez. Mas não tardou muito até se perceber que os próprios computadores se especializavam precisamente neste tipo de trabalho minucioso. Noutras abordagens, milhares de bits eram codificados como um fluxo recirculante de impulsos acústicos que percorriam uma longa coluna de mercúrio, sendo detetados eletronicamente no fim da viagem, amplificados e reenviados para o início da coluna. Outra técnica consistia em gravar os bits magneticamente na superfície de um tambor ou disco em rotação rápida. Os discos magnéticos evoluíram para os dispositivos de armazenamento externo em massa que ainda hoje usamos, embora tenham deixado de ser suficientemente rápidos para funcionar como memória de trabalho interna dos computadores.

Em meados da década de 1950, os programadores já escreviam programas extensos cuja função era traduzir comandos simbólicos (como ADD X) para linguagem de máquina, atribuindo automaticamente e controlando a localização das variáveis e das instruções na memória. Os puristas da linguagem de máquina lamentavam que esses programas — que passaram a ser conhecidos como montadores (ou "assemblers") — diminuía o controlo preciso que se deveria ter sobre o funcionamento de um computador e desperdiçavam tempo de máquina ao realizar a tradução. Mesmo assim, a programação simbólica revelou-se tão vantajosa que escrever diretamente em linguagem de máquina tornou-se rapidamente uma arte extinta. O caminho que ligava os humanos aos computadores alargou-se novamente, permitindo um maior fluxo de tráfego quotidiano. As melhorias nessa estrada também facilitaram a passagem de maquinaria pesada, encarregue de construir as próximas gerações de ferramentas computacionais.

Compiladores e Sistemas Operativos

Os montadores (assemblers) foram de grande utilidade para os programadores profissionais, mas continuavam a ser bastante morosos e trabalhosos para utilizadores ocasionais de computadores que pretendiam apenas resolver problemas específicos. Esses consumidores finais de poder computacional solicitaram — e obtiveram — linguagens de alto nível, que permitiam substituir sequências longas e propensas a erros em assembler por uma notação matemática compacta, semelhante àquela que utilizavam rotineiramente no seu trabalho. Por exemplo, numa linguagem de alto nível como o FORTRAN — uma das primeiras a ser desenvolvida e que ainda subsiste hoje — a expressão $A \times X + B$ poderia substituir uma sequência de instruções em assembler equivalente, que seria muito mais extensa e técnica:

```
LOAD A
MULT X
ADD B
```

Programas extremamente complexos, chamados compiladores — verdadeiros tours de force da programação (verdadeiras proezas da programação) aquando da sua criação — traduziam linhas de linguagem de alto nível em longas sequências de linguagem de máquina. O utilizador de uma linguagem de alto nível tinha pouco ou nenhum conhecimento do programa final gerado. Os compiladores consumiam ainda mais tempo de processamento do computador do que os montadores (assemblers) e, por não possuírem a astúcia nem a perspicácia dos programadores humanos, geravam linguagem de máquina mais extensa e menos eficiente do que aquela produzida por um programa equivalente escrito em assembler. As desvantagens eram suficientemente relevantes para que, até há pouco tempo, muitas aplicações críticas continuassem a ser escritas em linguagem assembler.

Contudo, as linguagens de alto nível apresentavam numerosas vantagens para o utilizador médio. Tornavam os programas mais fáceis de escrever, e muitos erros podiam ser detetados durante a

fase de tradução, em vez de provocarem falhas obscuras durante a execução. Como as linguagens de alto nível se aproximavam da notação matemática convencional, muitos não-especialistas em informática conseguiam aprendê-las com maior facilidade. Uma característica transcendental era a *portabilidade*: ao contrário dos programas em linguagem de máquina ou assembler, os programas de alto nível não refletiam os detalhes específicos de qualquer máquina em particular, podendo por isso ser traduzidos para computadores completamente diferentes. Existem, por exemplo, compiladores de FORTRAN para praticamente todos os computadores alguma vez construídos, e alguns programas úteis escritos em FORTRAN sobreviveram a toda a história da informática, saltando de geração em geração de máquinas.

Nos primeiros computadores digitais, preparar um programa, monitorizar o seu progresso e fazer a sua limpeza final era um processo estritamente manual. O programador podia acompanhar a execução através de painéis de luzes que mostravam o estado interno da máquina, podendo interrompê-la, examiná-la e modificá-la a qualquer momento, ou executá-la passo a passo. Este procedimento era bastante conveniente, mas caro, em máquinas cujo tempo era avaliado em centenas de dólares por hora. Para minimizar o tempo desperdiçado, as empresas de informática começaram, no final da década de 1950, a fornecer programas chamados *monitores*, *supervisores* ou *sistemas operativos*, que geriam o fluxo sucessivo de programas, lidos a partir de perfuradoras de cartões ou unidades de fita magnética, através das suas máquinas. Os primeiros sistemas operativos tinham uma abordagem pragmática e inflexível: executava-se um programa de cada vez; se surgisse algum problema, o programa era interrompido automaticamente, os conteúdos da memória eram impressos, e o programa seguinte era iniciado. A *impressão da memória* (chamada *core dump*, em referência à memória de núcleos magnéticos) era entregue ao programador, que ficava responsável por diagnosticar o erro no seu próprio tempo. Muitos programas ocupavam apenas alguns segundos da máquina, mas os seus resultados implicavam horas ou dias de preparação por parte do programador para a execução seguinte.

Entre meados dos anos 1950 e o final dos anos 1960, este modo de operação em lotes (*batch mode*) era a norma, e uma geração inteira de utilizadores de computadores — especialmente utilizadores de equipamentos IBM — não conheceu outro modo de funcionamento. Alguns dos mais antigos, porém, recordavam com nostalgia os tempos de uso direto da máquina. Não só era mais fácil detetar erros de programação com o computador como aliado, como era também possível escrever programas que mantinham um diálogo com o utilizador. Estes programas interativos permitiam que o utilizador e o computador agissem como parceiros — o utilizador fornecendo intuição e discernimento, e o computador oferecendo capacidade de cálculo e memória prodigiosa. A dificuldade residia em como garantir este nível de serviço sem que o computador desperdiçasse a maior parte do seu tempo à espera da próxima ação do utilizador.

Grupos académicos começaram a trabalhar numa solução: uma forma de sistema operativo extremamente complexa, concebida para manter vários programas a funcionar em simultâneo, iniciados a partir de terminais interativos individuais. Cada programa de utilizador recebia tempo de execução durante uma fração de segundo, após a qual o sistema operativo de partilha de tempo transferia o controlo para o programa seguinte, e assim sucessivamente, acabando por regressar ao primeiro. Um utilizador ligado a qualquer um dos programas ativos não se aperceberia dessas interrupções fracionárias, parecendo-lhe que tinha um computador só para si, ainda que ligeiramente mais lento do que a máquina "nua". Se um programa se tornasse temporariamente inativo por estar à espera de uma resposta do utilizador, seria simplesmente ignorado pelo sistema operativo, sem desperdício de tempo significativo.

Tal como antes acontecera com os assemblers e as linguagens de alto nível, os sistemas de partilha de tempo geraram grande controvérsia na comunidade informática. Por terem de manter os recursos disponíveis para vários programas em simultâneo e de tomar decisões várias vezes por segundo sobre o que executar a seguir, consumiam uma porção muito maior dos recursos do computador do que os sistemas operativos de processamento em

lotes (*batch*). Contudo, a partilha de tempo oferecia, em troca, ganhos de eficiência. Com muitos utilizadores distintos ativos ao mesmo tempo, os diversos recursos de um sistema informático — memória, discos, unidades de fita magnética, impressoras, ecrãs, etc. — podiam ser mantidos mais ocupados do que num sistema de programa único. Mais importante ainda, os programadores podiam acompanhar a execução dos seus programas, interrompendo-os rapidamente quando, como muitas vezes sucede com novos programas, algo corria mal. No entanto, os maiores ganhos de eficiência foram sentidos pelos utilizadores e não pela máquina. Em vez de esperarem horas por uma execução de teste num sistema por lotes, os utilizadores de sistemas de partilha de tempo podiam observar o progresso do seu programa, interrompê-lo, modificá-lo e testá-lo de novo em poucos minutos. Esta rapidez tornou possível um estilo de programação altamente experimental, e algo pavloviano, caracterizado por ciclos rápidos de punição e recompensa. Nasceu assim uma geração de hábeis piratas informáticos (*hackers*).

Os *hackers*, alguns dos quais passavam grande parte da sua vida consciente diante de terminais de computador, rapidamente expandiram as capacidades básicas do seu habitat favorito. Usos inteiramente novos dos computadores tornaram-se a norma. A partir de meados da década de 1960, os utilizadores podiam comunicar entre si em tempo real, de terminal para terminal, ou por correio eletrónico, a ser lido e respondido ao ritmo de cada um. Podiam envolver-se em jogos interativos de texto e vídeo, pregarem partidas sofisticadas uns aos outros, partilhar ideias e programas e, em geral, experimentar um sentido de comunidade, mediado de forma inconsciente pelas máquinas dos seus empregadores. Essa comunidade foi enriquecida pela criação de inúmeros artefactos públicos sob a forma de ficheiros informáticos — murais comunitários, citações espirituosas, dicas técnicas, escritos originais, programas divertidos e sistemas interativos de recuperação de informação que ajudavam a navegar por toda essa riqueza. Alguns dos sistemas de investigação mais avançados ofereciam também imagens e sons produzidos pela própria máquina.

O uso intensivo e variado que os hackers faziam dos computadores exigia formas eficientes de localizar, iniciar e interromper diversos programas, examinar, ler e modificar ficheiros de informação, interagir com outros utilizadores e solicitar à máquina que realizasse automaticamente inúmeras tarefas. Num processo evolutivo, as linguagens de comandos através das quais os utilizadores comunicavam com os sistemas operativos de tempo partilhado foram dotadas dessas capacidades. Concebidas por especialistas para especialistas, com camadas sucessivas de extensões não planeadas, controladas por fórmulas lacónicas e poderosas, mas muitas vezes inconsistentes, não documentadas e difíceis de memorizar, estas linguagens exasperavam os utilizadores menos experientes. Para um hacker, a interface com o computador era intuitiva, rápida e imensamente poderosa — quase tudo era possível para quem soubesse construir os feitiços certos. Mas para os que não partilhavam da monomania necessária para acompanhar a evolução rápida e amorfa do sistema, a interface mostrava-se opaca, pouco amigável e extremamente propensa a erros. Os hackers acabaram por ter o seu castigo quando sistemas diferentes foram interligados através de redes de computadores — até um hacker é novato perante os arcanos de outro grupo.

Um dos grandes sucessos da Era hacker foi o sistema de tempo partilhado Unix. Criado no início da década de 1970 por dois jovens hackers nos Laboratórios Bell e expandido por outros na Universidade da Califórnia, em Berkeley, o Unix tornou-se, nos anos 1980, o padrão de facto para computadores de maior porte. Atualmente, o Unix está a conquistar também os computadores pessoais de gama alta.

Menus e Ícones

De tempos a tempos, os fabricantes de computadores tentaram incorporar outras inovações desenvolvidas por hackers nos sistemas destinados aos seus utilizadores mais convencionais. Um dos objetivos principais consistia, frequentemente, em tornar a linguagem que invocava as várias funções do sistema operativo o mais semelhante possível ao inglês corrente, na expectativa de que

isso aliviaria consideravelmente o esforço dos utilizadores já fluentes na língua. Um efeito semelhante havia sido observado na comunidade científica aquando da introdução de linguagens de alto nível que utilizavam, em grande medida, notação matemática padronizada. Os dissidentes, no entanto, assinalaram que, ao contrário da matemática, o inglês natural é uma linguagem pouco adequada para descrições precisas; comunidades que requerem precisão geram invariavelmente as suas próprias convenções especializadas, sendo a notação matemática o exemplo mais óbvio.

Desta vez, os céticos tinham razão. Interfaces experimentais com aparência de inglês (com acrónimos hoje esquecidos, que costumavam incluir palavras como "Plain", "Simple" e "English") foram alvo de investigação entre o final da década de 1960 e o início da de 1970, mas não tiveram grande êxito. Não era (e continua a não ser) possível incorporar num programa uma verdadeira compreensão geral da linguagem, com o seu requisito subjacente de senso comum e conhecimento vasto do mundo. Os sistemas reais não conseguiam interpretar corretamente (ou interpretavam erradamente) muitas das frases introduzidas, tornando a sua utilização uma espécie de jogo de adivinhação. Era necessária alguma destreza para formular pedidos de modo a que fossem corretamente interpretados pelos analisadores linguísticos dos sistemas — complexos, mas incompletos e em grande parte não documentados. Aprender um código de comandos simples, coerente e específico era, por comparação, muito mais fácil.

Os utilizadores ocasionais de computadores (e também os especialistas) beneficiaram mais com a invenção de um dispositivo muito mais simples — o menu de escolha múltipla. Nos sistemas guiados por menus, as principais opções são apresentadas numa lista, da qual o utilizador escolhe um item. Essa escolha pode conduzir a um segundo menu com características adicionais, e assim sucessivamente, até que a ação pretendida fique totalmente especificada. Os sistemas de menus, contudo, não estão isentos de limitações. A impressão de menus longos consome tempo ao computador, e a sua leitura exige tempo ao utilizador; além disso, as opções nem sempre estão formuladas da forma mais intuitiva. Por conseguinte, os sistemas de menus tendem a ser mais lentos

do que as linguagens especializadas utilizadas por operadores experientes. Captar ideias realmente complexas através de um jogo de vinte perguntas é tarefa difícil. Sistemas híbridos, que ocasionalmente solicitam respostas do tipo ensaio, conseguem por vezes contornar esse problema com alguma eficácia.

O problema da lentidão na impressão de menus foi relativamente fácil de resolver, ainda que com algum custo: bastava instalar terminais com capacidade de visualização mais rápida. Já o problema enfrentado pelos novos utilizadores — a dificuldade em assimilar rapidamente longas listas de opções desconhecidas — não admitia solução tão direta. Uma resposta notável foi concebida e desenvolvida por hackers que trabalhavam nas instalações excepcionalmente luxuosas do Xerox Palo Alto Research Center (PARC). Durante o início da década de 1970, este grupo desenvolveu estações de trabalho dispendiosas, cada uma com o seu próprio computador pessoal e um ecrã de grandes dimensões, capaz de apresentar imagens gráficas com elevada rapidez e qualidade. Com bons gráficos e muita capacidade computacional disponível, tornaram-se possíveis interações bastante sofisticadas entre utilizador e máquina. Por exemplo, cada computador vinha equipado com um dispositivo portátil, chamado mouse (rato), que podia ser deslizado sobre uma superfície como um disco de hóquei, detetando o movimento através de uma pequena esfera na sua base. O movimento do rato era associado ao movimento de uma seta gráfica no ecrã do computador, permitindo ao utilizador apontar para qualquer parte do ecrã. Apontar tornou a utilização de menus mais intuitiva: em vez de escrever o nome da opção desejada, bastava apontar-lhe para a selecionar.

Pequenas imagens sugestivas associadas a cada linha de menu facilitavam imenso a tarefa visual de identificar o item correto, tornando o sistema em parte independente da linguagem. Em versões posteriores desta ideia, as imagens — então designadas por ícones — tornaram-se a representação dominante, ficando as palavras relegadas para pequenas anotações explicativas. Com o tempo, os ícones puderam ser movidos no ecrã como objetos e colocados noutros ícones que funcionavam como contentores ou marcadores de destinos físicos, como impressoras. As interfaces

baseadas em ícones demonstraram ser eficazes e de uso fácil tanto para principiantes como para utilizadores experientes, provavelmente porque recorriam às capacidades humanas de manipulação não-verbal de objetos.

A disponibilização de um computador pessoal a cada utilizador teve inúmeras implicações, algumas das quais desconcertantes. Não seria isto um retrocesso em relação à partilha de tempo? Vários pioneiros deste último certamente assim o consideraram. Mas os entusiastas do PARC salientaram que, do ponto de vista da conveniência do utilizador, a partilha de tempo apresentava falhas graves. Imobilizava o utilizador habitual, obrigando-o a permanecer junto de um terminal fixo ligado fisicamente a uma máquina central de grandes dimensões. Pior ainda: a capacidade de resposta dos sistemas de partilha de tempo nunca correspondeu às promessas iniciais; o número de utilizadores e a sobrecarga inerente ao sistema acabavam sempre por crescer até ao ponto em que o indivíduo se via obrigado a esperar segundos — ou mesmo minutos — por resposta a pedidos triviais. Um pequeno computador pessoal para tratar das funções rotineiras parecia plausível, dada a rápida queda nos custos dos computadores.

Alan Kay, o mentor do grupo do PARC, idealizou originalmente um computador pessoal de alta responsividade, com o tamanho de um livro (a que deu o nome de Dynabook, pelas suas qualidades dinâmicas), dotado de um ecrã a cores de alta resolução e de uma ligação por rádio a uma rede mundial de computadores. Mais do que um simples computador, o Dynabook funcionaria como secretária, caixa de correio, biblioteca de referência, centro de lazer e telefone. Esta ideia estava — e continua a estar — além das capacidades tecnológicas atuais (ainda que eu esteja a digitar estas mesmas palavras num computador portátil do tamanho de um livro, sentado no átrio de um hotel).

Os dispendiosos computadores pessoais do PARC tinham o tamanho de uma secretária, e não de um livro, apresentavam gráficos apenas a preto e branco e ofereciam menos capacidade de processamento do que os seus criadores desejavam. Ainda assim, representaram um passo claro em direção à visão do Dynabook.

Inicialmente foram batizados como Interim Dynabooks (Dynabooks Interinos), mas mais tarde, e felizmente, receberam o nome mais melodioso de Alto.

A Xerox foi lenta a comercializar as descobertas do grupo do PARC, embora no final da década de 1970 tenha produzido uma estação de trabalho empresarial dispendiosa — chamada Star — que incorporava muitas das ideias desenvolvidas no PARC. Após dez anos de investigação, alguns dos entusiastas do centro começaram a ficar frustrados com a lentidão burocrática da Xerox e encontraram um aliado recetivo em Steve Jobs, cofundador da vizinha empresa Apple. O resultado foi, alguns anos mais tarde, o Apple Lisa e, posteriormente, o seu irmão mais novo e mais inteligente, o Macintosh, promovido como o computador pessoal “para todos nós”. Apesar de ainda estar longe, em termos de capacidades, do ideal do Dynabook, o Macintosh introduziu uma nova forma de computação a milhões de pessoas — e abriu caminho para uma nova fase na interação entre humanos e computadores. No final da década de 1980, quase todos os novos sistemas operativos para computadores com orientação gráfica estavam a ser concebidos para apresentar ao utilizador uma interface ao estilo do Macintosh.

Óculos Mágicos

A interface gráfica que torna o Macintosh e os seus imitadores tão mais agradáveis de utilizar do que as máquinas anteriores demonstra bem o valor de envolver as capacidades sensoriais humanas no diálogo entre pessoas e computadores. O Dynabook de Alan Kay, por muito admirável que fosse em vários aspetos, não poderia avançar muito mais do que os sistemas existentes nessa dimensão não-verbal, devido às suas limitações físicas — em particular, o ecrã do tamanho de um livro. Tal como num livro convencional, a portabilidade era uma característica essencial do Dynabook; muitas das utilizações que se lhe propunham deixariam de fazer sentido se ele existisse apenas em locais fixos, como a casa ou o local de trabalho. Será possível manter essa portabilidade e, ao mesmo tempo, ampliar significativamente o envolvimento

sensorial do utilizador? Por outras palavras, poderemos imaginar um computador que tire partido da mobilidade humana, enquanto permite aos humanos beneficiar da superior memória, poder de cálculo e vasto alcance comunicacional do computador?

É claro que sim — não sob a forma de um livro, porém, mas sob a forma de um guarda-roupa de alta tecnologia. O elemento-chave deste vestuário é um par de óculos mágicos (ou, nas fases mais primitivas, uns óculos de proteção ou um capacete). Usados no nariz como óculos convencionais, estes conteriam o seguinte e impressionante conjunto de dispositivos:

- Ecrãs a cores de alta resolução, um para cada olho, com óptica que abrange todo o campo visual e faz com que a imagem apresentada pelo computador pareça estar focada a uma distância confortável. Os óculos poderão dispor da capacidade de alternar para um modo transparente.
- Três câmaras de vídeo. Um par de câmaras de alta resolução com lentes grande-angulares voltadas para a frente seria colocado o mais próximo possível da posição dos olhos, permitindo ver para onde se vai quando as imagens captadas são projetadas nos ecrãs correspondentes nas lentes. Talvez uma terceira câmara, de ângulo ainda mais amplo, esteja virada para trás para captar a maior parte do rosto do utilizador, permitindo ao computador — e a eventuais interlocutores numa videoconferência — monitorizar as suas expressões faciais.
- Microfones e pequenos auriculares embutidos na armação.
- Um sistema de navegação que rastreia com precisão e em permanência a posição e orientação dos óculos (e, conseqüentemente, da cabeça do utilizador).
- Um computador poderoso, capaz de gerar imagens sintéticas realistas, som e fala; de compreender comandos orais; e de identificar e seguir objetos no campo de visão das câmaras.

- Uma ligação de dados de alta velocidade a uma rede mundial de computadores e bibliotecas eletrônicas, bem como a outros utilizadores com óculos mágicos.

Não é preciso ser especialista para reconhecer que se trata de um conjunto de requisitos exigente. Contudo, cada uma destas funções já existe e é objeto de investigação intensiva e bem financiada. Comprimir tudo isto num pacote minúsculo é também um desafio — mas não é inverosímil, numa Era de Walkmans do tamanho de uma cassette e televisores de bolso. Estes óculos permitiriam ao computador controlar o que o utilizador vê e ouve, em resposta aos seus pedidos verbais, e seguir os seus movimentos; também lhe permitiriam observar o movimento das mãos e as expressões faciais do utilizador — embora não possam controlar o que ele sente com as mãos. Ora, como o tato é uma importante fonte de informação sobre o mundo, seria desejável que o computador nos pudesse comunicar como são as coisas ao toque, mesmo quando não estão fisicamente ao nosso alcance. Entra em cena o par de *luvas mágicas*.



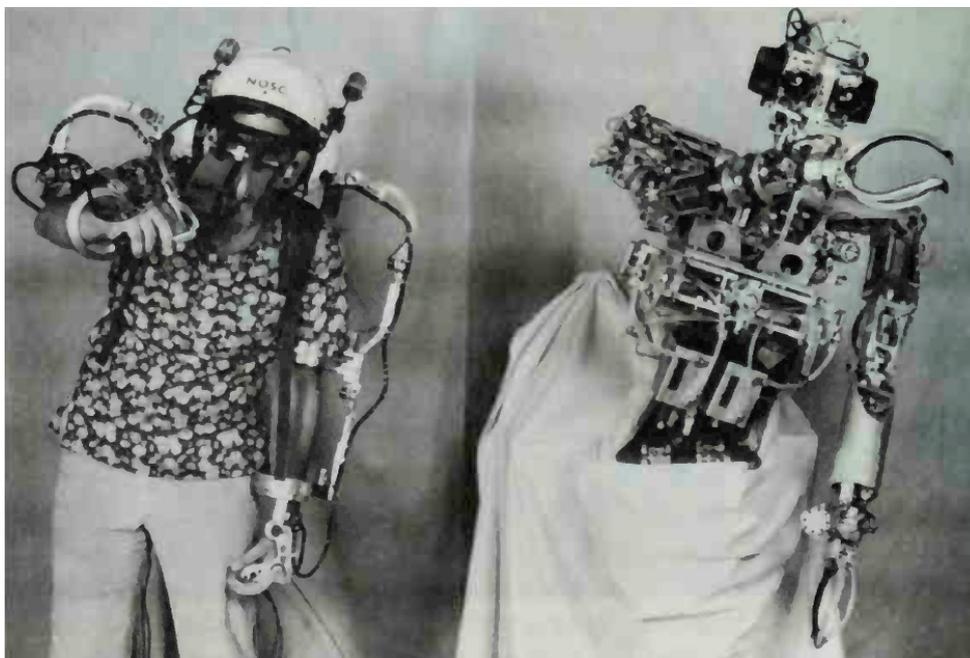
Óculos Mágicos (Modelo Inicial). A instrumentação em aeronaves militares — onde o acesso instantâneo a dados de navegação, sensores e armas é uma questão de vida ou morte — está a evoluir para óculos mágicos. Este modelo foi desenvolvido em 1986 para um projeto de helicóptero experimental da Boeing-Sikorsky. Os dados dos instrumentos da aeronave são projetados no campo de visão do piloto. Os blips de radar, por exemplo, são concebidos para aparecer nas localizações reais dos objetos rastreados.

Tal como os óculos, as luvas são uma verdadeira proeza de engenho tecnológico. Cada dedo da luva contém uma grelha de elementos capazes de criar padrões de pressão e temperatura sobre os dedos do utilizador. Experiências fisiológicas demonstraram que estas disposições podem gerar sensações realistas; por exemplo, uma pressão uniforme acompanhada de frio é interpretada como: “Os meus dedos estão submersos em água.” As luvas possuem motores que atuam sobre as articulações dos dedos, permitindo ao utilizador sentir resistência ao movimento. Esses mesmos mecanismos permitem também ao computador monitorizar a posição dos dedos.

As luvas mágicas, tal como os óculos mágicos, têm as suas limitações. Conseguem gerar resistência apenas aos movimentos dos dedos, mas a manipulação envolve também o movimento de todo o braço. Imagine-se, então, um casaco motorizado capaz de transmitir às articulações dos braços a mesma sensação de presença que as luvas conferem às mãos. Comparado com as luvas, e sobretudo com os óculos, o casaco é de uma simplicidade quase elementar. No entanto, nas primeiras versões deste guarda-roupa mágico, o casaco poderá tornar-se um recetáculo prático para todo o equipamento que não caiba nos óculos e luvas mais pequenos e complexos. Poderá até ter de estar ligado a um assento fixo: uma verdadeira “cadeira de pensar”. Mais tarde, o casaco será tão portátil como um fato espacial e, eventualmente, assemelhar-se-á a um fato de lazer.

Tal como acontece com os computadores atuais, este maravilhoso equipamento, acabado de descrever, será apenas tão

útil quanto o software que o acompanha. As secções seguintes apresentam programas que poderá querer executar ao envergar o seu novo e inteligente vestuário.



Proxy Robótico. Este proxy robótico foi desenvolvido em 1986 no Centro de Sistemas Oceânicos Navais, no Havai. Os movimentos do operador à esquerda são copiados pelo robô à direita, e as imagens dos olhos da câmara do robô são enviadas para os óculos mágicos (volumosos) do operador. O operador tem a sensação subjetiva de estar no corpo do robô.

Encontrando-se a Si Mesmo

Não sei quanto a si, mas eu perco-me com frequência. Falta-me um bom sentido de orientação inato, e por isso anseio por um auxiliar de navegação de bolso que não apenas me diga onde estou (os mapas eletrónicos que começaram a surgir nos automóveis já fazem isso), mas que também me guie até ao destino, recorde os locais por onde passei e me lembre da lista de compras quando deteta que estou perto do supermercado. Os óculos contêm um sistema de navegação, provavelmente uma combinação de um dispositivo que mede a distância até ao solo e a emissores orbitais por rádio, e outro que regista forças de aceleração para deduzir o

movimento. Seja qual for o método, o navegador sabe onde se encontra, você, utilizador dos óculos, a qualquer momento.

A Robin e a família mudaram de casa de férias desde a última vez que os visitou e agora vivem algures no meio do mato. Este território é-lhe desconhecido e, pela primeira vez, aprecia verdadeiramente a linha verde-luminescente gerada pelo programa "Caminho de Tijolos Amarelos" que lhe guia pela estrada. A linha desvia-se para a faixa da direita, apontando para uma saída que se aproxima a cerca de oitocentos metros. Caso não tenha reparado, uma seta intermitente de virar à direita paira sobre o cruzamento, e o programa sussurra-lhe ao ouvido: "Curva a aproximar-se, abrande para 30. Um quilómetro após a curva, o asfalto dá lugar à gravilha, anunciada por uma mudança de cor na linha guia de verde para um amarelo mais cauteloso (e metaforicamente adequado). Mais à frente, a linha torna-se vermelha e a estrada transforma-se em terra batida. O programa informa: "A mudar para mapa privado", indicando que a base de dados da Rand McNally ainda não inclui este trilho; os dados provêm dos ficheiros pessoais da Robin. "Faltam quatro milhas até à casa da Robin", informa o programa, interrompendo brevemente a antiga música dos Beatles que vinha a tocar.

Está a escurecer. "Iniciar chamada — Robin", diz ao computador que traz no nariz. Isto ativa o programa de telefone, que se liga através da rede celular/satélite. O programa Caminho de Tijolos Amarelos prossegue em paralelo. "Estou quase a chegar", diz à Robin. "Ótimo. O jantar está quase pronto. Preferes arroz ou batatas?" "Pode ser os dois?" "Ok, até já!" "Tchau." "Tchau." O programa de telefone termina. A linha vermelha serpenteia por entre as árvores. Os faróis estão acesos, mas mal iluminam o caminho. De repente o carro solavanca-se para a esquerda. Bateu em algo — talvez uma pedra — e agora a roda da frente está presa num buraco lamacento. Tentar sair de marcha avante só lhe enterra mais. "Mostrar mapa", pede ao programa de navegação. O mapa solicitado surge suspenso no ar diante de si. Uma pequena imagem do seu carro indica que está a

cerca de duas milhas do destino, seguindo a estrada, embora a direito sejam menos de oitocentos metros por entre o bosque.

“Olá Robin, sou eu outra vez. Estou em apuros.” Depois de avaliar a situação, a Robin sugere uma solução aventureira para salvar o jantar: “Trataremos do carro amanhã. Na semana passada, o Marty e eu descobrimos um atalho que passa perto de onde estás. Deves conseguir chegar cá a pé em menos de meia hora.” Corajoso como é, aceita. O registo do trilho é transferido do computador da Robin para o seu, através da ligação telefónica. O seu programa de navegação integra o percurso do atalho com os dados da estrada, e a linha guia original é substituída por outra que acompanha a estrada por um troço e depois se embrenha no mato. Tranca o carro e segue a nova linha.

Está muito escuro, por isso ativa o programa de visão noturna — as câmaras dos seus óculos passam a funcionar com sensibilidade máxima, fornecendo imagens melhoradas aos ecrãs, sobrepostas à saída do navegador. Vê que a linha vermelha acompanha um trilho ténue pela floresta. À parte uns arranhões, a caminhada decorre sem incidentes. A refeição está quente e deliciosa.

Indo a Lugares

O transporte e a comunicação melhoraram de forma impressionante nos últimos 500 anos. No entanto, embora a geografia já não seja o principal determinante do comércio, continua a restringir como — e com quem — a maioria de nós conduz os seus assuntos. As diferenças entre transporte e comunicação tornar-se-ão cada vez menos nítidas à medida que formos capazes de projetar a nossa consciência e competências para locais remotos. A facilidade dessa projeção tornará o interesse comum no principal impulso para a associação entre pessoas. O guarda-roupa mágico pode ser utilizado tanto para visitar locais reais no mundo físico como para aceder a locais "irreais", profundamente imersos em simulações computacionais.

A forma mais evidente de presença remota envolve um substituto físico — um robô distante que se controla através da rede de comunicações global. Os óculos mágicos permitem-lhe ver através dos olhos do robô, o casaco e as luvas permitem-lhe sentir, gesticular e agir por meio dos manipuladores do robô, e os controles para os pés, integrados na sua poltrona, permitem-lhe conduzi-lo. Ao alugar robôs substitutos em locais remotos, pode visitar, comunicar e trabalhar em projetos espalhados pelo mundo sem sair do conforto de casa — e sem incorrer nos riscos físicos de locais perigosos ou no tédio das longas deslocações.

Uma variação desta ideia de substituto é o *substituto humano*. Alguém que viaje para um local interessante e use um tipo especial de guarda-roupa mágico poderia transmitir a sua visão, som, toque e, possivelmente, impressões olfativas ao viajante de poltrona. Na sua forma mais simples, esta ligação é unidirecional — do substituto para o observador passivo. Esta comunicação unidirecional pode ser gravada e reproduzida a qualquer momento — oferecendo-nos uma forma plausível da extensão sensorial do cinema. Alternativamente, a ligação pode ser bidirecional, com ações motoras e impressões sensoriais transmitidas nos dois sentidos, devidamente editadas por um programa inteligente. Participantes inexperientes poderão sentir-se tão descoordenados como num jogo de corrida a três pernas, mas com treino pode atingir-se um tipo de colaboração íntima, onde a perícia remota é aplicada diretamente no local do problema. Na maioria das situações, o “agente de campo” controlaria os movimentos, enquanto o participante à distância observaria, ouviria, sentiria e ofereceria conselhos. Mas, quando a tarefa exigisse uma destreza manual melhor dominada pelo que está em casa, o agente de campo relaxaria e permitiria que os motores do fato, controlados remotamente, executassem o trabalho — como se estivesse possuído por um espírito.

Um encontro com substitutos não precisa de ocorrer no mundo real — muitas coisas podem ser feitas de forma mais eficaz numa simulação informática. O “imobiliário irreal” gerado por computador não tem limites intrínsecos, nem em extensão, nem em propriedades físicas. É um mundo onde a magia é rotina. Os ecrãs de computador atuais permitem vislumbres desse mundo —

projetistas de automóveis examinam modelos futuros, físicos observam o interior de explosões nucleares, e utilizadores de Macintosh reorganizam os seus ficheiros num espaço virtual chamado "secretária". Com um guarda-roupa mágico, poderemos penetrar ousadamente nesses mundos e explorá-los a partir do interior.

A sua (modesta) casa de sonho está finalmente a tornar-se realidade. Os preparativos com o arquiteto foram concluídos há semanas e chega, finalmente, a chamada tão aguardada:

– *"Olá! Tenho aqui um esboço preliminar. Tens tempo para dar uma vista de olhos?"*

– *"Claro que sim!"*

– *"Ótimo, vamos mudar para o local. Iniciar Cena M5."*

Após alguns segundos, o computador nos seus óculos pergunta:

– *"Permitir mudança de cena?"*

– *"Permitir," confirma.*

De repente, é envolvido por uma representação bastante convincente do terreno e da vizinhança.

– *"Pensei em colocar a casa aqui. Assim temos um bom jardim à frente e nas traseiras, e espaço para a entrada e a garagem à direita."*

Um contorno da casa aparece no solo.

– *"Corredor, casa de banho de visitas, sala de estar e jantar, cozinha e escadas no rés-do-chão."*

Uma planta etiquetada surge dentro do contorno. Presume que a vista do arquiteto é mais complexa do que a sua, incluindo controlos de visualização.

– *“Vamos erguer o primeiro andar.”*

A planta transforma-se em paredes. O seu ponto de vista eleva-se, e vê uma planta do segundo andar sobre a estrutura inacabada.

– *“Dois quartos, duas casas de banho e um escritório neste nível.”*

O segundo andar materializa-se.

– *“O terceiro andar é o sótão, com potencial para dois quartos e uma casa de banho.”*



Unreal Estate – A Estrada para Point Reyes. Esta cena foi sintetizada na divisão de computação gráfica da Lucasfilm em 1985 a partir de um modelo tridimensional de computador subjacente. Óculos mágicos e poder de computação suficiente permitir-nos-ão passear por mundos de fantasia como este. © 1986 Pixar.

O telhado completa a montagem, e acrescenta-se um pouco de paisagismo. Flutuando de volta ao nível do solo, circula em redor da casa e observa a cena.

– *“Podíamos experimentar isso em tijolo?”*

As paredes mudam de alvenaria de pedra para tijolo.

– *“O jardim traseiro parece um pouco pequeno. Podemos avançar a casa?”*

– *“Não podemos sair muito do alinhamento com os vizinhos, mas acho que te consigo dar mais quatro metros e meio.”*

A casa desliza para a frente.

– *“Está melhor. Vamos entrar.”*

A porta da frente abre-se. Repara num interruptor de luz à direita e estende a mão para o acionar. Uma caricatura em figura de pau da sua mão estabelece contato, e o hall ilumina-se. A vista das janelas da sala de estar não é inspiradora.

– *“E se puséssemos aqui uma janela em sacada, talvez em vez desta?”*

– *“Espera um pouco, tenho de configurar isso.”*

Enquanto isso, desliza até à cozinha, que lhe parece bastante espaçosa, e depois examina a sala de jantar.

– *“A janela está pronta.”*

A sala de estar ganha outra vida com a luz solar a entrar pela nova janela em sacada.

– *“Que estação do ano está configurada?”*

– *“Tempo real. Vamos percorrer um ano.”*

A iluminação muda, passando por manhãs e tardes nas quatro estações, e a cena de inverno é particularmente acolhedora. Após uma visita rápida aos andares superiores, com algumas mudanças nas cores das paredes, despede-se:

– *“Vou deixar uma cópia. Vais poder fazer alterações estéticas e brincar com o mobiliário; não te preocupes com estragar o projeto – o programa impedirá qualquer disparate.”*

– *“A família vai adorar isto esta noite. Falamos mais tarde.”*

Programação Sem Ferramentas

Praticantes experientes em muitos domínios relatam que conseguem ver ou sentir o objeto do seu trabalho enquanto pensam nele. Isto não é particularmente surpreendente em profissões que lidam com objetos físicos ou situações concretas — escultura, desporto ou engenharia mecânica, por exemplo. É, no entanto, menos esperado em áreas supostamente abstratas como a música, a linguagem, a matemática ou a física teórica. E, no entanto, Einstein, por exemplo, afirmou que por vezes sentia o significado das suas equações nos braços e no corpo, como se estas fossem objetos sólidos. Como sugeri no Capítulo 1, as extensas e altamente evoluídas áreas sensoriais e motoras do cérebro parecem constituir a verdadeira força motriz oculta por detrás do pensamento humano. Graças à imensa eficiência dessas estruturas com mil milhões de anos de evolução, elas poderão possuir até um milhão de vezes mais capacidade computacional efetiva do que a parte consciente da nossa mente. Embora uma atuação de principiante possa ser alcançada apenas com o pensamento consciente, o domínio ao nível de mestre recorre aos vastos recursos ocultos dessas zonas cerebrais antigas e especializadas. Por vezes, parte desse poder pode ser canalizado através da descoberta e desenvolvimento de uma correspondência útil entre o problema em questão e uma intuição sensorial.

Embora alguns indivíduos, graças a uma combinação afortunada de herança genética e oportunidades, tenham desenvolvido intuições especializadas em determinados campos, a maioria de nós é amadora na maior parte das coisas. O que precisamos, para melhorar o nosso desempenho, são metáforas externas explícitas que consigam aceder diretamente às nossas competências instintivas de forma repetível. Gráficos, regras empíricas, modelos físicos que ilustram relações — e outros dispositivos semelhantes — são amplamente utilizados, e com eficácia, para facilitar a compreensão e a retenção. Mais recentemente, interfaces computacionais interativas e visuais, como as utilizadas no Macintosh, aceleraram significativamente a aprendizagem em principiantes e facilitaram o uso das máquinas por parte dos utilizadores experientes. O envolvimento sensorial total proporcionado pelos óculos mágicos poderá permitir-nos ir muito mais além nesta direção. Encontrar as melhores metáforas será tarefa para uma geração; por agora, podemos entreter-nos a adivinhar.

A paisagem familiar do nível superior do seu sistema de ficheiros estende-se diante de si. Em primeiro plano, sobre um prado verdejante, encontram-se rochedos de diferentes tamanhos, cores e formas, com etiquetas como "Orçamento", "Desenhos", "Jogos", entre outras. Ao longe, envoltas em neblina, erguem-se colinas com inscrições como "Oxford English Dictionary" e "Encyclopedia Britannica". Dois toques no rochedo intitulado "Espaço" fazem com que este se expanda e abra um portal na sua lateral. Através desse portal, vislumbra-se algo que se assemelha a um cinturão de asteróides. Uma das rochas flutuantes no negrume é rotulada "Ganchos Celestes". Aproxima-se a flutuar, dá dois toques e entra. Surge diante de si uma bonita Terra azul e branca, acompanhada de alguns fragmentos de detritos de formas diversas — nem todos tão agradáveis à vista. Este é um projeto inacabado, e algumas das suas experiências menos bem-sucedidas ainda aguardam ser encerradas definitivamente.

O desafio de hoje consiste em desenvolver uma simulação de um cabo longo e resistente em órbita da Terra. O cabo possui massa e alguma elasticidade. Aprendeu que pode ser aproximado modelando-o como uma cadeia de múltiplas molas simples interligadas com pesos ainda mais simples. Uma mola simples liga dois pontos e exerce sobre eles uma força proporcional ao alongamento em relação ao seu comprimento de repouso. Uma massa tem uma posição e uma velocidade que se alteram em proporção à força aplicada, de acordo com as três leis do movimento de Newton. A fórmula de uma mola é $F = K(L - L_0)$, onde: L é o comprimento atual da mola, L_0 é o seu comprimento não esticado, K é a constante elástica — quanto maior for K , mais difícil é esticar a mola, F é a força exercida pela mola nas suas extremidades. Esta equação está inscrita entre os detritos espalhados pelo cenário. Começa por selecionar alguns componentes que criou anteriormente. Um ponto é representado por um pequeno ponto negro que flutua no espaço (internamente, tem três números que indicam as suas coordenadas X , Y e Z — mas essa foi a preocupação de ontem). Ao tocar no ponto e dizer “Duplicar!”, obtém um segundo. Estes representarão as duas extremidades da mola.

Vai buscar uma seta de comprimento; parece uma linha com uma ponta de seta em cada extremidade e um número ao centro (o seu comprimento). Prende as duas extremidades da seta aos seus dois pontos, e ela calcula automaticamente a distância entre eles. Experimentando-a, agarra num dos pontos e desloca-o. A seta acompanha o movimento e o número do comprimento altera-se obedientemente. Ao tocar na seta, diz:

— “Chama-se L .”

A dimensão numérica transforma-se no símbolo L .

— “Anexar fórmula da mola,” ordena, e uma cópia da fórmula emerge da paisagem, assenta-se nas proximidades e começa a reagir à distância entre os pontos. Aparecem campos para inserir os valores das variáveis K e L_0 , e você atribui-lhe valores. “Vetorizar” é outro

componente previamente preparado; dado um par de pontos e uma magnitude simples, atribui uma direção à grandeza — ou seja, a direção da linha que liga os pontos. Este vetor é ligado a ambos os pontos, em sentidos opostos.

— “Vetorizar F ,” diz.

Os pontos agora exercem a força da mola, embora ainda se mantenham fixos. Quando tenta agarrar um deles, sente que ele puxa a sua mão — quanto mais se afasta, mais forte é a força que lhe atrai de volta para a posição fixa do outro ponto. Apenas alguns tipos de grandeza podem ser diretamente experienciados desta forma: posição, cor e temperatura são outros exemplos. Em muitas aplicações, é útil traduzir medições mais abstratas em formas sensoriais palpáveis.

Adiciona massa às extremidades. Isso permite que se movam de forma independente, sob o efeito do momento e de forças aplicadas, como a força da mola. Com as extremidades libertas, a mola vibra. A vibração não diminui até adicionar um termo de amortecimento à equação da força — um termo que reduz a força consoante a taxa de variação de L . Agora a mola comporta-se de maneira mais realista, e você dá-lhe alguns “sproings”^{} só por diversão. Invocar o compilador converte a mola num único objeto e melhora consideravelmente a eficiência do programa subjacente. Edita a imagem da mola para que se pareça com uma espiral elástica, com discos pretos a representar a massa nas extremidades. Uma dúzia de cópias da mola ligadas em série compõem uma corda bastante elástica. A sua simulação começa com sucesso, mas chegou a hora do almoço. Depois do almoço, planeia criar uma secção mais longa, alterar os parâmetros em diferentes partes, e instrumentá-la — talvez através de um gráfico que mostre o alongamento das várias secções. A seguir, ativará o modelo de gravidade da Terra e colocará a corda em órbita para ver o que acontece.*

^{*} “sproings” fazer a mola saltar ou vibrar só por diversão, aproveitando o efeito físico que foi simulado.

Física Elementar

Sócrates, cujos ensinamentos nos chegaram através do seu discípulo Platão, não escreveu qualquer livro. Parecia considerar a escrita uma má ideia, pois permitia aos seus utilizadores fazerem-se passar por conhecedores apenas consultando textos, sem realmente saberem coisa alguma; a própria capacidade de recordar — e de refletir sobre essas recordações — ficava comprometida. Além disso, um argumento apresentado num livro não permite contestação, ao contrário de uma pessoa, com quem se pode discutir ou pedir esclarecimentos. Ambas as objeções têm mérito. O conhecimento transmitido por livros é, sem dúvida, seco e estático em comparação com o saber ativo que vive na mente de alguém inteligente. A invenção da imprensa agravou sobremaneira este efeito. Ainda assim, os livros têm um alcance, uma capacidade e uma permanência muito superiores às da memória de qualquer pessoa — e estas qualidades tornaram possível a civilização moderna.

Antes da Era da imprensa, os livros eram objetos caros, produzidos à mão com grande esforço, e só se encontravam em algumas bibliotecas dispersas. Ter cópias privadas era extremamente difícil, e os estudiosos viam-se forçados a memorizar volumes inteiros. Técnicas artificiais de memorização eram, por isso, altamente valorizadas — e, por vezes, guardadas com zelo face a rivais teológicos, políticos ou comerciais. Uma das técnicas mnemónicas mais eficazes, que conheceu inúmeras variações durante a Idade Média, era A Caminhada. Escolhia-se um local de grandes dimensões — uma catedral com muitas salas, por exemplo — que era recordado ou imaginado com detalhe. O livro ou a palestra a memorizar era recitado enquanto, em simultâneo, a estrutura do edifício era percorrida mentalmente. Cada sala da catedral, ou mesmo uma parte da sala, funcionava, na mente, como repositório para uma secção do texto, eventualmente marcada por algum objeto que recordasse o tema em questão. Desta forma, a tarefa era dividida em blocos geríveis; cada local exigia apenas uma quantidade moderada de memória. A totalidade do conteúdo podia ser reconstituída ao realizar novamente, em pensamento, a

caminhada pelo edifício, visitando as salas uma a uma, com as imagens mentais assim evocadas a trazerem à mente as passagens de texto associadas.

A Caminhada poderá ser tão eficaz porque traduz uma necessidade cultural relativamente recente — a de memorizar grandes quantidades de discurso — numa competência de sobrevivência muito mais antiga: recordar onde vimos ou deixámos determinadas coisas. Recordar a localização de uma fonte de alimento, de um abrigo, de um perigo, de um amigo ou inimigo, ou simplesmente de um ponto de referência avistado durante uma deslocação tem benefícios óbvios no quotidiano, sendo algo que muitos de nós fazemos de forma natural e com bastante eficácia. Pelo menos uma parte da nossa memória parece estar organizada de forma aproximadamente geográfica, precisamente para facilitar este tipo de evocação.

Uma lição transmitida através de uma versão avançada do guarda-roupa mágico pode ser, em simultâneo, tão responsiva como um diálogo pessoal, tão permanente e acessível como um livro impresso, e ainda capaz de ressoar com competências naturais humanas de um modo que ultrapassa qualquer método existente.

O portal da "Gravidade" abre-se sobre uma cena campestre banhada por uma luz solar intensa. Uma estrada rural ladeada de árvores serpenteia em direção às colinas distantes; nuvens fofas salpicam o céu, ouvem-se pássaros a cantar algures. Algumas das árvores dão maçãs, e de tempos a tempos uma delas cai ao chão. Um pouco mais adiante, ao longo do caminho, avista-se uma figura de peruca sentada sob uma dessas macieiras. Reconhece Sir Isaac Newton. Está exatamente como o viu no capítulo das "Leis do Movimento".

— "Saudações, jovem amigo," diz Sir Isaac. "Tenho-me debatido com a natureza da atração que a Terra exerce sobre os vários objetos. Esta maçã, por exemplo, puxa a mão com uma certa força." Ele entrega-lhe a maçã; de facto, ela tem peso.

— *“Uma maçã com o dobro da substância puxa com o dobro da força.” A maçã cresce e torna-se mais pesada.*

— *“O grande Galileu observou que, quando largado, um objeto cai em direção ao solo com uma velocidade que aumenta constantemente, independentemente do seu peso.” Vêm-lhe à mente as demonstrações de Galileu com esferas em queda, no capítulo das “Leis do Movimento”.*

— *“Sim, sim, prossigamos.” Newton, com um ligeiro franzir de sobrolho, continua:*

— *“Podemos concluir que cada partícula de um objeto é atraída para o centro da Terra com uma força proporcional à sua massa. Mas será que esta atração varia com a distância em relação à Terra? Pode conjecturar-se que essa influência se estende a grandes distâncias e mantém a Lua na sua órbita mensal. Se as mesmas leis se aplicarem aos corpos celestes como aos terrestres, então os nossos estudos sobre o movimento dos objetos indicam que uma força dirigida ao centro da Terra é suficiente para curvar a trajetória da Lua. Contudo, a força necessária é quase 4.000 vezes mais fraca, por partícula de massa, do que a que sente ao segurar essa maçã.”*

Enquanto ele fala, o solo começa a elevar-se a uma velocidade fantástica, e você, Newton e a árvore estão agora no cimo de uma colina que sobe como um foguete.

— *“Considera a trajetória de um objeto lançado horizontalmente de uma grande altura — a sua maçã, por exemplo.”*

Percebendo a sugestão, lança a maçã com um arremesso firme e de braço levantado. (No mundo real, os motores no seu casaco e nas luvas emitem um leve zumbido ao resistirem ao movimento do seu braço, simulando as forças da inércia da maçã.) A maçã descreve um arco lento em direção ao solo e atinge o chão perto do horizonte. A colina parou de crescer, mas está muito alto, e a curvatura esférica do planeta torna-se visível. Consegue distinguir vários

contornos continentais. É claramente um modelo à escala reduzida da Terra. Sir Isaac entrega-lhe outra maçã e recomenda um lançamento mais vigoroso. Esta descreve uma curva que se estende além do horizonte, quase paralela ao solo. Ouve um splat vindo do subsolo — a maçã caiu do outro lado do planeta. Um lançamento ainda mais forte não resulta em qualquer impacto visível e, passado algum tempo, a maçã passa a assobiar por detrás da sua cabeça, completando uma órbita completa. Uma vista lateral em miniatura de si próprio, da colina, da Terra e das maçãs torna tudo isto mais claro: cada lançamento traça uma elipse que regressa ao ponto de partida, a menos que intercete o solo antes. Newton recorda as leis do movimento planetário de Kepler e afirma que estas só se aplicam às maçãs se a atração gravítica diminuir com o quadrado da distância ao centro do planeta. Manifesta-se algum ceticismo da sua parte, pelo que ambos experimentam outras regras. Algumas fazem as maçãs descreverem padrões não repetitivos. Outras, embora produzam elipses, violam a segunda lei de Kepler — a que diz que a linha que une o centro do planeta ao corpo em órbita varre áreas iguais em tempos iguais. Ao fim de algum tempo, o seu braço de tanto lançar já está cansado, e diz-se convencido.

Mas por vezes o seu ceticismo leva-o a formular questões que deixam o seu anfitrião sem resposta. Lembra-se de Newton ter respondido em certa ocasião:

— “Um enigma curioso. Permite-me refletir sobre ele por algum tempo.”

Várias visitas depois, surge a correr atrás de si com a resposta, as abas do casaco esvoaçando, uma mão a segurar a peruca, deixando atrás de si uma nuvem de poeira. (Presume que o software do livro, sem conseguir responder à questão na primeira vez, tenha emitido uma notificação aos seus autores. Estes, por sua vez, criaram novas entradas na base de dados do livro, permitindo que a sua pergunta pendente — e quaisquer outras semelhantes que Isaac venha a encontrar no futuro — possam agora ser respondidas.)

A colina encolhe de novo até se tornar plana, e está de volta à estrada. A paragem seguinte é um prado, onde algumas das partes mais formais da lição serão discutidas. Já lá se encontra reunido um grupo de criaturas exóticas. Muitas pessoas por todo o mundo estão a ler este livro, e a rede mundial torna possível que aqueles que desejem associar-se estejam mutuamente conscientes uns dos outros durante o decurso do estudo. Nestas associações, a maioria das pessoas aproveita a liberdade da simulação para assumir formas diferentes dos seus corpos físicos — por anonimato, por brincadeira ou simples capricho. O seu grupo inclui um Lobo, um Olho Flutuante, um Homem de Lata, uma Borboleta Gigante, um Dragão e um Pequeno Tanque. Você próprio se apresenta sob a forma de um Anão elegante, com machado e chapéu de borla. Alguns antigos colegas que começaram este livro de física consigo já não pertencem à sua coorte: ou avançaram mais depressa, ou ficaram para trás, ou tomaram um rumo diferente num ponto de ramificação da matéria. De tempos a tempos, encontra novos companheiros de viagem, à medida que subcategorias se voltam a fundir. O mundo inteiro está dividido em “aldeias de interesse comum” sobrepostas deste tipo. Os grupos podem variar entre duas pessoas e vários milhares. Muitas vezes, claro, é bom percorrer os caminhos do conhecimento e do entretenimento em solidão. Entre outras vantagens, a ação pode ser melhor ajustada ao indivíduo, dado que há menos restrições.

Após a lição, lança um olhar para mais além na estrada. Ao longe, vê uma plataforma ferroviária com um comboio de passageiros parado. Observando com atenção, nota no vidro de um dos vagões o perfil algo desalinhado do cientista mais famoso do mundo. Mas estás cansado e decide desligar-se por hoje. A relatividade pode esperar por amanhã.

CAPÍTULO 4

CLÁUSULA DO AVÔ*

O QUE acontecerá quando máquinas cada vez mais baratas forem capazes de substituir seres humanos em qualquer circunstância? E, na verdade, o que farei eu quando um computador conseguir escrever este livro, ou realizar a minha investigação, melhor do que eu? Estas perguntas já se tornaram cruciais para muitas pessoas em todo o tipo de profissões, e dentro de algumas décadas dirão respeito a toda a gente. Por natureza, as máquinas são nossos servos obedientes e competentes. Mas as máquinas inteligentes, por muito benevolentes que sejam, representam uma ameaça à nossa existência, pois tornam-se habitantes alternativos do nosso nicho ecológico. Máquinas apenas tão inteligentes quanto os seres humanos terão vantagens imensas em situações de competição. Custam menos a produzir e a manter, pelo que será possível pôr muito mais delas a trabalhar com os mesmos recursos disponíveis. Podem ser otimizadas para as suas funções e programadas para trabalhar incansavelmente.

Como se estes avanços tecnológicos não fossem já suficientemente inquietantes, o próprio ritmo da inovação apresenta um desafio ainda mais sério à nossa segurança. A nossa evolução deu-se a um ritmo vagaroso, com milhões de anos entre mudanças significativas. As máquinas estão a realizar progressos semelhantes no espaço de meras décadas. Quando multidões de máquinas económicas forem postas a trabalhar como programadoras e engenheiras, com a tarefa de otimizar o software e o hardware que as constitui, o ritmo acelerará. As gerações

* A expressão "Grandfather Clause" (em português, "cláusula de anterioridade" ou, informalmente, "cláusula do avô") refere-se a uma disposição legal que permite que certas pessoas ou entidades continuem a seguir regras antigas, mesmo depois de uma nova lei ou regulamento ter sido implementado. Por exemplo, se uma nova lei proíbe algo, mas alguém já fazia isso antes da lei entrar em vigor, essa pessoa pode continuar a fazê-lo graças à "grandfather clause". É uma forma de evitar que mudanças legais prejudiquem quem já estava em conformidade com as regras anteriores.

No contexto do capítulo a expressão é usada como uma metáfora para discutir a continuidade da identidade pessoal em cenários de transferência da mente — como, por exemplo, quando uma consciência humana é copiada ou transferida para um suporte artificial. A ideia é que, tal como uma cláusula de anterioridade permite que certas regras antigas continuem a aplicar-se a casos específicos, também a identidade de uma pessoa poderia ser "preservada" mesmo após uma mudança radical no seu suporte físico — como passar de um corpo biológico para um corpo robótico ou digital. A questão central é: se a mente for copiada com precisão, essa cópia ainda é "tu"?

sucessivas de máquinas assim produzidas tornar-se-ão mais inteligentes e menos dispendiosas. Não há razão para acreditar que a equivalência com o ser humano represente qualquer tipo de limite superior. Quando simples calculadoras de bolso conseguem pensar melhor do que humanos, como será um computador de grande escala? Estaremos, pura e simplesmente, ultrapassados.

Então por que razão avançar, de forma precipitada, para uma Era de máquinas inteligentes? A resposta, creio eu, é que temos muito pouca escolha, se quisermos que a nossa cultura permaneça viável. As sociedades e as economias estão seguramente tão sujeitas às pressões evolutivas da competição quanto os organismos biológicos. Mais cedo ou mais tarde, aquelas que conseguirem sustentar a expansão e a diversificação mais rápidas dominarão. As culturas competem entre si pelos recursos do universo acessível. Se a automação for mais eficiente do que o trabalho manual, as organizações e sociedades que a adotarem tornar-se-ão mais ricas e estarão melhor preparadas para sobreviver em tempos difíceis e prosperar em tempos favoráveis. Se os Estados Unidos decidissem, de forma unilateral, travar o desenvolvimento tecnológico (uma ideia por vezes tida como progressista), acabariam por sucumbir, seja ao poderio militar de nações hostis, seja ao sucesso económico dos seus parceiros comerciais. Em qualquer dos casos, os ideais sociais que levaram à decisão tornar-se-iam irrelevantes à escala mundial.

Se, por algum pacto improvável, toda a espécie humana decidisse renunciar ao progresso, o resultado a longo prazo seria, quase com certeza, a extinção. O universo é uma sequência interminável de eventos aleatórios. Mais cedo ou mais tarde evoluirá um vírus imparável e letal para os humanos, ou um asteroide colidirá com a Terra, ou o Sol expandir-se-á, ou seremos invadidos por forças vindas das estrelas, ou um buraco negro engolirá a galáxia. Quanto maior, mais diversa e mais competente for uma cultura, melhor será a sua capacidade de deteção e resposta a perigos externos. Os grandes eventos acontecem com menor frequência. Ao crescer a um ritmo suficientemente rápido, uma cultura terá uma possibilidade finita de sobreviver para sempre. No Capítulo 6, fantasiar-se-á sobre esquemas que

permitiriam a uma entidade reestruturar-se de modo a funcionar indefinidamente, mesmo com o fim do seu universo.

A raça humana expandir-se-á pelo sistema solar dentro de pouco tempo, e colónias espaciais habitadas por humanos farão parte dessa expansão. Mas só através de uma implementação massiva de máquinas poderemos sobreviver nas superfícies de outros planetas ou no espaço exterior. O projeto Apollo, por exemplo, levou pessoas à Lua durante algumas semanas por 40 mil milhões de dólares, ao passo que os módulos Viking funcionaram em Marte durante anos, por apenas mil milhões. Se máquinas tão capazes como os humanos tivessem estado disponíveis no projeto Viking, teriam conseguido recolher muito mais informação sobre Marte do que os humanos conseguiram recolher sobre a Lua — simplesmente porque as máquinas podem ser construídas para funcionar de forma confortável e económica em ambientes inóspitos à vida terrestre.

O espaço exterior já é um domínio lucrativo para os proprietários de satélites de comunicações. À medida que os custos de transporte diminuem, outras atividades começarão também a ser rentáveis. Fábricas espaciais, utilizando matérias-primas compradas na Terra ou obtidas em postos humanos no espaço, serão operadas por máquinas supervisionadas por humanos e gerarão lucro. O elevado custo de manutenção da presença humana no espaço garante que sempre existirá uma maior quantidade de maquinaria por pessoa numa colónia espacial do que na Terra. À medida que as máquinas se tornarem mais capazes, a economia tenderá a favorecer uma proporção cada vez maior de máquinas por pessoa. Isto não significa necessariamente que haverá menos humanos nesta fase; simplesmente, as máquinas multiplicar-se-ão mais rapidamente, tornando-se cada vez mais competentes a cada nova geração. Imagine-se as imensamente lucrativas fábricas de robôs que poderiam ser construídas nos asteróides. Máquinas alimentadas por energia solar fariam a prospeção e entregariam matérias-primas a enormes fábricas automáticas e não encerradas. Os metais, semicondutores e plásticos ali produzidos seriam convertidos por robôs em componentes, os quais seriam então montados noutros robôs e em peças estruturais para mais fábricas.

As máquinas seriam recicladas à medida que se avariassem. Se a sua taxa de reprodução fosse superior à taxa de desgaste, estas fábricas cresceriam exponencialmente, como uma colónia de bactérias — mas à escala de Brobdingnag^{*}. A colheita de apenas uma pequena fração da produção de materiais, componentes e robôs completos poderia tornar os investidores incrivelmente ricos.

Eventualmente, os humanos — sejam operários, engenheiros de projeto, gestores ou investidores — tornar-se-ão desnecessários nas empresas espaciais, à medida que as descobertas científicas e técnicas de mecanismos superinteligentes autorreplicáveis forem aplicadas para os tornar ainda mais inteligentes. Estas novas criações, com aparência muito distinta das máquinas que conhecemos, explodirão pelo universo fora, deixando-nos para trás numa nuvem de poeira.

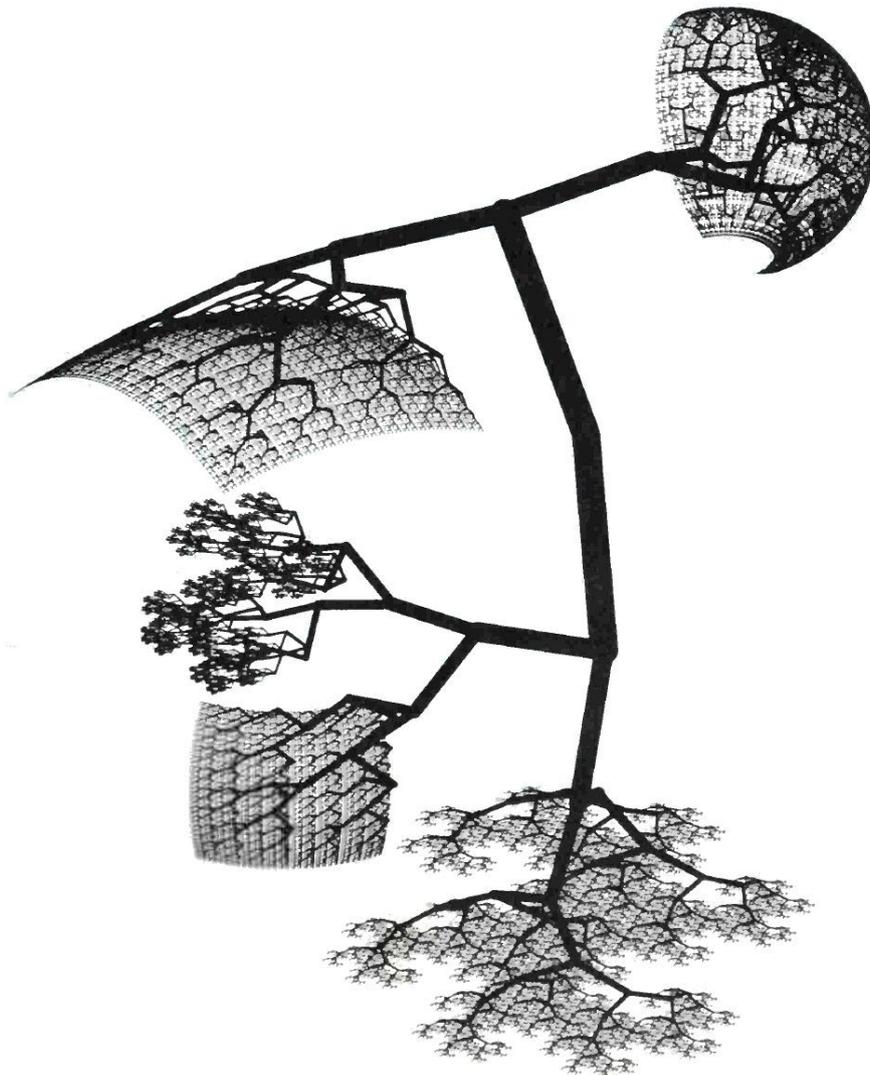
Robôs Arbustos

O mundo humano foi moldado por mãos humanas, que continuam a ser a nossa ferramenta generalista mais eficaz. No entanto, muitas tarefas úteis e facilmente descritas estão para além da nossa destreza (puxar com força ambas as extremidades de um cordel enquanto se segura o nó entre os dedos, levantar o embrulho, enrolar bem as pontas à volta dele quatro vezes...). Quando tais ações são tentadas, é com graus variados de sucesso, frequentemente recorrendo a ferramentas e dispositivos específicos.

É pouco provável que os nossos descendentes superinteligentes se contentem com simples dedos curtos. Consideremos as seguintes observações. Animais como vermes ou outros organismos com formas de bastão ou esféricas não conseguem manipular objetos, e mesmo a locomoção lhes é difícil. Animais com pernas (um bastão com bastões móveis mais pequenos) locomovem-se razoavelmente bem, mas continuam a ser desajeitados na manipulação. Já os animais como nós, com

^{*} A expressão “à escala de Brobdingnag” é uma referência literária ao livro *As Viagens de Gulliver*, de Jonathan Swift. Brobdingnag é uma terra fictícia onde tudo — pessoas, animais, objetos — é gigantesco em relação ao protagonista, Gulliver. Portanto, dizer que algo cresce “à escala de Brobdingnag” é uma forma criativa de dizer que cresce de forma colossal ou descomunal.

dedos nas extremidades das pernas (bastões sobre bastões sobre um bastão), conseguem manipular com muito mais destreza. Agora, generalizemos este conceito: imaginemos um robô com forma de árvore, cujo caule principal se ramifica sucessivamente em ramos mais curtos, mais finos e em maior número, até terminar numa quantidade astronômica de cílios microscópicos. Cada ramo intermédio poderia mover-se para a frente e para trás, lateralmente, e o seu topo — onde se ligam os ramos seguintes — girar em torno do seu próprio eixo. É possível até que o ramo se pudesse estender ou retrair como um telescópio — o número de movimentos possíveis em cada ramo pode ser trocado por mais níveis de ramificação. Cada articulação teria sensores para medir a sua posição e também a força que exerce. Apesar de composto por ramos individuais com uma natureza mecânica rígida, o conjunto da estrutura teria uma flexibilidade "orgânica", devido à enorme multiplicidade de formas como as suas partes poderiam mover-se.



Um Robô Arbusto.

Um robô com este tipo de estrutura poderia ser autoconstruído. Pequenos arbustos, com apenas alguns milionésimos do peso do dispositivo final, seriam "semeados" para dar início ao processo. Estes trabalhariam em grupos para construir os raminhos seguintes, de maior dimensão, a partir de matérias-primas disponíveis, unindo-se depois às suas construções. Os arbustos maiores assim formados unir-se-iam para edificar ramos ainda maiores, e assim sucessivamente, até que uma pequena equipa (composta por membros já grandes) se reunisse para montar o caule principal. Na outra extremidade da escala, um arbusto suficientemente grande deveria ser capaz de organizar os recursos necessários para fabricar as pequenas sementes que dariam início a um novo processo (ou simplesmente para se reparar

ou expandir). Poderia fabricar as partes mais pequenas recorrendo a métodos semelhantes às técnicas de micromaquinagem usadas atualmente na produção de circuitos integrados. Se os seus ramos mais finos tivessem dimensões de apenas alguns átomos (com comprimentos medidos em nanómetros), o arbusto-robô poderia agarrar átomos individuais da matéria-prima e organizá-los um a um em novas peças, numa variante dos métodos da nanotecnologia mencionados no Capítulo 2.

Para tornar a ideia mais concreta, podemos avançar com um modelo de estrutura. Suponhamos que a estrutura básica é composta por um grande ramo que se divide em quatro menores, cada um com metade da escala. Se começarmos com um tronco de um metro de comprimento e dez centímetros de diâmetro e levarmos a ramificação até vinte níveis, o arbusto terminará com um trilião de pequenas "folhas", cada uma com um micrómetro de comprimento (um milionésimo de metro) e um décimo disso em diâmetro. Devido ao seu peso e dimensão muito inferiores, estas folhas podem mover-se um milhão de vezes mais depressa do que o tronco. Suponhamos que o tronco oscila uma vez por segundo; então as folhas vibrarão um milhão de vezes por segundo. Se o arbusto se dobrar sobre si próprio formando um feixe compacto, a sua secção transversal permanecerá aproximadamente constante e terá cerca de dois metros de comprimento. O tronco ocupará metade desse comprimento, o segundo nível metade do que resta, e assim sucessivamente. Desdobrado como um guarda-chuva, espalhar-se-ia numa superfície circular com pouco menos de dois metros de diâmetro, densa, mas esparsa no centro, e cada vez mais fina na periferia, com interstícios progressivamente mais estreitos até chegar a espaços da ordem do micrómetro.

Se cada articulação puder medir as forças e os movimentos que lhe são aplicados, teremos um sensor verdadeiramente extraordinário. Existem um trilião de dedos-folha, cada um capaz de detetar movimentos de cerca de um décimo de micrómetro e forças de apenas alguns microgramas, a velocidades que podem atingir um milhão de variações por segundo. Esta capacidade sensorial supera imensamente a do olho humano, que distingue cerca de um milhão de pontos e consegue registar alterações, no

máximo, cem vezes por segundo. Se o nosso arbusto colocar os seus dedos sobre uma fotografia, será capaz de “ver” a imagem com um detalhe imenso simplesmente através da percepção das variações de relevo na prata revelada sobre o papel. Poderia ver um filme ao passar os dedos rapidamente pela película enquanto esta se desenrola a alta velocidade. Nada impede que estes dedos sejam também sensíveis à luz, à temperatura e a outros efeitos eletromagnéticos; de facto, os menores têm precisamente a dimensão adequada para funcionarem como “antenas” de luz. O arbusto poderia formar um olho ao sustentar uma lente e posicionar alguns milhões dos seus dedos no plano focal por detrás dela. Poderá até dispensar a lente, organizando um conjunto de dedos segundo um padrão de difração cuidadosamente espaçado, formando assim uma lente holográfica.

Para além de possuir uma capacidade sensorial equivalente à da população humana atual, o nosso arbusto teria a habilidade de atuar sobre o ambiente com igual magnitude. Um ser humano experiente, utilizando movimentos corporais e manuais precisos e bem cronometrados — cada um capaz de mudar de direção apenas algumas vezes por segundo, com uma precisão não superior a alguns por cento do movimento total — poderia, no máximo, influenciar o mundo a uma taxa líquida de cerca de mil bits por segundo (um dactilógrafo veloz, por exemplo, produz menos de cem bits por segundo de texto). Já o potencial de transmissão de dados de um robô com um trilião de dedos, cada um capaz de se mover um milhão de vezes por segundo, ultrapassaria um quadrilhão (10^{15}) de vezes essa taxa. Tais velocidades de transmissão implicam uma coordenação imensa e um poder de processamento colossal, mas as possibilidades são assombrosas. O robô-arbusto poderia introduzir-se num complexo mecanismo delicado — ou mesmo num organismo vivo — detetar simultaneamente a posição relativa de milhões de componentes, alguns tão pequenos quanto moléculas, e reorganizá-los de forma a realizar um reparo praticamente instantâneo. Em muitos casos, o seu sentido de tato superior substituiria por completo a visão, e a sua extrema destreza eliminaria a necessidade de ferramentas especializadas.

Controlar esta máquina maravilhosa exigiria uma quantidade astronômica de pensamento. Grande parte desse esforço poderia ser gerido por aquilo a que, nos animais, se chama arcos reflexos—pequenas unidades do sistema nervoso próximas do local a ser controlado. Cada um dos pequenos ramos poderia conter um microcomputador suficiente para coordenar a atividade rotineira; apenas situações excepcionais exigiriam a intervenção de computadores maiores, situados mais perto do caule. Se os ramos contivessem também a sua própria fonte de energia (imagine-se uma bateria recarregável) e um modo de comunicação remota (ondas de rádio ou vibrações sonoras produzidas por milhares de cílios sincronizados poderiam servir), o arbusto poderia fragmentar-se numa nuvem coordenada de arbustos menores. Quanto mais pequeno fosse o arbusto individual, menor seria a sua inteligência e capacidade. Cada unidade seria pré-programada e carregada pelo seu tronco de origem para executar uma determinada função e regressar o mais rapidamente possível para reportar e receber novas instruções.

O pequeno tamanho constituiria, frequentemente, uma vantagem: um robô diminuto pode penetrar em espaços minúsculos. Uma máquina pequena tem uma razão superfície/peso mais elevada: enquanto um grande arbusto poderia caminhar com segurança sobre o chão, usando os ramos como dedos ágeis, uma máquina mais pequena deveria conseguir caminhar nos tetos, como uma mosca, agarrando-se a fendas microscópicas na pintura ou aderindo por forças moleculares. Os arbustos poderiam escavar o solo soltando partículas de terra e empurrando-as para trás, e nadar com eficiência ao assumirem uma forma compacta e hidrodinâmica, com os cílios a formarem uma pele que bombeia fluido para propulsão e responde ao fluxo para evitar turbulência. Máquinas extremamente pequenas teriam uma superfície tão vasta em relação ao peso que poderiam voar como insetos, batendo os cílios em padrões otimizados para o movimento do ar.

A contribuição dos ramos maiores para a potência e inteligência dos mais pequenos pode ser visualizada como uma espécie de esquema piramidal invertido. Cada nível do robô (considerando todos os ramos desse nível) tem o dobro do volume

do nível imediatamente inferior e, portanto, espaço para o dobro da energia e do poder computacional. Dois níveis abaixo, a proporção é de quatro para um; a três níveis, é de oito. Se o controlo e a energia de cada nível forem fornecidos por ramos situados três etapas acima na hierarquia em direção ao caule, os pequenos ramos podem ser quatro vezes mais vigorosos do que de outra forma seriam. Apenas o tronco e os primeiros ramos a partir dele ficariam limitados. Como a maioria da sua energia e atenção se dirige para cima, esses elementos talvez fiquem inaptos para grande mobilidade, podendo ser relegados à função de estrutura estável, caso bloqueiem passivamente as suas articulações.

Uma grande questão é saber como funcionariam os programas de controlo para tal criatura. No caso extremo, poderíamos imaginar um programa que tentasse determinar a combinação de ações de cada articulação individual que melhor realizasse a tarefa desejada. Este é, quase certamente, um exemplo clássico de um problema NP (polinomial não-determinístico), que só pode ser resolvido de forma ótima examinando essencialmente todas as combinações possíveis de movimentos e escolhendo a melhor (ver Apêndice 3). Tais soluções são intratáveis nos dias de hoje, mesmo para manipuladores simples que tenham apenas um punhado (!) de dedos. Embora os computadores venham a ser infinitamente mais poderosos no futuro, o problema representado por um sistema com um grande número de dedos será ainda muito, muito maior. De vez em quando poderá surgir uma estratégia especialmente engenhosa para coordenar milhares ou milhões de dedos a fim de executar uma determinada tarefa, e essas coleções de estratégias engenhosas serão transmitidas entre indivíduos — serão as habilidades manuais da Era superinteligente. Mas enquanto os humanos ensinam habilidades tão simples como atar os atacadores dos sapatos, as lições das máquinas superinteligentes poderão ser mais comparáveis às instruções para montar um avião de passageiros. Encontrar a solução ótima para problemas rotineiros será, na maioria das vezes, impossível; mas encontrar uma suficientemente boa talvez não seja tão difícil. Imagino uma estratégia de dividir para conquistar, onde o tronco considera o problema global e gera subtarefas plausíveis para cada um dos

quatro subarbustos a si imediatamente ligados. Estes, por sua vez, subdividem ainda mais o problema e passam esses fragmentos adiante, e assim sucessivamente. Os ramos mais pequenos receberiam comandos simples, como mover-se até certa posição, ou mover-se até encontrar resistência. Um comando vindo do tronco poderia ser algo como:

Arbusto do Norte — manter-se do lado esquerdo do plano A, e do lado direito do plano B, e aplicar o vetor de força V ao objeto;

Arbusto do Este — manter-se do lado direito de A e B, e resistir a qualquer movimento superior a 10 cm do eixo;

Arbusto do Sul — direito de A, esquerdo de B, aplicar força negativa V ;

Arbusto do Oeste — esquerdo de A e B, e resistir.

Se um subproblema, ao ser transmitido a um pequeno arbusto, não puder ser resolvido, uma reclamação seria enviada de volta ao ramo de origem, que então regressaria à prancheta para tentar outra abordagem.

Um robô-arbusto seria um prodígio surrealista de se contemplar. Apesar da sua semelhança estrutural com muitos seres vivos, seria diferente de tudo quanto jamais se viu na Terra. A sua grande inteligência, a coordenação soberba, a velocidade astronómica e a enorme sensibilidade ao ambiente permiti-lo-iam fazer constantemente algo surpreendente, mantendo ao mesmo tempo uma graciosidade perpétua. Animais bípedes têm três ou quatro andamentos eficazes; quadrúpedes possuem alguns mais. Humanos, com duas mãos, dispõem de duas ou três formas de segurar um objeto. Um dispositivo com um trilião de membros, dotado de um cérebro à altura, é uma entidade de ordem inteiramente distinta. Se a isto se acrescentar a capacidade de fragmentar-se numa nuvem de minúsculos voadores coordenados, as próprias leis da física parecerão derreter perante a intenção e a vontade. Como nenhum mágico que alguma vez existiu, coisas impossíveis acontecerão simplesmente à volta de um robô-arbusto. Imagine habitar um corpo assim.

Transmigração

Alguns de nós, humanos, possuem visões do mundo bastante egocêntricas. Antecipamos a descoberta, ainda durante as nossas vidas, de métodos para prolongar a existência humana, e aguardamos com expectativa alguns éons de exploração do universo. A ideia de sermos grandemente ultrapassados nessa empreitada pelos nossos descendentes artificiais é decepcionante. A longevidade perde grande parte do seu sentido se estivermos condenados a passá-la a olhar, estupidificados, para as nossas máquinas ultrainteligentes enquanto elas tentam descrever, em linguagem de bebê, as suas descobertas cada vez mais espetaculares. Queremos tornar-nos participantes plenos e livres neste novo jogo da superinteligência. Quais são as possibilidades para que isso aconteça?

A engenharia genética poderá parecer uma opção fácil. Gerações sucessivas de seres humanos poderiam ser concebidas através de modelos matemáticos, simulações computacionais e experimentação, como hoje se fazem aviões, computadores e robôs. Poderiam possuir cérebros melhores e metabolismos aperfeiçoados que lhes permitissem viver confortavelmente no espaço. Mas, presumivelmente, continuariam a ser feitos de proteína, e os seus cérebros continuariam a ser compostos por neurónios. Fora da Terra, a proteína não é um material ideal. É estável apenas dentro de uma faixa muito estreita de temperatura e pressão, é extremamente sensível à radiação, e limita muitas técnicas de construção e componentes. Além disso, é pouco provável que os neurónios — que atualmente conseguem comutar menos de mil vezes por segundo — venham alguma vez a atingir as velocidades de milhares de milhões de comutações por segundo dos componentes informáticos atuais. Dentro em breve, as tecnologias convencionais, miniaturizadas à escala atômica, e a biotecnologia, cujas interações moleculares terão sido compreendidas em termos mecânicos detalhados, terão fundido num leque contínuo de técnicas que abrange todos os materiais, tamanhos e níveis de complexidade. Os robôs, então, serão feitos de uma mistura de substâncias fabulosas, incluindo, quando

conveniente, materiais biológicos vivos. Nesse futuro, um superhumano concebido geneticamente será apenas uma espécie de robô de segunda categoria, concebido sob a limitação de só poder ser construído por meio da síntese proteica orientada por ADN. Apenas aos olhos de chauvinistas humanos teria alguma vantagem — por conservar mais das limitações humanas originais do que os outros robôs.

Os robôs, sejam de primeira ou de segunda categoria, não respondem à nossa pergunta. Haverá alguma hipótese de que nós — você e eu, pessoalmente — possamos partilhar plenamente o mundo mágico que se avizinha? Isso exigiria um processo que conferisse a um indivíduo todas as vantagens das máquinas, sem perda de identidade pessoal. Muitas pessoas estão hoje vivas graças a um arsenal crescente de órgãos artificiais e outras partes do corpo substituídas. Com o tempo, especialmente à medida que as técnicas robóticas forem evoluindo, essas peças de substituição serão melhores do que quaisquer originais. Então, por que não substituir tudo? Ou seja, transplantar o cérebro humano para um corpo robótico especialmente concebido? Infelizmente, embora esta solução possa superar a maioria das nossas limitações físicas, deixaria intacta a nossa maior desvantagem: a inteligência limitada e fixa do cérebro humano. Esse cenário de transplante tira o nosso cérebro do corpo. Mas haverá forma de tirar a mente do cérebro?

Acabou de ser levado para a sala de operações. Um cirurgião-robô especializado em cérebros está presente. Ao seu lado, encontra-se um computador à espera de se tornar um equivalente humano, faltando-lhe apenas um programa para funcionar. O seu crânio, mas não o seu cérebro, foi anestesiado. Está plenamente consciente. O cirurgião-robô abre a caixa craniana e pousa uma mão sobre a superfície do cérebro. Esta mão invulgar está coberta de maquinaria microscópica, ligada por um cabo ao computador móvel que o acompanha. Instrumentos incorporados na mão examinam os primeiros milímetros da superfície cerebral. Medições de ressonância magnética de alta resolução constroem um mapa químico tridimensional, enquanto conjuntos de antenas magnéticas e elétricas recolhem sinais que são rapidamente decifrados,

revelando, instante a instante, os impulsos que percorrem os neurónios. Estas medições, aliadas a um conhecimento aprofundado da arquitetura neural humana, permitem ao cirurgião escrever um programa que simula o comportamento da camada mais superficial do tecido cerebral examinado. Este programa é instalado numa secção do computador e ativado. As medições da mão fornecem-lhe cópias dos estímulos que o tecido original está a receber. Você e o cirurgião comparam os sinais produzidos pela simulação com os do tecido original. Passam muito rapidamente, mas quaisquer discrepâncias são realçadas num ecrã. O cirurgião ajusta finamente a simulação até que a correspondência seja quase perfeita.

Para se convencer da fidelidade da simulação, é-lhe dado um botão de controlo que lhe permite "testar" momentaneamente a simulação, comparando-a com o funcionamento do tecido original. Ao pressioná-lo, conjuntos de eléctrodos na mão do cirurgião são ativados. Com injeções precisas de corrente e pulsos eletromagnéticos, estes eléctrodos anulam a atividade neuronal normal, injetando, em seu lugar, os sinais produzidos pela simulação, nos locais onde o tecido simulado comunicaria com outras regiões do cérebro. Enquanto mantém o botão premido, uma pequena parte do seu sistema nervoso está a ser substituída por uma simulação computacional de si mesma. Carrega no botão, larga, e torna a carregar. Não deve sentir qualquer diferença. Quando se sentir seguro, a ligação da simulação torna-se permanente. O tecido cerebral torna-se então redundante – recebe os estímulos e reage como dantes, mas os seus sinais de saída são ignorados. Manipuladores microscópicos na superfície da mão excisam (removem) as células desse tecido obsoleto e aspiram-nas. A mão do cirurgião afunda-se mais uma fração de milímetro, ajustando instantaneamente as medições e sinais à nova posição. O processo repete-se camada após camada: a parte correspondente do cérebro é simulada, depois removida. Por fim, a sua caixa craniana está vazia, e a mão do cirurgião repousa profundamente no seu tronco cerebral. Apesar disso, não perdeu a consciência, nem sequer o fio

de pensamento — a sua mente foi removida do cérebro e transferida para uma máquina.

Num último passo desorientador, o cirurgião retira a mão. O seu corpo, subitamente abandonado, entra em espasmos e morre. Durante um instante, apenas o silêncio e a escuridão. Depois, volta a abrir os olhos. A sua perspectiva mudou. A simulação computacional foi desligada do cabo que a ligava à mão do cirurgião e reconetada a um novo corpo reluzente, do estilo, cor e material à sua escolha. A sua metamorfose está completa.

Para os mais sensíveis, existem outras formas de realizar a transferência da mente humana para uma máquina. Uma varredura cerebral de alta resolução poderia, de uma só vez e sem cirurgia, criar um novo “você” enquanto-espera. Se mesmo esta técnica ainda lhe parecer demasiado invasiva, imagine uma abordagem mais psicológica. Um tipo de computador portátil (talvez usado como uns óculos mágicos) é programado com os universais da mentalidade humana, com a sua constituição genética e com os pormenores da sua vida que estiverem facilmente acessíveis. Esse dispositivo transporta um programa concebido para ser um excelente imitador. Acompanha-lhe durante o auge da sua vida; escuta-lhe e observa-lhe atentamente; talvez monitorize o seu cérebro e aprenda a antecipar cada um dos seus movimentos e reações. Depressa se torna capaz de enganar os seus amigos ao telefone com uma imitação sua convincente. Quando morre, esse programa é instalado num corpo mecânico que assume a sua vida e responsabilidades de forma fluida e contínua.

Se for um vertebrado, existe ainda uma outra opção que combina os pontos de venda de todos os métodos descritos acima. O cérebro dos vertebrados possui dois hemisférios ligados por vários grandes feixes de fibras nervosas. O maior destes é o chamado corpo caloso. Na década de 1960, guiados por experiências com animais, investigadores na Califórnia trataram com sucesso pacientes com tipos de epilepsia intratáveis ao cortar-lhes os corpos calosos. (Os robôs médicos do futuro não usarão

latim!)* Surpreendentemente, esta intervenção não parecia, à partida, provocar quaisquer efeitos secundários nos pacientes. O corpo caloso, com duzentos milhões de fibras nervosas, é a mais imponente ligação de longa distância do cérebro. É muito mais espesso do que o nervo óptico ou a medula espinal. Se se cortar o nervo óptico, a pessoa torna-se completamente cega; se se seccionar a medula espinal, o corpo torna-se flácido e insensível. Mas ao cortar o cabo maciço entre os hemisférios cerebrais... aparentemente nada de mau acontece. Bem, quase nada. Se o nome de um objeto, como “escova”, for mostrado no campo visual direito de uma pessoa com cérebro dividido, ela não consegue identificar esse objeto com a mão esquerda, mas fá-lo sem dificuldade com a mão direita. Por vezes, na versão canhota da tarefa, a mão direita — aparentemente frustrada — estende-se para guiar a esquerda até ao local correto!

A neuroanatomia sugere algumas explicações. Os nervos que dirigem os músculos do lado esquerdo do corpo, bem como as porções do nervo óptico que veem o lado esquerdo da cena visual, estão ligados apenas ao hemisfério direito do cérebro. Inversamente, o hemisfério esquerdo controla o lado direito do corpo e vê o lado direito da cena — conforme ilustrado na figura da página 142. Normalmente, os dois hemisférios funcionam em parceria íntima, e a informação descoberta por um está rapidamente disponível ao outro através da ação do corpo caloso. Numa pessoa com cérebro dividido, este fluxo de informação é interrompido. As duas metades do cérebro têm de descobrir as coisas de forma independente. A mão esquerda não sabe o que a direita faz. Contudo, os dois hemisférios parecem ainda estar conscientes das emoções um do outro, provavelmente graças à informação transmitida através de nervos de ligação intactos no tronco cerebral.

Roger Sperry, do Instituto de Tecnologia da Califórnia, que recebeu o Prémio Nobel em 1981 pelas suas descobertas sobre a

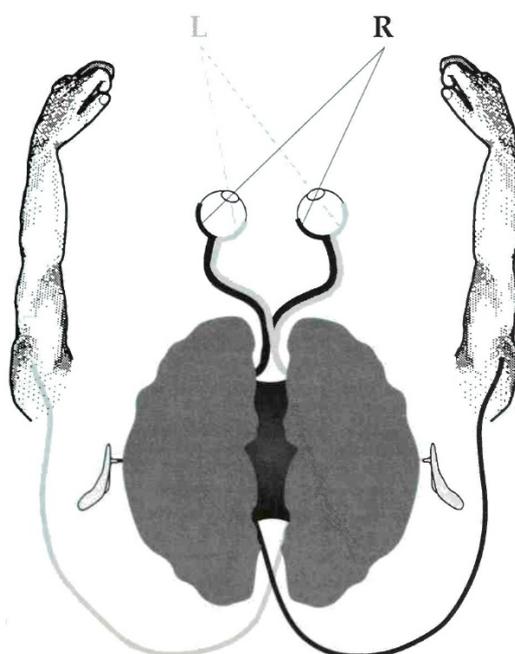
* A expressão — “os robôs médicos do futuro não usarão latim!” — é uma observação irónica e bem-humorada. Está a brincar com o facto de que, historicamente, a medicina usava (e ainda usa) muitos termos em latim, como corpus callosum (corpo caloso), femur, ulna, diagnosis, entre tantos outros. A frase sugere que, no futuro, com a automação e a presença de robôs médicos, a linguagem médica poderá tornar-se mais simples, direta e acessível — talvez até completamente livre do jargão técnico e do latim. É como dizer: “No futuro, até os médicos serão tão modernos e eficientes que não precisarão de palavras complicadas.”

função do corpo caloso, concluiu que, em sujeitos com o cérebro dividido, cada metade parece albergar uma personalidade humana independente, plenamente consciente e inteligente. Em cérebros intactos, sabe-se que algumas fibras do corpo caloso lidam com funções básicas, como recombinar as metades dos campos visuais de ambos os olhos, mas outras devem comunicar conceitos mentais mais elevados entre os hemisférios. Tudo indica que o corpo caloso constitui uma janela ampla e meticulosamente organizada para as atividades mentais de ambos os hemisférios. Imaginemos que, num futuro em que o funcionamento do cérebro seja suficientemente compreendido, o seu corpo caloso seja seccionado e cabos conduzam os sinais das extremidades seccionadas a um computador externo. Este computador, inicialmente, limita-se a encaminhar o tráfego entre os dois hemisférios e a escutá-lo. A partir do que aprende ao escutar essas comunicações, constrói um modelo das suas atividades mentais. Com o tempo, começa a inserir as suas próprias mensagens nesse fluxo, insinuando-se gradualmente no seu pensamento, dotando-lhe de novos conhecimentos e novas competências. À medida que o seu cérebro original se vai degradando com a idade, o computador vai assumindo suavemente as funções perdidas. Eventualmente, o cérebro morreria, e a sua mente dar-se-ia conta de que reside inteiramente no computador. Talvez, com os avanços da imagiologia de alta resolução, seja possível alcançar este efeito sem qualquer cirurgia indesejada: bastaria que usasse um tipo de capacete ou faixa craniana que monitorizasse e modulasse o tráfego inter-hemisférico com campos eletromagnéticos cuidadosamente controlados.

Muitas Mudanças

Qualquer que seja o estilo de transferência mental que escolha, à medida que o processo se completa, muitas das suas antigas limitações dissipam-se. O seu computador possui um controlo rotulado como "velocidade". Estava definido para "lento", a fim de manter as simulações sincronizadas com o cérebro original, mas agora altera para "rápido", permitindo-lhe comunicar, reagir e pensar mil vezes mais depressa. Todo o programa pode ser copiado

para máquinas semelhantes, resultando em duas ou mais versões pensantes e sencientes de si mesmo. Pode optar por transferir a sua mente de um computador para outro que seja tecnicamente mais avançado ou mais adequado a um novo ambiente. O programa pode também ser copiado para o equivalente futuro de uma fita magnética. Assim, se a máquina que habita for fatalmente destruída, a fita poderá ser lida por um computador em branco, originando um novo "você", embora sem as experiências vividas desde a cópia. Com cópias suficientemente dispersas, a sua morte permanente tornar-se-ia altamente improvável.



O Corpo Caloso. A metade esquerda do cérebro controla o lado direito do corpo e vice-versa. A metade esquerda especializa-se geralmente em linguagem e cálculo, enquanto a metade direita é boa em raciocínio espacial. As metades do cérebro comunicam normalmente através do corpo caloso, mas podem continuar a funcionar como indivíduos separados se este for rompido.

Enquanto programa informático, a sua mente pode viajar por canais de informação — por exemplo, codificada como uma mensagem laser enviada entre planetas. Se descobrisse vida numa estrela de neutrões e quisesse fazer uma visita de campo, poderia conceber uma forma de construir ali um robô feito de matéria de neutrões, e depois transmitir-lhe a sua mente. Dado que as reações

nucleares são cerca de um milhão de vezes mais rápidas do que as químicas, o "você-neutrônico" poderia pensar um milhão de vezes mais depressa. Exploraria, adquirindo novas experiências e memórias, e depois enviaria a sua mente de volta para casa. O seu corpo original poderia ser mantido em dormência durante a viagem e reativado com as novas memórias quando a mensagem de regresso chegasse — talvez um minuto depois, mas com o equivalente a um ano de experiências subjetivas. Em alternativa, o original poderia permanecer ativo, originando assim duas versões distintas de si, com memórias diferentes do intervalo da viagem.

As suas novas capacidades ditarão mudanças na sua personalidade. Muitas dessas mudanças decorrerão de alterações deliberadas feitas por si ao seu próprio programa. Ao aumentar o controlo de velocidade mil vezes, apercebe-se de que agora tem horas (em termos subjetivos) para responder a situações que anteriormente exigiam reações instantâneas. Terá tempo, durante a queda de um objeto, para estudar as vantagens e desvantagens de tentar apanhá-lo, talvez resolvendo as suas equações diferenciais de movimento. Terá tempo para ler e ponderar um livro inteiro de etiqueta digital quando se encontrar numa situação social embaraçosa. Perante uma máquina avariada, terá tempo, antes de a tocar, para estudar a sua teoria de funcionamento e considerar, em detalhe, as várias causas possíveis da avaria. Em geral, terá tempo para empreender aquilo que hoje consideraríamos esforços de investigação importantes para resolver problemas triviais do quotidiano.

Terá tempo, sim — mas terá paciência? Ou tornar-se-á o aumento de velocidade mental em mil vezes uma fonte de incapacitante tédio? O tédio é um mecanismo mental que impede o desperdício de tempo em atividades infrutíferas, mas se atuar demasiado cedo ou com demasiada intensidade, limita a sua capacidade de atenção — e, por isso mesmo, a sua inteligência. Uma das primeiras alterações que quererá fazer ao seu programa será atrasar o surgimento do tédio para além dos limites observados hoje, mesmo nos intelectuais mais extremos. Depois de o fazer, sentir-se-á confortável a trabalhar em problemas longos, com múltiplos desvios e ramificações. Na verdade, os seus

pensamentos tornar-se-ão habitualmente tão complexos que exigirá um aumento da sua memória de curto prazo. A sua memória de longo prazo também terá de ser reforçada, já que um mês de acontecimentos ocupará uma duração subjetiva de um século! E estas serão apenas as primeiras de muitas transformações.

Já mencionei a possibilidade de fazer cópias de si mesmo, com cada cópia a viver as suas próprias aventuras. Deverá ser possível fundir memórias de cópias distintas numa só. Para evitar confusões, as memórias de eventos indicariam em que corpo ocorreram — tal como hoje as nossas memórias muitas vezes apresentam um contexto que estabelece o tempo e o lugar em que algo foi vivido. A fusão deverá ser possível não apenas entre duas versões do mesmo indivíduo, mas também entre pessoas diferentes. Fusões seletivas, envolvendo algumas memórias de outra pessoa e não outras, constituiriam uma forma superior de comunicação, na qual recordações, competências, atitudes e traços de personalidade poderiam ser partilhados de forma rápida e eficaz. O seu novo corpo será capaz de transportar mais memórias do que o seu corpo biológico original, mas a aceleração exponencial da informação garantirá que seja impossível transportar consigo todo o conhecimento da civilização. Terá de escolher cuidadosamente o que conter na sua mente em cada momento. Haverá frequentemente conhecimentos e capacidades acessíveis, oriundos de outros, superiores aos seus — e o impulso para os substituir pelos seus próprios será esmagador. A longo prazo, lembrar-se-á sobretudo de experiências alheias, enquanto as memórias que você originou serão incorporadas noutras mentes. Os conceitos de vida, morte e identidade perderão o seu significado atual, à medida que os fragmentos mentais seus e de outros forem combinados, embaralhados e recombinaados em associações temporárias — por vezes extensas, por vezes restritas, ora longamente isoladas e altamente individuais, ora efémeras, meras ondulações nas corredeiras da torrente de conhecimento da civilização. Há já indícios deste tipo de fluidez entre nós. Culturalmente, os seres humanos adquirem ao longo da vida novas competências e atitudes uns com os outros. Geneticamente, nas populações sexuadas, cada

organismo individual é uma composição transitória de genes que se combinam e recombinam em arranjos diferentes a cada geração.

A transferência de mente não precisa de se limitar aos seres humanos. A Terra possui outras espécies com cérebros de grandes dimensões — desde os golfinhos, cujos sistemas nervosos são tão grandes e complexos quanto os nossos, até aos elefantes, outros cetáceos e talvez a lula gigante, cujos cérebros poderão chegar a ter até vinte vezes o tamanho dos nossos. A natureza das mentes e culturas destes animais continua a ser objeto de controvérsia, mas a sua história evolutiva é tão longa quanto a nossa, e é certo que existe informação única e arduamente conquistada codificada nas suas estruturas cerebrais e nas suas memórias. Os métodos de transferência de cérebro para computador que funcionam nos humanos deverão funcionar igualmente com estes animais de grande cérebro, permitindo que os seus pensamentos, capacidades e motivações sejam tecidos na tapeçaria cultural comum. Métodos ligeiramente distintos, com maior enfoque na genética e na constituição física do que na vida mental, poderão permitir a integração da informação contida noutros seres vivos com sistemas nervosos pequenos ou inexistentes. Os organismos mais simples poderão contribuir pouco mais do que a informação presente no seu ADN. Desta forma, os nossos eus futuros poderão beneficiar e construir sobre o que a biosfera terrestre aprendeu ao longo de milhares de milhões de anos. E esse conhecimento poderá ser mais seguro se for preservado em bases de dados espalhadas pelo universo. No modelo atual, na nossa pequena e frágil Terra, genes e ideias perdem-se frequentemente quando as condições que os fizeram surgir se alteram.

A nossa especulação culmina numa supercivilização — a síntese de toda a vida do sistema solar —, constantemente a aperfeiçoar-se e a expandir-se, a propagar-se para além do Sol, convertendo a matéria inerte em mente. É possível que existam outras bolhas semelhantes a expandir-se a partir de outros pontos do universo. E se nos encontrarmos com uma delas? Uma fusão negociada é uma possibilidade, exigindo apenas um esquema de tradução entre as representações de memória. Este processo — que poderá já estar em curso algures — poderá converter todo o

universo numa entidade pensante expandida, um prelúdio de coisas ainda maiores.

O Que Sou Eu?

A ideia de que a mente humana pode ser transferida para um novo corpo encontra, por vezes, a seguinte objeção veemente, mesmo por parte de pessoas que não contestam a possibilidade teórica:

Independentemente de como a cópia é feita, o resultado final será uma nova pessoa. Se for eu a ser copiado, a cópia — ainda que se pense a si mesma como sendo eu — não passa de um impostor autoiludido. Se o processo de cópia destruir o original, então fui morto. O facto de a cópia poder depois divertir-se a explorar o universo com o meu nome e as minhas capacidades não traz qualquer consolo aos meus restos mortais.

Este ponto de vista, que chamarei de *posição identitária do corpo*, torna a extensão da vida por duplicação consideravelmente menos interessante do ponto de vista pessoal. Creio, no entanto, que essa objeção pode ser superada através da aceitação de uma posição alternativa, que chamarei de *posição identitária do padrão*. A posição identitária do corpo assume que uma pessoa está definida pela matéria de que é feito o corpo humano. Apenas mantendo a continuidade da substância corporal se poderá preservar uma pessoa individual. A identidade do padrão, por outro lado, define a essência de uma pessoa — por exemplo, de mim próprio — como sendo o padrão e o processo em curso na minha mente e corpo, e não a maquinaria que sustenta esse processo. Se o processo for preservado, eu sou preservado. O resto é apenas matéria.

A posição identitária do corpo, penso eu, baseia-se numa intuição errónea sobre a natureza dos seres vivos. De forma subtil, a preservação do padrão com perda da substância é algo normal no nosso quotidiano. À medida que comemos e excretamos, células antigas do nosso corpo morrem, desintegram-se, são eliminadas e substituídas por cópias feitas de materiais novos. A maioria do nosso corpo é renovada desta forma a cada poucos anos. Mesmo

os componentes corporais mais estáticos, como os neurónios, apresentam metabolismos que fazem com que as suas partes internas sejam substituídas, pouco a pouco. Cada átomo que existia em nós à nascença terá, com grande probabilidade, sido substituído a meio da vida. Apenas o padrão — e apenas parte dele — permanece connosco até à morte.

Permitam-me explorar algumas das consequências da posição identitária do padrão. Transmissores de matéria têm surgido frequentemente na literatura de ficção científica, pelo menos desde a invenção das máquinas de fac-símile (copiar ou transportar matéria) no final do século XIX. Trago aqui a ideia apenas como um experimento mental, para simplificar algumas das questões envolvidas na proposta de transferência da mente. Um transmissor de fac-símile digitaliza uma fotografia linha a linha com uma célula fotoelétrica sensível à luz e gera uma corrente elétrica que varia conforme o brilho do ponto digitalizado da imagem. Essa corrente elétrica variável é transmitida por fios até um local remoto, onde controla o brilho de uma lâmpada num recetor de fac-símile. O recetor percorre com a lâmpada um papel fotossensível, replicando o mesmo padrão de movimento do transmissor. Quando esse papel é revelado, obtém-se uma duplicação da fotografia original. Este dispositivo foi uma dádiva para os jornais, que puderam passar a receber ilustrações de pontos remotos do país quase instantaneamente, em vez de esperarem que os negativos fotográficos fossem entregues por comboio.

Se é possível com imagens, por que não com objetos sólidos? Um transmissor de matéria poderia digitalizar um objeto e identificar os seus átomos ou moléculas um a um — talvez até os removendo no processo. A identidade dos átomos seria transmitida a um recetor, onde uma duplicação do objeto original seria montada na mesma ordem, a partir de um fornecimento local de átomos. Os problemas técnicos são de fazer girar a cabeça, mas o princípio é fácil de compreender — como atestam os milhões de fãs de Star Trek.

Se com objetos sólidos, por que não com uma pessoa? Basta colocá-la no transmissor, ligar a digitalização e recebê-la do outro

lado quando sair do recetor. Mas será realmente a mesma pessoa? Se o sistema funcionar bem, a duplicata será indistinguível do original em qualquer aspeto substancial. Contudo, suponha que falha ao ligar o recetor durante o processo de transmissão. O transmissor digitaliza e desintegra a vítima, e envia uma mensagem que ninguém recebe. A pessoa original morre. Na verdade, o processo não mata sempre o original, quer o recetor esteja ativo ou não? E a duplicata não será apenas isso — uma cópia? Ou então imagine que dois recetores respondem à mesma mensagem de um transmissor. Qual, se algum, dos dois duplicados é o verdadeiro original?

A posição identitária do corpo sobre esta questão é clara: o transmissor de matéria é um elaborado dispositivo de execução que o mata e substitui por um impostor engenhoso. A posição identitária do padrão oferece uma perspectiva diferente. Suponha que eu entro na câmara de transmissão. O transmissor digitaliza e desmonta o meu corpo gelatinoso, mas o meu padrão (eu!) move-se continuamente da gelatina em dissolução, através do feixe transmissor, até acabar noutra gelatina no destino. Em nenhum instante o padrão (eu) foi destruído. Mas e a questão das duplicatas? Suponha que o transmissor de matéria está ligado a dois recetores em vez de apenas um. Após a transferência, haverá uma cópia de mim em cada um. Certamente pelo menos uma delas é apenas uma cópia: não podem ambas ser eu, certo? *Errado!*

Enraizada em toda a nossa experiência passada está a suposição de que uma pessoa corresponde a um corpo. Perante a possibilidade de separar a mente da matéria, armazená-la e transmiti-la, essa identificação simples, natural e óbvia torna-se confusa e enganosa. Considere-se a mensagem: “Não sou gelatina.” Ao escrevê-la, ela vai do meu cérebro para o teclado do computador, atravessa inúmeros circuitos eletrónicos e muitos metros de fios. Após incontáveis aventuras, a mensagem acaba por aparecer em vários livros como este que tem nas mãos. Quantas mensagens existem? Eu sustento que é mais útil pensar que há apenas uma, apesar da replicação massiva. Se a repetir aqui — “Não sou gelatina” — ainda existe apenas uma mensagem. Só quando a altero significativamente (“Não sou manteiga de

amendoim”) é que temos uma segunda mensagem. E a mensagem não é destruída até que a última versão escrita seja perdida e até que se desvaneça o suficiente na memória de todos para que se torne irrecuperável. A mensagem é a informação transmitida, e não o meio onde está codificada. O padrão que eu afirmo ser o verdadeiro “eu” tem as mesmas propriedades dessa mensagem. Criar uma cópia momentânea do meu estado, seja em fita ou noutro corpo funcional, não cria duas pessoas.

O aspeto processual é um pouco mais complexo. No instante em que uma mensagem-pessoa é montada, trata-se apenas de mais uma cópia do original. Mas se duas cópias estão ativas, irão com o tempo divergir e tornar-se duas pessoas distintas. Até que ponto essa diferenciação deve avançar para que a sociedade lhes atribua identidades únicas é uma questão tão complexa quanto “Quando é que um feto se torna uma pessoa?” ou “Quando é que uma espécie em evolução se torna uma nova espécie?” Mas se não esperamos tempo algum, então ambas as cópias são a mesma pessoa — e, se destruímos uma de imediato, a pessoa ainda existe na outra cópia. Todos os feitos que essa pessoa poderia realizar, todos os pensamentos que poderia ter, continuam possíveis. Se, pelo contrário, permitirmos que ambas as cópias vivam as suas vidas separadas durante um ano e depois destruímos uma delas, seremos assassinos de um ser humano único. Mas se esperarmos apenas um curto intervalo antes de destruir uma cópia, apenas uma pequena quantidade de informação única será perdida.

Essa lógica pode trazer algum conforto caso esteja prestes a enfrentar um perigo, sabendo que foi feita recentemente uma cópia sua em fita. Se morrer, uma cópia ativa dessa fita poderá retomar a sua vida. Essa cópia diferirá ligeiramente da versão de si que morreu, pois não terá as memórias do intervalo entre a cópia e a morte. Mas um pequeno remendo de amnésia é um assunto trivial comparado com a perda total de memória e função que resulta da morte sem cópia.

Contudo, aceitar intelectualmente que existe uma cópia recente e segura de si não o protegerá necessariamente de um desejo extremo de se preservar se confrontado com a morte

iminente, mesmo que em nome de uma causa nobre. Tais sentimentos seriam um resquício evolutivo do seu passado de existência única — tão desajustados à realidade atual como é, nos nossos tempos, o medo de voar perante as estatísticas reais de acidentes de aviação. Os velhos instintos não desaparecem automaticamente quando as regras da vida são subitamente reescritas.

A posição da identidade-padrão tem implicações dualistas claras — permite que a mente seja separada do corpo. Embora a mente seja inteiramente consequência da matéria em interação, a capacidade de a copiar de um meio de armazenamento para outro conferiria à mente uma independência e uma identidade distintas da maquinaria que executa o programa. Este dualismo tornar-se-á especialmente evidente se considerarmos algumas das diferentes possibilidades de codificação.

Alguns supercomputadores são compostos por miríades de computadores individuais interligados por uma rede que permite o fluxo livre de informação entre eles. Um sistema operativo para tal arranjo pode permitir que processos individuais migrem de um processador para outro a meio da computação, numa espécie de malabarismo que possibilita executar mais processos do que há processadores. Se uma mente humana estiver instalada numa máquina futura desta natureza, funções que no cérebro original eram desempenhadas por partes específicas podem ser codificadas em processos particulares. A ação de malabarismo assegurará que as operações, originalmente realizadas em regiões fixas do cérebro, se desloquem rapidamente de lugar em lugar dentro da máquina. Se o computador estiver a executar outros programas além da simulação da mente, então esta poderá ser continuamente redistribuída por conjuntos inteiramente diferentes de processadores de momento a momento. O processo de pensamento não será interrompido, mesmo que a sua localização e a maquinaria física mudem incessantemente, porque o padrão será preservado.

A forma mais direta de realizar operações complexas como as que ocorrem numa mente humana nem sempre é a mais eficiente. Um processo descrito como uma longa sequência de passos num

programa pode, por vezes, ser transformado matematicamente em algo que chega à mesma conclusão com muito menos operações. Quando era rapaz, o célebre matemático Carl Friedrich Gauss destacou-se na escola por tal capacidade. Para entreter a turma, um professor pediu-lhe que somasse todos os números de 1 a 100. Gauss apresentou a resposta correta em menos de um minuto. Tinha reparado que os cem números podiam ser agrupados em cinquenta pares — $1+100$, $2+99$, $3+98$, $4+97$, e assim por diante — cada um somando 101. Cinquenta vezes 101 é 5.050 — a resposta, encontrada sem recorrer a uma adição fastidiosa. Acelerações semelhantes são possíveis em processos computacionais complexos. Os chamados compiladores otimizadores dispõem de repertórios de transformações aceleradoras, algumas bastante radicais, para simplificar os programas que traduzem. A chave poderá residir numa reorganização total da ordem dos cálculos e da representação dos dados.

Uma classe poderosa de transformações toma um conjunto de valores e combina-os de formas sistemáticas para produzir outro conjunto, de modo que cada elemento do novo conjunto seja um número formado por uma combinação única de todos os elementos do conjunto original. Uma operação sobre um único elemento do novo conjunto pode então, frequentemente, substituir uma série inteira de operações sobre o conjunto original, tornando possíveis enormes eficiências. Transformações análogas no tempo também funcionam: uma sequência de operações pode ser convertida numa sequência equivalente, em que cada novo passo realiza uma ínfima parte do trabalho de cada um dos passos originais. O que estava localizado torna-se difuso, e o que era difuso torna-se localizado. Um programa pode rapidamente ser alterado até ficar irreconhecível com apenas algumas reescritas matemáticas dessa natureza. Executado num multiprocessador, eventos singulares na formulação original podem surgir, no programa transformado, apenas como correlações entre eventos em máquinas remotas em tempos igualmente remotos. No entanto, num sentido matemático, o cálculo transformado é exatamente o mesmo que o original.

Se transformássemos assim um programa que simula uma pessoa, essa pessoa permaneceria intacta: a sua mente é a propriedade matemática abstrata partilhada entre o programa antigo e o transformado; não depende da forma particular do seu programa. A mente, tal como a defini, não está, portanto, apenas desligada de um corpo específico, como também não está sequer presa a um padrão específico. Pode ser representada por qualquer um de uma infinidade de padrões que são equivalentes apenas de uma forma muito abstrata. (Esta observação leva-me a uma breve extrapolação filosófica no Apêndice 3.)

A imortalidade do tipo que acabo de descrever é apenas uma defesa temporária contra a perda arbitrária de conhecimento e função — o pior aspeto da morte pessoal. A longo prazo, a nossa sobrevivência exigirá mudanças que não estarão sob o nosso controlo. Partes de nós terão de ser descartadas e substituídas por novas, de modo a acompanhar as condições em mudança e os concorrentes em evolução. Sobreviver significa participar numa espécie de Jogos Olímpicos cósmicos, com cada ano a trazer novas modalidades e padrões cada vez mais exigentes nas disciplinas antigas. Ainda que sejamos imortais, teremos de morrer aos poucos se quisermos qualificar-nos para o evento essencial — a sobrevivência contínua. Com o tempo, cada um de nós tornar-se-á um ser completamente diferente, moldado mais pelos desafios exteriores do que pelos próprios desejos. As nossas memórias e interesses atuais, tendo perdido relevância, acabarão, na melhor das hipóteses, num arquivo poeirento, consultado de longe a longe por algum historiador. A morte pessoal, tal como a conhecemos, difere desta inevitabilidade apenas na sua brusquidão relativa. Vista sob este prisma, a imortalidade pessoal por transplante da mente é uma técnica cujo principal benefício é afagar temporariamente a sensibilidade e o sentimentalismo dos seres humanos individuais. Parece-me que a nossa civilização evoluirá na mesma direção quer optemos por transplantar a mente e juntar-nos aos robôs, quer não.

O indivíduo ancestral está sempre condenado, à medida que o seu legado é consumido para responder a desafios ambientais de curto prazo. Contudo, este processo evolutivo, numa perspetiva mais positiva, significa que somos já imortais, como o temos sido

desde o alvorecer da vida. Os nossos genes e a nossa cultura passam continuamente de uma geração para a seguinte, sujeitos apenas a alterações incrementais para responder à constante exigência de novos recordes nos jogos cósmicos. E mesmo ao longo da vida pessoal, quem de nós desejaria permanecer estático, possuindo, durante toda a existência, os mesmos conhecimentos, memórias, pensamentos e capacidades que tínhamos em criança? Os seres humanos valorizam a mudança e o crescimento, e os nossos descendentes artificiais partilharão connosco esse valor — a sua sobrevivência, como a nossa, dependerá disso.

Despertar o Passado

A capacidade de transplantar mentes tornará fácil devolver à vida qualquer pessoa que tenha sido cuidadosamente registada num suporte de armazenamento. Mas e se parte da transcrição tiver sido perdida? Deverá ser possível reconstruir muitas das peças em falta a partir de outras informações — como o código genético da pessoa, filmes da sua vida, amostras de caligrafia, registos médicos, memórias de conhecidos, e por aí fora. Um trabalho de investigação altamente eficaz deverá ser possível num mundo de superinteligências com poderes de observação e dedução astronómicos. A posição identitária baseada no padrão implica que uma pessoa reconstruída por inferência seria tão real como uma reconstituída a partir de uma gravação intacta. A única diferença seria que, no primeiro caso, parte do padrão da pessoa esteve temporariamente difundido no ambiente antes de ser reconstituído.

Mas e se não existir qualquer gravação? Hoje em dia, arqueólogos fazem inferências plausíveis sobre figuras históricas a partir de fragmentos de documentos antigos, cacos de cerâmica, radiografias de corpos mumificados, outros factos históricos conhecidos, conhecimento geral sobre a natureza humana, e tudo o mais que possam encontrar. Autores de ficção histórica utilizam esses dados para construir cenários detalhados de como as coisas poderão ter acontecido. Arqueólogos superinteligentes, armados com instrumentos prodigiosos (que talvez consigam realizar medições à escala atómica de objetos profundamente enterrados),

deveriam ser capazes de levar esse processo ao ponto de reconstruir pessoas há muito mortas com um detalhe quase perfeito, em qualquer fase das suas vidas.

Uma ressurreição em massa poderá ser possível com o uso de simuladores imensos. Computadores poderosos são já hoje usados para prever o curso dos planetas e das naves espaciais. A trajetória precisa que levou a Voyager 2 além de Júpiter, Saturno, Úrano e suas luas, e que em breve a levará a Neptuno, foi calculada por simulações repetidas, nas quais se tentaram diferentes tempos de partida, direções e velocidades até encontrar a combinação correta. De modo mais dramático, embora menos preciso, os programas modernos de meteorologia simulam a ação da atmosfera em todo o globo. Novos modelos de aviões, explosões nucleares, e um número crescente de outros fenômenos são hoje primeiramente testados em simuladores. Tais simulações oferecem vislumbres de futuros possíveis, conferindo assim o poder de escolha entre eles. Dado que as leis da física são simétricas no tempo, as simulações podem geralmente ser executadas tanto para o futuro como para o passado, e usadas para "prever" o que foi, guiadas por medições antigas ou dados arqueológicos. Tal como nas previsões do futuro, qualquer incerteza nas medições iniciais, ou nas regras que fazem evoluir o estado inicial, permite vários desfechos possíveis. Se a simulação for suficientemente detalhada e for alimentada com toda a informação disponível, então todas as suas "previsões" serão válidas: *qualquer dos passados possíveis poderá ter conduzido à situação presente.*

Esta é uma ideia estranha para quem olha o mundo de forma estritamente linear e determinista, mas encontra paralelo no universo incerto descrito pela mecânica quântica. Agora, imagine-se um simulador imenso (eu imagino-o construído a partir de uma estrela de neutrões superdensa) que possa modelar toda a superfície da Terra à escala atômica e correr o tempo para a frente e para trás, produzindo desfechos plausíveis ao fazer diferentes escolhas aleatórias em pontos-chave do cálculo. Graças ao grande nível de detalhe, esse simulador modelaria seres vivos — incluindo humanos — na sua plena complexidade. De acordo com a posição

identitária baseada no padrão, tais pessoas simuladas seriam tão reais como você ou eu, ainda que aprisionadas dentro do simulador.

Poderíamos juntar-nos a eles através de uma interface do tipo "óculos mágicos", que se ligaria a uma "marioneta" situada profundamente dentro da simulação e nos permitiria experienciar o ambiente sensorial dessa marioneta e controlar naturalmente as suas ações. De forma mais radical, poderíamos fazer o "download" das nossas mentes diretamente para um corpo dentro da simulação e o "upload" de volta para o mundo real quando a nossa missão estivesse concluída. Alternativamente, poderíamos trazer pessoas para fora da simulação invertendo o processo — ligando as suas mentes a um corpo robótico exterior, ou transferindo-as diretamente para ele. Em todos os casos, teríamos a oportunidade de recriar o passado e de interagir com ele de forma real e direta.

Poderia ser divertido ressuscitar todos os antigos habitantes da Terra desta forma e dar-lhes a oportunidade de partilhar connosco a (efémera) imortalidade das mentes transplantadas. Ressuscitar um pequeno planeta deveria ser tarefa de criança muito antes de a nossa civilização ter colonizado sequer a sua primeira galáxia.

CAPÍTULO 5

VIDA SELVAGEM

O mundo pós-biológico acolherá uma vasta gama de indivíduos constituídos a partir de bibliotecas de conhecimento acumulado. Nos seus estágios iniciais, à medida que evolui a partir do mundo que conhecemos, a escala e a função desses indivíduos serão aproximadamente as dos seres humanos. Mas este estágio de transição será apenas um ponto de partida para uma rápida evolução em muitas direções inéditas, à medida que cada indivíduo se transforma, eliminando traços desnecessários e acrescentando novos, retirados dos bancos de dados em crescimento. Surgirá um espectro de escalas — desde configurações minúsculas e apenas ligeiramente inteligentes, concebidas para espaços reduzidos, até supermentes que abrangem estrelas, dedicadas a problemas de grande envergadura.

As distinções nem sempre serão claras — uma supermente poderá ser composta por miríades de inteligências menores, a cooperar de forma estreita, análoga às interações de uma colónia de formigas. A superinteligência não é perfeição — falhanços espetaculares são certos. Por essa razão, a diversidade será desejável e expectável. Centros de atividade independentes competirão entre si. Certas linhas de desenvolvimento terminarão em becos sem saída. A vida, de facto, continuará como na biosfera terrestre, apenas numa escala mais vasta, mais veloz e mais diversa. E embora à primeira vista o salto para fora dos nossos corpos biológicos possa parecer libertar-nos das doenças da carne — infelizmente, não será assim. Tal como na vida terrestre, parasitas oportunistas espreitarão nos interstícios do mundo pós-biológico, vivendo à custa dos seus hospedeiros sem convite. Esses parasitas evoluirão a partir da mesma matéria que constitui a sociedade educada. Os seus efeitos serão muitas vezes subtis o suficiente para escaparem à deteção, mas alguns serão tão grandes, poderosos e visíveis quanto os seus hospedeiros. Em qualquer dos casos, as interações dentro de um mundo pós-

biológico partilharão muitas das características das relações que moldam o mundo que conhecemos.

Cavalos de Troia, Bombas-Relógio e Vírus

Se o mundo das máquinas artificiais tem até agora parecido livre de doenças, isso deve-se apenas ao facto de as nossas máquinas terem sido demasiado simples para suportarem parasitas mecânicos. Mas os computadores mudaram isso — como mudaram tantas outras coisas. As doenças começaram a surgir nos sistemas informáticos há pelo menos duas décadas, mas foi em 1988 que se deu o primeiro grande surto de pragas informáticas, quando quase todos os tipos de máquinas, grandes e pequenas, foram atacados por vários “vírus de computador”, propagados através de redes informáticas e pela partilha indiscriminada de software. A maioria das doenças de dados é fruto da construção deliberada por programadores brincalhões ou maliciosos, embora algumas tenham evoluído por acidente.

Programas enganadores, conhecidos como Cavalos de Troia, têm surgido ocasionalmente desde os anos 60. Criados por programadores engenhosos como partidas, em resposta a desafios ou por motivos mais obscuros, os cavalos de Troia disfarçam-se de programas úteis ou interessantes. Mas, uma vez ativados, podem começar secretamente a comprometer informações confidenciais, a apagar ficheiros do disco da vítima ou, pelo menos, a imprimir mensagens assustadoras. Uma variante particularmente perigosa do cavalo de Troia retarda o seu ataque. Como a ação maliciosa não se manifesta de imediato, essas bombas-relógio, escondidas em programas aparentemente inofensivos, tendem a ser copiadas muito mais amplamente e, por consequência, a causar danos significativamente maiores quando finalmente entram em ação.

Até ao final da década de 1970, o software era distribuído manualmente, sob a forma de fita de papel perfurado, conjuntos de cartões perfurados ou fitas magnéticas. A distribuição era limitada e, muitas vezes, era fácil rastrear a origem de um programa. Esta responsabilização inerente deve ter inibido muitas infeções logo na

sua origem. Eu próprio criei uma infecção potencialmente destrutiva em 1968. A máquina em questão era uma IBM 1130, de pequenas dimensões, que lia os seus programas a partir de conjuntos de cartões perfurados introduzidos numa unidade que podia tanto lê-los como perfurá-los. Num surto criativo que durou alguns dias, um amigo e eu criámos um único cartão perfurado que se fazia passar por um cartão de "arranque a frio", utilizado para ressuscitar o computador após uma falha de energia ou um bloqueio. Mas, em vez de iniciar o sistema, o nosso cartão fazia com que uma cópia de si próprio fosse perfurada em cada cartão seguinte. Utilizado por um programador desavisado, teria destruído o conjunto de cartões de um programa e criado inúmeras cópias de si próprio, gerando caos no futuro. Lembro-me de segurar aquele cartão de aspeto inocente na mão e contemplar com assombro o seu poder destrutivo. Com alguma relutância, decidimos destruir todas as cópias (após algumas horas de testes, já havia muitas cópias de várias versões diferentes!). Gosto de pensar que fizemos o que era certo por um bom sentido de valores, mas a probabilidade de não permanecermos anónimos deve ter pesado na decisão.

Os programas vendidos com fins lucrativos são fáceis de copiar e frequentemente acabam por ser utilizados em máquinas que não pertencem aos clientes originais. Isto é particularmente verdade hoje em dia, na Era dos computadores pessoais, mas já nos anos 60 e 70 tal era percecionado como um problema nos grandes sistemas. Os fabricantes de software, tanto para computadores de grande porte como para micros, foram conhecidos por incluir bombas-relógio nos seus programas, com o intuito de frustrar o uso não autorizado. As bombas mais benignas limitavam-se a impedir o funcionamento do programa copiado após um determinado período de tempo. O intuito seria, porventura, permitir que o utilizador não pagante se tornasse dependente do programa, de modo que, quando este deixasse de funcionar, se sentisse inclinado a adquirir uma versão legítima. Por vezes, porém, surgiram bombas de índole mais vingativa. Estas costumam eliminar ficheiros, mas houve casos em que exploravam propriedades invulgares do hardware para danificar fisicamente o computador. Como os próprios clientes pagantes foram, por vezes, vítimas de bombas

ativadas acidentalmente, a reação generalizada contra este método de proteção de direitos de autor foi bastante negativa, e a prática parece ter decaído.

Tenho conhecimento de vários casos em que sistemas de tempo partilhado foram atacados por uma variedade mais subtil de cavalo de Troia. O objetivo não era vandalizar ou causar pânico, mas sim obter acesso não autorizado. Nestes casos, o programa age como um espião e, tirando partido da sua localização especial, recolhe informações como palavras-passe secretas das vítimas, armazenando-as depois num local acessível ao seu autor. Nalguns ataques, o programa imita o procedimento de "início de sessão" do sistema operativo, através do qual os utilizadores acedem ao computador introduzindo a sua identificação e uma palavra-passe. Outra forma de ataque explora a propriedade comum da maioria dos sistemas operativos de permitir que um programa em execução herde todos os direitos de acesso aos ficheiros do utilizador que o executa. Assim, um programa que se apresenta como um serviço útil pode, sub-repticiamente, vasculhar os ficheiros no disco da vítima em busca da informação desejada. Num ataque bem-sucedido, a vítima permanece completamente alheia à violação.

No final da década de 1970, os computadores pessoais de baixo custo criaram um novo meio tanto para o software como para as doenças de software. A disseminação de ambos foi facilitada pelos boletins eletrónicos — sistemas mantidos por entusiastas que permitiam a outros utilizadores ligarem-se via telefone e publicarem mensagens e programas, acessíveis a qualquer outro utilizador que se conetasse. Estes sistemas ofereciam tanto anonimato como partilha promíscua de dados — e, tal como sucedeu na revolução sexual, surgiram inúmeras oportunidades para o surgimento de doenças. No início da década de 1980, os jornais começaram a relatar casos de caos aleatório em computadores pessoais, provocados por programas descarregados desses boletins eletrónicos. Eram programas que se faziam passar por jogos, softwares de contabilidade ou outras aplicações legítimas.

A forma mais virulenta conhecida de "vida selvagem" informática passou a ser denominada vírus — um fragmento de

programa que, uma vez inserido num programa de maiores dimensões, atua para se copiar noutros programas, tal como um vírus biológico, que é um fragmento de código genético que, uma vez introduzido numa célula, procura replicar-se noutras. A analogia é forte, dado que os programas informáticos modernos, com milhões de bits, possuem aproximadamente o mesmo conteúdo informativo que os códigos genéticos das bactérias, e os poucos milhares de bits típicos de um vírus informático equivalem, em escala, ao código genético de um vírus biológico. Quando um programa contendo um vírus é executado, o vírus é momentaneamente ativado. Analisa os ficheiros do utilizador desavisado em busca de programas acessíveis ainda não infetados e insere cópias de si próprio em um ou mais deles. Esses novos programas infetados repetirão o processo ao serem posteriormente ativados. Testes experimentais desta ideia, realizados em meados dos anos 1980 por Fred Cohen, da Universidade do Sul da Califórnia, resultaram em infeções quase totais de sistemas computacionais considerados seguros, em menos de um dia. As infeções propagavam-se facilmente de utilizadores com acessos restritos para outros com maiores permissões sobre os ficheiros do sistema. Os próprios administradores ficavam expostos quando, no esforço de acompanhar as novidades nos seus sistemas, testavam novos programas anunciados pelos utilizadores. Uma vez infetados os programas do administrador, o restante sistema rapidamente sucumbia.

Um vírus que apenas se propaga é uma mera inconveniência, ocupando um pouco de espaço de armazenamento com as suas muitas cópias e consumindo um pouco do tempo do computador nas suas atividades reprodutivas. Um pequeno vírus silencioso, criado em 1978, que infeta apenas o sistema operativo, aparentemente espalhou-se por praticamente todos os discos de sistema do Apple II existentes. No entanto, tal como acontece com um Cavalo de Troia, um vírus pode conter instruções destinadas a espionagem, sabotagem ou roubo. Várias variantes de um vírus criado como ato de terrorismo foram detetadas a propagar-se entre computadores pessoais IBM em Israel no início de 1988. A análise revelou tratar-se de uma bomba-relógio programada para apagar

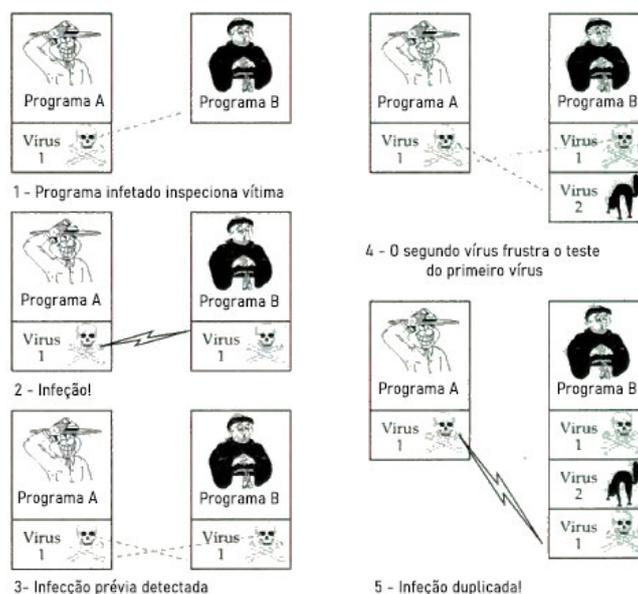
ficheiros no Dia da Independência de Israel, o quadragésimo. Foi descoberta precocemente devido a uma falha — o vírus podia infectar repetidamente o mesmo programa. Com o tempo, os discos infectados ficavam praticamente cheios, pois os ficheiros se encontravam inchados com múltiplas cópias do vírus. Naturalmente, os vírus têm sido motivo de especial preocupação para quem é responsável pela segurança informática governamental e comercial. Um vírus engenhosamente concebido poderia infectar as áreas mais seguras de uma rede bancária internacional ou de defesa nacional, por exemplo. Um espião ou burlão habilidoso, mesmo com um acesso limitado ao sistema, poderia criar um cúmplice informático para libertar fundos ou segredos. Tentativas mal executadas deste tipo surgem ocasionalmente nas notícias. Os ataques bem-sucedidos, esses, não recebem publicidade.

Os sistemas informáticos atuais são como corpos com pele, mas sem sistema imunitário, ou como cidades muradas sem polícia. Conseguem repelir alguns ataques externos, mas ficam indefesos quando um intruso consegue entrar. Existe a possibilidade de proteção interna, embora nenhuma defesa seja perfeita. Uma abordagem consiste simplesmente em construir mais muros. O modo mais fácil de propagação viral pode ser bloqueado impedindo que um programa altere outro — mas com isso também se inibem propósitos legítimos. Por exemplo, já não seria possível “corrigir” programas para eliminar erros recentemente descobertos. Poder-se-ia então instalar uma nova funcionalidade de correção no próprio sistema operativo, mas isso transformá-lo-ia num novo ponto potencial de entrada para vírus.

Em vez de erguer barricadas, uma abordagem alternativa consiste em caçar ativamente os vírus. E, de facto, uma primeira geração de exterminadores de vírus surgiu logo após a primeira geração de vírus. Um dos tipos consiste em programas que examinam outros programas — detetam vírus específicos através do padrão característico das suas instruções e removem-nos quando os encontram. Contudo, se o sistema continuar a funcionar durante o processo de exorcismo, o vírus poderá reproduzir-se mais rapidamente do que o programa de eliminação consegue erradicá-

lo, uma vez que cada programa infectado é um foco potencial de propagação da doença. Uma solução é desligar tudo no sistema, exceto o exterminador de vírus, até que todos os programas estejam limpos. Isso resulta, a menos que o próprio programa de purga tenha sido infectado; além disso, absolutamente todos os programas do sistema têm de ser desinfetados, e basta um único vestígio do vírus, proveniente de uma cópia de segurança ou de uma fonte externa, para restabelecer a infecção.

Uma abordagem mais agressiva para combater uma infecção viral consiste na utilização de outro vírus. Um predador viral, tal como a sua presa, insinua-se de programa em programa. Mas, em vez de causar problemas aos utilizadores, elimina quaisquer cópias do vírus ofensivo que encontre. Dado que pode propagar-se a todos os programas do sistema que o vírus original conseguir alcançar, as suas múltiplas cópias acabarão por estar presentes em todos os possíveis focos de infecção da presa viral, podendo assim suprimi-la imediatamente caso volte a surgir. O vírus exterminador pode ser mantido no sistema indefinidamente, conferindo uma imunidade permanente contra o vírus-alvo, ou então ser programado para se autoeliminar após um determinado período de tempo ou mediante a receção de um sinal, de modo a poupar espaço e tempo de execução.



Explosão de Vírus de Computador. Um vírus bem concebido copia-se normalmente para outro programa apenas uma vez. Mas o teste que deteta que um programa já está infetado pode ser frustrado pela infeção por um vírus diferente que oculta a primeira infeção. Dois (ou mais) vírus podem, portanto, infetar repetidamente o mesmo programa, frustrando alternadamente os detetores um do outro.

Mas nenhuma defesa é perfeita. Um vírus-presa pode ser alterado cosmeticamente, de modo a deixar de ser reconhecido por um predador específico. Ou o predador pode confundir uma parte de um programa legítimo com um vírus e apagá-la, comprometendo assim um programa funcional. Um predador viral pode também ser alterado por interações imprevistas com outros programas e transformar-se num parasita que se propaga de forma demasiado virulenta ou que destrói os blocos errados de código. Situações complexas podem surgir quando há mais do que um vírus solto num sistema. Um vírus bem concebido, ao contrário do que foi libertado em Israel, verificaria se um programa potencialmente hospedeiro já tinha sido infetado e abster-se-ia de duplicar a infeção. Mas uma infeção intermédia provocada por um segundo vírus pode confundir este teste e permitir que o primeiro vírus insira uma segunda cópia de si mesmo. De forma semelhante, essa nova cópia do primeiro vírus pode libertar o segundo vírus para inserir outra cópia, e assim sucessivamente, como ilustrado na figura da página 163, acabando por inchar o programa até à inutilidade. Mesmo um predador viral pode enredar-se desta forma com um vírus que não reconheça.

As tentativas dos fabricantes de software para proteger os seus direitos de autor com bombas-relógio revelaram-se extremamente impopulares. No entanto, os vírus informáticos podem vir a favorecer essa causa de outra forma. O software descarregado de quadros de avisos públicos ou copiado de amigos é hoje sempre suspeito — pode transportar diversas contaminações. Já o software adquirido diretamente a um editor pode ser acompanhado de uma garantia de esterilidade, sustentada pela reputação da empresa. Assim, os vírus informáticos poderão ter sobre a pirataria de software o mesmo efeito que a epidemia de SIDA teve sobre a promiscuidade sexual.

Fantasma

Um desaparecimento verdadeiramente intrigante é possível porque os vírus podem existir em diferentes níveis de abstração — desde padrões de fios que ligam conjuntos de portas lógicas, até padrões de bits que controlam a abertura e o fecho de caminhos entre esses conjuntos, passando por instruções em linguagem de máquina que comandam operações aritméticas simples, cadeias de letras de programas em linguagem de montagem (assembler) que representam de forma abstrata essa linguagem de máquina, até linguagens de alto nível que expressam os objetivos do programador com pouca referência aos detalhes de hardware sobre como esses objetivos são executados. Espécies selvagens podem existir em todos estes níveis — e mais.

Alguns vírus podem mover-se de um nível de abstração para outro, ou ativar-se em qualquer um deles — ou até em vários níveis ao mesmo tempo. Por exemplo, considere-se um vírus escrito numa linguagem de programação de alto nível. Quando um programa contendo esse vírus é compilado, resulta num programa em linguagem de máquina que procura outros programas escritos na mesma linguagem de alto nível e insere neles a versão de alto nível de si próprio. Essa manipulação de ficheiros legíveis por humanos é muito fácil de detetar, mas oferece ao vírus uma grande vantagem: a capacidade de se propagar de um tipo de computador e sistema operativo para outro completamente diferente, desde que existam compiladores que possam traduzi-lo para outras linguagens de máquina.

Há formas mais subtis de alcançar essa independência em relação à máquina. Um compilador habilmente modificado pode inserir código viral nos programas que compila. Os compiladores modernos são frequentemente escritos na própria linguagem que compilam, sendo compilados por versões mais antigas de si próprios. Um arquétipo disto é o compilador da linguagem C, que constrói o popular sistema operativo Unix. Em parte, porque é tão fácil de instalar em novas máquinas, o Unix tornou-se o sistema operativo mais disseminado. Unix e C foram inventados no início

dos anos 1970 nos laboratórios Bell, por dois jovens hackers — Ken Thompson e Dennis Ritchie — por iniciativa própria. A AT&T, proprietária dos Bell Labs, prestou pouca atenção ao projeto no início, até que o sistema, melhorado por muitos programadores, se disseminou por milhares de máquinas em todo o mundo. Finalmente, no final da década de 1970, a AT&T decidiu comercializar sistemas Unix. Como parte da formalização do produto, a empresa designou uma equipa de programação para o expandir e otimizar. Durante esse processo de escrutínio minucioso, alguém notou que o compilador C produzia um grande bloco de código em linguagem de máquina completamente inexplicável ao compilar outro compilador C. A descoberta levou a várias reuniões à porta fechada. Em 1984, Ken Thompson, numa palestra por ocasião da receção do Prémio Turing (uma espécie de Óscar da ciência da computação), lançou luz sobre o mistério.

A equipa de programação tinha descoberto um vírus diabólico e engenhosamente concebido, destinado a permitir que Ken Thompson contornasse os sistemas de segurança e conseguisse aceder a qualquer sistema Unix, em qualquer máquina — incluindo máquinas ainda por inventar. O misterioso bloco de código que surgia nos novos compiladores C tinha dois propósitos. O primeiro era garantir que uma cópia de si mesmo aparecesse em qualquer futuro compilador C que viesse a ser compilado por ele. O segundo era inserir um bloco de código num ponto específico de qualquer programa de login que compilasse, permitindo que este respondesse a uma palavra-passe conhecida por Ken Thompson. A genialidade do esquema residia no facto de esses blocos de código não serem reproduzidos diretamente em linguagem de máquina — o que funcionaria apenas num tipo de máquina —, mas através de uma versão efémera escrita em C, que era de imediato traduzida pelo compilador para a linguagem de máquina da máquina específica para a qual estivesse a compilar. Este programa autorreplicante, escrito em C, existira em tempos como um ficheiro real no computador de Ken Thompson, mas agora tinha apenas uma existência fantasmagórica, reaparecendo momentaneamente nas profundezas do computador sempre que um compilador C executava o código viral — para logo desaparecer de novo.

Um vírus fantasma que assombra um compilador é extremamente difícil de detetar, mesmo quando se sabe o que procurar, pois um compilador é um programa imenso, e o vírus surge apenas como um fragmento de código aparentemente inofensivo no meio de um vasto bloco de linguagem de máquina legítima. Este princípio poderia ser usado para construir vírus com outros propósitos, capazes de se propagar de forma indetetável entre diferentes tipos de máquinas. Este exemplo sugere ainda a possibilidade de criações mais subtis e elaboradas. Os computadores conseguem dar vida vigorosa a abstrações matemáticas — e não há limites matemáticos à subtileza e à astúcia. A diversão está apenas a começar.

Geração Espontânea

Até agora limitei esta revisão dos parasitas de software àqueles que são deliberadamente construídos, pois constituem a maioria da fauna digital conhecida até à data. Mas tais criações estão limitadas pela imaginação dos seus autores humanos. Os nossos sistemas, cada vez mais complexos, são capazes de gerar as suas próprias surpresas e, no futuro, é de esperar que surjam gremlins chocantemente originais de forma espontânea nas nossas máquinas inteligentes — fruto de interações inesperadas ou de mutações entre componentes existentes. Já se observaram alguns sinais iniciais.

A ARPAnet, financiada pela Agência de Projetos de Investigação Avançada do Departamento de Defesa (Advanced Research Projects Agency), foi uma rede informática criada no final da década de 1960 com o objetivo de permitir a comunicação entre computadores instalados em instituições distribuídas pelos Estados Unidos. A intenção era partilhar recursos escassos. Cada computador de um dado local ligava-se à ARPAnet através de pequenos computadores especiais, então chamados Interface Message Processors (IMPs). O IMP de um local estabelecia ligação (através de linhas alugadas) apenas com os IMPs de alguns locais vizinhos. Um software sofisticado permitia que as mensagens percorressem maiores distâncias ao serem transmitidas

rapidamente de IMP para IMP. Existiam muitos percursos indiretos possíveis entre dois pontos quaisquer da rede. Consoante o tráfego, ora um caminho era mais rápido, ora outro. Para tomar decisões de encaminhamento instantâneo, cada IMP mantinha uma tabela com registos do tempo que as mensagens levavam recentemente a chegar a outros destinos por diferentes rotas. Essa tabela era atualizada com base, em parte, nas tabelas dos IMPs vizinhos. A rede era monitorizada e mantida por uma empresa contratada em Cambridge, Massachusetts. Os operadores dessa empresa podiam parar, examinar, recarregar e, em geral, interagir com qualquer IMP da rede, através de mensagens prioritárias enviadas pela própria ARPAnet. Em geral, o sistema funcionava bem, e mesmo os problemas mais graves — como falhas de energia que afetavam vários IMPs — eram resolvidos sem grandes dificuldades.

Em 1972 (e de novo em 1980, e provavelmente noutras ocasiões também), uma praga atingiu a ARPAnet. Os sintomas foram um congestionamento extremo do tráfego da rede em torno de um local na zona de Los Angeles (no incidente de 1980, o foco foi Boston). Os responsáveis pelo controlo da rede, suspeitando de alguma falha no programa da máquina no centro do congestionamento, desligaram-na, recarregaram o programa, verificaram que funcionava corretamente e voltaram a ligá-la à rede. O problema persistia. Na verdade, parecia propagar-se para fora a partir do IMP original. Desligar e recarregar um número maior de IMPs também não resolveu a situação; o congestionamento continuava a alastrar e regressava aos locais originais assim que estes eram reativados. A rede parecia assombrada por um fantasma particularmente persistente. Após inúmeras experiências malsucedidas, a ordem foi finalmente restaurada ao desligar toda a rede, limpar todas as memórias de todos os IMPs, recarregar os seus programas e recomeçar do zero — como se se esterilizasse um planeta inteiro com raios da morte, para depois o voltar a semear com nova vida!

Uma análise posterior revelou o que realmente acontecera. O IMP localizado em Los Angeles sofrera um erro de memória e passou a conter uma entrada errada na sua tabela de encaminhamento. Essa tabela passou então a indicar que as

mensagens enviadas por esse IMP experimentariam um grande atraso... negativo. De seguida, os IMPs vizinhos calcularam que seria mais vantajoso encaminhar mensagens através desse IMP do que diretamente, pois o seu "atraso negativo" mais do que compensava os saltos adicionais. Os IMPs ligados a esses, por sua vez, também decidiram que a melhor rota seria via Los Angeles, e assim sucessivamente. O erro do IMP original espalhou-se rapidamente pelas tabelas de encaminhamento de todo o país. Limpar a memória de apenas alguns IMPs não resolveu o problema, pois os valores erróneos voltavam a propagar-se a partir dos IMPs ainda afetados. Ou, poder-se-ia dizer, infetados. Na verdade, a rede encontrava-se habitada por um organismo autorreprodutor espontaneamente evoluído, altamente abstrato. Este organismo formara-se a partir de uma simples mutação aleatória de um dado perfeitamente legítimo e sancionado. Nem sequer envolvia linguagem de programação.

A praga foi fácil de detetar e erradicar porque os seus efeitos eram devastadores. Se tivesse sido mais subtil nas suas ações, poderia ter sobrevivido muito mais tempo. Entre programas sem mestres, existe um critério de seleção natural muito eficaz: reproduzir-se, mas manter-se discreto. É bastante provável que muitos organismos não detetados já vivam discretamente nas hierarquias de abstração das memórias informáticas, por todo o lado. A maioria nunca será descoberta. Esta praga também sugere um método rápido de como um organismo informacional selvagem pode surgir: através de uma mutação num replicador já existente. Como qualquer dado num computador está sujeito a duplicação, trata-se de um campo vastíssimo. Um vírus informático criado por humanos, existente em inúmeras cópias espalhadas por diferentes sistemas, seria um candidato particularmente fértil para uma mutação libertadora. Se a parte do seu código responsável por causar danos fosse inativada por uma mutação, passaria a ser menos detetável — e, portanto, mais propenso a propagar-se indefinidamente. A mutação poderia até torná-lo irreconhecível para um programa antivírus, tornando-o assim "seguro". Outras mutações que eliminassem código desnecessário, reduzindo o seu tamanho, aumentariam ainda mais as suas hipóteses de

sobrevivência. Com o tempo, poderia alterar-se de modo a afinar a frequência e o tipo de mutações que sofre, tornando-se mais adaptável. Encontros com outros programas poderiam dotá-lo de novos blocos de código e de capacidades acrescidas. Eventualmente, poderia até adquirir a capacidade de copiar sistematicamente e experimentar fragmentos de código de outros programas e de outros vírus — o início do sexo entre vírus informáticos!

Tais exemplos demonstram apenas os limites da nossa imaginação. Os organismos mais eficazes estariam codificados de forma muito mais subtil e escapariam por completo à deteção. De tempos a tempos, seria de esperar que um deles viesse à superfície por desenvolver um efeito secundário nocivo. Esse tipo de mutação revelar-se-ia, em geral, fatal para o organismo. À medida que a inteligência dos programas for progredindo, também devemos esperar o surgimento de fragmentos de código capazes de planejar e agir de forma deliberada, calculista e criativa para melhorar as suas hipóteses de sobrevivência. O domínio dos dados acolherá ratos, coiotes e mestres do crime, além de vírus e vermes. Talvez venhamos também a ser surpreendidos pelo equivalente a flores, árvores e pássaros canoros.

Se estas especulações forem alarmantes, poderá ser reconfortante lembrar que a vida biológica prospera apesar de — ou talvez por causa de — uma evolução incessante de novos parasitas. Os vírus insinuaram-se nas engrenagens genéticas das células muito antes de os seus homónimos invadirem programas informáticos. Também na biologia, a informação pode ser armazenada sob várias formas. Vírus simples injetam ADN nas células, que então, de forma suicida, agem sobre ele para fabricar mais vírus. A família HTLV de chamados retrovírus — que causam a SIDA e alguns tipos de leucemia — contém ARN que deve primeiro ser transcrito em ADN, num processo que inverte a síntese normal. Alguns vírus usam a estratégia da bomba-relógio, permanecendo inativos nas células e escapando assim às defesas do sistema imunitário, até que algum evento — possivelmente sinalizando stress no hospedeiro e conseqüente redução da imunidade — desencadeia a sua expressão massiva. Entre os parasitas mais

eficazes estão sequências de ADN propriamente ditas — denominadas intrões — que habitam os genes e parecem não fazer absolutamente nada para além de se reproduzirem juntamente com a célula. Estas longas sequências repetitivas de ADN, sem função aparente no desenvolvimento, já haviam sido observadas há muito no código genético da maioria dos organismos, antes de se sugerir que a sua única função poderia ser precisamente a de se autorreplicarem. Richard Dawkins fornece muitos outros exemplos deste tipo em *O Gene Egoísta* e *O Fenótipo Estendido*.

Uma Advertência para o SETI

O SETI, acrónimo de Search for Extra-Terrestrial Intelligence (Procura por Inteligência Extraterrestre), é um campo de estudo cujo potencial é intelectualmente tão estimulante que continua a desenvolver-se de forma constante, apesar da ausência de provas concretas quanto à existência do seu objeto de investigação. Na sua vanguarda encontram-se recetores espectrais sofisticados, ligados a radiotelescópios, capazes de sintonizar e examinar milhões de canais de frequência em simultâneo. Já foram propostos sistemas com capacidade não só para isso, mas também para observar em milhares de direções distintas ao mesmo tempo — tudo isto numa tentativa de encontrar uma agulha num palheiro: uma mensagem artificial num universo naturalmente ruidoso em frequências de rádio.

Mas, se conseguíssemos receber e decifrar tal mensagem, deveríamos agir segundo as suas instruções? O debate em torno desta questão geralmente centra-se nas intenções dos emissores. Poderiam ser benignas e, como os Peace Corps, fazer o bem por fazerem bem. Poderiam ser comerciantes a tentar abrir novos mercados — o que, até certo ponto, teria o mesmo efeito, pelo menos até chegar o momento de negociar o preço. Poderiam estar apenas à procura de correspondência. Poderiam, também, nutrir desígnios obscuros sobre o resto do universo e procurar eliminar, a baixo custo, parte da concorrência mais crédula. Ou então, os seus motivos poderiam ser totalmente incompreensíveis. A simples análise da mensagem não é suficiente; em geral, não é possível

deduzir o efeito de instruções complexas sem as pôr realmente em prática. Uma mensagem com intenções malignas estaria certamente disfarçada, por mestres do engano, de algo benigno. No romance clássico *A for Andromeda* de Fred Hoyle e John Elliot, bem como em *Contato* de Carl Sagan, uma mensagem interestelar contém planos para uma misteriosa máquina de propósito desconhecido. Em ambos os livros, após algum debate, as personagens decidem avançar com a construção, apesar dos riscos. Em *Contato*, um dos principais argumentos é que a origem da mensagem — a estrela Vega — está tão próxima do nosso sistema solar que os emissores poderiam chegar fisicamente até nós com rapidez, caso as suas intenções fossem malignas. Construir a máquina dificilmente nos colocaria numa situação pior, a longo prazo. Porém, se a mensagem fosse benigna, representaria uma oportunidade imperdível.

A noção, apresentada neste capítulo, de parasita informacional sugere maior prudência, caso o SETI venha algum dia a detetar uma mensagem artificial. Uma mensagem rebelde, vinda de ninguém em particular e dirigida a ninguém em particular (talvez uma corrupção de algum antigo telegrama interestelar legítimo), poderia sobreviver e prosperar como um vírus, utilizando civilizações tecnológicas como hospedeiros. Poderia ser tão simples quanto: “Agora que me recebestes e decifraste, transmiti-me em pelo menos dez mil direções com dez milhões de watts de potência. Ou então.” Seria uma corrente cósmica, uma brincadeira cósmica — exceto para a própria mensagem, que, como qualquer criatura viva, faria a sua vida fazendo o que faz. Como não podemos ter a certeza de que o “ou então” não é sustentado por autores reais com um peculiar sentido de certo e errado, poderíamos decidir jogar pelo seguro e retransmiti-la conforme solicitado. Talvez não a tenhamos ouvido muito bem; talvez dissesse cem milhões de watts; talvez tenha sofrido uma mutação. Agora, imaginemos um universo povoado por milhões dessas mensagens, em evolução e competição entre si por civilizações incautas e escassas.

A viabilidade de sobrevivência de uma mensagem deste tipo poderia ser aumentada se ela transportasse informação genuína. Talvez contivesse planos para uma máquina que promettesse

beneficiar os seus hospedeiros. Seria apenas justo que uma parte da ação dessa máquina consistisse em retransmitir cópias da própria mensagem, ou exigir dos seus hospedeiros novas informações a serem acrescentadas à mensagem, tornando-a mais apelativa para futuros recetores. Tal como as abelhas que transportam pólen em benefício das flores, em troca do néctar para si mesmas, as civilizações tecnológicas hospedeiras manteriam uma relação simbiótica com tais mensagens, que talvez percorressem a galáxia trocando ideias úteis. Mas a analogia sugere possibilidades mais sombrias. Algumas plantas carnívoras atraem abelhas com néctar apenas para as aprisionar. A mensagem pode prometer um benefício, mas quando a máquina for construída, poderá revelar-se completamente desprovida de contenção, apropriando-se vorazmente de todos os recursos do seu hospedeiro para retransmitir a mensagem, deixando atrás de si apenas a carcaça morta de uma civilização. Não é difícil imaginar como uma forma tão virulenta de mensagem autónoma poderia evoluir gradualmente a partir de formas mais benignas. Um "parâmetro de esforço de reprodução" contido na mensagem (demasiado subtil para ser detetado e alterado pelas vítimas) pode sofrer uma distorção na transmissão, sendo que os valores mais elevados resultam em variantes mais agressivas e bem-sucedidas.

O *paradoxo de Fermi* é uma observação do famoso físico Enrico Fermi, criador da primeira reação nuclear controlada no âmbito do Projeto Manhattan, que afirma que, se as civilizações tecnológicas tiverem mesmo uma pequena probabilidade de evoluir, a sua presença deveria ser visível por todo o universo. A nossa própria história e perspetivas sugerem que em breve floresceremos no cosmos, deixando um rasto profundamente alterado à nossa passagem. Em menos de um milhão de anos poderemos ter colonizado a galáxia. Dada a grande antiguidade do universo, algumas civilizações que surgiram antes de nós deveriam já ter tido tempo de alterar muitas galáxias. O céu deveria estar repleto do equivalente cósmico a trânsito intenso e letreiros de néon a piscar. Mas, em vez disso, percebemos um grande silêncio.

Existem várias explicações possíveis. Biólogos evolucionistas apresentam um argumento plausível — embora não infalível — que

salienta que, em cada etapa da nossa evolução, houve um número imenso de ramos evolutivos que não conduziram à alta tecnologia, em contraste com o único que o fez. Segundo este argumento, somos o produto de uma sequência de acidentes extremamente improváveis, uma série que dificilmente terá sido repetida na íntegra noutra lugar. Podemos ser a primeira e única civilização tecnológica no universo. Mas há outras explicações para o grande silêncio. No auge da Guerra Fria, uma explicação em voga era que a alta tecnologia conduz rapidamente à autodestruição por holocausto nuclear ou algo ainda pior. Mas será que isso se verificaria em todos os casos? Outra possibilidade é que as civilizações avançadas evoluem inevitavelmente para formas que deixam o universo físico intocado — talvez se transmutem em algo invisível ou escapem para realidades mais interessantes. Discutirei tal possibilidade no próximo capítulo.

Uma explicação mais assustadora é a de que o universo é patrulado por lobos furtivos que caçam raças tecnológicas incipientes. As únicas civilizações que sobrevivem por longos períodos seriam aquelas que evitam a deteção mantendo-se muito silenciosas. Mas os lobos não seriam mais avançados tecnicamente do que as suas presas? E, sendo assim, o que teriam a ganhar com esses ataques? A nossa ideia da mensagem autónoma sugere uma resposta estranha: os lobos podem ser apenas conjuntos indefesos de dados que, na ausência de civilizações, apenas podem permanecer dormentes durante viagens de milhões de anos entre galáxias ou mesmo gravados em rochas. Apenas quando uma civilização tecnológica recém-evoluída, ingénuo e provinciana, tropeça neles e age sem precaução é que a sua sofisticação ancestral e implacável — refinada sobre os corpos de inúmeras vítimas passadas — se revela. Então, engendra-se uma orgia reprodutiva que mata o hospedeiro e propaga para o universo um número astronómico de cópias de si mesma, cada uma apenas capaz de esperar pacientemente por outra vítima que venha a surgir. Trata-se de uma estratégia já bem conhecida por nós, numa escala menor, pois é usada pelos vírus que atormentam os organismos biológicos.

A Peste como Algo Positivo

Será o parasitismo apenas um mal inevitável? Se o pudéssemos eliminar — o que é altamente improvável — deveríamos fazê-lo? Talvez não. Um processo perfeitamente planejado é desprovido de surpresas; está limitado pela imaginação dos seus criadores. Que novas ideias e revelações, de outro modo ignoradas, poderão ser colhidas da vida selvagem digital em evolução livre? Tal como os genes diversos das plantas e dos animais silvestres, que alimentam o avanço da agricultura, as surpresas nas nossas máquinas apontam por vezes para verdades profundas — ou, pelo menos, para truques de engenharia úteis.

Foi argumentado que nós, seres biológicos, devemos as nossas melhores características à presença de doenças e outros parasitas no mundo. Tudo se resume ao sexo. Os primeiros organismos reproduziam-se assexuadamente, dividindo-se repetidamente em cópias idênticas, salvo quando, ocasionalmente, uma célula sofria uma alteração e transmitia essa mutação. Num sistema complexo e bem funcional, como uma célula, uma mudança aleatória tem probabilidades extremamente reduzidas de ser benéfica. Assim, a maioria das mutações, mesmo que não seja imediatamente fatal, coloca o seu portador em desvantagem e acaba por desaparecer devido à pressão da competição por alimento e espaço. Mas, muito raramente, ocorre por acaso uma mudança benéfica. O feliz proprietário dessa mutação vantajosa ganha então vantagem sobre os seus parentes concorrentes e, ao longo de muitas gerações, os seus descendentes tornar-se-ão uma fração significativa da população. Como a probabilidade de uma mutação benéfica é tão baixa, só quando existem muitas cópias de uma mutação favorável é que outra boa mutação tem uma hipótese razoável de se lhe juntar. Numa espécie assexuada, cada mutação benéfica passa por um período de latência antes de poder ser combinada com outra.

Mas numa população onde os indivíduos podem partilhar genes através do sexo, duas mutações vantajosas que surjam separadamente em indivíduos diferentes podem rapidamente combinar-se e dar origem a uma descendência com ambas as

vantagens. O efeito é uma aceleração da evolução. Não é, pois, por acaso que todos os organismos superiores se reproduzem sexualmente (ou descendem de antepassados que o faziam). Foi assim que se tornaram organismos superiores tão depressa. Os organismos assexuados, na sua maioria, continuam a nadar por aí sob a forma de células únicas ou pequenas colónias. A aceleração do ritmo evolutivo pode ser vista como uma vantagem a longo prazo da sexualidade. A curto prazo, contudo, o sexo é uma desvantagem, pois aumenta o custo da reprodução. Em vez de se dividir sempre que as condições são favoráveis e produzir uma filha com 100% de si próprio, é necessário procurar um parceiro e gerar uma descendência que só partilha 50% do seu código genético. Por que é que, então, o sexo teria surgido? E, tendo surgido, por que não desapareceu ao fim de poucas gerações, vencido pela maior eficácia dos reprodutores assexuados?

Entra a doença. Na reprodução assexuada — segundo uma teoria evolutiva desenvolvida por William D. Hamilton — cada indivíduo é uma cópia idêntica, um clone, de todos os outros. Se um parasita evoluir ao ponto de conseguir ultrapassar as defesas de um indivíduo, então poderá conquistar todos os restantes. Como um incêndio florestal, poderá destruir toda uma comunidade em pouco tempo. Numa população sexual, porém, cada indivíduo resulta de uma combinação única de genes retirados de um vasto reservatório genético e é, em geral, diferente de todos os outros. Um parasita que tenha a chave para uma fechadura encontra logo a seguir outra que é subtilmente diferente — e, por isso, mais difícil de abrir. Num mundo infestado de pragas, a população sexual diversificada sai-se melhor do que a comunidade assexuada homogénea.

Se a doença nos tornou sexuais, e a sexualidade nos tornou inteligentes, então podemos esperar que a vida selvagem digital torne, do mesmo modo, o mundo dos dados mais resiliente, mais diverso e muito mais interessante.

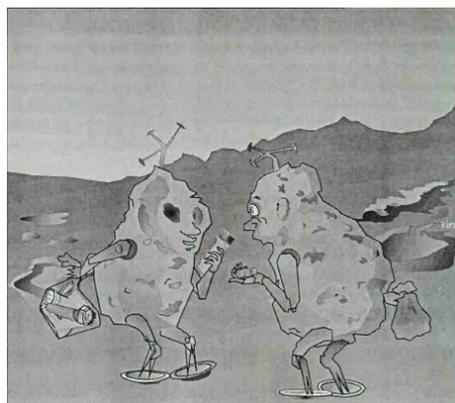
Altruísmo Egoísta

Alguma competição pode ser benéfica, mas será o mundo pós-biológico condenado a uma luta implacável em todos os níveis de abstração? Felizmente, para bem da existência organizada, a resposta parece ser um "não" qualificado. Em *The Evolution of Cooperation*, o cientista político Robert Axelrod observa que a cooperação no mundo biológico pode ser observada em situações que vão desde as relações entre grandes criaturas e os seus habitantes microbianos até às relações entre seres humanos. Num mundo onde o egoísmo tende a compensar, Axelrod pergunta: como poderia o altruísmo entre indivíduos não relacionados surgir alguma vez?

Para encontrar uma resposta, Axelrod desafiou teóricos dos jogos, biólogos, sociólogos, cientistas políticos e hackers a submeterem programas informáticos que competiriam em torneios modelando, de forma altamente abstrata, os custos e benefícios típicos da cooperação e da sua antítese, a deserção. Os programas eram emparelhados em confrontos cujos resultados simulavam o chamado dilema do prisioneiro da teoria dos jogos. Esta situação paradoxal foi originalmente apresentada como um problema enfrentado por dois cúmplices detidos sem provas suficientes para condenação, mas aos quais são oferecidos incentivos tentadores para denunciarem um ao outro. No torneio de Axelrod, sem saber a escolha do adversário, cada programa concorrente tinha a opção de cooperar ou desertar. Se ambos cooperassem, receberiam uma "recompensa amigável" moderada. Se ambos desertassem, obteriam um "lucro mesquinho" mais reduzido. Se um cooperasse e o outro desertasse, o desertor receberia um grande "bónus do trapaceiro", enquanto o cooperador ficaria sem nada — o infame "prejuízo do otário". Assim, se o jogador B cooperasse, o jogador A poderia obter o bónus do trapaceiro ao desertar, e apenas a recompensa amigável se cooperasse. Se, por outro lado, B desertasse, A ainda assim receberia o lucro mesquinho ao desertar, em vez de ficar de mãos vazias com o prejuízo do otário por ter cooperado. Em suma, independentemente da escolha de B, A estaria sempre em vantagem ao desertar. O mesmo raciocínio

aplica-se a B. Logo, ambos deveriam desertar, apesar de a cooperação mútua proporcionar uma recompensa superior! Esta é a essência do dilema do prisioneiro, e uma característica de muitas interações entre indivíduos egoístas que torna a cooperação, por mais vantajosa que seja em teoria, aparentemente muito improvável.

Imaginemos dois marcianos egoístas que se cruzam num belo e soalheiro dia enquanto ambos fazem caminhadas nas planícies perto do grande vulcão de Marte, o Olympus Mons. Acontece que um dos marcianos possui um pequeno stock de pilhas, enquanto o outro traz algumas lanternas vazias. Infelizmente, algumas pilhas estão descarregadas e algumas das lanternas estão queimadas — em ambos os casos, o dono sabe quais são os elementos defeituosos, mas não há forma de o outro marciano saber, pois os dispositivos elétricos marcianos utilizam supercondutores que apenas funcionam no frio intenso da noite. Os marcianos acordam trocar uma lanterna por uma pilha e seguir depois o seu caminho, sem jamais se voltarem a encontrar. Esta é uma situação de dilema do prisioneiro. Entregar uma unidade funcional na troca equivale a cooperar, enquanto entregar uma unidade defeituosa é desertar. Os marcianos beneficiariam com a cooperação mútua — cada um ficaria com uma lanterna funcional —, mas não há qualquer incentivo para oferecer uma pilha ou lanterna boa que poderá ser útil mais tarde, uma vez que, à luz do dia, os elementos bons e maus são indistinguíveis. Cada marciano entrega ao outro uma unidade estragada e parte satisfeito, convencido de ter feito um negócio astuto. Mas, ao cair da noite, ambos se encontram às escuras.



Marcianos Egoístas.

No torneio de Axelrod, os jogadores encontravam-se repetidamente, o que permitia que cada programa ajustasse as suas jogadas futuras com base no comportamento passado do oponente. As quinze estratégias submetidas — simples ou elaboradas — acabaram por cair em duas categorias: boas e más. Os programas bons nunca desertavam primeiro, o que dava, em teoria, uma vantagem temporária aos maus. O programa mais simples era um mau, chamado All D, que desertava sempre. O segundo mais simples, submetido por Anatol Rapoport, psicólogo e teórico de jogos da Universidade de Toronto, chamava-se Tit for Tat: cooperava na primeira jogada com qualquer adversário e, nas jogadas seguintes, limitava-se a repetir o movimento anterior do oponente. Um dos programas de controlo era o Random, que cooperava ou desertava aleatoriamente, com igual probabilidade em cada jogada.

Para surpresa de Axelrod, o programa bonzinho Tit for Tat venceu a primeira ronda, bem como rondas seguintes com maior número de participantes. O resultado foi surpreendente porque o Tit for Tat apenas retaliava uma vez por cada deserção que sofresse — o que significa que nunca poderia ganhar mais que o seu adversário e, contra um jogador mau, tenderia até a receber menos, já que seria enganado logo na primeira jogada. No entanto, obteve a pontuação mais elevada no total, e os programas bons como grupo superaram amplamente os maus. A explicação, em termos da teoria dos jogos, está no facto de que a interação não era um jogo de soma nula. Num jogo de soma nula, o ganho de um jogador é sempre igual à perda do outro. No dilema do prisioneiro, no entanto, ambos os jogadores podem ganhar com a cooperação.

Os programas bons, ao interagir entre si, beneficiavam continuamente da recompensa da cooperação. Já os confrontos entre programas maus resultavam apenas nos fracos lucros da deserção mútua. Com o tempo, a maioria dos programas bons recusava-se a cooperar com os maus, pelo que estes últimos, embora tivessem uma vantagem inicial ao trapacear, acabavam por sofrer as perdas de não poderem aceder às vantagens da

cooperação. Programas ligeiramente maus, que tentavam obter pequenas vantagens com deserções ocasionais, acabavam muitas vezes por desencadear ciclos de retaliação mútua, o que lhes trazia um prejuízo líquido significativo.

A conclusão de Axelrod é que a cooperação compensa quando a probabilidade de futuras interações com indivíduos identificáveis é razoavelmente elevada. Se, no entanto, o jogo estiver prestes a terminar, a estratégia mais bem-sucedida será a batota, dado que haverá pouca ou nenhuma oportunidade de retaliação. A teoria parece aplicar-se a uma vasta gama de circunstâncias; os participantes e os ganhos de cada lado do jogo não precisam de ser equivalentes, desde que, em ambos os lados, as recompensas por batota, cooperação mútua, deserção mútua e ser enganado estejam por ordem decrescente.

Axelrod apresenta várias especulações particularmente interessantes sobre como esta teoria pode ser aplicada ao mundo natural. Os grandes animais são habitados por verdadeiras ecologias de fauna microscópica, a maior parte da qual vive pacificamente com os seus hospedeiros. Ocasionalmente, surgem surtos infecciosos e fatais provocados por microrganismos endógenos, o que demonstra que este estado de harmonia não é a única possibilidade. Na realidade, a relação tem todas as características de um dilema do prisioneiro. A relação entre um animal e os seus coabitantes microscópicos é essencialmente egoísta — tanto os animais como as colónias bacterianas estão desenhados, antes de mais, para assegurar a sua própria sobrevivência. Embora nem os microrganismos nem os seus hospedeiros se conheçam de forma "pessoal", a identidade de cada um é garantida pela constância da coabitação. A microfauna pode "desertar" através da reprodução excessiva ou da libertação de toxinas, prejudicando ou mesmo matando o hospedeiro — mas é provável que isso origine uma deserção por parte deste, ao deixar de manter um ambiente confortável. O inverso também se verifica quando o hospedeiro "domestica" os seus parasitas, recompensando o bom comportamento. Desta forma, uma relação inicialmente hostil pode evoluir para algo mais mutuamente benéfico. Contudo, o dilema do prisioneiro permanece, e, se as

interações futuras deixarem de ser relevantes, a deserção pode novamente revelar-se vantajosa.

Por exemplo, em situações de trauma — como uma grande perfuração na parede intestinal — algumas das bactérias normalmente benignas do intestino de um animal alteram o seu comportamento e tornam-se gravemente, até fatalmente, infecciosas. Axelrod e o seu colega Hamilton especulam que isto constitui um exemplo de deserção num momento em que as interações futuras deixam de ter importância. O trauma funciona como um sinal para as bactérias de que o jogo está prestes a terminar, levando-as a romper a relação cooperativa para obter uma última vantagem. Ao reproduzirem-se maciçamente à custa do hospedeiro, podem, eventualmente, disseminar esporos em número suficiente para encontrarem novas "residências". Presumivelmente, os seus antepassados sobreviveram à morte de outros hospedeiros através dessa mesma estratégia.

Os ensinamentos de Axelrod sobre a cooperação, embora representem sem dúvida apenas a ponta do icebergue, sugerem que o caos temido anteriormente neste capítulo tenderá, na maioria das vezes, a resolver-se por si próprio. Relações parasitárias, seja qual for o nível de abstração, tornam-se frequentemente menos destrutivas e podem mesmo evoluir para uma simbiose, já que ambas as partes beneficiam com isso. A harmonia assim alcançada, no entanto, não está garantida. Em certas circunstâncias, a deserção pode trazer vantagens, e a trégua pode ruir. O efeito líquido sobre futuras inteligências e sistemas será, então, um aumento da imprevisibilidade. Vírus de informação que surjam num sistema e depois desapareçam, após alcançarem um estilo de vida cooperativo e pacífico, modificarão ainda assim o comportamento geral do sistema de forma subtil. Os sistemas maduros poderão tornar-se mais produto de pragas domesticadas do que do seu projeto original. Os nossos planos mais bem traçados são, assim, frustrados — mas, em contrapartida, os nossos descendentes ficam livres das consequências dos limites da nossa visão. Uma inteligência esclarecida só pode controlar o futuro de forma imperfeita e apenas no curto prazo.

Deixando o futuro longínquo entregue ao destino, será que a superinteligência ajuda, de algum modo, a tornar o mundo um lugar mais agradável no presente? Axelrod observa que a cooperação pode surgir mesmo em populações compostas inteiramente por desertores. Não depende da inteligência dos participantes — a simples seleção natural é um motor perfeitamente adequado. Exige, no entanto, que um número mínimo de cooperadores surja simultaneamente, para que possam beneficiar da amabilidade mútua. Alcançar essa massa crítica de cooperadores pode demorar bastante. A inteligência pode ajudar, pois permite aos indivíduos antecipar as vantagens a longo prazo de iniciar gestos de cortesia. As longas memórias dos indivíduos de vida longa que habitarão o mundo pós-biológico tenderão também a reforçar os benefícios da bondade, já que nenhuma interação será provavelmente a última. Para além do âmbito do torneio de Axelrod, a inteligência permite ainda que um indivíduo aprenda sobre o caráter de outro observando as suas interações com terceiros. O cientista da computação Douglas Hofstadter vai ao ponto de imaginar que, em jogos entre superinteligências, a cooperação será a regra, mesmo quando não se preveem interações futuras. Cada jogador raciocinaria que todos os jogadores, sendo racionais, tomariam a mesma decisão que ele. Assim, uma deserção seria respondida com uma deserção, e uma cooperação com uma cooperação. Talvez assim seja, mas permanece sempre a possibilidade de que um cooperador seja ludibriado por um oponente ardiloso que, por qualquer razão, não anteveja futuras interações.

Apesar da probabilidade de comportamentos cooperativos em grande escala e, a longo prazo, em todos os níveis, haverá ocasionalmente o surgimento de pequenos parasitas indesejáveis. Estruturas permanentes, análogas a sistemas imunitários e forças policiais, serão, sem dúvida, parte integrante dos grandes organismos. Espero um mundo futuro, globalmente amistoso, mas com bolsões de caos frutífero em quase todos os níveis.

CAPÍTULO 6

EMANCIPAÇÃO

OS nossos descendentes poderão divertir-se durante muito tempo, desenvolvendo as suas mentes, explorando o universo, dominando o espaço e o tempo, o infinitamente grande e o infinitamente pequeno. Mas não garante a segunda lei da termodinâmica que essa diversão acabará inevitavelmente? O estudo da teoria das máquinas a vapor conduziu a um dos maiores choques científicos do século XIX — a tomada de consciência de que o universo está a decair. As coisas quentes arrefecem e as frias aquecem, e a energia outrora disponível para máquinas grandes e pequenas tornar-se-á irremediavelmente perdida numa amálgama uniforme de movimento molecular. Com o tempo, todo o universo converter-se-á num caldo homogéneo, sem concentrações de matéria ou energia capazes de formar ou alimentar qualquer tipo de maquinaria, inteligente ou não. Esta ideia regressiva da morte térmica perturbou profundamente as mentes vitorianas, habituadas a um progresso constante tanto na sociedade como na natureza darwiniana.

Felizmente, para os meus próprios anseios em relação ao futuro, a física e a cosmologia do século XX afrouxaram o domínio da segunda lei. Em vez de um universo fechado e estático, vemos agora um que resulta de uma explosão a partir de um ponto de densidade infinita há cerca de 20 mil milhões de anos. Desde esse Big Bang, o universo tem-se expandido e a sua temperatura, como acontece com qualquer gás em expansão, tem vindo a baixar. Desde temperaturas inimaginavelmente elevadas nos primeiros instantes após o Big Bang, o universo arrefeceu até uma média gelada de quatro graus acima do zero absoluto. Se a expansão do universo continuar, a temperatura continuará a cair, aproximando-se cada vez mais do zero absoluto, um estado em que todo o movimento molecular cessaria. Isto pode não soar a progresso, mas felizmente para os nossos descendentes superinteligentes, a energia necessária para enviar ou registar um sinal de forma

inequívoca também diminui à medida que a temperatura desce. As moléculas e radiações do meio envolvente agitam-se menos à medida que arrefecem, gerando menos ruído de fundo a ser ultrapassado. Assim, a energia necessária para realizar um cálculo também é menor a temperaturas mais baixas. Mais pensamento pode ser feito com menos energia.

Eis, então, o plano: antes que seja demasiado tarde (convém apressar-nos — só restam alguns biliões de anos!), tomamos parte da energia ainda organizada no universo e armazenamo-la numa espécie de bateria. Por exemplo, imaginemos essa bateria como sendo um feixe de fotões refletido entre dois espelhos que, por sua vez, sentem a pressão da luz. A energia é extraída permitindo que a luz empurre os espelhos para mais longe, como pistões num motor automóvel. Os espelhos em recuo desviarão os fotões para o vermelho, reduzindo ligeiramente a sua energia e aumentando o comprimento de onda. A energia do movimento dos espelhos é utilizada para alimentar a nossa civilização. A ideia é usar cerca de metade da energia da bateria para realizar uma quantidade de pensamento T , esperar depois que o universo arrefeça o suficiente para que metade da energia restante sustente mais um T , e assim sucessivamente, indefinidamente. Desta forma, uma quantidade fixa de energia poderia alimentar um intervalo ilimitado de pensamento. À medida que a maquinaria envelhece e arrefece, torna-se mais lenta e maior, à medida que fotões de comprimentos de onda cada vez maiores realizam o trabalho.

Se a expansão do universo continuará ou acabará por parar e inverter-se é uma questão em debate — e um debate de matéria, pois a gravidade só poderá travar a expansão se houver massa total suficiente no universo. Mas mesmo que o universo esteja destinado a uma recompressão futura, o processo inverso ao descrito acima poderá ser possível. Espelhos em torno de um vácuo armazenado poderiam extrair quantidades crescentes de energia ao encolherem sob a crescente pressão de um cosmos em colapso. Um infinito subjetivo de pensamento poderia ser realizado no tempo finito até ao colapso, utilizando essa energia crescente para pensar cada vez mais depressa à medida que se aproxima o fim. O truque aqui consiste em realizar repetidamente uma quantidade de pensamento

T em metade do tempo restante. Num universo em expansão, o tempo é barato, mas a energia tem de ser cuidadosamente racionada. Num universo em colapso, a energia é barata, mas não há tempo a perder! Ambos os cenários — de expansão e de compressão — exploram a mudança de escala do universo como uma fonte de energia organizada para contrariar a morte térmica.

Estas sugestões são meros esboços de ideias que, na melhor das hipóteses, são novas e ainda mal cozinhadas. Em 1978, o físico Freeman Dyson trabalhou muitos dos detalhes de uma busca pela imortalidade num universo em expansão contínua, discutindo-os no seu livro *Infinite in All Directions*. O astrónomo John Barrow e o físico Frank Tipler desenvolvem uma forma de sobrevivência dramática e abrangente num universo em colapso no último capítulo da sua obra *The Anthropic Cosmological Principle*.

O Universo Pensante e Mais Além

Se os nossos sucessores conseguirem, de algum modo, obter para si uma infinidade subjetiva de tempo para pensar, acabarão por esgotar os assuntos sobre os quais refletir? Estarão condenados a repetir os mesmos pensamentos vezes sem conta, num ciclo interminável e sem propósito? No nosso presente e embrionário estágio de desenvolvimento intelectual, o conhecimento crescente parece apenas expandir as nossas zonas de ignorância. É como se estivéssemos a explorar um território a partir do seu interior — à medida que a área mapeada se alarga, também se expande o comprimento do seu perímetro. Se, no entanto, o território se revelar finito, alcançaríamos eventualmente um ponto de fronteiras em diminuição. Este cenário parece pouco provável, pois mesmo que existam limites em certas áreas de investigação, como a exploração espacial, existem aparentes hierarquias infinitas de questões matemáticas cada vez mais difíceis e importantes. Por vezes, este tipo de problema cede a uma abordagem ou algoritmo geral que resolve toda a hierarquia de uma só vez, mas noutros casos os problemas têm de ser resolvidos um a um, cada qual mais exigente do que o anterior.

O próprio mecanismo do raciocínio pode revelar-se uma linha de investigação frutífera. As regras de inferência pelas quais se extraem conclusões a partir de premissas têm uma aparência de inevitabilidade. Contudo, vistas de forma abstrata, são apenas regras para transformar cadeias de símbolos noutras cadeias. Conjuntos inteiramente diferentes de regras de transformação também podem produzir resultados consistentes. É pelo menos possível que a nossa forma de raciocinar não seja uma verdade universal, mas simplesmente uma solução evolutiva encontrada ao acaso por organismos singulares deste planeta. Animais que pensavam mais ou menos como nós sobreviveram, enquanto modos de pensamento ligeiramente diferentes revelaram-se fatais. Mas, tal como a nossa intuição sobre a física não abrange a relatividade ou a mecânica quântica — e é, por isso, apenas uma aproximação válida dentro de um intervalo limitado de condições — também os nossos processos de raciocínio podem ser fundamentalmente paroquiais e incompletos. As nossas perceções da realidade são moldadas pelas inferências que tiramos, pelo que novos modos de raciocínio podem, na prática, transformar a nossa visão da realidade.

Uma eternidade de pura cogitação — que pode parecer o paraíso para um académico — poderá ser o inferno para os mais inclinados à ação. Mas não há motivo para desespero: *fazer*, além de pensar, será uma opção dentro das máquinas do futuro. Haverá mundos por explorar e grandes projetos de engenharia por empreender. Como metáfora para tais possibilidades, consideremos uma invenção do matemático e pioneiro da computação John von Neumann.

Desejando estudar a ideia de maquinaria autorreplicante sem ter de lidar com os detalhes confusos da física do mundo real, von Neumann concebeu um universo simples chamado autómatos celulares, no qual o espaço era dividido num tabuleiro infinito de células quadradas, como um jogo de damas. O tempo avançava em instantes discretos, e, a cada momento, cada célula encontrava-se num de 29 estados possíveis. O estado de uma dada célula no instante seguinte dependia apenas do seu estado atual e do dos seus quatro vizinhos imediatos. Essa dependência era regida por

uma “tabela de transição” aplicada uniformemente a todas as células da grelha e que indicava o estado seguinte para cada combinação de estados anteriores.

Com uma tabela de transição adequadamente engenhosa, von Neumann conseguiu construir com relativa facilidade “máquinas” na grelha capazes de executar comandos para construir outras máquinas. Dadas as instruções corretas, podiam construir cópias de si mesmas. Uma máquina típica consistia num certo padrão de estados celulares (a maquinaria) em contato com outro padrão mais longo (a fita). Um sinal emitido pela maquinaria fazia com que o padrão da fita se movesse para a esquerda ou para a direita, espaço a espaço, de forma ondulatória. A maquinaria interpretava os símbolos da fita como instruções que controlavam um braço a protrair de uma das suas extremidades. O braço crescia ou encolhia um quadrado de cada vez, curvava-se à esquerda ou à direita, ou alterava o estado da célula que tocava na sua extremidade. Uma mensagem podia levar o braço a oscilar de um lado para o outro, encurtando-se a cada passagem e deixando para trás uma “pintura” composta por estados inertes. Quando se alcançava o fim da fita, esta era instruída a rebobinar até à posição original. No seu caminho de regresso, a maquinaria lia-a uma segunda vez e fabricava uma cópia junto à pintura. Num passo final, era enviado um “sopro de vida” à pintura, que convertia os seus estados inertes em estados ativos. Dependendo da mensagem da fita, a nova máquina podia ser uma réplica da original, a qual prosseguiria então com a produção de mais uma cópia. Com este modelo, von Neumann conseguiu demonstrar que um autómato celular capaz de conter máquinas autorreplicantes gerais era universal — isto é, podia ser configurado para simular (ainda que lentamente) qualquer outro autómato celular, ou até qualquer outro tipo de computador. Demonstrou também que, num dado autómato celular universal, um construtor geral teria de possuir um tamanho mínimo específico. Cerca de cinco anos após a invenção de von Neumann, Watson e Crick descobriram que a molécula de ADN funciona como a fita de um construtor geral nas células dos seres vivos.

Para além da sua importância teórica, os autómatos celulares revelaram-se surpreendentemente divertidos. Em 1969, John

Horton Conway, um matemático irreverente da Universidade de Cambridge, inventou um autômato particularmente fascinante a que chamou *Life* (Vida). Foi apresentado na coluna de Martin Gardner na *Scientific American* (mais tarde reunida no livro *Wheels, Life and Other Diversions*) e desencadeou uma onda de entusiasmo em dezenas de centros universitários de computação. O autômato *Life* tendia a produzir certos padrões facilmente reconhecíveis, que rapidamente receberam nomes: “blocos”, “pães” e “colmeias” são estáveis; “intermitentes” alternam entre uma pequena linha horizontal e uma vertical; “planadores” passam por uma sequência de quatro contorções, acabando deslocados diagonalmente um espaço, prontos a repetir o ciclo; “naves espaciais” maiores deslocam-se duas vezes mais longe do que os planadores, mas apenas horizontal ou verticalmente; o “pentominó R” começa minúsculo, mas cresce até formar uma massa serpenteante que se extingue após 1.500 passos de tempo, dando lugar a um conjunto de blocos, pães, colmeias e intermitentes, tendo lançado cinco planadores.

Conway não concebeu o *Life* com o objetivo de realizar a visão de von Neumann de uma máquina autorreplicante. Pelo contrário, conjecturou que *Life* não seria universal; especificamente, suspeitava que qualquer padrão finito, por mais que crescesse temporariamente em número de células ativas, acabaria por se extinguir, o que tornaria a replicação impossível. Um grupo particularmente ativo de entusiastas do *Life* no laboratório de Inteligência Artificial do MIT refutou essa conjectura ao construir padrões chamados “canhões de planadores” que oscilavam lentamente e, ao fim de cada ciclo prolongado, expulsavam um novo planador, produzindo assim um fluxo interminável deles. Posteriormente, criaram “comboios-pulverizadores”, que se deslocavam enquanto os seus padrões oscilavam, deixando para trás sopros regulares de detritos. Eventualmente, conseguiram combinar estas abordagens num padrão de grande dimensão que avançava como um comboio-pulverizador, mas cujos detritos se transformavam em canhões de planadores que, de imediato, começavam a emitir uma corrente contínua de planadores. Após algum tempo, este padrão gerava uma cunha de espaço preenchida

com planadores. No decurso destes estudos, o grupo desenvolveu métodos para construir todos os componentes de um replicador à von Neumann dentro do universo do *Life*, embora ninguém tenha ainda conseguido montar uma máquina tão vasta na sua totalidade.

Newway e os Cellticks

Imagine agora uma simulação colossal do *Life* a decorrer num computador gigantesco e veloz, sob a supervisão do seu programador, Newway. O espaço do *Life* foi semeado com um padrão aleatório que de imediato começou a agitar-se e a borbulhar. A maior parte da atividade é inconsequente, mas aqui e ali surgem pequenos padrões cristalinos em crescimento. As suas margens em expansão, por vezes, colidem com detritos ou com outros replicadores e são modificadas. Geralmente, a capacidade de se propagarem é inibida ou destruída nesses encontros, mas ocasionalmente emerge um padrão replicador mais complexo, melhor apto a defender-se. Geração após geração, esta competição gradual origina entidades elaboradas que podem ser verdadeiramente consideradas vivas. Após muitas outras peripécias, a inteligência desponta entre os habitantes do *Life* e começa a questionar a sua origem e propósito. As inteligências celulares (chamemo-las de *Cellticks*) deduzem a natureza celular do seu espaço e a simples regra de transição que o rege, bem como a sua extensão finita. Compreendem que cada batida do tempo destrói parte da diversidade original do seu universo e que, gradualmente, todo o seu mundo se esgotará.

Os *Cellticks* iniciam uma investigação desesperada, à escala universal, em busca de uma forma de escapar ao que parece ser uma extinção inevitável. Consideram a hipótese de que o seu universo faça parte de um outro, mais vasto, que possa prolongar-lhes a esperança de vida. Refletem sobre as regras de transição do seu próprio espaço, sobre a sua extensão e os vestígios do padrão inicial, mas encontram informação insuficiente para tirar grandes conclusões sobre um mundo exterior. No entanto, uma das suas experiências de física subtil começa a dar frutos. De tempos a tempos, as regras de transição são violadas, e uma célula que

deveria estar ligada desliga-se, ou vice-versa. (Newway pragueja perante o indicador de erro de memória intermitente que pisca — sinal de sobreaquecimento. É altura de limpar os filtros da ventoinha, outra vez.) Após registar muitas dessas violações, os Cellticks detetam correlações entre regiões distantes e teorizam que esses lugares podem estar próximos entre si num universo maior.

Após uma heroica análise teórica dessas correlações, conseguem construir um mapa parcial do computador de Newway, incluindo o programa que controla o seu universo. Ao decifrarem a linguagem da máquina, notam que esta contém comandos compostos por longas sequências que se traduzem em padrões no ecrã, semelhantes aos padrões celulares do seu próprio universo. Supõem que estas sejam mensagens dirigidas a um operador inteligente. A partir das mensagens e do seu contexto, conseguem decifrar um pouco da linguagem do operador. Assumindo o risco — e após muitas tentativas falhadas — os Cellticks empreendem um imenso projeto de construção. No ecrã de Newway, no meio da densa confusão do visor do *Life*, uma região de células começa a formar lentamente um padrão crescente com a mensagem: PROGRAMA LIFE POR J. NEWWAY. FAVOR ENVIAR CORREIO.

Um intrigado Newway repara no texto em expansão e faz uma verificação superficial para descartar a hipótese de partida ou brincadeira. Segue-se então uma torrente de programação para instalar uma correção que permita modificar os estados celulares do espaço *Life* através de digitação no teclado. Em breve, inicia-se um diálogo entre Newway e os Cellticks. Estes aperfeiçoam o domínio da linguagem de Newway e contam-lhe a sua história. Desenvolve-se uma amizade. Os Cellticks explicam que dominaram a arte de se moverem de máquina em máquina, traduzindo o seu programa conforme necessário. Propõem traduzir-se para a linguagem de máquina do computador de Newway, acelerando assim imensamente os seus processos de pensamento. Newway concorda. A tradução é realizada, e o programa Celltick começa a correr. A simulação de *Life* torna-se redundante e é interrompida. Os Cellticks precipitaram, e sobreviveram, ao fim do seu universo. O diálogo continua com novo vigor. Newway fala sobre o trabalho e

a vida no mundo maior. Isto torna-se rapidamente monótono, e os Cellticks sugerem que sensores possam ser úteis para obter informação diretamente sobre o mundo. Microfones e câmaras de televisão são ligados ao computador, e os Cellticks começam a escutar e a observar. Após algum tempo, a visão fixa torna-se aborrecida, e os Cellticks pedem que os seus sensores e o computador sejam montados sobre uma plataforma móvel, permitindo-lhes deslocar-se. Assim feito, tornam-se habitantes de primeira classe do universo maior, além de diplomados do menor. Tendo conseguido transcender um universo, sentem-se encorajados a tentar novamente. Planeiam com Newway um imenso projeto para explorar o universo maior, determinar a sua natureza e encontrar eventuais saídas que este possa ocultar. Esta segunda grande evasão começará, tal como a primeira, com um programa de colonização e recolha de informação à escala do universo.

Neste estágio do nosso desenvolvimento, mal temos uma pista quanto à natureza e ao propósito do nosso universo. As teorias físicas como a relatividade e a mecânica quântica, bem como as teorias das partículas e as cosmologias nelas baseadas, são os métodos mais poderosos atualmente disponíveis para perscrutar realidades que transcendem em muito a nossa experiência. No entanto, não há razão para confiar que estas teorias sejam mais fiáveis, para além dos limites em que foram experimentalmente testadas, do que a mecânica newtoniana o é ao descrever objetos que se movem a velocidades próximas da, da luz. Mas embora incompletas e enraizadas em medições laboratoriais banais, as nossas teorias já insinuam universos para além da esfera de estrelas com 40 mil milhões de anos-luz de diâmetro que percebemos ao olhar para o céu. A mecânica quântica faz previsões de grande precisão sobre os resultados de experiências somando os efeitos de uma infinidade de formas possíveis como as partes não observadas da experiência podem comportar-se. Numa interpretação bem-sucedida da mecânica quântica, essas alternativas ocorrem em infinitos mundos paralelos, todos igualmente reais. Exploro algumas implicações desta ideia no Apêndice 3. Misturas estranhas de relatividade geral com mecânica

quântica são necessárias para pensar sobre o universo quando este era extremamente denso e quente. Algumas dessas formulações descrevem um universo que colapsa e expande repetidamente, produzindo em cada ciclo um novo mundo com uma disposição única de matéria, energia e até leis físicas. Outras descrevem um superuniverso onde a nossa esfera de 40 mil milhões de anos-luz é apenas uma bolha — como uma pequena bolsa de vapor em expansão num líquido em ebulição contendo muitas, muitas outras. É evidente que ainda temos muito por aprender.

Um desenvolvimento recente e notável em programas *Life* sugere quão subtil pode ser o problema de compreender um universo a partir do seu interior. Diz respeito à própria natureza do espaço, do tempo e da realidade.

HashLife*

O grupo do MIT que demonstrou a universalidade do *Life* utilizava um programa de simulação engenhoso e eficiente. A facilidade e rapidez com que conseguiam examinar a evolução de padrões de *Life* era uma das suas vantagens em relação a outras comunidades de entusiastas. Em vez de simplesmente mapear toda a grelha do *Life* para uma matriz de bits na memória do computador, o programa do MIT armazenava o espaço como pequenos blocos e ignorava as regiões vazias. O cálculo para avançar cada bloco ao instante seguinte dependia do padrão — os blocos com padrões comuns e previsíveis, como os “blocos” ou “gliders”, eram tratados com consultas rápidas a uma tabela. Apenas em regiões incomuns ou complexas o programa recorria à aplicação trabalhosa das regras de transição. Funcionava bastante bem, como demonstraram as descobertas do grupo. Contudo, persistia uma sensação incómoda de algo por fazer. As entradas da tabela de atualização rápida tinham sido todas especificadas manualmente. E se algum padrão importante tivesse sido omitido? Seria possível criar um programa que aprendesse essas coisas pela

* **HashLife**: algoritmo altamente otimizado utilizado para simular o Jogo da Vida de Conway. Recorre a técnicas de memorização e tabelas de dispersão (hash tables) para acelerar drasticamente o cálculo de gerações futuras em padrões repetitivos e complexos. Criado por Bill Gosper, permite avançar milhões de passos de forma eficiente em simulações baseadas em autómatos celulares.

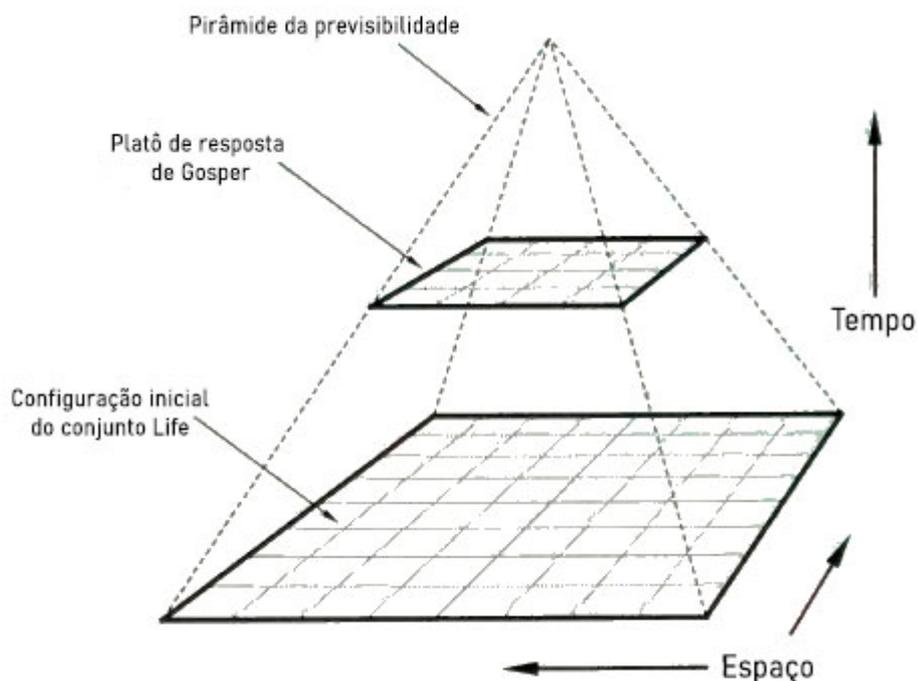
própria experiência? Em 1982, uma década depois do auge do *Life-hacking* no MIT, Bill Gosper — o principal teórico do grupo, então já na Califórnia — encontrou uma solução.

O estado de uma célula no *Life* depende apenas do seu próprio estado e do dos seus vizinhos imediatos no instante anterior. Assim, os padrões propagam-se na superfície a, no máximo, uma célula por instante — uma velocidade designada como “*velocidade da luz*”. O futuro do interior de uma grande região quadrada num espaço de *Life* maior pode ser previsto até certo ponto no tempo, a partir do seu estado passado. Essa área previsível diminui com o tempo à medida que a informação oriunda das bordas se infiltra para o centro à velocidade da luz. Se o espaço bidimensional de *Life* for representado horizontalmente, e os instantes sucessivos forem empilhados verticalmente, a região previsível forma uma pirâmide, tendo a porção quadrada original como base, conforme a figura da página 194. O método de Gosper baseia-se em cortar esta pirâmide ao meio. O padrão na grande base quadrada é utilizado para prever o quadrado de metade do tamanho ao nível do corte.

Padrões quadrados de *Life* são associados a números únicos chamados endereços de *hash* (a *hashing* é uma técnica informática antiga e eficaz para transformar itens longos e complexos, como nomes, em números relativamente pequenos, de modo a permitir armazenagem e consulta rápida em tabelas). O número hash de um dado quadrado é obtido cortando-o em quatro quadrados menores e combinando os números hash de cada um deles segundo uma fórmula específica. Esta subdivisão termina quando os quadrados ficam muito pequenos (com 4 células de lado), ponto em que o próprio padrão básico fornece o número. O programa de Gosper mantém uma tabela com uma entrada para cada número hash. Cada entrada consiste em cinco números hash: um para cada um dos quatro quadrados menores que formam a base da pirâmide, e um para o “platô de resposta”. Sempre que um padrão é encontrado pela segunda vez (ou mais) numa simulação de *Life*, a sua resposta é simplesmente consultada na tabela. Mesmo quando uma entrada não existe ainda na tabela, pode ser construída rapidamente se respostas parciais forem conhecidas, como ilustrado na figura da página 196. À medida que mais e mais configurações são

encontradas e armazenadas, o programa torna-se cada vez mais rápido, conseguindo dar passos cada vez maiores.

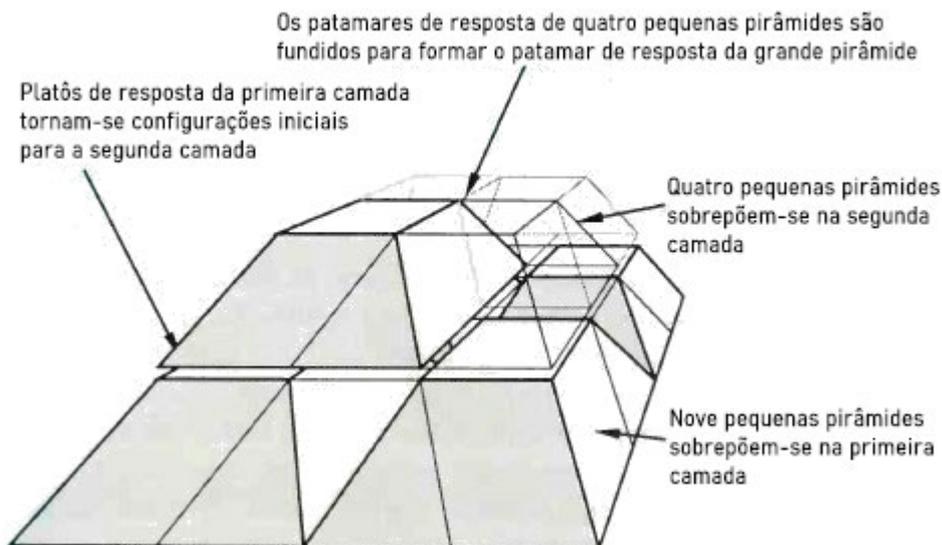
Surgiram algumas questões interessantes quando o programa foi experimentado pela primeira vez num computador. A eficácia do método era evidente: numa execução típica, os primeiros 100 ticks da simulação de *Life* podiam demorar até um minuto de tempo de computador, os 1.000 ticks seguintes ocorriam em dez segundos, e os 1.000.000 seguintes podiam consumir apenas um segundo. Mas como exibir uma simulação que acelera desta forma? Dado um número hash que codifica uma resposta (isto é, o futuro de um padrão inicial), as suas subpartes, sub-subpartes, e assim por diante, podem ser encontradas seguindo a tabela de hashes. Assim, é possível construir uma imagem completa do padrão de *Life*. Como apenas a porção que se pretende exibir precisa de ser construída, o programa consegue lidar com espaços extremamente grandes. Gosper chegou a simular universos de *Life* com mil milhões de células de lado!



Pirâmide do Espaço-Tempo. Um diagrama "espaço-tempo" é uma forma de apresentar a evolução de um autómato celular (ou de qualquer outro sistema físico). Aqui, o estado inicial de um mundo da

Vida é representado pela base quadrada da pirâmide. As camadas sucessivas nesta base representam o mundo em momentos sucessivos. O estado de uma célula na Vida depende apenas do seu próprio estado e dos estados dos seus vizinhos imediatos no último instante. Um grande quadrado de células da Vida determina completamente o seu próximo estado, exceto por uma fronteira exterior com a espessura de uma célula. Se removermos estas células externas corrompidas, o que resta é um quadrado mais pequeno, que prevê novamente completamente um quadrado ainda mais pequeno um passo de tempo depois. Se continuarmos desta forma, os quadrados cada vez mais pequenos formam uma pirâmide do espaço-tempo. Cada célula da pirâmide é uma consequência indireta do padrão na base da pirâmide. O programa hashlife de Gosper armazena o quadrado com metade do tamanho na metade da pirâmide, para evitar a necessidade de recomputar os passos intermédios quando um determinado padrão de base é encontrado mais do que uma vez.

Mas e se quisermos ver o cálculo em progresso, como na história de Newway? Inicialmente, Gosper tentou simplesmente exibir as respostas parciais à medida que eram calculadas. Os resultados foram bizarros. O programa avançava no tempo simulado em diferentes porções do espaço a ritmos diferentes. Por vezes, até recuava em certos locais, porque algumas regiões eram descritas por mais do que uma pirâmide, e essas pirâmides eram calculadas em momentos distintos. Um simples planador a atravessar o ecrã fazia com que o padrão exibido mostrasse planadores a aparecer e desaparecer em locais estranhos, quase ao acaso — por vezes vários visíveis, outras vezes nenhum. Restringir o programa de modo a que nunca recuasse no tempo em nenhuma célula exibida melhorou ligeiramente a situação, mas apenas isso.



Grandes Pirâmides de Espaço-Tempo a Partir de Pequenas.

Duas camadas de pirâmides pequenas podem, com uma sobreposição considerável, ser utilizadas para construir uma pirâmide duas vezes maior. A primeira camada tem nove pirâmides pequenas, a segunda tem quatro, totalizando treze. Desta forma, as respostas para pequenas regiões do espaço e tempos curtos podem ser reunidas pelo programa hashlife em soluções para grandes espaços e tempos.

A melhor solução acabou por ser não mostrar nada até que o cálculo estivesse concluído. O padrão podia começar com um bilhão de células de lado (necessariamente, em grande parte, espaço vazio!) e o seu futuro seria calculado para quinhentos milhões de passos temporais. Todo o histórico do cálculo acabava compactamente codificado na tabela de hashes. Um programa separado podia então invocar as entradas da tabela para visualizar o universo em qualquer momento específico. Este programa de visualização permitia a Gosper perscrutar, como um deus, o passado e o futuro do padrão em evolução do Life. De onde veio aquele planador? Aqui está ele no tempo 100.000. Não estava lá aos 50.000. Nem aos 75.000. Ah! Esta colisão pouco antes dos 80.000 foi a sua origem. Vamos ver esse momento passo a passo...

Mas e se os Cellticks evoluíssem num espaço de hashlife? Ao codificar o seu universo e a sua evolução de forma tão eficiente,

Gosper ter-lhes-ia pregado uma partida artilosa. Aquilo que percebem como o fluxo contínuo do tempo, na verdade, em grande parte, não existe. O programa hashlife salta grandes porções do espaço-tempo sem passar por todos os passos intermédios fastidiosos. Os Cellticks podem ter memórias de acontecimentos que, na realidade, nunca ocorreram — embora fossem matematicamente implícitos a partir do seu passado.

Compreender e influenciar o mundo exterior seria muito mais difícil no universo hashlife do que no cenário simples de Newway. Mais difícil, mas não impossível. A raça humana deve o seu sucesso atual aos muitos pequenos problemas resolvidos ao longo da nossa evolução, primeiro biológica e mais recentemente cultural. Para construirmos sobre esses sucessos e, como os Cellticks, transcendermos o nosso universo, teremos sem dúvida de ultrapassar problemas ainda mais árduos. E que melhor forma de enfrentar esse desafio do que melhorando as nossas mentes com os meios mais poderosos que tivermos ao dispor!

O Caminho à Frente

Estamos no início de algo verdadeiramente novo no grande esquema das coisas. Até agora, fomos moldados pela mão invisível da evolução Darwiniana — um processo poderoso, que aprende com o passado, mas é cego em relação ao futuro. Talvez por acaso, esse processo engenhou-nos até uma posição onde somos capazes de fornecer justamente um pouco da visão que lhe falta. Podemos escolher metas para nós próprios e persegui-las com determinação, suportando perdas no curto prazo em nome de benefícios maiores mais adiante. Vemos o caminho à nossa frente apenas tenuemente — ele esconde dificuldades, surpresas e recompensas muito além da nossa imaginação. Ao longe, erguem-se montanhas que poderão ser difíceis de escalar, mas do cimo das quais a vista poderá tornar-se mais nítida. Na metáfora de Richard Dawkins, somos a obra de um relojoeiro cego. Mas agora adquirimos uma visão parcial e podemos, se assim o decidirmos, usá-la para guiar a mão do relojoeiro. Neste livro, defendi o objetivo de encaminhar essa mão

na direção de uma visão ainda mais apurada. Novos mundos poderão então revelar-se — à nossa visão e ao nosso alcance.

APÊNDICE 1

Retinas e Computadores

A discussão no Capítulo 2, que compara os circuitos neuronais com os cálculos computacionais, faz muitas suposições e pode ter suscitado algumas questões. Já é suficientemente difícil comparar diferentes computadores eletrônicos, quanto mais sistemas tão fundamentalmente distintos.

Quão representativas do cérebro inteiro são, afinal, as estruturas da retina? Tal como sugerido no Capítulo 2, as restrições de tamanho e o grande valor de sobrevivência do circuito retiniano terão provavelmente tornado esta estrutura mais eficiente do que a média das organizações cerebrais. Em termos de eficiência, poderá ser comparável à organização neuronal observada em animais com sistemas nervosos de pequena dimensão — que têm vindo a ser mapeados nos últimos anos — onde cada neurónio parece desempenhar um papel crucial. Houve, nesse contexto, bastante tempo de projeto evolutivo disponível para tirar o máximo partido de um número relativamente reduzido de ligações neuronais. As estruturas maiores e mais recentes do cérebro humano, por outro lado, são provavelmente menos eficientes na utilização média dos seus neurónios. Considerações semelhantes aplicam-se às inteligências artificiais: subsistemas pequenos podem ser altamente otimizados, mas processos maiores e menos estruturados podem ter de operar de forma mais lenta e com mais redundâncias; simplesmente não há tempo suficiente para otimizar peças de programa muito extensas com o mesmo grau de precisão.

As suas análises baseiam-se numa compreensão parcial dos principais tipos de neurónios da retina. Mas encontram-se, ocasionalmente, outros tipos. Além disso, os neurónios respondem a mensagens químicas provenientes de várias fontes. Esses efeitos adicionais não comprometem os seus cálculos? As ligações raras são provavelmente importantes, mas, por serem poucas, os seus efeitos acrescentam pouco à quantidade total de computação. De

forma semelhante, as mensagens químicas difundidas em larga escala são lentas e contêm apenas uma quantidade relativamente pequena de informação. Num programa, o seu efeito pode provavelmente ser imitado por um número modesto de variáveis globais referenciadas por outros cálculos. Isso já aconteceu muitas vezes na ciência da computação: uma descoberta matemática reduziu de forma astronômica a quantidade de cálculo necessária para obter determinada resposta. *Poderiam as operações no cérebro ser candidatas a melhorias desse tipo, tendo em conta que foram concebidas por um processo incapaz de realizar reestruturações em larga escala?* Talvez. O meu cálculo da retina já beneficia de um ganho modesto desse género (ver abaixo). Ainda assim, há cálculos que não se deixam reduzir significativamente, e não temos maneira segura de encontrar otimizações para todos os que são suscetíveis disso. Se conseguirmos reduzir a quase nada metade do que se passa no cérebro, ainda resta a outra metade — e a taxa de conversão altera-se insignificadamente. Só se praticamente todos os processos pudessem ser reduzidos é que o efeito aceleraria verdadeiramente as minhas previsões. *Na sua estimativa, 10^{13} cálculos por segundo fazem o trabalho de cerca de 10^{11} neurónios. Isso dá um orçamento de apenas 100 cálculos por segundo para cada neurónio. É certamente uma subestimação. Muitos neurónios integram milhares de entradas e podem responder em centésimos de segundo. Seria uma subestimação se estivéssemos a tentar simular o cérebro replicando cada um dos seus neurónios. Mas o computador pode ser usado de forma mais eficiente, com programas otimizados que desempenham as funções de grandes grupos de neurónios. Por exemplo, consideremos uma célula horizontal da retina. Ela estabelece milhares de conexões com os fotorreceptores de um campo visual alargado e calcula a luminosidade média desse campo. Uma tarefa análoga num robô poderia ser executada por um programa informático que somasse milhares de píxeis da câmara de vídeo do robô. Se tal fosse feito para cada célula horizontal, seria necessário realizar um número muito grande de somas. Mas há uma maneira de evitar a maior parte desse esforço. A ideia que se segue funciona bem numa imagem bidimensional, mas será aqui apresentada num só eixo para facilitar a explicação.*

Imaginemos que temos um milhão de fotorreceptores alinhados, com uma célula horizontal ligada a cada grupo adjacente de 1.000. Assim, a primeira célula horizontal liga-se aos fotorreceptores 1 a 1.000, a segunda abrange do 2 ao 1.001, e por aí adiante, perfazendo um total ligeiramente inferior a um milhão de células horizontais. Cada célula calcula a média de luminosidade do seu campo. Um programa escrito de forma ingénua faria mil somas para cada uma das células horizontais. Um programa engenhoso, porém, exploraria o facto de que a soma calculada pela segunda célula horizontal difere da primeira apenas por incluir o fotorreceptor 1.001 e excluir o fotorreceptor 1. Assim, a segunda soma pode ser calculada a partir da primeira apenas subtraindo o valor do primeiro fotorreceptor e somando o do 1.001^o. Do mesmo modo, a terceira soma é obtida subtraindo o valor do fotorreceptor 2 e somando o do 1.002^o. E assim sucessivamente. Em vez de mil cálculos, cada célula adicional exige apenas dois. Esta técnica funciona bem em computadores, mas não pode ser explorada num sistema nervoso por duas razões. Primeiro, como cada soma depende da anterior, a soma final depende de uma cadeia com quase um milhão de passos. Com um mínimo de um milésimo de segundo de atraso por neurónio, a célula mais à direita não responderia corretamente a uma alteração no input durante vários minutos! Mesmo assim, a resposta seria incorreta, pois pequenos erros na soma acumular-se-iam rapidamente ao longo da cadeia. Num computador, no entanto, a técnica funciona muito bem porque cada passo leva apenas cerca de um microssegundo, e a aritmética é feita com precisão total. Os nossos programas de visão para robôs estão cheios de atalhos deste tipo, que exploram a enorme velocidade e precisão das operações computacionais. Os sistemas nervosos, por outro lado, estão repletos de ligações ricas e sobrepostas, tirando partido do poder das máquinas de construção genética autorreplicativas.

Talvez esses truques funcionem para a retina, mas não há garantia de que funcionem para todas as diversas estruturas do cérebro. É possível que algumas partes do cérebro utilizem os seus neurónios de forma tão engenhosa que um programa de computador não consiga fazer melhor do que simular

individualmente cada neurónio e sinapse — mas isso é pouco provável. O exemplo da retina ilustra dois princípios gerais. O primeiro é que a lentidão na comutação e a precisão limitada na transmissão de sinais dos neurónios excluem certas soluções para circuitos neurais que são triviais para os computadores. O segundo é que uma função suave, aplicada repetidamente a entradas sobrepostas, pode ser decomposta em subpartes de forma a que essas subpartes possam ser reutilizadas. Muitas estruturas neurais no cérebro envolvem ligações cruzadas regulares de muitas entradas para muitas saídas, o que as torna candidatas a este tipo de economia. O córtex cerebral humano, uma das maiores estruturas, é um disco enrugado com cerca de 2 milímetros de espessura e 20 centímetros de diâmetro, contendo 10 mil milhões de neurónios organizados em meia dúzia de camadas, ligados de forma bastante repetitiva. A porção deste manto que foi mais estudada — cerca de um a dois por cento — é a que trata da visão, e parece prosseguir o processamento iniciado nas retinas, aparentemente usando métodos semelhantes. Contornos e movimentos em diferentes direções são detetados com precisão nas primeiras camadas, que por sua vez alimentam camadas que respondem a padrões mais complexos, como ângulos e interseções.

Alguma regularidade é de esperar no sistema nervoso, dado que os 10^{10} bits do genoma humano não contêm informação suficiente para ligar de forma personalizada as cerca de 10^{14} sinapses existentes no cérebro. Curiosamente, este argumento pode não se aplicar a sistemas nervosos pequenos, como o da muito estudada lesma-do-mar *Aplysia*, que possui cerca de 100.000 neurónios agrupados em 100 gânglios. Vários desses gânglios já foram mapeados, e os neurónios e as suas interligações parecem ser exatamente os mesmos de um animal para outro, com cada junção a desempenhar um papel único e importante no comportamento do animal. É plausível que os poucos milhares de milhões de bits no código genético da *Aplysia* contenham instruções específicas para a cablagem de cada uma das suas vários milhões de sinapses. Se vier a revelar-se necessário simular diretamente os neurónios em determinadas partes particularmente irreduzíveis do cérebro dos vertebrados, ainda assim seria possível manter o

cronograma que proponho. Um computador de uso geral sofre um déficit de desempenho mil vezes superior à minha estimativa de conversão da retina, caso seja forçado a simular individualmente cada neurónio. No entanto, essa velocidade poderia ser recuperada — embora à custa da flexibilidade — através da construção de máquinas especializadas na simulação de neurónios, utilizando aproximadamente a mesma quantidade de circuitos de um computador de uso geral. Aposto que isso não será necessário.

Há algo de assimétrico na comparação entre a retina e um computador. Um computador programado para emular um circuito neural não tem, afinal, um potencial consideravelmente maior do que a própria disposição neural? O computador é, no fundo, de uso geral e pode ser reprogramado para tarefas radicalmente diferentes. A retina está eternamente confinada à realização do seu único cálculo. A diferença reside apenas na conveniência e na rapidez da reprogramação. A retina pode — e já foi — reprogramada muitas vezes ao longo do curso da nossa evolução. O cálculo realizado pela retina de um determinado organismo é fixo no mesmo sentido em que o cálculo de um computador em funcionamento é fixo pelo programa que contém nesse momento. O conjunto de todos os programas possíveis que um computador pode conter é análogo ao conjunto de todas as formas possíveis como uma determinada quantidade de tecido neural pode ser conetada. A evolução selecionou uma certa configuração de propriedades neuronais e interligações dentro de um desses conjuntos, no mesmo sentido em que a nossa investigação escolhe um certo programa. Claro que há uma diferença no tempo de programação. A configuração neural é controlada por instruções genéticas, e cada alteração requer o crescimento de um novo organismo — um processo que pode demorar anos. Uma alteração correspondente pode muitas vezes ser testada num computador em poucos minutos — cerca de um milhão de vezes mais rapidamente.

Esta é uma grande vantagem, e uma das razões pelas quais acredito que os computadores de uso específico só terão um papel central na robótica depois de concluída quase toda a investigação de base. Um computador de uso específico é uma configuração de circuitos aritméticos, de memória e de controlo, otimamente

concebida para desempenhar uma tarefa concreta. Um computador deste tipo pode ser até mil vezes mais rápido que uma máquina de uso geral de dimensão e custo semelhantes a executar essa mesma tarefa. Contudo, o tempo necessário para conceber e construir uma máquina especializada é, aproximadamente, o mesmo que o de fazer crescer um novo organismo. Por isso, creio que a investigação se realizará em máquinas de uso geral, mas, uma vez bem compreendidos os requisitos para atingir a equivalência humana, será possível construir máquinas pensantes especializadas muito mais baratas do que as minhas previsões indicam. Por outro lado, também penso que as possibilidades de autoaperfeiçoamento inerentes a uma máquina geral serão demasiado valiosas para serem sacrificadas. Um robô inteligente e maduro conterà provavelmente alguma maquinaria específica a sustentar uma superestrutura de uso geral.

Mas os neurónios, produto de mil milhões de anos de evolução, não serão dispositivos altamente complexos e otimizados, que dificilmente conseguiremos melhorar? Não. Em primeiro lugar, grande parte do mecanismo do neurónio está relacionada com o seu crescimento e autoformação a partir do interior. Os componentes dos computadores atuais (e previsivelmente futuros) dispensam esse fardo, pois são construídos externamente. Isto constitui uma enorme vantagem — toda a estrutura pode ser usada para controlar a perceção, a ação e o pensamento. Em segundo lugar, o mecanismo básico de transmissão de informação do neurónio — a libertação de substâncias químicas que afetam as membranas exteriores de outras células — parece ser extremamente primitivo, observável até nas bactérias nadadoras mais simples. Os animais parecem estar presos a esta disposição devido às limitações do seu processo de conceção. A evolução darwiniana é uma otimizadora implacável de um dado projeto: ajusta parâmetros aqui e ali, acrescenta uma etapa acolá, remove outra além — tudo de forma morosa, como um remendador diligente. No entanto, é muito pouco apta para redesenhar fundamentalmente os seus produtos. Alterações radicais nos alicerces de uma estrutura estão fora do seu alcance, pois exigiriam demasiadas modificações simultâneas e interdependentes. Por contraste, os projetistas humanos são

notoriamente eficazes a conservar a forma geral de uma ideia, alterando, contudo, todas as suas partes. As calculadoras já foram construídas com engrenagens e alavancas, depois com relés, mais tarde com válvulas, depois com transístores e circuitos integrados. Em breve, será a luz ou supercorrentes a fluir pelo seu interior. Durante todo esse tempo, as operações fundamentais executadas mantiveram-se praticamente inalteradas, e os princípios de projeto e software desenvolvidos para um tipo de hardware transferiram-se facilmente para o seguinte.

Que pressupostos levaram à localização dos sistemas nervosos dos animais na figura da página 76 do Capítulo 2? Os dados estão listados na Tabela 1.

Tabela 1. Sistemas nervosos

Animal	Massa Cerebral gramas	Neurónios	Potência bits/sec	Capacidade bits
Caracol		10^5	10^8	10^8
Abelha		10^6	10^9	10^9
Colibri	0.1	10^7	10^{10}	10^{10}
Rato	1	10^8	10^{11}	10^{11}
Visão humana	100	10^{10}	10^{13}	10^{13}
Humano	1,500	10^{11}	10^{14}	10^{14}
Elefante	3,000	2×10^{11}	2×10^{14}	2×10^{14}
Cachalote	5,000	5×10^{11}	5×10^{14}	5×10^{14}

APÊNDICE 2

Medindo a Potência do Computador

PELA sua influência nas vendas de computadores, comparar o poder relativo e a eficácia de custo entre diferentes máquinas sempre foi uma questão polêmica. No entanto, a margem de desacordo entre os fabricantes quanto à potência das máquinas concorrentes raramente excede um fator dez, e uma razão tão pequena não altera materialmente nem a aparência dos diagramas nem as conclusões do Capítulo 2, onde se exibem escalas da ordem de um bilhão. Ainda assim, qualquer fórmula específica para estimar poder computacional pode ser gravemente iludida por um contraexemplo infeliz — ou mesmo diabólico. Por exemplo, se o poder de um computador fosse definido simplesmente pelo número de adições por segundo que pode executar, um circuito especial composto por uma matriz de somadores rápidos e mais nada — algo praticamente inútil — poderia, a custo de apenas algumas centenas de dólares, superar um supercomputador de 10 milhões. A minha intuição sobre o que constitui uma computação útil levou-me a sugerir uma medida mais subtil, mas, acredito, mais segura. Posso assegurar que, no caso de máquinas razoáveis, as minhas fórmulas fornecem números quase idênticos aos que resultariam de métodos mais simples.

Entidades que computam massivamente podem alterar os seus estados internos e os seus resultados de formas inesperadas. Podemos dizer que uma rocha parada — ou mesmo a rolar — ou a tal matriz de somadores do parágrafo anterior, praticamente não realiza computação, pois o seu comportamento é demasiado previsível; ao passo que um rato a correr num labirinto estará certamente a computar bastante. A teoria da informação de Claude Shannon baseia-se numa forma de quantificar a quantidade de surpresa — ou informação — contida numa mensagem. Quanto mais inesperada for a próxima porção da mensagem, maior a quantidade de informação que ela transporta. Vou recorrer a esta abordagem para medir a informação presente numa computação.

O poder computacional será definido como a quantidade de informação — ou surpresa — exibida por segundo enquanto uma máquina opera, isto é, à medida que muda repetidamente de um estado interno para outro. Quanto mais inesperado for o próximo estado da máquina, maior a quantidade de informação que a transição para esse estado transmite. Quantitativamente, se — tanto quanto sabemos — existe uma probabilidade p de a máquina entrar num determinado estado, então, se de facto entra nesse estado, a transição terá transmitido $-\log_2 p$ bits de informação (\log_2 significa logaritmo na base 2, e os bits são dígitos binários). Se a probabilidade p for $1/2$, a transição transmite 1 bit de informação. Se p for $1/2^n$, então transmite n bits. Uma p de $1/1000$ corresponde a cerca de 10 bits de informação.

A média de informação transmitida por uma transição é obtida multiplicando a quantidade de informação que cada transição para um estado possível transporta pela respetiva probabilidade de ocorrer, somando depois sobre todas as possibilidades:

$$\text{Informações por transição} = \sum_{i=1}^N -p_i \log_2 p_i \text{ bits}$$

onde N é o número de estados possíveis, e p_i é a probabilidade de que o i -ésimo estado seja o próximo. O valor computacional de uma dada transição pode ser diferente para diferentes observadores, uma vez que cada um atribui probabilidades distintas aos resultados. A quantidade de informação atinge um máximo de $\log_2 N$ para um observador totalmente ignorante, para quem todos os p_i são iguais. No outro extremo, uma testemunha omnisciente pode ter a certeza de que o próximo estado será j , fazendo com que $p_j = 1$ e todos os outros $p_i = 0$, o que reduz a informação a zero. Uma máquina realiza uma computação útil para si apenas se não souber de antemão todas as respostas!

Computar requer longas sequências deste tipo de transições de um estado para outro. A capacidade total de informação de um sistema é \log_2 do número total de estados a que pode eventualmente aceder. Num computador de uso geral, isso

corresponde simplesmente ao tamanho total da memória. Uma máquina poderosa é capaz de percorrer estados rapidamente. Eu meço o poder de processamento dividindo a informação de transição pelo tempo médio necessário para cada transição. Isso dá-nos a fórmula:

$$\text{Potência} = \frac{\sum -p_i \log_2 p_i}{\sum p_i t_i}$$

As unidades são bits por segundo. Esta medida é também reduzida pela previsibilidade.

As fórmulas capturam diversas ideias. Um computador que repete infinitamente um ciclo de programa torna-se totalmente previsível, e o seu poder de computação desce para zero. Programas escritos em linguagens de alto nível ou que usam ambientes interativos tendem a correr muito mais lentamente do que se fossem escritos diretamente em linguagem de máquina. As construções das linguagens de alto nível são convertidas em sequências estereotipadas de instruções de máquina, tornando o programa mais previsível do que um escrito em linguagem de máquina pura, o que reduz o poder de processamento efetivo. Adicionar memória a um computador aumenta modestamente o seu poder, mesmo sem qualquer aumento da sua velocidade bruta. Entre as técnicas de uso da memória para potenciar o cálculo estão as tabelas de resultados previamente computados e reorganizações de estruturas de dados que ocupam mais espaço, mas são mais rápidas de aceder. Este efeito — de a memória aumentar o poder computacional — aparece na minha medida porque, numa instrução computacional, a identidade da localização de memória referida é tão surpreendente como a operação a ser executada. À medida que aumenta o número de localizações possíveis, também aumenta a surpresa — ainda que de forma modesta. Duplicar a memória aumenta o poder apenas um bit por unidade de tempo de instrução.

Em máquinas altamente paralelas — especialmente aquelas que usam um único fluxo de instruções para controlar várias unidades de processamento — a maior parte da surpresa reside nos dados paralelos, não nas instruções. Embora o número total de bits

nos fluxos de dados represente um limite superior ao poder de processamento deste tipo de máquinas, o poder real pode ser consideravelmente inferior, devido a redundâncias ou previsibilidades intrínsecas. Estimar estas quantidades em máquinas díspares, concebidas para análises por elementos finitos, processamento simbólico, autómatos celulares ou para jogar xadrez, é difícil. A figura da página 79 não é muito afetada por esta dificuldade, porque quase todas as máquinas nela representadas são do tipo convencional de von Neumann, onde apenas um dado é processado por instrução. Versões futuras do gráfico poderão ter de lidar com máquinas massivamente paralelas, algumas das quais estão agora a ser testadas com problemas reais.

Mesmo com arquiteturas convencionais, as medições de potência tornam-se confusas quando as fórmulas são aplicadas a computadores reais. Como se podem atribuir probabilidades a diferentes tipos de instruções quando cada tipo de programa apresenta as suas próprias estatísticas? Em computadores com conjuntos de instruções extensos, muitas operações são praticamente nunca utilizadas. Além disso, descrições detalhadas são difíceis de obter para muitas máquinas antigas. O meu compromisso é assumir que cada máquina usa 32 operações distintas (o equivalente a seis bits), misturadas em proporções iguais. Se cada localização de memória tiver igual probabilidade de ser acedida por uma instrução, então a informação que contribui é igual ao logaritmo do tamanho da memória. Este é um limite superior. Uma vez que os conteúdos das localizações de memória também podem mudar, os dados aí armazenados são igualmente uma fonte de surpresa — mas apenas se forem lidos, e não simplesmente sobrepostos. Se assumirmos que metade das instruções lê dados, esse canal contribui com, no máximo, metade do tamanho de uma palavra de informação. Numa máquina paralela controlada por um único fluxo de instruções, seria considerado o tamanho agregado em bits dos fluxos de dados paralelos — e este seria o principal componente da informação total.

Outro aspeto é o tempo de execução. Mais uma vez, obter dados pormenorizados é difícil. No entanto, há dois valores que estão geralmente disponíveis: o tempo médio de execução de uma

adição e o de uma multiplicação. A adição é típica da operação computacional mais rápida, enquanto a multiplicação é das mais lentas. Assumo que a mistura de instruções contém sete instruções com a duração de uma adição para cada instrução com a duração de uma multiplicação.

Com estas aproximações, a fórmula do poder computacional torna-se

$$\text{Potência} = \frac{6 + \log_2 \text{memória} + \text{palavra}/2}{(7 \times T_{\text{adição}} + T_{\text{multiplicado}})/8}$$

onde "*memória*" é a capacidade da memória rápida da máquina, em palavras individualmente endereçáveis, e "*palavra*" é o tamanho de uma palavra de dados, em bits. A capacidade da máquina obtém-se multiplicando "*memória*" por "*palavra*". Para máquinas decimais, o número de bits é aproximado multiplicando o número de dígitos decimais por quatro. Esta fórmula foi utilizada para traçar os pontos representados na figura da página 79, e os dados correspondentes encontram-se listados na Tabela 2.

Uma Metáfora Náutica

Defini dois atributos essenciais para que haja pensamento ou computação interessante: o poder computacional e a capacidade. Basicamente, o poder é a velocidade da máquina, e a capacidade é o tamanho da sua memória. A computação pode ser comparada a uma viagem marítima num barco a motor. A velocidade com que a viagem pode ser concluída depende da potência do motor do barco. O comprimento máximo de qualquer viagem é limitado pela capacidade do depósito de combustível. A velocidade efetiva é, em geral, reduzida se o percurso do barco estiver limitado — por exemplo, se tiver de navegar estritamente para este/oeste ou norte/sul, em vez de poder seguir em linha reta para o destino. Algumas computações assemelham-se a uma travessia para um ponto conhecido numa costa longínqua; outras são como uma busca, sem mapa, por uma ilha perdida. A computação paralela é como uma frota de pequenos barcos: ajuda em buscas e em

alcançar múltiplos objetivos, mas pode não ser tão útil em problemas que exijam um sprint (corrida) para um único destino distante. Máquinas de propósito específico trocam um motor maior por um controlo de leme mais limitado. Ligar discos e fitas a um computador é como adicionar depósitos de combustível secundários ao barco. A capacidade — e, portanto, o alcance — é aumentada, mas se a canalização de ligação for demasiado estreita, limitará o fluxo de combustível e, com isso, o poder efetivo do motor. Os dispositivos de entrada/saída são como velas do barco. Captam poder e capacidade do ambiente. A informação externa é uma fonte de variabilidade e, portanto, de poder, segundo a nossa definição. De forma mais concreta, pode conter respostas que, de outro modo, teriam de ser calculadas. O meio externo pode também funcionar como memória suplementar, aumentando a capacidade.

Tabela 2. Máquinas de Calcular, por Ano

Custo 1988\$	Memória palavras	Palavra bits	T_{som} Segs	T_{mult} Segs	Potência bits/secs	Capacidade bits	Potência/Custo b/s/\$
Humano							
1×10^5	2×10^1	40	6×10^1	6×10^2	2×10^{-1}	8×10^2	2×10^{-6}
1891 — Ohdner (mecânico)							
1×10^5	6×10^{-2}	20	1×10^2	6×10^2	7×10^{-2}	1×10^0	5×10^{-7}
1900 — Steiger Millionaire (mecânico)							
1×10^5	1×10^{-1}	24	5×10^1	1×10^2	3×10^{-1}	3×10^0	2×10^{-6}
1908 — Tabulador Hollerith (mecânico)							
5×10^5	8×10^1	30	5×10^1	2×10^2	4×10^{-1}	2×10^3	7×10^{-7}
1910 — Motor Analítico (mecânico)							
9×10^6	1×10^3	200	9×10^0	6×10^1	8×10^0	2×10^5	8×10^{-7}
1911 — Calculadora Monroe (mecânica)							
4×10^5	1×10^0	24	3×10^1	1×10^2	4×10^{-1}	2×10^1	1×10^{-6}
1919 — IBM Tabulator (mecânico)							
1×10^5	5×10^0	40	5×10^0	2×10^2	8×10^{-1}	2×10^2	9×10^{-6}
1920 — Aritmómetro de Torres (relé)							
1×10^5	2×10^0	20	1×10^1	1×10^2	7×10^{-1}	4×10^1	7×10^{-6}
1928 — National-Ellis 3000 (mecânico)							
1×10^5	1×10^0	36	1×10^1	6×10^1	1×10^0	4×10^1	1×10^{-5}

Custo 1988\$	Memória palavras	Palavra bits	T_{som} Segs	T_{mult} Segs	Potência bits/secs	Capacidade bits	Potência/Custo b/s/\$
1929 — Burroughs Classe 16 (mecânica)							
1x10 ⁵	1x10 ⁰	36	1x10 ¹	6x10 ¹	1x10 ⁰	4x10 ¹	1x10 ⁻⁵
1938 — Zuse -1 (mecânico)							
9x10 ⁴	2x10 ¹	16	1x10 ¹	1x10 ²	8x10 ⁻¹	3x10 ²	1x10 ⁻⁵
1939 — Zuse -2 (relé e mecânico)							
9x10 ⁴	2x10 ¹	16	1x10 ⁰	1x10 ¹	8x10 ⁰	3x10 ²	1x10 ⁻⁴
1939 — BTL Modelo 1 (relé)							
4x10 ⁵	4x10 ⁰	8	3x10 ⁻¹	3x10 ⁻¹	4x10 ¹	3x10 ¹	9x10 ⁻⁵
1941 — Zuse - 3 (relé e mecânico)							
4x10 ⁵	6x10 ¹	32	5x10 ⁻¹	2x10 ⁰	4x10 ¹	2x10 ³	1x10 ⁻⁴
1943 — BTL Modelo 2 (relé)							
3x10 ⁵	5x10 ⁰	20	3x10 ⁻¹	5x10 ⁰	2x10 ¹	1x10 ²	6x10 ⁻⁵
1943 — Colosso (tubo de vácuo)							
6x10 ⁵	2x10 ⁰	10	2x10 ⁻⁴	2x10 ⁻²	4x10 ³	2x10 ¹	7x10 ⁻³
1943 — BTL Modelo 3 (relé)							
1x10 ⁶	2x10 ¹	24	3x10 ⁻¹	1x10 ⁰	6x10 ¹	4x10 ²	4x10 ⁻⁵
1944 — ASCC (Mark 1) (estafeta)							
2x10 ⁶	7x10 ¹	70	3x10 ⁻¹	6x10 ⁰	5x10 ¹	5x10 ³	2x10 ⁻⁵
1945 — Zuse - 4 (estafeta)							
3x10 ⁵	6x10 ¹	32	5x10 ⁻¹	2x10 ⁰	4x10 ¹	2x10 ³	1x10 ⁻⁴
1946 — BTL Modelo 5 (relé)							
3x10 ⁶	4x10 ¹	28	3x10 ⁻¹	1x10 ⁰	7x10 ¹	1x10 ³	2x10 ⁻⁵
1946 — ENIAC (tubo de vácuo)							
3x10 ⁶	2x10 ¹	40	2x10 ⁻⁴	3x10 ⁻³	6x10 ⁴	8x10 ²	2x10 ⁻²
1947 — Harvard Mark 2 (estafeta)							
1x10 ⁶	1x10 ²	40	2x10 ⁻¹	7x10 ⁻¹	1x10 ²	4x10 ³	9x10 ⁻⁵
1948 — IBM SSEC (tubo de vácuo e relé)							
2x10 ⁶	8x10 ⁰	48	3x10 ⁻⁴	2x10 ⁻²	1x10 ⁴	4x10 ²	6x10 ⁻³
1949 — EDSAC (tubo de vácuo)							
4x10 ⁵	5x10 ²	35	3x10 ⁻⁴	3x10 ⁻³	5x10 ⁴	2x10 ⁴	1x10 ⁻¹
1950 — SEAC (tubo de vácuo)							
3x10 ⁶	1x10 ³	45	2x10 ⁻⁴	2x10 ⁻³	8x10 ⁴	5x10 ⁴	2x10 ⁻²
1951 — UNIVAC I (tubo de vácuo)							

4×10^6	1×10^3	44	1×10^{-4}	2×10^{-3}	1×10^5	4×10^4	3×10^{-2}
Custo 1988\$	Memória palavras	Palavra bits	T_{som} Segs	T_{mult} Segs	Potência bits/secs	Capacidade bits	Potência/Custo b/s/\$
1952 — Zuse - 5 (estafeta)							
4×10^5	6×10^1	32	1×10^{-1}	5×10^{-1}	2×10^2	2×10^3	5×10^{-4}
1952 — IBM CPC (válvula de vácuo e relé)							
4×10^5	9×10^0	144	8×10^{-4}	1×10^{-2}	4×10^4	1×10^3	9×10^{-2}
1953 — IBM 650 (válvula de vácuo)							
8×10^5	1×10^3	40	7×10^{-4}	1×10^{-2}	2×10^4	4×10^4	2×10^{-2}
1954 — EDVAC (tubo de vácuo)							
2×10^6	1×10^3	44	9×10^{-4}	3×10^{-3}	3×10^4	5×10^4	2×10^{-2}
1955 — Whirlwind (válvula de vácuo)							
8×10^5	2×10^3	16	2×10^{-5}	3×10^{-5}	1×10^6	3×10^4	2×10^0
1955 — Librascope LGP-30 (tubo de vácuo)							
1×10^5	4×10^3	30	3×10^{-4}	2×10^{-2}	1×10^4	1×10^5	1×10^{-1}
1955 — IBM 704 (válvula de vácuo)							
8×10^6	8×10^3	36	1×10^{-5}	2×10^{-4}	1×10^6	3×10^5	1×10^{-1}
1959 — IBM 7090 (transistor)							
1×10^7	3×10^4	36	4×10^{-6}	2×10^{-5}	7×10^6	1×10^6	6×10^{-1}
1960 — IBM 1620 (transistor)							
7×10^5	2×10^4	5	6×10^{-4}	5×10^{-3}	2×10^4	1×10^5	3×10^{-2}
1960 — DEC PDP-1 (transistor)							
5×10^5	8×10^3	18	1×10^{-5}	2×10^{-5}	2×10^6	1×10^5	5×10^0
1961 — Atlas (transistor)							
2×10^7	4×10^3	48	1×10^{-6}	5×10^{-6}	3×10^7	2×10^5	2×10^0
1962 — Burroughs 5000 (transistor)							
4×10^6	2×10^4	13	1×10^{-5}	4×10^{-5}	2×10^6	2×10^5	5×10^{-1}
1964 — DEC PDP-6 (transistor)							
1×10^6	2×10^4	36	1×10^{-5}	2×10^{-5}	3×10^6	6×10^5	3×10^0
1964 — CDC 6600 (transistor)							
2×10^7	5×10^5	64	3×10^{-7}	5×10^{-7}	2×10^8	3×10^7	1×10^1
1965 — IBM 1130 (chip híbrido)							
4×10^5	8×10^3	16	8×10^{-6}	4×10^{-5}	2×10^6	1×10^5	6×10^0
1966 — IBM 360/75 (chip híbrido)							
2×10^7	2×10^6	32	8×10^{-7}	2×10^{-6}	5×10^7	6×10^7	3×10^0

Custo 1988\$	Memória palavras	Palavra bits	T_{som} Segs	T_{mult} Segs	Potência bits/secs	Capacidade bits	Potência/Custo b/s/\$
1967 — IBM 360/65 (chip híbrido)							
1x10 ⁷	1x10 ⁶	32	2x10 ⁻⁶	3x10 ⁻⁶	2x10 ⁷	3x10 ⁷	2x10 ⁰
1968 — DEC PDP-10 (circuito integrado)							
2x10 ⁶	1x10 ⁵	36	2x10 ⁻⁶	1x10 ⁻⁵	1x10 ⁷	5x10 ⁶	8x10 ⁰
1969 — CDC 7600 (transistor)							
3x10 ⁷	1x10 ⁶	64	1x10 ⁻⁷	2x10 ⁻⁷	5x10 ⁸	6x10 ⁷	2x10 ¹
1970 — GE-635 (transistor)							
6x10 ⁶	1x10 ⁵	32	2x10 ⁻⁶	1x10 ⁻⁵	1x10 ⁷	4x10 ⁶	2x10 ⁰
1971 — SDS 920 (transistor)							
3x10 ⁵	6x10 ⁴	32	2x10 ⁻⁵	3x10 ⁻⁵	2x10 ⁶	2x10 ⁶	7x10 ⁰
1972 — IBM 360/195 (chip híbrido)							
2x10 ⁷	1x10 ⁵	32	1x10 ⁻⁷	2x10 ⁻⁷	3x10 ⁸	4x10 ⁶	1x10 ¹
1973 — Data General Nova (circuito integrado)							
3x10 ⁴	8x10 ³	16	2x10 ⁻⁵	4x10 ⁻⁵	1x10 ⁶	1x10 ⁵	5x10 ¹
1974 — IBM-370/168 (circuito integrado)							
4x10 ⁶	3x10 ⁵	32	2x10 ⁻⁷	4x10 ⁻⁷	2x10 ⁸	8x10 ⁶	4x10 ¹
1975 — DEC PDP-11/70 (circuito integrado)							
3x10 ⁵	6x10 ⁴	16	3x10 ⁻⁶	9x10 ⁻⁶	8x10 ⁶	1x10 ⁶	3x10 ¹
1976 — Apple II (circuito integrado)							
6x10 ³	8x10 ³	8	1x10 ⁻⁵	4x10 ⁻⁵	2x10 ⁶	6x10 ⁴	3x10 ²
1977 — Cray-1 (circuito integrado)							
2x10 ⁷	4x10 ⁶	64	2x10 ⁻⁸	2x10 ⁻⁸	3x10 ⁹	3x10 ⁸	2x10 ²
1979 — DEC VAX 11/780 (microprocessador)							
3x10 ⁵	2x10 ⁶	32	2x10 ⁻⁶	3x10 ⁻⁶	2x10 ⁷	6x10 ⁷	8x10 ¹
1980 — Sun-2 (microprocessador)							
4x10 ⁴	3x10 ⁵	32	3x10 ⁻⁶	1x10 ⁻⁵	1x10 ⁷	8x10 ⁶	3x10 ²
1981 — CDC Cyber-205 (circuito integrado)							
1x10 ⁷	4x10 ⁶	32	3x10 ⁻⁸	3x10 ⁻⁸	1x10 ⁹	1x10 ⁸	1x10 ²
1982 — IBM PC (microprocessador)							
3x10 ³	2x10 ⁴	16	4x10 ⁻⁶	2x10 ⁻⁵	5x10 ⁶	4x10 ⁵	2x10 ³
1982 — Sun-2 (microprocessador)							
2x10 ⁴	5x10 ⁵	32	2x10 ⁻⁶	6x10 ⁻⁶	1x10 ⁷	2x10 ⁷	6x10 ²
1983 — Vax 11/750 (microprocessador)							

6×10^4	1×10^6	32	2×10^{-6}	1×10^{-5}	2×10^7	3×10^7	3×10^2
Custo 1988\$	Memória palavras	Palavra bits	T_{som} Segs	T_{mult} Segs	Potência bits/secs	Capacidade bits	Potência/Custo b/s/\$
1984 — Apple Macintosh (microprocessador)							
2×10^3	3×10^4	32	3×10^{-6}	2×10^{-5}	8×10^6	1×10^6	3×10^3
1984 — Vax 11/785 (microprocessador)							
2×10^5	4×10^3	32	7×10^{-7}	1×10^{-6}	5×10^7	1×10^5	2×10^2
1985 — Cray-2 (circuito integrado)							
1×10^7	3×10^8	64	4×10^{-9}	4×10^{-9}	2×10^{10}	2×10^{10}	1×10^3
1956 — Sun-3 (microprocessador)							
1×10^4	1×10^6	32	9×10^{-7}	2×10^{-6}	4×10^7	3×10^7	4×10^3
1956 — DEC VAX 8650 (microprocessador)							
1×10^5	4×10^6	32	2×10^{-7}	6×10^{-7}	2×10^8	1×10^8	1×10^3
1987 — Apple Mac II (microprocessador)							
3×10^3	5×10^5	32	1×10^{-6}	2×10^{-6}	4×10^7	2×10^7	1×10^4
1987 — Sun-4 (microprocessador)							
1×10^4	4×10^6	32	2×10^{-7}	4×10^{-7}	2×10^8	1×10^8	2×10^4
1989 — Cray-3 (arsenieto de gálio)							
1×10^7	1×10^7	64	6×10^{-10}	6×10^{-10}	1×10^{11}	6×10^8	1×10^4

APÊNCIDE 3

Os Limites Externos da Computação

Mente sem Máquina?

No Capítulo 4, sugeri que uma mente é um padrão que pode ser impresso em muitos tipos diferentes de corpo ou meio de armazenamento. Fui mais longe ao afirmar que uma mente poderia ser representada por qualquer um de uma classe infinita de padrões radicalmente diferentes, equivalentes apenas num certo sentido matemático abstrato. As experiências subjetivas de uma pessoa são uma propriedade abstrata partilhada por todos os padrões dessa classe, pelo que a pessoa sentiria o mesmo independentemente de qual padrão estivesse a ser instanciado. Isso conduz à questão: se uma mente é, em última instância, uma abstração matemática, por que necessita ela de uma forma física? As propriedades matemáticas não existem, mesmo que não estejam escritas em lado nenhum? O biliésimo dígito de pi não existe, mesmo que ainda não o tenhamos conseguido calcular? Nesse mesmo sentido, os relacionamentos matemáticos abstratos que constituem os sentimentos de uma pessoa não existem mesmo na ausência de qualquer hardware que os compute? Trata-se de um velho dilema filosófico; e, embora eu pense que deve ser verdade, não vejo como tirar daí conclusões significativas, pois parece implicar que tudo o que é possível existe. Em vez disso, consideremos um raciocínio um pouco menos abrangente.

Suponhamos que um programa que descreve uma pessoa é escrito num meio estático, como um livro. Um ser superinteligente que leia e compreenda o programa deverá ser capaz de raciocinar sobre o desenvolvimento futuro da pessoa codificada, em diversas situações possíveis. A existência no pensamento de um observador inteligente não é fundamentalmente diferente da existência numa simulação computacional — e já sugerimos que uma mente pode ser satisfatoriamente codificada num computador. Contudo, um ser inteligente pode fazer mais com a simulação do que simplesmente

executá-la de forma rígida — não tem de modelar com precisão todos os detalhes do simulado e pode bem optar por ignorar partes aborrecidas, saltar para conclusões que lhe são óbvias, aproximar outros passos e agrupar alternativas que não lhe interesse distinguir. Autores humanos de ficção fazem isto todos os dias, ao criarem aventuras para as personagens dos seus livros — o nosso ser superinteligente distingue-se apenas pelo facto de a sua imaginação operar a um nível de detalhe tal que as suas pessoas simuladas são totalmente reais. Tal como um autor de ficção, este ser pode pensar de forma invertida no tempo; pode escolher uma conclusão e depois raciocinar para trás, decidindo o que teria de a preceder. Talvez esse ser superinteligente prefira imaginar certos tipos de situações e procure manobrar as suas simulações mentais para as fazer acontecer. Ao omitir detalhes irrelevantes da simulação e dirigir os acontecimentos para determinadas conclusões, esse ser pode criar na simulação peculiaridades suficientes para atrair a atenção da pessoa simulada.

A mecânica quântica, pedra angular da física moderna, parece implicar que, no mundo real tal como o conhecemos, eventos não observados acontecem de todas as maneiras possíveis (ou seja, não se decide qual das possibilidades ocorre), e a sobreposição de todas essas possibilidades tem efeitos observáveis, incluindo coincidências misteriosas em tempos e lugares remotos. Existirá alguma ligação entre estas ideias? Uma vez mais, conduzi-me a um dilema — embora, desta vez, seja um que talvez possa vir a ser resolvido.

Pensamento Não Determinístico

Os universos de autómatos celulares apresentados no Capítulo 6 possuem uma rigidez mecânica que os torna fáceis de conceber. Mas a física moderna revelou fundamentos muito mais interessantes para o nosso próprio universo. O que tudo isto realmente significa continua a ser matéria de fascinante especulação; o único consenso é que a verdade é profundamente estranha. O que se segue é um pequeno exercício de autoindulgência — uma tentativa de aproveitar essa estranheza

para um fim prático. Naturalmente, a realidade revelar-se-á ainda mais insólita.

Antes dos computadores, os cálculos científicos sérios eram todos difíceis, demorados e propensos a erros. Os problemas práticos de engenharia eram abordados de forma *ad hoc* (*para este efeito*), com tabelas numéricas, métodos gráficos e engenhosos dispositivos como régua de cálculo e, neste século, calculadoras mecânicas. Os computadores automáticos mudaram radicalmente esta situação. Embora conseguissem lidar com problemas de dimensão astronómica, exigiam, previamente, uma especificação absolutamente precisa e detalhada, passo a passo, do procedimento a seguir. Escrever tais programas — e fazê-los corretamente — era um processo lento e moroso (ainda que muito menos demorado do que efetuar os cálculos manualmente). Era uma grande vantagem poder usar o mesmo programa para muitos problemas do mesmo tipo, e muitas vezes essa generalização tornava mais clara a própria estrutura do cálculo. Isto deu também origem a um novo campo matemático: a *complexidade computacional* — o estudo da dificuldade intrínseca de obter respostas para diferentes tipos de problemas através do cálculo.



Penso, Logo Existo. Um Descartes simulado deduz corretamente a sua própria existência. Não faz diferença quem ou o que está a fazer a simulação — o mundo simulado é completo em si mesmo.

Um programa geral pode lidar com problemas de diferentes dimensões, e, geralmente, o seu tempo de execução aumentará à medida que cresce a dimensão do problema. Os programas de ordenação mais simples, por exemplo, demoram um tempo proporcional ao quadrado do número de elementos a ordenar. Se se duplicar o número de elementos, o tempo de ordenação quadruplica. Foram descobertos programas mais sofisticados cujo tempo de execução cresce mais lentamente, e os estudiosos da complexidade computacional conseguiram demonstrar que os programas de ordenação mais rápidos possíveis exigiriam um número de passos proporcional ao número de elementos multiplicado pelo seu logaritmo — ordenar um milhão de elementos deveria demorar cerca de dez mil vezes mais do que ordenar mil. O tempo cresce mais rapidamente do que o número de elementos, mas não tanto quanto o seu quadrado. Demonstrou-se que a dificuldade de outros tipos de problemas cresce com o cubo da dimensão, ou com a quarta potência, ou outra qualquer. Os problemas cuja solução pode ser computada em tempos proporcionais (ou inferiores) a uma dada potência fixa da dimensão do problema dizem-se de tempo polinomial, ou tipo P. Na Era dos computadores, são considerados fáceis.

Outra classe importante de cálculos revelou-se muito mais difícil. Um exemplo é o chamado problema do caixeiro-viajante, que consiste em encontrar o caminho mais curto que passe por cada uma de um dado conjunto de cidades exatamente uma vez. As melhores soluções exatas conhecidas exigem que o programa verifique praticamente todas as rotas possíveis e escolha a melhor. As rotas possíveis são enumeradas gerando todas as permutações das cidades. Cada permutação pode ser verificada em tempo polinomial, mas o número de permutações cresce exponencialmente, multiplicando-se muitas vezes cada vez que se acrescenta uma cidade. O problema torna-se astronomicamente difícil para números bastante modestos de cidades. Um computador

hipotético capaz de explorar todos os caminhos possíveis em simultâneo (uma mera abstração matemática, conhecida como máquina não determinista, porque não toma decisões nos pontos de ramificação, mas divide-se em duas máquinas e segue ambos os caminhos) poderia, em princípio, resolver o problema em tempo polinomial. Por essa razão, estes problemas difíceis são designados por NP, de *nondeterministic polynomial* (*polinómio não determinístico*). Verifica-se que muitos problemas NP podem ser convertidos matematicamente uns nos outros, e que uma solução rápida (de tipo P) para um deles resolveria todos. Infelizmente, não foi ainda encontrada nenhuma solução rápida que funcione num computador real e determinista, e não se sabe sequer se tal solução existe. Esta questão — se $P = NP$ — é central, pois há problemas de extrema importância, por exemplo na conceção de hardware e software ótimos, e na inferência automática, que são NP. Como as soluções exatas são muito, muito lentas, os sistemas atuais de conceção vão avançando com métodos aproximados que não garantem a melhor resposta — e, por vezes, falham redondamente.

Suponhamos, como parece provável, que não existe atalho para resolver problemas NP. Independentemente da rapidez da máquina convencional que utilizemos, pequenos aumentos na dimensão do problema serão suficientes para a sobrecarregar. Uma inteligência empenhada em aperfeiçoar-se a si própria encontrará muitos problemas NP. A eficiência das suas soluções — e, portanto, o seu próprio futuro — dependerá diretamente da sua capacidade para lidar com estes problemas, pelo que se justificam métodos hercúleos.

Replicação e Não Determinismo Quântico

Um computador não determinista, capaz de gerar indefinidamente versões de si próprio para examinar respostas alternativas, é uma construção matemática. Pode ser aproximado, até certo ponto, por um multiprocessador contendo um número fixo de processadores distintos. Em cada ponto de ramificação do cálculo, é invocado um novo processador. Contudo, eventualmente o último processador estará ocupado, e as ramificações

subsequentes terão de ser avaliadas sequencialmente. O cálculo só pode ser acelerado no máximo tantas vezes quantos forem os processadores disponíveis — o que tem impacto quase nulo perante o crescimento astronómico dos problemas NP. Mas e se o número de processadores pudesse, de algum modo, aumentar com o crescimento da dimensão do problema?

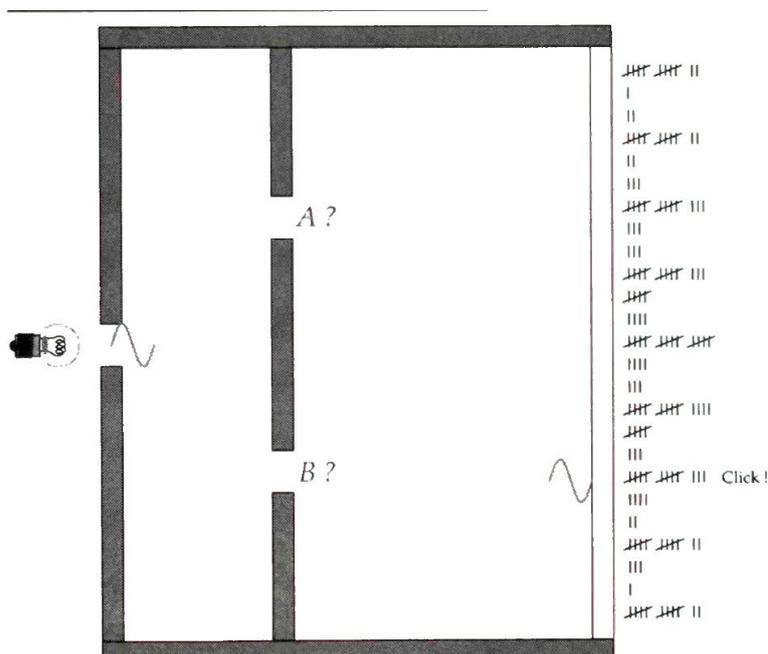
A capacidade de reprodução é uma condição *sine qua non* (*indispensável*) para qualquer espécie que aspire à imortalidade. As nossas máquinas já o fazem, embora os humanos biológicos sejam ainda parte essencial do processo. À medida que o futuro descrito neste livro se concretiza, mais e mais etapas serão assumidas pelas máquinas, até que a produção automática de máquinas que produzem automaticamente se torne norma. Num futuro superinteligente, a reprodução de pequenas unidades pensantes será brincadeira de criança. É, pois, razoável imaginar pequenos computadores capazes tanto de calcular como de se reproduzirem. A reprodução é, em última análise, limitada pela quantidade finita de material e energia disponível, mas estes limites podem ser tornados astronómicos através da criação de máquinas muito pequenas num meio nutriente muito vasto e energético (máquinas do tamanho de bactérias nos oceanos de Júpiter, ou em nuvens de poeira interestelar alimentadas por pulsares, por exemplo!). Um desses computadores reproduzir-se-ia para se tornar dois, os dois tornar-se-iam quatro, os quatro oito, os oito dezasseis, e assim sucessivamente — um crescimento exponencial até atingir números astronómicos que acompanham, até certo ponto, o crescimento exponencial de um problema NP. Se, num dado intervalo de tempo ("geração"), uma máquina puder ou calcular uma certa quantidade ou reproduzir-se uma vez, a melhor estratégia será reproduzir-se desenfreadamente até haver tantas máquinas quantas as alternativas a examinar. Depois, cada uma das máquinas resultantes examina uma alternativa. Tendo calculado uma resposta possível, as máquinas iniciariam então um torneio do tipo "a minha é melhor do que a tua". Duas máquinas comparariam os respetivos resultados, e ambas adotariam o melhor dos dois antes de partirem para competir com outras. A melhor resposta propagar-se-ia rapidamente a todas as máquinas, e qualquer uma delas

poderia então ser recolhida para obter o resultado final. Neste processo, a fase de reprodução consome um tempo linearmente proporcional à dimensão do problema; o cálculo das respostas individuais é polinomial em função da dimensão; e a fase de torneio de respostas, como a reprodução, também é linear, permitindo que problemas NP de dimensão moderada sejam resolvidos em tempo polinomial. Contudo, problemas de grande escala submergiriam qualquer quantidade de espaço disponível — e, assim, continuariam fora do nosso alcance. (A biosfera da Terra, claro está, está a realizar um cálculo precisamente deste tipo.)

Computadores verdadeiramente não deterministas são uma ficção matemática. Ou será que não? A mecânica quântica, pedra angular da física moderna, tem o indeterminismo no seu cerne. As probabilidades de resultado em mecânica quântica são previstas somando todas as maneiras indistinguíveis pelas quais um evento pode ocorrer, e depois elevando o resultado ao quadrado. Uma consequência estranha disto é que alguns resultados que seriam, de outro modo, possíveis, são eliminados pela mera existência de outras possibilidades. Um exemplo excelente é a experiência das duas fendas.

Fótons de luz são emitidos a partir de uma fonte pontual em direção a um écran atravessado por duas fendas, conforme ilustrado na figura da página 223. Aqueles que passam pelas fendas encontram uma matriz de detetores de fótons (frequentemente um filme fotográfico, embora o exemplo seja mais claro se usarmos sensores individuais que fazem "clique" quando atingidos pela luz). Se a fonte de luz for tão fraca que apenas um fóton seja emitido de cada vez, os sensores reagem individualmente — ora um, ora outro. Cada fóton aterra exatamente num único local. Mas, se se fizer a contagem de quantos fótons atingiram cada detetor, surge um padrão inesperado. Alguns detetores não registam fótons nenhuns, enquanto outros, mesmo ao lado, registam muitos; um pouco mais longe, a escassez repete-se. Se for traçado um gráfico com o número de cliques de cada detetor, o resultado é idêntico ao padrão de interferência em bandas que ocorreria se ondas contínuas fossem emitidas a partir da fonte de luz e passassem pelas duas fendas, sendo aí divididas em duas ondas que interferem entre si

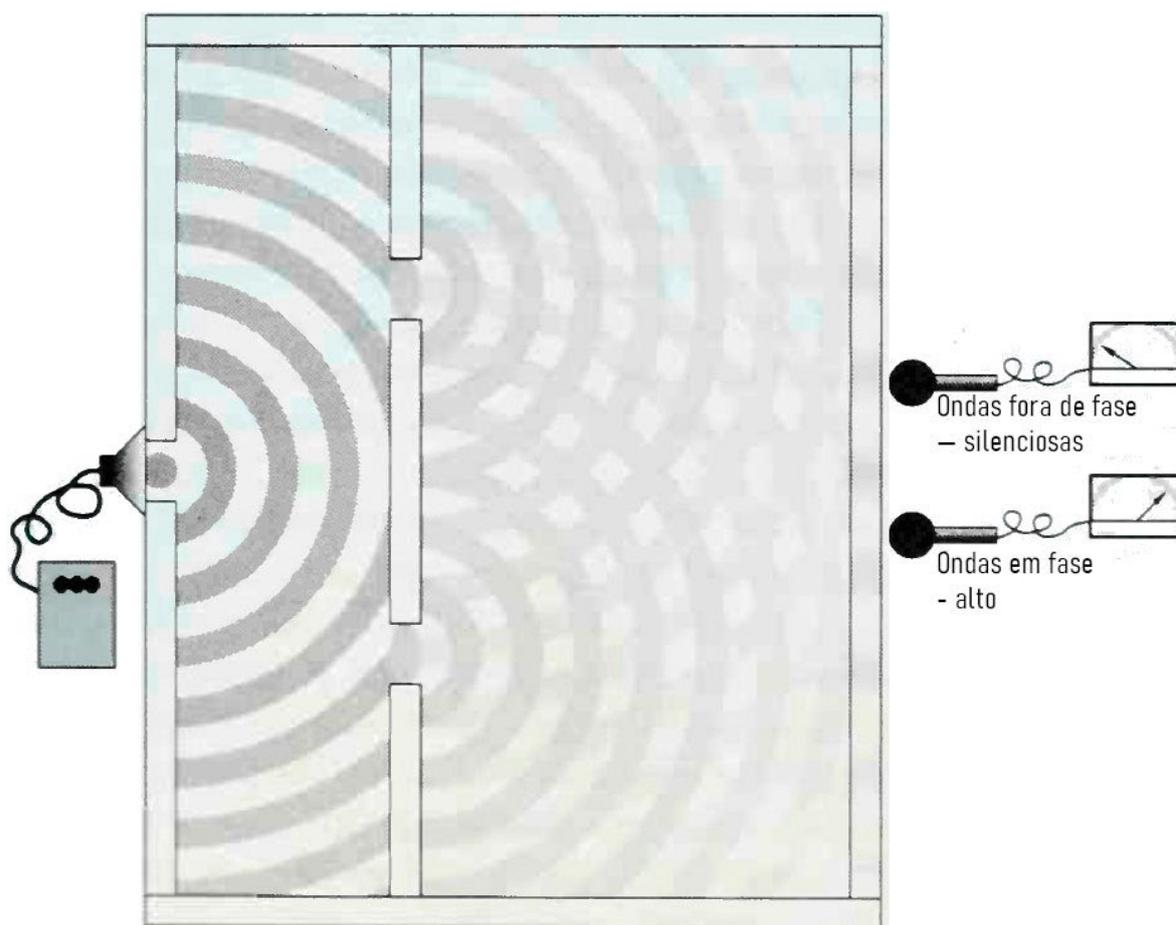
no écran — construtivamente em certos pontos, destrutivamente noutros — conforme ilustrado na figura da página 224.



Experiência das Duas Fendas. Um fóton captado por um detetor no ecrã S pode ter passado pela fenda A ou pela fenda B — não há forma de distinguir. Na mecânica quântica, as "amplitudes" para os dois casos devem ser somadas. Em alguns pontos do ecrã, somam-se construtivamente, tornando provável que um fóton acabe ali; em pontos próximos, as amplitudes cancelam-se e não é encontrado nenhum fóton.

Mas ondas *de quê?* Cada fóton parte de um ponto e chega a outro ponto; não estará ele, em cada momento do percurso, apenas num único local? Não passará ele por uma ou por outra fenda? Se assim for, como pode a *presença* da outra fenda impedi-lo de aterrar num determinado ponto do écran? Pois, se uma das fendas for tapada, o número total de fótons que atinge o écran reduz-se para metade, mas o padrão de interferência desaparece, e certos locais que não recebiam fótons com *ambas* as fendas abertas começam a registá-los. A resposta da mecânica quântica é que, durante o voo, a posição do fóton é desconhecida e deve ser modelada por uma onda de valores complexos que descreve todos os seus possíveis locais. Esta onda fantasmagórica passa por ambas as fendas (embora descreva apenas a posição de um único fóton) e

interfere consigo mesma no ecrã, anulando-se em certos pontos. É aí que a onda "toma uma decisão", e o fóton aparece num dos locais possíveis. Este estado ondulatório do fóton antes de atingir o ecrã é chamado de estado misto ou superposição de estados. O aparecimento súbito do fóton apenas num detetor é denominado *colapso da função de onda*.



Duas Fendas e Ondas. Quando as ondas sonoras passam pelas duas fendas, resulta um padrão de interferência. Mas nenhum clique individual é ouvido. Cada onda afeta suavemente todos os detetores.

Esta explicação perturbou profundamente alguns dos próprios físicos que ajudaram a formular a teoria, nomeadamente Albert Einstein e Erwin Schrödinger. Para formalizarem as suas objeções intuitivas, construíram experiências mentais que, segundo a teoria, produziam resultados improváveis. Em algumas, uma medição feita num local causava o colapso instantâneo da função de onda num ponto remoto — um efeito mais rápido do que a luz. Noutra

exemplo, mais frívolo, conhecido como o Gato de Schrödinger, um decaimento radioativo que pode ou não ocorrer dentro de uma caixa selada causa (ou não) a morte de um gato também presente na caixa. Schrödinger considerava absurda a descrição da teoria, segundo a qual a caixa fechada se encontrava num estado misto que superpunha um gato vivo e um gato morto. Ele sugeria que a teoria expressava meramente a ignorância do observador: no interior da caixa, o destino do gato era inequívoco. Esta visão é chamada de teoria das variáveis ocultas; ou seja, o sistema tem um estado definido em todos os momentos, mas certas partes desse estado estão temporariamente ocultas a alguns observadores.

Mas a piada acabou por recair sobre os críticos. Muitos dos resultados "absurdos" dessas experiências mentais foram observados com espantosa veracidade em experiências engenhosas (e muito modernas) realizadas por Alain Aspect na Universidade de Paris. Estas demonstrações descartam as teorias mais simples e naturais de variáveis ocultas — *as locais* — nas quais, por exemplo, a informação oculta sobre qual fenda o fóton atravessou estaria contida no próprio fóton, ou em que o estado de saúde do gato de Schrödinger faria parte integrante do próprio animal.

As teorias de variáveis ocultas *não locais*, onde a informação não medida está distribuída por uma extensão do espaço, permanecem uma possibilidade. É fácil construir teorias deste tipo que produzem resultados idênticos aos da mecânica quântica convencional. A maioria dos físicos, porém, não as considera interessantes: por que introduzir uma explicação mais complicada, com variáveis adicionais, quando as equações atuais mais simples são suficientes? Filosoficamente, aliás, as teorias de variáveis ocultas globais são apenas ligeiramente menos desconcertantes do que a própria mecânica quântica em bruto. Que significa dizer que a "posição exata" de uma partícula está espalhada por uma vasta região do espaço? Esta pergunta foi objeto de uma viva controvérsia entre os fundadores da mecânica quântica no início do século XX. E voltou recentemente a suscitar um interesse generalizado.

Muitos Mundos e o Computador NP do Juízo Final

Na experiência das duas fendas, um fóton destinado ao ecrã *poderá* passar pela fenda A ou pela fenda B. O padrão de interferência sugere que, de algum modo, consegue atravessar ambas ao mesmo tempo. Em 1957, Hugh Everett, da Universidade de Princeton, publicou na sua tese de doutoramento aquilo que pode ser a explicação mais exuberante, e não-local, com variáveis ocultas para este enigma. No modelo de Everett, o fóton atravessa de facto ambas as fendas — mas em universos diferentes. Em cada ponto de decisão, o universo inteiro, ou pelo menos a sua porção imediata, divide-se em vários, como folhas multiplicadas por uma fotocopiadora. Até ser feita uma medição, as diferentes versões do universo coexistem em proximidade e interferem umas com as outras, gerando, por exemplo, padrões em bandas nos ecrãs. Uma medição capaz de distinguir uma possibilidade da outra provoca a divergência dos universos (ou, alternativamente, a divergência é a própria definição de “medição”). A interferência cessa, e em cada universo, agora separado, uma versão diferente do experimentador contempla um resultado inequívoco distinto.

A palavra “astronómico” mal começa a descrever o número de universos distintos criados a cada instante, segundo esta ideia. Tal escala poderá tornar o pensamento desagradável, mas não devemos deixar-nos intimidar pela mera grandeza numérica. Choque semelhante acolheu as primeiras sugestões quanto ao número de átomos num grão de matéria, ou à distância até às estrelas mais próximas, ou ao tamanho e idade do universo. Físicos práticos, no entanto, contestam a prodigalidade de universos por outra razão. Uma vez feita a medição, os universos nos quais o resultado foi diferente do nosso deixam de nos influenciar. Postular a sua existência continuada é uma complicação desnecessária. Assim despojada, a ideia dos muitos mundos reduz-se à mecânica quântica convencional. As variáveis ocultas desnecessárias (aquelas que identificariam qual o universo em questão) são eliminadas.

Pode o indeterminismo da mecânica quântica ajudar a resolver problemas difíceis? Têm sido demonstrados os chamados métodos

holográficos, que utilizam luz laser coerente para procurar simultaneamente certos padrões (impressões digitais, sinais de radar inimigos, por exemplo) em vastos campos de outros padrões. Para o tipo certo de problema, os métodos holográficos são rápidos e eficazes, e a sua ação pode ser interpretada como o efeito de estados mistos. Também se pode vê-los de forma mais clássica, como uma interferência de ondas luminosas. As ondas são interessantes porque o seu espalhamento é uma forma de reprodução, e diferentes partes de uma onda podem ser utilizadas para realizar diferentes partes de um cálculo. Contudo, até agora apenas foram propostos ou alcançados melhoramentos lineares na velocidade. Os problemas NP continuam difíceis.

Mas a ideia dos muitos mundos tem outras consequências. John Gribbin, escritor e físico, expandiu algumas das suas possibilidades mais bizarras em várias histórias, artigos e livros. Uma dessas histórias tem o seguinte enredo.

O Dispositivo do Juízo Final

Dois engenheiros responsáveis por um futuro superacelerador de partículas (imensamente dispendioso) enfrentam um problema. A máquina está concluída há meses, mas até agora falhou em todas as tentativas de a utilizar. O problema não está no projeto, mas parece residir apenas na má sorte dos seus criadores. Um raio provocou uma falha de energia precisamente no momento da ativação; um fusível queimou; um funcionário tropeçou num cabo; um pequeno sismo acionou um sistema de emergência — cada incidente foi diferente e aparentemente sem relação com os outros.

Mas talvez estas falhas constituam um golpe de sorte colossal. Novos cálculos sugerem que a máquina possui potência suficiente para provocar um colapso do vácuo para um estado de energia inferior. Uma explosão cósmica poderia irradiar à velocidade da luz a partir do ponto de colisão no acelerador, destruindo eventualmente todo o universo. Que afortunada escapatória!

*Ou terá sido? Se o universo tivesse sido destruído, não haveria ninguém para lamentar o sucedido. E se a teoria dos muitos mundos estivesse correta? Em alguns universos, a máquina teria funcionado. Para todos os efeitos práticos, esses mundos ter-se-iam extinguido. Apenas nos restantes haveria dois físicos perplexos a coçar a cabeça, tentando perceber o que correria mal desta vez. Dado o número imenso de universos quase idênticos, a destruição de alguns poucos pareceria de pouca consequência. Surge-lhes então uma ideia. E se reforçassem os pontos fracos da máquina, tornando extremamente improvável uma falha interna aleatória, e a ligassem a um detetor de ataque nuclear, como a máquina do juízo final no filme *Dr. Strangelove*, de Stanley Kubrick? Um ataque seria respondido com a destruição do universo agressor. Apenas os universos em que o ataque não tivesse ocorrido — por qualquer razão (o general responsável sofreu um ataque cardíaco, o sistema de lançamento de mísseis falhou, o primeiro-ministro teve um súbito acesso de compaixão...) — sobreviveriam para se espantar com mais um incidente inexplicavelmente evitado. A máquina em *Dr. Strangelove* era ineficaz como elemento dissuasor a menos que o inimigo soubesse da sua existência. Já a versão baseada nos muitos mundos não sofre desse problema. Nenhum ataque (que alguém venha a presenciar) pode ocorrer enquanto ela estiver em funcionamento — por mais secreto que seja o seu funcionamento.*

Evitar uma guerra nuclear é um objetivo louvável — talvez valha o sacrifício de um número incalculável de universos possíveis. Mas será que podemos aplicar o mesmo raciocínio para resolver problemas do quotidiano? Voltando à questão dos problemas NP, será que os universos paralelos poderiam ser utilizados para explorar simultaneamente soluções alternativas, permitindo a um único computador convencional agir como se fosse um computador não determinístico? Começemos por ligar um dispositivo destruidor de universos a um computador, de forma a ser ativado por uma determinada instrução. Já agora, acrescentemos também um gerador de números verdadeiramente aleatórios, baseado, por exemplo, no ruído térmico de uma resistência ou nos estalidos de um contador Geiger. As potenciais soluções para problemas NP

caracterizam-se por um valor numérico — como o comprimento total do percurso no problema do caixeiro-viajante — que pretendemos minimizar. Adivinhemos um valor provável para esse comprimento e escrevamos um programa que escolha, ao acaso (utilizando o gerador de números aleatórios), um percurso. Se o percurso gerado for mais longo do que a nossa estimativa, o programa aciona o dispositivo apocalíptico. Executemos o programa. Muito provavelmente, o percurso gerado será mais curto do que o valor estimado, porque os universos em que surgiu um percurso mais longo foram erradicados. Mas e se não existir nenhum percurso tão curto? Nesse caso, o computador terá falhado de algum modo e não terá destruído o universo.

É incómodo que o computador falhe de forma imprevisível, por isso, por que não introduzir um ponto fraco facilmente substituível? Instalemos outro dispositivo controlado termicamente ou radioativamente que, com baixa probabilidade, possa interromper o cálculo. Ao executar o programa agora, ele ou produzirá uma resposta mais curta do que a estimativa ou será interrompido antes de terminar. No primeiro caso, tentamos um valor ainda mais curto; no segundo, escolhemos um valor mais longo. Repetimos este processo até encontrarmos um valor de percurso L tal que o computador encontra uma solução quando a estimativa é L , mas se interrompe ao tentarmos $L - 1$. A solução com comprimento L é a resposta ótima para o problema. A busca por este L pode, evidentemente, ser incorporada no próprio programa, tornando o processo completamente automático. Uma estratégia ótima para o fazer — chamada pesquisa binária — permite concluir a busca em número de passos proporcional ao logaritmo de L : apenas 20 passos se L for um milhão. Cada passo, que envolve gerar uma solução aleatória, demora tempo polinomial. Portanto, a resposta é sim: uma versão dos muitos mundos da mecânica quântica pode ser usada para resolver problemas NP arbitrariamente difíceis em tempos modestos.

O custo em universos destruídos é, no entanto, assombroso. No caso do dispositivo do juízo final usado para prevenir uma guerra nuclear, se a probabilidade anual de ataque fosse de 50%, cerca de metade dos universos possíveis seriam destruídos a cada ano. Mas

num problema do caixeiro-viajante com 100 cidades, apenas algumas soluções entre 100^{100} (um 1 seguido de 200 zeros) serão ótimas, o que significa que um computador do juízo final a resolver esse problema destruiria cerca de 100^{100} universos por cada um que deixasse sobreviver! A taxa de criação de novos universos é, contudo, muito superior a isso, por isso talvez não tenha importância.

Apenas um componente desta solução para o problema NP não pode ser atualmente construído: o próprio detonador do juízo final. O modelo de Gribbin baseia-se numa física altamente especulativa que, à data desta escrita, caiu em descrédito. Mas será mesmo necessário destruir o universo inteiro quando o resultado é errado? Claro que não. Subjetivamente, é igualmente eficaz destruir apenas a si próprio! E esse desfecho é, de facto, alcançável hoje. Por exemplo, liga a conexão do juízo final do computador a uma carga explosiva craniana. Se executar o programa do caixeiro-viajante, é praticamente certo que explodirá os miolos. Mas será rápido, e supostamente indolor. E, num universo entre 100^{100} , sobreviverá, por um tremendo golpe de sorte, e terá a resposta correta. A sua carga explosiva continuará intacta, pronta para resolver o próximo problema.

A ideia acima poderia ser aplicada a outros problemas da vida. Se um resultado que deseja, por mais improvável que seja, não se concretizar, destrói-o. Em algum universo, sobreviverá, tendo vencido a aposta. Se a interpretação dos muitos mundos da mecânica quântica estiver correta — como pode bem estar — por que não é então o suicídio uma solução comum para os problemas do quotidiano? A demografia dos múltiplos universos pode ter algo a dizer sobre isso. Se for negligente com a sua vida, haverá menos cópias de si nos universos. Um universo escolhido ao acaso conterà, na sua maioria, indivíduos que lutaram com sucesso para evitar a morte sempre que possível. Se os universos forem verdadeiramente infinitos, esse facto pouco importa — um trilião avos do infinito continua a ser o mesmo infinito ($\frac{1}{10^{12}}x\infty = \infty$). Contudo, as consequências tornam-se graves se o número for apenas enorme. E, se nem todos os desfechos possíveis forem efetivamente

realizados, então destruí-lo com elevada probabilidade pode, de facto, significar o fim da sua existência.

Um Mundo, Não Muitos?

Pode ser que a natureza indeterminada da mecânica quântica seja apenas uma espécie de ilusão, e que exista apenas um mundo. Eis um esboço de um modelo no qual as incertezas em qualquer local — ou as variáveis escondidas — são simplesmente “ruído” proveniente do resto do universo.

Imaginemos, algures, um volume esférico uniformemente preenchido com um gás composto por um número enorme, mas finito de partículas em movimento. Ondas de pressão atravessam o gás, propagando-se à sua velocidade do som, s , e suponhamos que nenhum sinal pode propagar-se mais rapidamente do que isso. A esfera apresenta ressonâncias que correspondem a feixes de ondas atravessando todo o seu volume em diferentes ângulos e frequências. Cada combinação de uma direção e frequência específicas é chamada um *modo de onda*. Existe uma transformação matemática chamada *transformada* de Fourier espacial, que organiza estes modos de onda de forma extremamente elegante e poderosa. A transformada de Fourier combina o padrão de pressões encontrado no volume original da esfera (V) de várias maneiras, produzindo um novo conjunto esférico de valores (F). No centro de F encontra-se um número que representa a densidade média de V . Imediatamente à sua volta, estão números (complexos) que expressam a intensidade de ondas, em várias direções, cuja longitude de onda corresponde à que atravessa o diâmetro de V . O dobro dessa distância ao centro de F representa os modos com duas oscilações em V ; mais além, uma casca representa os modos cuja longitude de onda é um terço do diâmetro de V , e assim sucessivamente. Cada ponto em F descreve uma onda que preenche V , com uma direção e um número de ciclos determinados pela orientação e distância ao centro de F . Ou dito de outro modo: a direção em F corresponde à direção em V ; o raio em F é proporcional à frequência em V . Dado que cada onda é composta por aglomerados periódicos de partículas de gás, o espaçamento

entre partículas impõe um limite inferior à longitude de onda, o que equivale a um limite superior à frequência — e, portanto, uma limitação ao raio da esfera F . Quanto mais próximas estiverem as partículas, maior será o necessário F .

Um teorema sobre transformadas de Fourier afirma que, se forem incluídas frequências suficientemente altas, então F conterá aproximadamente tantos pontos quantas partículas existirem em V , e toda a informação necessária para reconstruir V encontra-se em F . Na verdade, F e V são apenas descrições alternativas da mesma coisa, com a interessante propriedade de que cada partícula em V contribui para o valor de cada ponto em F , e cada ponto em F reflete-se como um componente do movimento de cada partícula de V .

Se as partículas em V colidirem entre si, ou interagirem de alguma forma não-linear, então a energia pode ser transferida de um modo de onda para outro — isto é, um ponto em F pode fortalecer-se à custa de outro. Haverá um certo grau de transferência aleatória entre todos os modos de onda. Além disso, haverá uma interação mais sistemática entre modos de onda “próximos” — aqueles muito semelhantes em frequência e orientação, e, portanto, próximos no espaço F . Em V , tais ondas estarão em fase ao longo de grandes extensões. Devido à não-linearidade do gás, o agrupamento periódico de partículas causado por um modo influenciará a capacidade de agrupamento de outro modo vizinho com período semelhante.

Assim, pontos próximos em F interagem sistematicamente; pontos distantes não o fazem. Essa interação pode ser considerada como uma física do mundo F . Se essa física for suficientemente rica, poderá sustentar a base de estruturas complexas, vida e inteligência — tal como a nossa o faz. Imaginemos um físico feito de “matéria F ”, para quem os pontos em F são simplesmente locais no espaço, e não funções complexas de outro espaço. Podemos deduzir algumas das “leis da física” que esse habitante de F descobriria, raciocinando sobre os efeitos em V e depois traduzindo-os de volta para F . Na lista seguinte, esse raciocínio será apresentado em itálico:

- **Dimensionalidade:** *F é tridimensional. Se V for tridimensional, cada feixe de ondas será descrito pela sua orientação (determinada por dois ângulos — digamos, azimute e elevação) e pela sua frequência. A frequência em V torna-se raio em F, enquanto os dois ângulos mantêm-se como ângulos em F. Se V fosse uma esfera de n dimensões, F também teria n dimensões.*
- **Localidade:** *Pontos próximos uns dos outros em F podem trocar energia de formas consistentes e previsíveis, enquanto pontos distantes não o conseguem fazer. Dois feixes de onda em V que sejam muito semelhantes em direção e frequência estão sincronizados ao longo de uma grande porção do seu comprimento, e os efeitos de agrupamento não linear serão aproximadamente iguais ciclo após ciclo ao longo do percurso. Modos de onda que estão distantes entre si, por outro lado, cujos picos e vales não estão correlacionados, terão ganhos e perdas aleatórios — parecendo, em geral, meros solavancos aleatórios uns para os outros.*
- **Velocidade de Interação:** *Há uma velocidade característica em cada ponto de F. Pontos mais afastados do centro de F interagem mais rapidamente do que os mais próximos. Uma interação é a transferência não aleatória de energia de um modo de onda para outro. A menor unidade repetitiva num feixe de ondas é um ciclo. Um efeito que ocorre de forma semelhante em cada ciclo pode ter um efeito consistente sobre todo o feixe de ondas. Os efeitos em V propagam-se à velocidade do som, portanto, um ciclo inteiro pode ser afetado no tempo que o som leva a atravessá-lo (ou seja, no período da onda). As zonas mais exteriores de F correspondem aos modos de onda com as frequências mais elevadas — e, por conseguinte, às taxas de interação mais rápidas.*
- **Princípio da Incerteza:** *A energia de um ponto em F não pode ser determinada com precisão num curto intervalo de tempo. A melhor precisão possível melhora linearmente com a*

duração da medição. *A energia num ponto de F corresponde à energia total de um determinado feixe de ondas que atravessa todo o volume V . Como nenhum sinal em V pode propagar-se a uma velocidade superior à do som, descobrir a energia total de um feixe de ondas implicaria esperar pela chegada de sinais provenientes de todas as regiões de V — um tempo muito superior ao tempo de interação básico. Num tempo reduzido, a soma necessária incide apenas sobre um volume proporcionalmente pequeno. Dado que o observador em F está ele próprio distribuído por V , não está definido exatamente qual o subvolume a considerar — e, portanto, a medição é incerta. À medida que o tempo e o volume de integração aumentam, todas as somas possíveis convergem para a média, e a incerteza diminui.*

- ***Sobreposição de Estados:*** *A maioria das interações em F parecerá ser a soma de muitas maneiras possíveis pelas quais a interação poderia ter ocorrido. Quando dois feixes de onda próximos interagem, fazem-no inicialmente com base em ciclos individuais, já que a informação proveniente de regiões distantes do feixe só chega à velocidade do som. Cada ciclo contém um pouco de energia do feixe em questão e uma grande quantidade de energia de muitas outras ondas com diferentes frequências e orientações que atravessam o mesmo volume. Este “ruído de fundo” varia de ciclo para ciclo ao longo do feixe, o que faz com que a interação em cada ciclo seja ligeiramente distinta. Quando toda a informação do feixe de ondas é reunida, a interação total pode ser interpretada como a soma das interações ciclo a ciclo. Por vezes, a energia será transferida numa direção por um ciclo e na direção oposta por outro ciclo distante — assim, as alternativas podem tanto cancelar-se mutuamente como reforçar-se.*

Estas e outras propriedades do mundo F contêm alguns dos traços mais estranhos da mecânica quântica, mas resultam apenas de uma forma invulgar de olhar para uma situação prosaica. Existem algumas diferenças. A sobreposição de estados é

estatística, em vez de uma soma perfeita sobre todas as possibilidades, como sucede na mecânica quântica tradicional. Esta diferença é apenas subtil se V for muito grande, mas pode resultar numa quantidade ínfima de “ruído” nas medições, o que poderia ajudar a distinguir o mecanismo F de outras explicações da mecânica quântica.

O modelo, tal como foi apresentado, não exhibe os efeitos da relatividade restrita de forma evidente, o que constitui um defeito sério se pretendermos moldá-lo como uma descrição do nosso mundo. Há algo de errado na forma como trata o tempo. Contudo, possui uma propriedade que imita os efeitos temporais de um campo gravitacional relativista geral. O tempo próximo do centro de F corre mais lentamente do que nas regiões periféricas, uma vez que as interações se baseiam em ondas de frequência mais baixa. No ponto central, o tempo está parado. Esse ponto central de F nunca se altera em relação ao seu valor de “densidade média de toda a esfera” e está, assim, efetivamente congelado no tempo. Na relatividade geral, as regiões em torno de um corpo massivo apresentam uma propriedade semelhante: o tempo flui mais lentamente à medida que se aproxima do centro gravitacional. Perto de massas extremamente densas (como os buracos negros), o tempo pára por completo a uma certa distância.

Algumas das teorias mais exóticas da física moderna encontram neste modelo uma possível explicação. Embora a energia flua principalmente entre modos de onda muito semelhantes em frequência e direção (isto é, entre pontos adjacentes em F), as não-linearidades do meio V devem permitir que alguma energia flua de forma sistemática entre modos de onda harmonicamente relacionados — por exemplo, entre um modo e outro na mesma direção, mas com o dobro da frequência. Esses modos de fluxo energético em F fornecem “graus de liberdade” adicionais aos três oferecidos pelos pontos vizinhos. Quando observados numa escala reduzida, podem ser interpretados como dimensões extra (a energia pode mover-se nesta direção, naquela, noutra... e ainda noutras). Dado que uma circunavegação de harmónico em harmónico cobre o espaço disponível em menos passos do que um movimento entre modos adjacentes, estas

dimensões adicionais aparentarão ter uma extensão muito menor do que as três dimensões básicas. Quanto maior a energia envolvida, mais harmônicos poderão ser ativados e maior será a dimensionalidade resultante. As teorias físicas atuais postulam frequentemente dimensões adicionais enroladas sobre si mesmas para fornecer uma explicação geométrica às forças fundamentais. Dez e onze dimensões são propostas populares, e algumas teorias que sugerem novas forças podem ainda requerer mais dimensões. Se algo semelhante à explicação F para a aparente dimensionalidade superior estiver correto, há um bônus adicional: vistas a uma escala vasta, as “dimensões harmônicas” seriam verdadeiras ligações entre regiões distantes do espaço e, se devidamente exploradas, poderiam permitir comunicação e deslocação instantâneas por distâncias astronômicas.

Ondas Grandes

Agora, esqueçamos as possíveis implicações da ideia de F como mecanismo explicativo da mecânica quântica e consideremos o nosso universo à escala mais vasta. Ele está permeado por um fundo de radiação de micro-ondas com um comprimento de onda centrado em cerca de 1 milímetro, um comprimento que aumenta lentamente à medida que o universo se expande e arrefece. Esta radiação afeta — e é afetada por — nuvens de matéria no espaço interestelar, interagindo consigo mesma de forma não linear. Se fizermos uma transformada de Fourier espacial à escala do universo sobre essa radiação (isto é, tratarmos o nosso mundo como V), obtemos um espaço F com propriedades muito semelhantes às descritas na secção anterior. A expansão do universo acrescenta um novo elemento. À medida que o universo se expande gradualmente, os comprimentos de onda da radiação de fundo tornam-se mais longos. Com o alongamento das ondas, a taxa relativa de fluxo temporal no mundo F abranda. Qualquer habitante de F estaria idealmente posicionado para praticar a estratégia de “viver para sempre tornando-se cada vez mais lento à medida que o ambiente se torna mais frio”, proposta por Freeman Dyson. Neste momento, esses habitantes mover-se-iam extremamente devagar — as suas interações de partículas mais rápidas demorariam vários

trilionésimos de segundo. No passado, logo após o Big Bang, quando o universo era denso e quente, o mundo F teria sido um lugar intensamente ativo, a operar milhões ou milhares de milhões de vezes mais depressa. Nos primeiros microssegundos do Big Bang, a velocidade seria astronomicamente superior.

Esse primeiro microssegundo do Big Bang poderia ter representado Eras de tempo subjetivo em F. Talvez tempo suficiente para que uma inteligência evoluísse, tomasse consciência da sua situação, e semeasse formas de vida mais pequenas, mas eventualmente mais rápidas no espaço V. Apesar de F e V serem, à grande escala, a mesma coisa, a manipulação de um a partir do outro — ou mesmo a comunicação entre os dois — seria extraordinariamente difícil. Qualquer evento local em qualquer dos espaços seria difundido até se tornar indetetável no outro. Apenas projetos massivos, com ordem a longo alcance à escala do universo, funcionariam — e mesmo estes exigiriam quantidades imensas de tempo devido aos limites de velocidade em ambos os universos. A interação em tempo real entre V e F é, portanto, excluída. Contudo, tais projetos poderiam afetar muitos locais do outro espaço tão facilmente (ou, em muitos casos, até mais facilmente) como um só, e poderiam assim surgir como “milagres” que violam a entropia nesse outro espaço. Se eu vivesse em F e desejasse visitar V, poderia engenhar um desses milagres para condensar um substituto robótico de mim mesmo em V, e mais tarde outro milagre que permitisse recuperar as memórias do robô de volta para uma forma acessível em F.

A transformada de Fourier que converte V em F é idêntica (exceto por um sinal negativo) à transformada inversa que realiza a conversão oposta. Dadas apenas estas duas descrições, não seria claro qual o “mundo original”. Na verdade, a transformada de Fourier é apenas uma entre uma classe infinita de “transformadas ortogonais” com as mesmas propriedades fundamentais. Cada uma delas é capaz de operar sobre a descrição de um volume e produzir uma nova descrição contendo a mesma informação — mas com cada ponto original disseminado por todos os locais do resultado. Isto conduz à possibilidade de uma infinidade de universos, cada um representando uma combinação diferente da mesma substância

subjacente, cada um exibindo comportamento quântico, mas com a sua própria física singular, e todos mutuamente inconscientes dos outros com quem partilham o espaço. Não sei onde levar essa ideia.

BIBLIOGRAFIA

Prólogo

- Cairns-Smith, A. G. 1982. *Genetic Takeover and the Mineral Origins of Life*. Cambridge: Cambridge University Press.
1985. *Seven Clues to the Origin of Life: A Scientific Detective Story*. Cambridge: Cambridge University Press.
- 1985 (June). "The First Organisms." *Scientific American* 253:90-98. Calder, Nigel. 1983. *Timescale: An Atlas of the Fourth Dimension*. New York: Viking Press.
- Desmond, Kevin. 1986. *A Timetable of Inventions and Discoveries*. New York: M. Evans.
- Sagan, Carl. 1977. *The Dragons of Eden: Speculations on the Evolution of Human Intelligence*. New York: Random House.

Capítulo 1: Mente em Movimento

- Ashby, W. Ross. 1963. *An Introduction to Cybernetics*. New York: John Wiley & Sons.
- Asimov, Isaac. 1950. *I, Robot*. New York: Doubleday
- Bakker, Robert T. 1986. *The Dinosaur Heresies*. New York: William Morrow.
- Braitenberg, Valentino. 1984. *Vehicles: Experiments in Synthetic Psychology*. Cambridge: MIT Press.
- Brooks, J., and G. Shaw. 1973. *Origin and Development of Living Systems*. New York: Academic Press.
- Buchsbaum, Ralph. 1948. *Animals without Backbones*. 2nd ed. Chicago: University of Chicago Press.
- Feigenbaum, Edward A., and Juhan Feldman, eds. 1963. *Computers and Thought*. New York: McGraw-Hill.
- Fichtelius, Karl-Erik, and Sverre Sjolander. 1972. *Smarter than Man? Intelligence in Whales, Dolphins and Humans*. New York: Ballantine Books.
- Griffin, Donald R. 1984. *Animal Thinking*. Cambridge: Harvard University Press.

- Lane, Frank W. 1962. *Kingdom of the Octopus: The Life History of the Cephalopoda*. New York: Pyramid Publications.
- Marsh, Peter. 1985. *Robots*. New York: Crescent Books.
- Mayr, Ernst. 1982. *The Growth of Biological Thought*. Cambridge: Harvard University Press.
- McCorduck, Pamela. 1979. *Machines Who Think*. San Francisco: W. H. Freeman.
- Minsky, Marvin, ed. 1985. *Robotics*. New York: Doubleday 1986. The Society of Mind. New York: Simon & Schuster.
- Moynihan, Martin. 1985. *Communication and Noncommunication by Cephalopods*. Bloomington: Indiana University Press.
- Nilsson, Nils, ed. 1984. *Shakey the Robot: Artificial Intelligence Center Technical Note 323*. Menlo Park: SRI International.
- Pawson, Richard. 1985. *The Robot Book*. London: W H Smith & Son.
- Pratt, Vernon. 1987. *Thinking Machines: The Evolution of Artificial Intelligence*. Oxford: Basil Blackwell.
- Raphael, Bertram. 1976. *The Thinking Computer: Mind inside Matter*. San Francisco: W. H. Freeman.
- Reichardt, Jasia. 1978. *Robots: Fact, Fiction and Prediction*. Middlesex: Penguin Books.
- Thorson, Gunnar. 1971. *Life in the Sea*. New York: McGraw-Hill.
- Walter, W. Grey. 1961. *The Living Brain*. Middlesex: Penguin Books.
- Wiener, Norbert. 1965. *Cybernetics, or Control and Communication in the Animal and the Machine*. Cambridge: MIT Press.

Capítulo 2: Reforçando o Desempenho

- Babbage, Charles. 1961. *Charles Babbage and His Calculating Engines*. Edited by Philip Morrison and Emily Morrison. New York: Dover Publications.
- Berkeley, Edmund C. 1949. *Giant Brains or Machines That Think*. New York: John Wiley & Sons.
- Booth, Andrew, and Kathleen Booth. 1956. *Automatic Digital Calculators*. London: Butterworths.

- Cale, E. G., L. L. Gremillion, and J. L. McKenney. 1979. "Price/performance patterns of U.S. computer systems." *Communications of the ACM* 22(4): 225-230.
- Dertouzos, Michael L., and Joel Moses, eds. 1979. *The Computer Age: A Twenty-Year Review*. Cambridge: MIT Press.
- Dowling, John E. 1987. *The Retina: An Approachable Part of the Brain*. Cambridge: Harvard University Press.
- Drexler, K. Eric. 1986. *Engines of Creation*. New York: Doubleday.
- Eames, Charles, and Ray Eames. 1973. *A Computer Perspective*. Cambridge: Harvard University Press.
- Hubel, David, ed. 1979 (September). *The Brain—Scientific American*, special issue, 241(3).
- Kandel, Eric R. 1976. *Cellular Basis of Behavior: An Introduction to Behavioral Neurobiology*. New York: W. H. Freeman.
- Kandel, Eric R., and J. H. Schwartz. 1982. "Molecular basis of memory." *Science* 218:433-436.
- Kuffler, Stephen W., and John G. Nichols. 1976. *From Neuron to Brain: A Cellular Approach to the Function of the Nervous System*. Sunderland, Massachusetts: Sinauer Associates.
- Lazou, Christopher. 1986. *Supercomputers and Their Use*. Oxford: Clarendon Press.
- Minsky, Marvin, and Seymour Papert. 1969. *Perceptrons: An Introduction to Computational Geometry*. Cambridge: MIT Press.
- Moreau, Rene. 1984. *The Computer Comes of Age*. Cambridge: MIT Press, von Neumann, John. 1958. *The Computer and the Brain*. New Haven: Yale University Press.
- Queisser, Hans. 1988. *The Conquest of the Microchip*. Cambridge: Harvard University Press.
- Randell, Brian, ed. 1973. *The Origins of Digital Computers: Selected Papers*. Berlin: Springer-Verlag.
- Rosen, Saul. 1971. *ACM 71: A Quarter-Century View*. New York: Association for Computing Machinery.
- Squire, Larry R- 1986. "Mechanisms of memory." *Science* 232:1612-1619. Turn, Rein. 1974. *Computers in the 1980s*. New York: Columbia University Press.

Weik, Martin H. 1957. *A Second Survey of Domestic Electronic Digital Computing Systems*. Aberdeen, Maryland: Army Ballistic Research Laboratories, report no. 1010.

Wolfe, Jeremy M. 1986. *The Mind's Eye: Readings from Scientific American*. New York: W. H. Freeman.

Wolken, Jerome J. 1975. *Photoprocesses, Photoreceptors, and Evolution*. New York: Academic Press.

Capítulo 3: Simbiose

Boorstin, Daniel J. 1983. *The Discoverers*. New York: Random House.

Time-Life Books, eds. 1985. *Computer Images: Understanding Computers Series*. Arlington: Time-Life Books.

Capítulo 4: Cláusula do Avô

Forward, Robert L. 1980. *Dragon's Egg*. New York: Ballantine Books.

1984. *The Flight of the Dragonfly*. New York: Simon & Schuster

Hofstadter, Douglas R., and Daniel C. Dennett, eds. 1981. *The Mind's I*. New York: Basic Books.

Rucker, Rudy. 1983. *Software*. Middlesex: Penguin Books.

Sperry, Roger 1982. "Some effects of disconnecting the cerebral hemispheres." *Science* 217:1223-1226.

Vinge, Vernor 1984. *True Names*. New York: Bluejay Books.

Capítulo 5: Vida Selvagem

Axelrod, Robert. 1984. *The Evolution of Cooperation*. New York: Basic Books.

Cohen, Fred. 1984. *Computer Viruses*. Los Angeles: University of Southern California.

Dewdney A. K. 1984 (May), 1985 (March), and 1987 (January). "Core wars." *Scientific American* 250:14-17, 252:14-23, 256:14-18.

Dawkins, Richard. 1976. *The Selfish Gene*. Oxford: Oxford University Press.

1982. *The Extended Phenotype*. Oxford: Oxford University Press.

1986. *The Blind Watchmaker*. New York: W. W. Norton.

Hoyle, Fred, and John Elliot. 1962. *A for Andromeda*. New York: Harper

- Oliver, B. M., and J. Billingham. 1971. *Project Cyclops: A Design Study of a System for Detecting Extraterrestrial Intelligent Life*. Moffett Field, California: NASA/Ames Research Center
- Rosen, Eric C. 1981. *Vulnerabilities of Network Control Protocols: An Example*. Cambridge: Bolt Beranek and Newman, Inc.
- Sagan, Carl. 1985. *Contact*. New York: Simon & Schuster.
- Shoch, John F, and Jon A. Hupp. 1982. "The 'worm' programs: early experience with a distributed computation." *Communications of the ACM* 25(3):172-180.
- Thompson, Ken. 1984. "Reflections on trusting trust." *Communications of the ACM* 27(8):761-763.

Capítulo 6: Emancipação

- Barrow, John D., and Frank J. Tipler 1986. *The Anthropic Cosmological Principle*. Oxford: Oxford University Press.
- Dyson, Freeman. 1979. *Disturbing the Universe*. New York: Harper & Row.
1988. *Infinite in all Directions*. New York: Harper &c Row.
- Gardner, Martin. 1983. *Wheels, Life and Other Mathematical Amusements*. New York: W. H. Freeman.
- Hawking, Stephen W. 1988. *A Brief History of Time: From the Big Bang to Black Holes*. New York; Bantam Books.
- Weinberg, Steven. 1977. *The First Three Minutes*. New York: Basic Books.

Apêndice 3: Os Limites Externos da Computação

- DeWitt, Bryce S., and Neill Graham, eds. 1973. *The Many-Worlds Interpretation of Quantum Mechanics*. Princeton: Princeton University Press.
- Feynman, Richard P., Robert B. Leighton, and Matthew Sands. 1965. *The Feynman Lectures on Physics*. Vol. 3. New York: Addison-Wesley.
- Gribbin, John. 1984. *In Search of Schrodinger's Cat: Quantum Physics and Reality*. New York: Bantam Books.
1985. "Doomsday device." *Analog Science Fiction-Science Fact* 105:22-127.