

- 13) W. Ross Ashby, *Design for a Brain*. Londres, Chapman & Hall, 1960.
- 14) *Communication*, n.º 18. Paris, Ecole Practique des Hautes Studes, 1972. Este número da revista é dedicado ao "evento".
- 15) Walter Buckley, *Sociology and the Modern System Theory*. New Jersey, Prentice-Hall, 1967. Tradução brasileira: *A Sociologia e a Moderna Teoria dos Sistemas*. S. Paulo, Cultrix-Ed. da USP, 1971.
- 16) "The Second Cybernetics. Deviation. Amplifying Mutual Causal Processes" in *American Scientist*, n.º 51, 1963, pp. 164-179.
- 17) Watzlawick, Beavin e Jackson, *Pragmatics of Human Communication*. Nova Iorque, Norton, 1967. Tradução brasileira: *Pragmática da Comunicação Humana*. S. Paulo, Cultrix, 1973.
- 18) David Cooper, *Psychiatry and Anti-Psychiatry*. Londres, Tavistock, 1967. Tradução brasileira: *Psiquiatria e Antipsiquiatria*, S. Paulo, Perspectiva, 1973.
- 19) Eliseo Verón, Luís Prieto e outros, *Lenguaje y Comunicacion Social*. Buenos Aires, Nueva Visión, 1968.
- 20) Anthony Wilden, "L'écriture et le bruit dans la morphogenese du system ouvert", in *Communication* n.º 18, cit.
- 21) Garret Hardin, *Nature and Man's Fate*. Tradução brasileira: *A Natureza e o Destino do Homem*. S. Paulo, Cia. Editora Nacional, 1969.
- 22) Ernst Mayr, *Populations, Species and Evolution*. Cambridge, Harvard University Press, 1970.

COMPUTADORES E INTELIGÊNCIA

A.M. TURING

NOTA PRÉVIA

Alan Mathison Turing, matemático inglês que faleceu em 1954 com 42 anos de idade, foi um dos cientistas mais criativos em computação. O seu conceito mais conhecido é o de Máquina de Turing. Esta é uma "caixa negra" (uma máquina com dispositivos não-especificados) capaz de percorrer uma fita infinita de espaços divididos em quadrados. A caixa, por sua vez, pode ter um número qualquer, porém finito, de estados. Quando a caixa esquadrinha um quadrado, ela pode deixar o símbolo inalterado, erradicá-lo e imprimir outro símbolo ou imprimir um símbolo num quadrado branco. A fita então se desloca um quadrado para a direita, para a esquerda, ou permanece fixa. A caixa, então, ou permanece no mesmo estado ou passa a outro estado. Cada máquina Turing é definida por uma tabela.

Imaginemos 13 quadrados numa fita de papel marcados 11111111 + 1111 significando a soma $8 + 4$ no sistema unitário, onde um algarismo n é simbolizado por n algarismos 1.

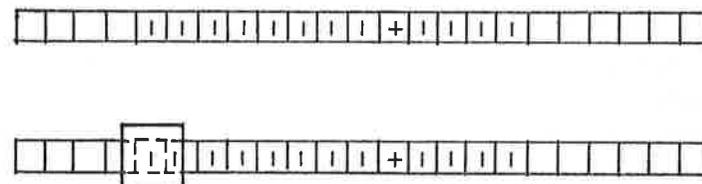


Fig. 10

A Fig. 10 representa uma máquina de Turing. Na fita de papel se insere uma cartolina que tem uma janela onde é visível apenas um quadrado. Ajustada a cartolina ao primeiro número 1, a tabela abaixo dá as regras do seu funcionamento:

	Estado A	Estado B
1	1. Erradique o 1 2. Esquadrinhe célula à direita 3. Vá ao estado B	1. Esquadrinhe célula à direita 2. Fique no estado B
+		1. Erradique o sinal + 2. Imprima 1 3. Pare

Começa-se assumindo que a máquina esteja no estado A. Consulta-se a tabela para verificar a combinação do símbolo 1 com o estado A: apaga-se o número 1 e move-se a fita para a esquerda de tal forma que a cartolina (a caixa negra) se mova para o quadrado à direita. A máquina vai ao estado B. Continue-se desta forma até a máquina parar. Se seguidas corretamente, as regras conduzirão ao resultado 1111111111, ou seja 12, pois apagou-se o primeiro 1 e substituiu-se o sinal + por 1. Esta idéia extremamente simples é básica para a teoria dos computadores digitais. Turing mostra que a sua máquina idealizada poderia ser programada para executar tudo o que o mais poderoso computador pudesse fazer.

O jogo da imitação descrito no artigo de Turing coloca, porém, a questão de saber se poderemos, em determinadas circunstâncias, distinguir o comportamento humano do comportamento da máquina. O problema é colocado como uma indagação para a qual precisamos partir desarmados das conotações usuais.

Deixando de lado como problema externo ao proposto a questão do caráter transcendente do homem, proposições mais simples podem ser colocadas.

Há máquinas que pensam?

O ato de pensar pode ser considerado como se estendendo num contínuo de n dimensões e será um vetor com n compo-

nentés. Se tomarmos apenas dois destes componentes, a saber a velocidade e a sofisticação dos processos envolvidos, teríamos o gráfico da Fig. 11.

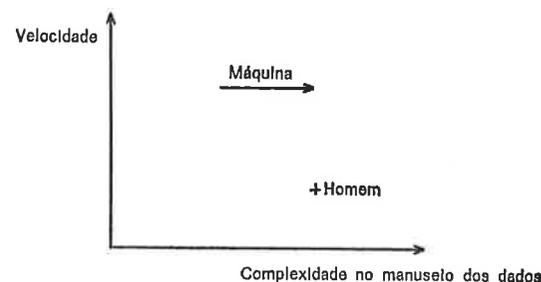


Fig. 11

Esta situação mostra uma tendência para a mudança, deslocando-se a máquina no sentido do aumento da complexidade no manuseio dos dados e processamento dos mesmos.

Mas não é um computador uma máquina que execute apenas aquilo para que é programada e não mais? Comentando tal pergunta, conhecido pesquisador diz:

“Esta afirmativa de que os computadores podem fazer aquilo para que são programados é intuitivamente óbvia e intuitivamente verdadeira, mas não apóia nenhuma das implicações que são ordinariamente daí extraídas. Um ser humano pode pensar, aprender e criar porque o seu programa biológico assim o dotou, conjuntamente com a capacidade para as alterações que este programa produza pela interação com o ambiente e que o habilitem a pensar, aprender e criar. Se um computador pensa, aprende e cria é porque um programa o dotou de tais capacidades. Este programa não será repetitivo e estereotipado, mas fará com que seu comportamento seja condicionado pelo meio ambiente. Terá capacidade de extrair indícios do meio ambiente, indicando seu próprio progresso em relação aos objetivos. Terá de ser um programa que analise por algum meio seu próprio desempenho, diagnostique suas falhas e providencie mudanças para melhorar sua eficiência futura.” (Simon, 1960).

Um problema existe quando várias alternativas são oferecidas, das quais uma deve ser escolhida. A cada alternativa,

o meio ambiente pode responder de forma diferente, resultando uma elaborada rede de alternativas combinadas com respostas. Temos então o que se chama "árvore de possibilidades".

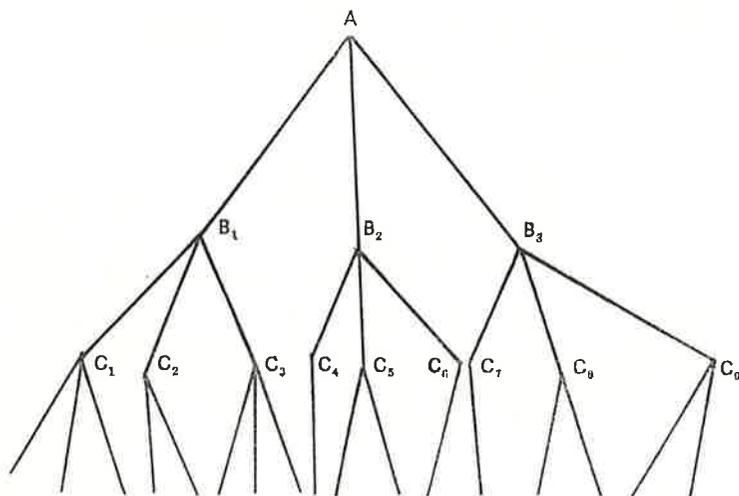


Fig. 12

Se explorássemos todos os ramos, seria possível saber se existe ou não uma solução, e, em caso afirmativo, qual o seu algoritmo. Existem, porém, árvores tão complexas que escapam inclusive a um exame feito pela máquina, muito mais veloz que a mente humana. Nestes casos, de que jogos como xadrez e damas são exemplos, torna-se necessário associar ao programa características de tentativa e erro, aprendizado e critérios heurísticos. Quando os critérios são judiciosamente combinados, os programas resultantes apresentam desempenhos surpreendentes.

Programas já foram elaborados que capacitam as máquinas a jogarem jogos como o de damas (Samuel), xadrez (Newell, Simon), demonstrarem teoremas de geometria (Gelernter), traduzirem textos ou simularem comportamentos sociais elementares (Gullahorn), comporem música (Hiller, Isaacson), desenharem motores (Goodwin) ou resolverem problemas de cál-

culo infinitesimal (Slagle). Um dos mais interessantes é o programa de A.L. Samuel para o jogo de damas. Em suas palavras:

"A atividade não pode ser determinística no sentido prático. Não há um algoritmo conhecido que garanta a vitória ou o empate no jogo de damas. A exploração de todos os caminhos possíveis num jogo de damas envolveria cerca de 10^{40} escolhas, que à razão de três escolhas por milimicrosegundos, levariam 10^{21} séculos a serem consideradas. O programa elaborado por Samuel (*IBM Journal of Research and Development*, julho 1959, pp. 211-229) joga verificando alguns lances futuros, avaliando o resultado pelas posições decorrentes e melhorando seu padrão pela prática constante, isto é, as soluções favoráveis serão mais e mais repetidas e as desfavoráveis eliminadas.

Neste tipo de programa, as alternativas são tão numerosas que nem o computador as pode exaurir à base do esgotamento de todas as possibilidades.

O programa então, tal como os seres humanos, lança mão de processos heurísticos. Certos caminhos e parâmetros são valorizados, num processo em que o resultado dá uma contínua retroação para um aperfeiçoamento constante à base da experiência. Poderíamos dizer que a máquina apresenta traços do que chamamos de intuição e aprendizado nos seres humanos. O programa original, é claro, definirá todas estas condições, mas permitirá mudanças à base dos resultados obtidos.

No verão de 1962, uma partida foi efetuada entre o programa de Samuel e um ex-campeão de damas de Connecticut, Robert W. Nealey, um dos melhores jogadores dos Estados Unidos. A máquina venceu e o ex-campeão comentou:

É interessante notar que o computador executou vários movimentos brilhantes, a fim de vencer, e eu tive várias oportunidades de empatar. Por isso continuei a jogar. A máquina jogou um final perfeito, sem erros. Em matéria de finais, eu não tive tal competição por parte de outro ser humano desde 1954, quando perdi meu último jogo.

ISAAC EPSTEIN

1. O jogo da imitação

Proponho a seguinte questão: "Podem as máquinas pensar?" A consideração disso deveria ser iniciada com definições do significado dos termos "máquina" e "pensar". As defini-

ções poderiam ser ordenadas de modo a refletir, na medida do possível, o uso comum das palavras, mas tal atitude é perigosa. Se os significados das palavras “máquina” e “pensar” tiverem de ser encontrados por meio de um exame de seu uso habitual, será difícil escapar à conclusão de que o significado da e a resposta à pergunta “Podem as máquinas pensar?” deverão ser procurados numa pesquisa estatística do tipo Gallup. Mas isto é absurdo. Em vez de tentar uma definição deste tipo, eu substituiria a questão por outra, que está relacionada de perto com ela e é expressa em palavras menos ambíguas.

A nova formulação do problema pode ser descrita em termos de um jogo a que nós chamamos “jogo da imitação”. É jogado por três pessoas: um homem (A), uma mulher (B), e um interrogador (C), que pode ser de qualquer dos sexos. O interrogador permanece num quarto, separado dos outros dois. O objetivo do jogo, para o interrogador, é determinar, em relação aos outros dois, qual o homem e qual a mulher. Ele os conhece por rótulos X e Y e no fim do jogo dirá ou “X é A e Y é B”, ou “X é B e Y é A”. É permitido ao interrogador fazer perguntas a A e B, tais como:

C: Será que X poderia me dizer qual o comprimento de seu cabelo?

Supondo-se agora que X seja realmente A, então A deverá responder. O objetivo do jogo para A é tentar induzir C a uma identificação errada. Sua resposta, portanto, poderia ser:

“Meu cabelo é curto, e os fios mais longos têm cerca de 20 centímetros de comprimento”.

Para que tons de vozes não ajudem o interrogador, as respostas deveriam ser escritas, ou ainda melhor, datilografadas. O arranjo ideal é um telegravador com comunicação entre os dois quartos.

Alternativamente, a pergunta e as respostas podem ser repetidas por um intermediário. O objetivo do jogo para a terceira jogadora (B) é ajudar o interrogador. Sua melhor estratégia será provavelmente dar respostas verdadeiras. Ela pode acrescentar frases como: “Eu sou a mulher, não escute a ele”. Mas isso será inútil, porque o homem pode dar respostas semelhantes.

Agora formulamos a questão: “O que acontecerá quando uma máquina ocupar o lugar de A nesse jogo?” Será que o interrogador decidirá erroneamente com a mesma frequência, quando o jogo é jogado dessa forma, do que quando o fazia ao tempo em que o jogo era jogado entre um homem e uma mulher? Estas questões substituem a pergunta original “Podem as máquinas pensar?”

2. Crítica do novo problema

Assim como se pergunta “Qual a resposta para essa nova forma de pergunta?”, pode-se perguntar “Essa nova pergunta é digna de ser investigada?” Tal questão nós a investigaremos de pronto, com isso atalhando uma regressão infinita.

O novo problema tem a vantagem de traçar uma linha bastante nítida entre as capacidades físicas e intelectuais de um homem. Nenhum engenheiro ou químico pode alegar ser capaz de produzir um material que seja indistinguível da pele humana. É possível que algum dia isso possa ser feito, mas mesmo supondo que tal invenção esteja disponível, deveríamos perceber que há pouca vantagem em tentar tornar uma “máquina pensante” humana vestindo-a com tal carne artificial. A forma na qual propusemos o problema reflete esse fato na condição que impede o interrogador de ver ou tocar os outros competidores, ou ouvir-lhes as vozes. Algumas outras vantagens do critério proposto podem ser demonstradas por amostras de perguntas e respostas, tais como:

P: Por favor, escreva-me um soneto cujo tema seja a “Forth Bridge”.

R: Poupe-me isso. Nunca consegui escrever poesia.

P: Some 34 957 e 70 764.

R: (Pausa de mais ou menos 30 segundos e depois como resposta) 105 721.

P: Você joga xadrez?

R: Sim.

P: Eu tenho R em meu R_1 , e nenhuma outra peça. Você tem somente R no R_0 e T no T_1 . É a sua vez. Qual o seu lance?

R: (Depois de uma pausa de 15 segundos) T- T_8 mate.

O método de pergunta e resposta parece ser adequado para uso em quase todos os campos de atividade humana que desejemos abarcar. Não queremos punir a máquina por sua inabilidade de brilhar em concursos de beleza, nem punir um homem por perder uma corrida contra um aeroplano. As condições de nosso jogo tornam essas inaptidões descabidas. As "testemunhas" podem vangloriar-se, se acharem conveniente, do seu fascínio, força ou heroísmo, mas o interrogador não pode pedir demonstrações práticas.

O jogo talvez possa ser criticado sob o pretexto de que as desvantagens pesam bastante contra a máquina. Se o homem fosse tentar fingir-se de máquina, iria certamente fazer uma demonstração muito pobre. Iria trair-se imediatamente por sua lentidão e imprecisão em aritmética. Não podem acaso as máquinas realizar algo que deveria ser descrito como pensamento, mas que é muito diferente do que um homem faz? Tal objeção é muito forte, mas ao menos podemos dizer que se, não obstante, puder-se construir uma máquina capaz de jogar o jogo da imitação satisfatoriamente, não precisaremos preocupar-nos com semelhante objeção.

Pode-se alegar que, ao jogar o "jogo da imitação", a melhor estratégia para a máquina será possivelmente algo que não seja a imitação do comportamento de um homem. Tal é possível, mas creio ser improvável algo dessa espécie. De qualquer modo, não intentaremos investigar aqui a teoria do jogo, e admitiremos que a melhor estratégia será tentar dar as respostas que seriam naturalmente dadas por um homem.

3. As máquinas implicadas no jogo

A pergunta que fizemos no § 1 não será totalmente definida antes que especifiquemos o que pretendemos dizer com a palavra "máquina". É natural que queiramos permitir que todo tipo de técnica de engenharia seja usado em nossas máquinas.

Também desejaríamos admitir a possibilidade de um engenheiro ou uma equipe de engenheiros construir uma máquina que funcione, mas cujo modo de operação não possa ser satisfatoriamente descrito por seus construtores, porque utilizaram um método em grande parte empírico. Finalmente, queremos excluir das máquinas todos os homens de carne e osso. É difícil formular as definições de modo que satisfaçam essas três condições. Poder-se-ia, por exemplo, insistir que os membros da equipe de engenheiros fossem todos do mesmo sexo, mas isso não seria de fato satisfatório, porque é provavelmente possível criar um indivíduo completo a partir de uma única célula de — digamos — pele humana. Realizar isso seria uma proeza da técnica biológica digna dos maiores elogios, mas não estaríamos dispostos a considerá-lo como um caso de "construção de uma máquina pensante". Isto nos induz a abandonar o requisito de que todo tipo de técnica deveria ser permitido. Estamos tanto mais prontos a abandoná-lo quanto se sabe que o atual interesse por "máquinas pensantes" foi despertado por um tipo particular de máquina, geralmente chamado de "computador eletrônico" ou "computador digital". Seguindo tal sugestão, só permitimos que computadores digitais tomem parte em nosso jogo.

Essa restrição parece, à primeira vista, muito drástica. Tentarei mostrar que não é assim. Mas isso exige uma breve explicação da natureza e propriedade desses computadores.

Também se poderia dizer que essa identificação de máquinas com computadores digitais, tanto quanto nosso critério de "pensar", só será insatisfatório se (contrariamente à minha crença), os computadores digitais se revelarem incapazes de uma boa demonstração no jogo.

Já há um certo número de computadores digitais em funcionamento, e pode-se perguntar "Por que não tentar a experiência imediatamente? Seria fácil satisfazer as condições do jogo. Certo número de interrogadores poderia ser usado e uma estatística compilada para mostrar a frequência com que a identificação certa fosse dada". A resposta imediata é que não estamos perguntando se todos os computadores digitais fariam boa figura no jogo nem se os computadores presentemente disponíveis teriam bom desempenho, e sim se existem computadores imagináveis capazes de tanto. Mas isso é unicamente a

resposta imediata. Veremos a questão mais tarde, sob um aspecto diferente.

4. Computadores digitais

A idéia que existe atrás de computadores digitais pode ser explicada, dizendo-se que essas máquinas são planejadas para realizar quaisquer operações passíveis de serem feitas por um computador humano. O computador humano deve seguir regras fixas; não tem autoridade para se desviar delas em nenhum detalhe. Podemos supor que essas regras sejam fornecidas por um livro, alterado sempre que ao operador se confie novo trabalho. O operador dispõe também de um suprimento ilimitado de papéis onde fazer seus cálculos. Ele também pode fazer suas multiplicações e adições numa máquina de calcular de mesa, mas isto não é importante.

Se usarmos a explicação acima como uma definição, estaremos em perigo de cair num argumento circular. Nós o evitaremos dando um esboço dos meios pelos quais o efeito desejado é alcançado. Um computador digital pode ser usualmente visto como consistindo de três partes:

- a) Memória
- b) Unidade executiva
- c) Controle

A memória é uma reserva de informação e corresponde ao papel utilizado pelo computador humano, seja este a folha de papel onde faz seus cálculos, ou o do livro onde as regras estão impressas. Na medida em que o computador humano faça cálculos de cabeça, uma parte da memória corresponderá à sua própria memória.

A unidade executiva é a parte que realiza as várias operações individuais envolvidas num cálculo. Quais sejam tais operações individuais é coisa que poderá variar de máquina para máquina. Normalmente, podem-se fazer longas operações, tais como "Multiplicar 3 540 675 445 por 7 076 345 687", mas em algumas máquinas somente algumas operações muito simples, tais como "Escreva 0", são possíveis.

Mencionamos que o "livro de regras" fornecido ao computador pode ser substituído na máquina por uma parte de sua memória. Chama-se então "tabela de instruções". É dever do controle verificar que essas instruções sejam obedecidas corretamente e na ordem certa. O controle é construído de tal forma que isso necessariamente acontece.

As informações na memória são via de regra fragmentadas em blocos de tamanho relativamente pequeno. Numa máquina, por exemplo, um bloco pode consistir em 10 algarismos decimais. Atribuem-se números às partes da memória nas quais os vários blocos de informações são estocados, de alguma forma sistemática. Uma instrução típica diria: "Adicione o número estocado na posição 6809 àquele da 4302 e ponha o resultado de volta na última posição da memória."

Desnecessário dizer que isso não ocorreria na máquina sob a forma de expressão lingüística. Seria mais provavelmente codificado numa forma tal como 6809430217. Aqui, o 17 nos diz qual das várias alterações possíveis será realizada nos dois números. Nesse caso, a operação é a descrita acima: "Adicione o número... "Notar-se-á que a instrução toma 10 algarismos e assim forma um bloco de informação muito conveniente. O controle normalmente tomará as instruções a serem obedecidas na ordem das posições nas quais elas estão memorizadas, mas ocasionalmente uma instrução tal como: "Agora obedeça à instrução memorizada na posição 5606 e continue daí", pode ser encontrada, ou então: "Se a posição 4505 contém 0, obedeça à instrução memorizada em 6707; do contrário, prossiga".

Instruções destes últimos tipos são muito importantes porque possibilitam que uma seqüência de operações seja refeita várias vezes até que se satisfaça alguma condição, mas de forma a obedecer, não instruções novas em cada repetição, mas as mesmas, seguidamente. Para recorrer a uma analogia doméstica, suponha-se que Mamãe queira que João passe pelo sapateiro toda manhã, em seu caminho para a escola, para verificar se os sapatos dela estão prontos; ela pode pedir-lhe isso todas as manhãs. Alternativamente, ela pode de uma vez por todas afixar um lembrete na sala de entrada, que ele verá quando for para a escola, e que o lembrará de perguntar pelos sapatos e também de destruir o lembrete quando trouxer os sapatos consigo.

O leitor deve aceitar como fato que computadores digitais podem ser construídos, e até já foram construídos, de acordo com os princípios que descrevemos, e que podem de fato muito aproximadamente imitar as ações de um computador humano. *

O livro de regras a que nos referimos, usado pelo nosso computador humano, é uma ficção conveniente. Os computadores humanos lembram, em realidade, o que tenham de fazer. Se alguém quiser construir uma máquina que imite o comportamento do computador humano em alguma operação complexa, terá de perguntar-lhe como ela é feita e então traduzir a resposta na forma de uma tabela de instruções. A construção de tabelas de instruções é comumente descrita como "programação". "Programar uma máquina que realize a operação A" significa colocar a tabela de instruções apropriada dentro da máquina, de tal forma que esta realize A.

Uma variante interessante da idéia de um computador digital é a do "computador digital com um elemento aleatório". Estes têm instruções que envolvem o lançamento de um dado ou algum processo eletrônico equivalente; uma instrução que tal poderia ser por exemplo: "Atire o dado e ponha o número que resultou na memória 1000". Algumas vezes, tal máquina é descrita como dotada de livre arbítrio (embora eu pessoalmente não usasse essa frase). É normalmente impossível determinar, observando-se uma máquina, se ela possui ou não um elemento aleatório, porque um efeito similar pode ser produzido por artifícios como fazer escolhas que dependam dos algarismos que compõem a seqüência decimal do número π .

A maioria dos computadores digitais atuais possui somente uma memória finita. Não há dificuldade teórica em relação à idéia de um computador com memória ilimitada. Naturalmente, só uma parte finita pode ter sido usada a qualquer tempo. Da mesma forma, somente uma soma finita pode ter sido construída, mas podemos imaginar mais e mais parcelas sendo adicionadas, se necessário. Tais computadores têm interesse teórico especial e serão chamados computadores de capacidade infinita.

(*) Este artigo foi publicado em 1950. A. M. Turing morreu subitamente em 1954, após uma brilhante, porém curta, carreira. Pouco antes de sua morte, ele completou o projeto de um dos primeiros computadores digitais de alta velocidade (N. do T.).

A idéia de um computador digital é antiga. Charles Babbage, professor de Matemática em Cambridge, de 1828 a 1839, planejou tal máquina, a chamada Máquina Analítica, que nunca foi completada. Embora Babbage tivesse todas as idéias essenciais, sua máquina, na época, não apresentava perspectivas atraentes. A velocidade então disponível seria certamente maior que a do computador humano, mas era cem vezes mais vagarosa que a Máquina de Manchester; esta por sua vez, é uma das mais vagarosas das máquinas modernas. A memória teria de ser puramente mecânica, por meio de rodas dentadas e cartões.

O fato de que a Máquina Analítica de Babbage tivesse de ser inteiramente mecânica, ajuda-nos a livrar-nos de uma superstição. Dá-se freqüentemente importância ao fato de que os computadores digitais modernos são elétricos e de que o sistema nervoso seja também elétrico. Como a máquina de Babbage não era elétrica, e como todos os computadores digitais são, em certo sentido, equivalentes, vemos que tal uso da eletricidade não pode ter importância teórica. É claro que a eletricidade normalmente aparece quando se trate de sinalização rápida, de forma que não surpreende a encontremos em ambos os casos. No sistema nervoso, os fenômenos químicos são tão importantes quanto os elétricos. Em determinados computadores, o sistema de memória é principalmente acústico. A característica de usar a eletricidade apresenta assim semelhança muito superficial. Se quiséssemos achar semelhanças que tais precisaríamos voltar-nos para as analogias matemáticas de função.

5. A universalidade dos computadores digitais

Os computadores digitais considerados na última seção podem ser classificados entre as "máquinas de estado discreto". Estas são as máquinas que se movimentam por pulos ou estados súbitos de um estado bem definido para outro. Tais estados são suficientemente diferentes para que não haja possibilidade de confusão entre eles. A rigor, tais máquinas não existem. Tudo, na realidade, se move continuamente. Mas há muitos tipos de máquinas que podem ser vantajosamente consideradas máquinas de estado discreto. Por exemplo, quando se consideram os interruptores de um sistema de iluminação, é

uma ficção conveniente admitir que cada interruptor tenha de estar definitivamente ligado ou desligado. Deve haver posições intermediárias, mas, para a maioria dos propósitos, podemos esquecê-las. Como um exemplo de máquina de estado discreto, podemos considerar uma roda cujas posições se alternem de 120° por segundo, mas que pode ser detida por uma alavanca operável de fora; ademais, uma lâmpada acende-se numa das posições da roda. Semelhante máquina pode ser descrita abstratamente como segue. O estado interno da máquina (que é descrito pela posição da roda) pode ser q_1 , q_2 ou q_3 . Há um sinal de entrada i_0 ou i_1 (posição da alavanca). O estado interno a qualquer momento é determinado pelo último estado, e sinal de entrada, de acordo com a tabela:

Último Estado

	q_1	q_2	q_3
i_0	q_2	q_3	q_1
i_1	q_1	q_2	q_3

Os sinais de saída, a única indicação externa visível do estado interno (a luz), são descritos pela tabela:

Estado	q_1	q_2	q_3
Saída	0_0	0_0	0_1

Este exemplo é típico das máquinas de estado discreto. Elas podem ser descritas por tais tabelas sob a condição de que tenham somente um número finito de estados possíveis.

Parecerá que, dado o estado inicial da máquina e os sinais de entrada, seja sempre possível prever todos os estados futuros. Isto faz lembrar a concepção de Laplace de que, a partir do estado completo do universo num dado momento de tempo, tal como descrito pelas posições e velocidades de todas as partículas, seria possível prever todos os estados futuros. A previsão que estamos considerando está, contudo, mais próxima da praticabilidade do que a considerada por Laplace. O sistema

do "universo como um todo" é tal que erros assaz pequenos nas condições iniciais podem ter um efeito esmagador num tempo posterior. O deslocamento de um simples elétron por um bilionésimo de centímetro, em determinado momento, pode representar a diferença entre um homem ser morto por uma avalanche, um ano mais tarde, ou escapar dela. Propriedade essencial dos sistemas mecânicos a que chamamos "máquinas de estado discreto" é a de que tal fenômeno não ocorra. Mesmo quando consideramos máquinas físicas reais, em vez de máquinas idealizadas, um conhecimento razoavelmente preciso do estado num determinado momento produz conhecimento razoavelmente preciso certo número de passos mais tarde.

Como dissemos, os computadores digitais pertencem à classe das máquinas de estado discreto. Mas o número de estados de que essa máquina é capaz é usualmente enorme. Por exemplo, para a máquina ora funcionando em Manchester, tal número é cerca de $2^{165.000}$, isto é, cerca de $10^{50.000}$. Compare-se isto com nosso exemplo da roda dentada descrita acima, que possui só três estados. Não é difícil de ver porque o número de estados deva ser tão imenso. O computador inclui uma memória que corresponde ao papel usado por um computador humano. Tem de ser possível escrever na memória qualquer uma das combinações de símbolos que poderia ter sido escrita no papel. Para simplificar, suponha-se que somente algarismos de 0 a 9 sejam usados como símbolos. Variações caligráficas são ignoradas. Suponha-se que se entregue ao computador 100 folhas de papel, cada qual com 50 linhas, cabendo em cada linha 30 algarismos. Então o número de estados é $10^{100 \times 50 \times 30}$ isto é, $10^{150.000}$. Este é aproximadamente o número de estados de três máquinas de Manchester reunidas. O logaritmo de base dois do número de estados é usualmente chamado "capacidade de memória" da máquina. Assim, a máquina de Manchester tem uma capacidade de memória de cerca de 165 000, e a máquina de roda dentada do nosso exemplo, cerca de 1,6. Se duas máquinas são acopladas, suas capacidades têm de ser somadas para obter-se a capacidade da máquina resultante. Isto leva à possibilidade de enunciados como "A máquina de Manchester contém 64 trilhas magnéticas, cada qual com uma capacidade de 2 560, oito válvulas eletrônicas com uma capacidade de 1 280.

A memória mista chega a 300, dando uma capacidade de 174 380.

Dada a tabela correspondente à máquina de estado discreto, é possível prever o que fará. Não há razão para que este cálculo não seja feito por meio de um computador digital. Conquanto que o cálculo seja feito com suficiente rapidez, o computador digital pode imitar o comportamento da máquina de estado discreto. O jogo de imitação poderia então ser jogado com a máquina em questão (como B) e o computador digital imitador (como A); o interrogador seria incapaz de distingui-los. Naturalmente, o computador digital tem de ter uma capacidade de memória adequada, bem como trabalhar com suficiente rapidez. Além disso, tem de ser programado novamente para cada nova máquina que se deseje imitar.

Descreve-se essa propriedade especial dos computadores digitais de imitar qualquer máquina de estado discreto dizendo-se que são máquinas universais. A existência de máquinas com tais propriedades tem a importante conseqüência de que, considerações de velocidade à parte, é desnecessário desenhar novas e diferentes máquinas para realizar diferentes processos de computação. Eles podem ser todos levados a cabo com um computador digital, adequadamente programado para cada caso. Ver-se-á que, em conseqüência disso, todos os computadores digitais são, em certo sentido, equivalentes.

Podemos agora considerar de novo o ponto suscitado no final do § 3. Sugeriu-se, conjecturalmente, que a questão "Podem as máquinas pensar?" fosse substituída por, "Existem computadores digitais imagináveis que tivessem bom desempenho no jogo de imitação?" Se quisermos, podemos generalizar e perguntar "Existem máquinas de estado discreto capazes de bom desempenho?" Mas, em vista da propriedade de universalidade, verificamos que qualquer uma dessas perguntas equivale a isto: "Fixemos nossa atenção num computador digital particular C. É verdade que modificando-se este computador para obter uma memória adequada, aumentando-lhe convenientemente a velocidade de ação, e provendo-o de um programa apropriado, C pode ser preparado para desempenhar satisfatoriamente o papel de A no jogo de imitação, sendo o papel B desempenhado por um homem?"

6. Opiniões contrárias acerca da questão principal

Podemos agora considerar que o terreno foi limpo e que estamos prontos para prosseguir no debate de nossa questão "Podem as máquinas pensar?" e da variante dela citada no final da última seção. Não podemos abandonar totalmente a forma original do problema, porque as opiniões vão diferir quanto a adequação da substituição, e precisamos, pelo menos, ouvir o que tem a ser dito neste particular.

As coisas se tornarão mais simples para o leitor se eu explicar primeiramente minhas próprias convicções a respeito. Considere-se primeiro a forma mais acurada da questão. Acredito que, dentro de cerca de 50 anos, será possível programar computadores, com uma capacidade de memória de cerca de 10^9 , para fazê-los jogar o jogo da imitação tão bem que um interrogador médio não terá mais de 70% de probabilidade de chegar à identificação correta, após 5 minutos de interrogatório. A pergunta original "Podem as máquinas pensar?" é, a meu ver, insignificante demais para merecer discussão. Contudo, acredito que no fim do século, o uso das palavras e a opinião geral esclarecida estarão tão mudados que se poderá falar de máquinas pensantes sem contradita. Acredito ainda que não há nenhum benefício em ocultar tais convicções. A noção popular de que cientistas avançam inexoravelmente de fatos bem estabelecidos para fatos bem estabelecidos, jamais sendo influenciados por conjecturas não provadas, está totalmente errada. Desde que se esclareça o que são fatos provados e o que são conjecturas, nenhum dano pode resultar. As conjecturas são de grande importância, pois sugerem úteis linhas de investigação.

Considerarei agora opiniões opostas às minhas.

(1) A objeção teológica

Pensar é uma função da alma humana imortal. Deus deu uma alma imortal a todo homem e a toda mulher, mas a nenhum outro animal ou máquina. Logo, nenhum animal ou máquina pode pensar.*

(*) Possivelmente essa visão é herética. S. Tomás de Aquino [*Suma Teológica*], citação de Bertrand Russell (1945, p. 458) afirma

Sou incapaz de aceitar mesmo em parte tal argumento, mas tentarei responder em termos teológicos. Acharia o argumento mais convincente se os animais fossem postos na mesma classe dos homens, porque há uma diferença bem maior, para mim, entre o ser animado e o inanimado típicos, do que entre o homem e outros animais. O caráter arbitrário da visão ortodoxa se evidencia se considerarmos como deve parecer a um membro de alguma outra comunidade religiosa. Como encaram os cristãos o pensamento muçulmano de que as mulheres não têm alma? Mas vamos deixar esse ponto de lado e retornar ao argumento principal. Parece-me que o argumento citado acima implica uma séria restrição à onipotência do Todo-poderoso. Admite-se que há certas coisas que Ele não pode fazer, tal como tornar um igual a dois, mas não deveríamos acreditar que tem a liberdade de conceber alma a um elefante, se quisesse? Poderíamos esperar que só exercesse tal poder por via de uma mutação que provesse o elefante de um cérebro devidamente aperfeiçoado para atender às necessidades de sua alma. Um argumento de todo semelhante pode ser invocado no caso das máquinas. Talvez pareça diferente porque é mais difícil de "engolir". Mas isso realmente significa apenas que acreditamos fosse menos provável que Ele considerasse as circunstâncias apropriadas para conferir uma alma. As circunstâncias em questão são discutidas no decorrer deste trabalho. Ao tentar construir tais máquinas não estaríamos irreverentemente usurpando-Lhe o poder de criar almas, como não o usurpamos ao procriar crianças: ao contrário, somos, em ambos os casos, instrumentos de Sua vontade providenciando moradas para as almas que Ele cria.

Entretanto, isto é mera especulação. Não estou muito impressionado com argumentos teológicos, quaisquer que sejam as circunstâncias nas quais são utilizados. Tais argumentos mostraram-se freqüentemente insatisfatórios no passado. Na época de Galileu, sustentou-se que os textos "O sol se deiteve no meio do céu e não se apressou a pôr-se, quase um dia inteiro." (Josué, cap. 10) e "Ele lançou os fundamentos da

que Deus não pode fazer um homem sem alma. Mas isto não pode ser uma restrição real de Seus poderes, mas unicamente resultado do fato de que as almas humanas são imortais, e portanto indestrutíveis.

terra, para que não se movesse em tempo algum." (Salmos, 104) eram refutações adequadas à teoria de Copérnico. Com o nosso conhecimento atual, semelhante argumento parece fútil. Quando esse conhecimento não existia, fazia uma impressão bem diferente.

(2) *A objeção das "Cabeças na Areia"*

"As conseqüências de máquinas pensantes seriam terríveis. Esperemos e confiemos em que não possam fazer isso."

Este argumento é raramente expresso de maneira tão ostensiva quanto na forma acima. Mas afeta a maioria de nós que pensa assim. Gostaríamos de acreditar que o Homem é, de alguma maneira sutil, superior ao resto da criação. É melhor que possa ser mostrado como *necessariamente* superior, pois aí não há perigo de ele perder sua posição de comando. A popularidade do argumento teológico está claramente ligada a esse sentimento. É provável que seja mais forte em pessoas intelectuais, pois elas valorizam o poder do pensamento mais que outras, e estão mais inclinadas a basear em tal poder a crença na superioridade do Homem.

Não creio que esse argumento seja suficientemente substancial para exigir refutação. Consolo seria mais apropriado: talvez deva ser buscado na transmigração das almas.

(3) *A objeção matemática*

Há certos resultados da lógica matemática que podem ser usados para mostrar que há limitações aos poderes das máquinas de estado discreto. O mais conhecido desses resultados é o teorema de Gödel (1931); mostra que em qualquer sistema lógico suficientemente poderoso podem-se formular enunciados que não são passíveis de prova ou refutação dentro do sistema, a menos que possivelmente o próprio sistema seja inconsistente. Há outros resultados semelhantes devidos a Church (1936), Kleene (1935), Rosser and Turing (1937).

Este último é o mais conveniente para considerar, de vez que diz respeito diretamente a máquinas, enquanto os outros só podem ser usados num argumento comparativamente indireto: por exemplo, para usar o argumento de Gödel, precisamos,

ademais, dispor de certos meios de descrever sistemas lógicos em termos de máquinas, e máquinas em termos de sistemas lógicos. O resultado em questão se refere a um tipo de máquina que é essencialmente um computador digital de capacidade infinita. Afirma este resultado que há certas coisas que uma máquina assim não pode realizar. Se ela estiver aparelhada para dar respostas a perguntas, como no jogo da imitação, haverá algumas perguntas às quais dará ou resposta errada ou nenhuma resposta, não importa quanto tempo se lhe conceda para responder. Haverá certamente muitas questões assim, que não podendo ser respondidas por uma máquina, poderão ser respondidas de modo satisfatório por outra. Claro que, de momento, estamos supondo que as perguntas sejam do tipo para o qual uma resposta "Sim" ou "Não" é adequada, e não perguntas como "Que acha de Picasso?". As perguntas que sabemos a máquina não conseguirá responder são deste tipo: "Considere a máquina a seguir especificada. (...) Responderá ela sempre "Sim" a qualquer pergunta?" Os pontos devem ser substituídos por uma descrição de alguma máquina, num modelo-padrão, que poderia ser semelhante ao usado no § 5. Quando a máquina descrita apresenta relação real comparativamente simples com a máquina que está sendo interrogada, pode-se demonstrar que a resposta ou é errada ou nula. Este é o resultado matemático: demonstra-se que prova uma incapacidade das máquinas à qual o intelecto humano não está sujeito.

A resposta mais simples a este argumento é a de que, embora esteja estabelecido que há limitações aos poderes de qualquer máquina específica, enunciou-se apenas, sem qualquer espécie de prova, que nenhuma limitação desse tipo se aplica ao intelecto humano. Mas não creio que tal concepção possa ser rejeitada tão levemente. Sempre que a qualquer uma dessas máquinas se faz a pergunta crítica devida, e ela dá uma resposta definida, sabemos que a resposta deverá estar errada, e isso nos infunde certo sentimento de superioridade. Será esse sentimento ilusório? É, sem dúvida, assaz genuíno, mas não creio que se lhe deva dar muita importância. Nós mesmos freqüentemente damos respostas erradas a perguntas, e assim não temos justificativa para estar muito satisfeitos com semelhante evidência de falibilidade por parte das máquinas.

Mais ainda: só nos podemos sentir superiores, nessa ocasião, relativamente à máquina específica sobre a qual conquistamos nosso insignificante triunfo. Não se trata de triunfar simultaneamente sobre *todas* as máquinas. Em resumo, haverá talvez homens mais inteligentes do que qualquer máquina dada, mas mesmo então poderão existir outras máquinas mais inteligentes, e assim por diante.

Os que defendem o argumento matemático estariam, creio eu, dispostos sobretudo a aceitar o jogo da imitação como base para discussão. Os que acreditam nas duas objeções anteriores provavelmente não estariam interessados em critério algum.

(4) O argumento da consciência

Este argumento está muito bem expresso no discurso "Lister" do Prof. Jefferson, de 1949, de onde transcrevo: "Somente quando uma máquina puder escrever sonetos ou compor concertos por causa de pensamentos ou emoções sentidas, e não por via de ocorrência casual de símbolos, é que concordaríamos em que a máquina iguala o cérebro — isto é, que não apenas os escreveu ou compôs como também sabia que os escrevera. Nenhum mecanismo poderia experimentar (e não meramente assinalar de modo artificial, por meio de uma engenhoca fácil) prazer pelos seus êxitos, tristeza quando suas válvulas queimam, deleite ante a lisonja; sentir-se infeliz por causa de seus erros, encantar-se com o sexo, ficar irritado ou deprimido por não poder alcançar o que deseja."

Este argumento parece ser uma negação da validade de nossa prova. De acordo com a forma mais extremada de tal concepção, a única maneira de a pessoa estar segura de que a máquina pensa é *ser* ela a própria máquina e sentir-se pensando. Poderia então descrever esses sentimentos ao mundo, mas naturalmente não se justificaria que alguém lhe desse atenção. De modo semelhante, de acordo com a mesma concepção, a única maneira de saber-se se um *homem* pensa é ser esse homem. Trata-se, de fato, do ponto de vista solipsista. Pode ser a concepção mais lógica a ser sustentada, mas dificulta a comunicação de idéias. *A* está sujeito a pensar que "*A* pensa, mas *B* não", mas da mesma maneira por que *B* acredita que "*B* pensa, mas *A* não". Em vez de discutir continuamente

tal questão, é de hábito aceitar-se a convenção cortês de que toda gente pensa.

Tenho a certeza de que o Prof. Jefferson não deseja adotar esse ponto de vista extremado e solipsista. Ele provavelmente concordaria em aceitar o jogo da imitação como uma prova. O jogo (com omissão do jogador B) é freqüentemente usado na prática sob o nome de *viva voce*, para descobrir se alguém de fato compreende algo ou se o "decorou como um papagaio". Vamos ouvir parte de um *viva voce*:

Interrogador: No primeiro verso do seu soneto, que diz "Devo eu te comparar a um dia de verão", "um dia de primavera" não estaria igualmente bem ou ainda melhor?

Testemunha: Não tem o número certo de sílabas.

Interrogador: Que tal "um dia de inverno"? Metricamente, dá certo.

Testemunha: Mas ninguém quer ser comparado a um dia de inverno.

Interrogador: Por exemplo, você diria que o Sr. Pickwick faz lembrar o Natal?

Testemunha: De certo modo.

Interrogador: Contudo, o Natal é um dia de inverno, e não creio que o Sr. Pickwick fizesse objeções a essa comparação.

Testemunha: Não creio que você esteja falando a sério. Quando se diz "um dia de inverno", quer-se dizer um dia típico de inverno, não um dia especial como o Natal.

E assim por diante. O que diria o Prof. Jefferson se a máquina de escrever sonetos tivesse capacidade para responder assim no *viva voce*? Não sei se ele consideraria que a máquina estava "apenas sinalizando artificialmente" tais respostas, se as respostas fossem tão satisfatórias e compreensivas como na passagem acima, mas não creio que ele a descrevesse como "uma engenhoca fácil". Esta frase, a meu ver, visa a designar dispositivos tais como a inclusão, na máquina, da gravação de alguém a ler um soneto, com um interruptor para ligá-la e desligá-la de quando em quando. Em suma, penso que a maioria dos que sustentam o argumento da consciência poderia ser persuadida a abandoná-lo em lugar de ver-se coagida a assumir uma

posição solipsista. Provavelmente, estaria então disposta a aceitar a nossa prova.

Não quero dar a impressão de que penso não exista nenhum mistério no que respeita à consciência. Existe, por exemplo, algo assim como que um paradoxo vinculado às tentativas de localizá-la. Mas não acredito que tais mistérios tenham de ser necessariamente resolvidos antes de podermos responder à pergunta que nos preocupa neste artigo.

(5) *Argumentos de várias incapacidades*

Esses argumentos assumem a forma de "Concordo em que você é capaz de construir máquinas que façam todas as coisas mencionadas, mas você nunca conseguirá construir uma máquina que faça X." Numerosas características X são sugeridas, neste particular. Eis uma seleção delas.

Seja bondosa, atilada, bela, amigável; tenha iniciativa; tenha senso de humor; distinga o certo do errado; cometa erros; apaixone-se; delicie-se com morangos com creme; faça alguém apaixonar-se por você; aprenda com a experiência; use corretamente as palavras; seja o tema de seus próprios pensamentos; tenha tanta diversidade de pensamento quanto um homem; faça algo de realmente novo.

Nenhuma justificativa é, no geral, oferecida para esses enunciados. Creio que estão, na maior parte, alicerçados no princípio da indução científica. Um homem viu milhares de máquinas durante a sua vida. Do que observou, tira certo número de conclusões gerais. As máquinas são feias; cada uma delas foi construída para um objetivo muito limitado; quando se lhes pede algo um pouco diferente, elas se tornam inúteis; a variedade de comportamento de qualquer uma delas é muito pequena; etc., etc. Naturalmente, a conclusão é a de que tais serão, necessariamente, propriedades das máquinas em geral. Muitas limitações advêm da capacidade de memória assaz reduzida dessas máquinas (dou por entendido que a idéia de capacidade de memória seja ampliada de alguma maneira, de forma a abranger outras máquinas que não as de estado discreto. A definição exata não importa, de vez que não se pretende nenhuma precisão matemática na presente discussão). Há alguns anos atrás, quando quase não se falava de computadores digi-

tais, era possível suscitar muita incredulidade a respeito deles caso se lhes mencionasse as propriedades sem descrever-lhes a construção. Isso se devia, presumivelmente, a uma aplicação semelhante do princípio da indução científica. Tais aplicações do princípio, na maioria dos casos, são evidentemente inconscientes. Uma criança que já se queimou teme o fogo, e mostra que o teme evitando-o; está aplicando, ao meu ver, o princípio da indução científica. (Claro que eu poderia também descrever o seu comportamento de muitos outros modos.) As obras e hábitos da humanidade não parecem ser material muito apropriado a que aplicar a indução científica. Grande parte do espaço-tempo deve ser investigada, se se quiserem resultados seguros. De outro modo, podemos decidir (como a maioria das crianças inglesas), que toda a gente fala inglês, e que é bobagem aprender francês.

Há, contudo, observações especiais a serem feitas acerca das incapacidades que foram mencionadas. A incapacidade de saborear morangos com creme talvez tenha impressionado o leitor como frívola. Possivelmente, pode-se fazer com que uma máquina saboreie esse prato delicioso, mas tal empenho seria idiota. O que é importante, no relativo a essa capacidade, é que ela contribua para algumas das outras capacidades, por exemplo: a dificuldade de ocorrer entre homem e máquina a mesma espécie de amizade existente entre um homem branco e outro, ou entre dois homens pretos.

A alegação de que "as máquinas não podem cometer erros" parece curiosa. Sente-se a tentação de responder: "Será que elas são piores por causa disso?" Mas vamos adotar uma atitude mais simpática e procurar ver o que isso realmente significa. Penso que tal crítica pode ser explicada em termos do jogo da imitação. Pretende-se que o interrogador possa distinguir a máquina do homem com simplesmente propor-lhes certo número de problemas de aritmética. A máquina seria desmascarada devido à sua precisão mortal. A resposta a esse argumento é simples. A máquina (programada para jogar o jogo) não procuraria dar respostas *corretas* aos problemas de aritmética. Introduziria deliberadamente erros, de um modo calculado a confundir o interrogador. Um defeito mecânico provavelmente se revelaria através de uma decisão inapropriada quanto a que espécie de erro de aritmética seria feito. Mes-

mo tal interpretação da crítica não é suficientemente simpática. Não dispomos, entretanto, de espaço bastante para continuar a examiná-la. A meu ver, semelhante crítica deriva de uma confusão entre dois tipos de erros. Podemos chamar-lhes "erros de funcionamento" e "erros de conclusão". Os erros de funcionamento são devidos a uma falha mecânica ou elétrica que faz com que a máquina se comporte de modo diferente do planejado. Nas discussões filosóficas, prefere-se ignorar a possibilidade de tais erros; discutem-se então "máquinas abstratas". Essas máquinas abstratas são antes ficções matemáticas que objetos físicos. Por definição, são incapazes de erros de funcionamento. Nesse sentido, podemos dizer realmente que "máquinas nunca podem errar". Erros de conclusão só podem surgir quando algum significado é atribuído aos sinais de saída da máquina. A máquina poderia, por exemplo, escrever equações matemáticas ou frases em inglês.

Quando uma proposição falsa é escrita, dizemos que a máquina cometeu um erro de conclusão. Claro que não existe razão alguma para afirmar que a máquina não possa cometer erro desse tipo. Ela poderia não fazer mais que escrever repetidamente " $0 = 1$ ". Para usar um exemplo menos despropositado, a máquina poderia ter algum método de tirar conclusões por indução científica. Devemos esperar que tal método leve ocasionalmente a resultados errôneos.

A alegação de que uma máquina não pode ser o tema de seu próprio pensamento só poderia ser respondida, evidentemente, se se lograsse demonstrar que a máquina tem *algum pensamento* sobre *algum tema*. Entretanto, o "tema das operações de uma máquina" parece de fato significar alguma coisa, pelo menos para as pessoas que lidam com ela. Se, por exemplo, a máquina estivesse procurando achar uma solução para a equação $X^2 - 40X - 11 = 0$, estaríamos tentados a descrever tal equação como parte do tema da máquina naquele momento. Neste sentido, uma máquina pode, sem dúvida, ser seu próprio tema. Pode ser usada para participar na criação dos seus próprios programas, ou para predizer o efeito das alterações de sua própria estrutura. Pela observação dos resultados de seu próprio comportamento, pode modificar seus próprios programas, de modo a alcançar algum objetivo mais eficazmente. Estas são possibilidades do futuro próximo, não sonhos utópicos.

A crítica de que uma máquina não pode ter muita diversidade de comportamento é simplesmente uma maneira de dizer que ela não pode ter muita capacidade de memória. Até bem recentemente, uma capacidade de memória de mesmo mil dígitos era coisa muito rara.

As críticas que estamos aqui considerando são amiúde formas disfarçadas no argumento da consciência. Em geral, se alguém sustentar que uma máquina *pode* fazer uma dessas coisas, e descrever o tipo de método que a máquina poderia usar, não causaria grande impressão. Pensa-se que o método (qualquer que seja, pois deve ser mecânico) é realmente pouco importante. Comparem-se os parêntesis no argumento de Jefferson, citado páginas atrás.

(6) *As objeções de Lady Lovelace*

Nossa informação mais pormenorizada sobre a máquina analítica de Babbage vem de uma dissertação de Lady Lovelace (1842). Nela, declara-se que “a Máquina Analítica não tem nenhuma pretensão de *criar* o que quer que seja. Pode fazer tudo quanto sabemos ordenar-lhe que faça (o grifo é de Lady Lovelace).” Esse enunciado é referido por Hartree (1949), que acrescenta: “Isso não implica que não seja possível construir equipamento eletrônico que “pense por si mesmo”, ou no qual, em termos biológicos, alguém possa estabelecer um reflexo condicionado, que serviria de base para a “aprendizagem”. A questão, muito estimulante, de se isso é ou não possível em princípio, foi sugerida por algum desses recentes desenvolvimentos. Mas não parece que as máquinas construídas ou projetadas naquela época tivessem essa propriedade.”

Estou de completo acordo com Hartree neste ponto. Deve-se observar que ele não sustenta que as máquinas em questão não tinham essa propriedade, mas antes que os indícios de que dispunha Lady Lovelace não a encorajavam a crer que tivessem. É muito possível que as máquinas em questão possuísem, de certo modo, tal propriedade. Suponhamos que uma máquina de estado discreto a possuía. A Máquina Analítica era um computador digital universal, de forma que, caso sua capacidade de memória e sua rapidez fossem adequadas ela poderia, por meio de uma programação conveniente, ser levada a imitar a máquina

em questão. Este argumento provavelmente não ocorreu à Condessa ou a Babbage. De qualquer modo, eles não tinham a obrigação de alegar tudo quanto pudesse ser alegado.

A questão toda pode ser considerada de novo sob a rubrica de máquinas que aprendem.

Uma variante da objeção de Lady Lovelace afirma que a máquina não pode nunca “fazer algo de realmente novo”. A tal argumento poderíamos opor o velho ditado de que “Não há nada de novo sob o sol”. Quem pode ter a certeza de que a “obra original” que fez não foi simplesmente o crescimento de uma semente em si plantada pelo ensino ou o efeito de seguir princípios gerais conhecidos? Uma variante melhor da objeção diz que a máquina jamais “nos pode pegar de surpresa”. Essa afirmativa é um desafio mais direto e pode ser enfrentada de modo direto. As máquinas me pegam de surpresa com muita freqüência. Isso ocorre em grande parte porque não faço cálculos suficientes para decidir o que devo esperar que façam, ou antes porque embora eu faça cálculos, faço-os apressadamente, assumindo riscos. Talvez eu diga comigo: “Suponho que a voltagem aqui deve ser a mesma de lá; de qualquer modo, vamos admitir que seja.” Naturalmente, engano-me amiúde e o resultado é uma surpresa para mim, porque, à altura em que o experimento se realiza, tais suposições foram esquecidas. Estas admissões podem tornar-me vulnerável a admoestações pelos meus vícios, mas não lançam dúvida alguma sobre a credibilidade do meu testemunho quanto às surpresas que experimentei.

Não espero que semelhante resposta possa silenciar o meu crítico. Ele dirá provavelmente que tais surpresas são devidas a algum ato mental criativo de minha parte e que não conferem maior mérito à máquina. Isto nos leva de volta ao argumento da consciência e para longe da idéia de surpresa. É uma linha de argumentação que devemos considerar encerrada, mas talvez valha notar que a apreciação de algo como surpreendente requer, de igual maneira, um “ato mental criativo”, quer o acontecimento surpreendente provenha de um homem, de um livro, de uma máquina ou de qualquer outra coisa.

A concepção de que as máquinas não podem suscitar surpresas se deve, creio eu, a uma falácia a que estão particularmente sujeitos filósofos e matemáticos. Refiro-me à suposição de que, tão logo um fato seja apresentado à mente, todas as con-

seqüências desse fato se impõem à mente simultaneamente com ele. Trata-se de uma suposição útil em muitas circunstâncias, mas esquece-se facilmente que é falsa. Uma consequência natural desse esquecimento é a de pensar que não há mérito algum na simples dedução de consequências a partir de dados e princípios gerais.

(7) *O argumento da continuidade do sistema nervoso*

O sistema nervoso não é certamente uma máquina de estado discreto. Um pequeno erro de informação acerca da grandeza de um impulso nervoso que atinja um neurônio pode influenciar seriamente a grandeza do impulso de saída. Pode-se argumentar que, sendo assim, não é de esperar seja possível imitar o comportamento do sistema nervoso com um sistema de estado discreto.

É verdade que uma máquina de estado discreto tem de ser diferente de uma máquina contínua. Mas se nos ativermos às condições do jogo da imitação, o interrogador não será capaz de tirar qualquer vantagem dessa diferença. A situação poderá ser aclarada se considerarmos outras máquinas contínuas mais simples. Um analisador diferencial servirá muito bem. (O analisador diferencial é um certo tipo de máquina que não a de estado discreto, usado em alguns tipos de cálculo.) Alguns analisadores fornecem suas respostas em forma datilografada, e assim mostram-se adequados para tomar parte no jogo. Não seria possível a um computador digital predizer exatamente que respostas o analisador diferencial daria a um problema, mas ele seria bem capaz de dar o tipo certo de resposta. Por exemplo, se requisitado a dar o valor de π (na verdade, cerca de 3,1416), seria razoável escolher ao acaso entre os valores 3,12, 3,13, 3,14, 3,15, 3,16, com as probabilidades de 0,05, 0,15, 0,55, 0,19, 0,06 (digamos). Nestas circunstâncias, seria muito difícil para o interrogador distinguir o analisador diferencial de um computador digital.

(8) *O argumento da informalidade do comportamento*

Não é possível produzir um conjunto de regras que pretenda descrever o que um homem deveria fazer em cada circuns-

tância imaginável. Alguém pode, por exemplo, ter como regra parar quando vê a luz vermelha do semáforo, e prosseguir quando vê o sinal verde, mas o que aconteceria se, por alguma falha, ambos os sinais aparecessem conjuntamente? Pode-se talvez decidir que é mais seguro parar. Mas alguma dificuldade posterior talvez resulte dessa decisão. Tentar oferecer regras de conduta para abarcar todas as eventualidades, mesmo as oriundas de semáforos, parece impossível. Com tudo isso eu concordo.

A partir de tal situação, sustenta-se que não podemos ser máquinas. Tentarei reproduzir o argumento, mas temo que dificilmente lhe farei justiça. Parece ser algo assim como: "Se cada homem tivesse um conjunto de regras de conduta definidas pelas quais regulasse sua vida, não seria melhor do que uma máquina. Mas não existem tais regras, de modo que os homens não podem ser máquinas." A premissa menor não é abrangente. Não acredito que o argumento seja sempre formulado exatamente dessa maneira, mas acredito, não obstante, que é o argumento usado. Pode, todavia, haver certa confusão entre "regras de conduta" e "leis de comportamento" a obscurecer a questão. Por "regras de conduta" quero dizer preceitos tais como "Pare quando avistar luz vermelha", com base nos quais se pode agir e dos quais se pode estar consciente. Por "leis de comportamento" quero dizer leis da natureza aplicadas a um corpo humano, tais como "Se você o beliscar, ele berrará". Se substituirmos "leis de comportamento que regulam sua vida" por "leis de conduta pelos quais ele regula sua vida", no argumento acima, a premissa menor não é mais insuperável. Pois acreditamos não apenas seja verdade que ser governado por leis de comportamento implica ser-se algum tipo de máquina (embora não necessariamente uma máquina de tipo discreto) como também, inversamente, que ser uma máquina assim implica ser governado por tais leis. Entretanto, não podemos convencer-nos facilmente da ausência de leis completas de comportamento, bem como da de regras completas de conduta. A única maneira que conhecemos de encontrar leis que tais é a observação científica, e não sabemos absolutamente de quaisquer circunstâncias em que pudéssemos dizer: "Já investigamos o suficiente. Não existem tais leis."

Podemos demonstrar mais convincentemente que uma afirmação desse tipo seria injustificada. Pois suponhamos que pu-

déssemos estar seguros de encontrar tais leis caso elas existissem. Então, dada uma máquina de estado discreto, seria certamente possível descobrir, por observação, o bastante sobre ela para predizer seu comportamento futuro, e isso dentro de um tempo razoável, digamos mil anos. Mas não parece ser esse o caso. Estabeleci no computador de Manchester um pequeno programa que usa somente 1 000 unidades de memória, e no qual a máquina, alimentada com um número de 16 algarismos, responde com outro em dois segundos. Eu desafiaria qualquer pessoa a descobrir, com base nessas respostas, o suficiente acerca do programa para poder predizer quaisquer respostas a valores não experimentados.

(9) *O argumento da percepção extra-sensorial*

Presumo que ao leitor seja familiar a idéia da percepção extra-sensorial e o significado de seus quatro itens: telepatia, clarividência, precognição psicocinética. Estes fenômenos perturbadores parecem negar todas as nossas idéias científicas habituais. Como gostaríamos de desacreditá-los! Infelizmente, os indícios estatísticos, pelos menos quanto à telepatia, são esmagadores. É muito difícil recompor nossas idéias para nelas encaixar esses novos fatos. Uma vez que tenhamos de aceitá-los, não será despropositado acreditar em fantasmas e espectros. A idéia de que nossos corpos se movem tão-somente de acordo com leis conhecidas da física, a par de outras ainda não descobertas, mas de alguma forma semelhantes, seria uma das primeiras a desaparecer.

Esse argumento é, para mim, muito forte. Pode-se dizer, em resposta, que muitas teorias científicas parecem permanecer viáveis na prática, a despeito de colidirem com a PES; que, de fato, se pode viver muito bem se a esquecermos. Isso é muito pouco confortador, sendo de temer que o pensamento seja exatamente o tipo de fenômeno em que a PES possa ser especialmente importante.

Um argumento mais específico, baseado na PES, é formulável como segue: "Jogamos o jogo da imitação usando como testemunha um homem que seja tão bom como receptor telepático quanto como computador digital. O interrogador pode fazer perguntas como: 'A que naipe pertence a carta na

minha mão direita?' O homem, por telepatia ou clarividência, dá a resposta certa 130 vezes em 400 tentativas. A máquina só pode adivinhar ao acaso e talvez conseguir 104 respostas corretas, pelo que o interrogador faz a identificação certa." Há uma possibilidade interessante, que se abre aqui. Suponhamos que o computador digital contenha um gerador de números aleatórios. Será natural então usá-lo para decidir qual a resposta a dar. Mas nesse caso o gerador de números aleatórios estará sujeito aos poderes psicocinéticos do interrogador. Talvez esta psicocinética possa fazer com que a máquina adivinhe com maior frequência do que a esperável num cálculo de probabilidades, de forma que o interrogador seria ainda incapaz de fazer a identificação correta. Por outro lado, talvez ele fosse capaz de adivinhar corretamente, sem fazer pergunta alguma, por meio da clarividência. Com a PES, tudo pode acontecer.

Se a telepatia for admitida, será necessário "apertar" a nossa prova. Poder-se-ia encarar a situação como análoga à que ocorreria se o interrogador falasse consigo próprio e um dos competidores estivesse à escuta, o ouvido colado à parede. Colocar os competidores numa "sala à prova de telepatia" satisfaria todos os requisitos.

7. Máquinas que aprendem

O leitor terá percebido que eu não disponho de muitos argumentos convincentes de natureza positiva para sustentar os meus pontos de vista. Se dispusesse, não teria tido tanto trabalho em apontar as falácias dos pontos de vista contrários. Enumerarei agora os indícios de que disponho.

Retornemos por um momento à objeção de Lady Lovelace, que sustenta ser a máquina capaz tão-somente de fazer o que lhe ordenarmos que faça. Digamos então que um homem poderia "injetar" uma idéia na máquina e esta responderia em certa medida, voltando em seguida à imobilidade, como uma corda de piano percutida por um martelo. Outra analogia seria uma pilha atômica, de tamanho menor que o crítico: uma idéia injetada corresponderá a um nêutron a entrar na pilha, vindo de fora. Cada nêutron causará certa perturbação, que por fim desaparecerá. Se, contudo, o tamanho da pilha for aumentado

o bastante, a perturbação causada pelo nêutron entrante muito provavelmente continuará a aumentar até que toda a pilha seja destruída. Existe, acaso, um fenômeno correspondente para mentes, e outro para máquinas? Parece haver um para a mente humana. A maioria delas parece ser "subcrítica", isto é, correspondente, nesta analogia, a pilhas de tamanho subcrítico. Uma idéia apresentada a uma mente tal dará, em média, menos de uma idéia em resposta. Pequena proporção é supercrítica. Uma idéia apresentada a uma mente dessas pode dar origem a toda uma "teoria", consistente de idéias secundárias, terciárias e mais remotas. As mentes animais parecem ser definitivamente subcríticas. Firmados nessa analogia, perguntamos: "Pode-se fazer com que uma máquina seja supercrítica?"

A analogia da "casca de cebola" também é útil. Ao considerar as funções da mente ou do cérebro, encontramos certas operações que podemos explicar em termos puramente mecânicos. Isto que dizemos não corresponde à mente real: é uma espécie de casca que temos de retirar se quisermos chegar à mente verdadeira. Mas então, no que resta, achamos outra casca a ser arrancada, e assim por diante. Procedendo dessa maneira, chegaremos à mente real ou acabaremos por chegar à casca que nada contém? Neste caso, a mente toda é mecânica. (Não seria entretanto uma máquina de estado discreto. Já discutimos isso.)

Estes dois últimos parágrafos não se pretendem argumentos convincentes. Deveriam antes ser descritos como "recitações destinadas a suscitar crença."

O único apoio realmente satisfatório que pode ser dado à concepção expressa no princípio do parágrafo 6 será esperar o fim do século para então realizar o experimento descrito. Mas o que podemos dizer, entretanto? Que providências devem ser tomadas desde agora para que o experimento seja bem sucedido?

Como já expliquei, o problema é principalmente de programação. Progressos de engenharia terão também de ser feitos, mas parece pouco provável que não venham a satisfazer as exigências. As estimativas da capacidade de memória do cérebro variam de 10^{10} a 10^{15} dígitos binários (bits). Inclino-me para valores mais baixos e acredito que apenas uma pequena

fração dessa capacidade seja usada para os tipos superiores de pensamento. A maior parte dela é provavelmente usada para a retenção de impressões visuais. Eu ficaria surpreso se mais que 10^9 bits fossem necessários para atuação satisfatória no jogo da imitação, pelo menos contra um homem cego. (Nota: a capacidade da Enciclopédia Britânica, 11.^a edição, é de 2×10^9). Uma capacidade de memória de 10^7 bits seria uma possibilidade muito prática, mesmo pelas técnicas atuais. Talvez não seja absolutamente necessário aumentar a velocidade de operação das máquinas. Setores de máquinas modernas que podem ser considerados como análogos das células nervosas trabalham cerca de 1 000 vezes mais rapidamente que elas. Isto proveria uma margem de segurança, que poderia cobrir perdas de velocidade de muitas origens. Nosso problema é então descobrir como programar tais máquinas para jogar o jogo da imitação. No meu ritmo de trabalho atual, produzo cerca de 1 000 dígitos de programação por dia, de modo que uns 60 operadores, trabalhando sem interrupção durante cerca de 50 anos, poderiam levar a cabo o trabalho, se não se desperdiçasse nada. Alguns métodos mais expeditos parecem ser desejáveis.

No processo de tentar imitar a mente humana adulta, temos de refletir bastante sobre o processo que a levou até o ponto onde se encontra. Cumpre atentar para três componentes:

- (a) O estado inicial da mente, isto é, ao nascer;
- (b) A educação que recebeu;
- (c) Outras experiências, que não as descritas como educação, a que foi submetida.

Em vez de tentar produzir um programa que simule a mente adulta, por que não tentar produzir um que simule a mente infantil? Se ele fosse então submetido à educação apropriada, ter-se-ia um cérebro adulto. Presumivelmente, o cérebro da criança é algo assim como um desses cadernos que se compram em papelaria. Pouco mecanismo e muitas folhas em branco. (Mecanismo e escrita são, do nosso ponto de vista, quase sinônimos.) Nossa esperança é a de que haja tão pouco mecanismo no cérebro da criança que algo que se lhe assemelhe possa ser programado. Numa primeira aproximação, a soma de trabalho gasto na educação pode ser considerada equivalente à gasta na educação da criança humana.

Dividimos assim nosso problema em duas partes: o programa infantil e o processo de educação. Estas duas partes permanecem intimamente ligadas. Não podemos esperar encontrar uma boa máquina-criança logo na primeira tentativa. Deve-se experimentar ensinar uma máquina que tal para ver como ela aprende. Pode-se então tentar ensinar outra para ver se se sai melhor ou pior. Há uma conexão óbvia entre esse processo e a evolução, por via destas identificações:

Estrutura da máquina-criança = material hereditário

Mudanças na máquina-criança = mutações

Seleção natural = juízo do experimentador

Pode-se esperar, contudo, que este processo seja mais expedito que a evolução. A sobrevivência do mais apto é um método vagaroso de medir vantagens. O experimentador, valendo-se da inteligência, deveria ser capaz de acelerá-lo. Igualmente importante é o fato de que o experimentador não está restrito às mutações casuais. Se conseguir determinar a causa de alguma fraqueza, poderá provavelmente pensar no tipo de mutação apto a superá-la.

Não será possível aplicar à máquina exatamente o mesmo processo de ensino que se aplica a uma criança normal. A máquina, por exemplo, não disporá de pernas; assim, não se pode pedir-lhe que vá buscar um balde de carvão lá fora. Possivelmente não terá olhos. Todavia, ainda que essas deficiências sejam superadas por uma engenharia inteligente, não se poderia mandar a criatura à escola sem que as outras crianças a escarnecessem. Melhor arranjar-lhe um professor particular. Não é preciso que nos preocupemos tanto com pernas, olhos etc. O exemplo de Helen Keller demonstra que a educação é possível desde que a comunicação em ambas as direções, entre professor e aluno, ocorra de alguma forma.

Normalmente, associamos punições e recompensas ao processo de ensino. Algumas máquinas-crianças simples podem ser construídas ou programadas de acordo com esse tipo de princípio. A máquina tem de ser construída de forma tal que os acontecimentos que antecedem imediatamente a ocorrência de um sinal de punição sejam de repetição improvável, ao passo que um sinal de recompensa aumente a probabilidade de repetição dos acontecimentos que levaram a ela. Estas definições

não pressupõem quaisquer sentimentos por parte da máquina. Fiz alguns experimentos com uma máquina-criança assim e conseqüi ensinar-lhe algumas coisas, mas o método de ensino era muito pouco ortodoxo para que o experimento pudesse ser considerado realmente bem sucedido.

O uso de punições e recompensas pode, no melhor dos casos, constituir uma parte do processo de ensino. *Grosso modo*, se o professor não dispõe de outro meio de comunicação com o aluno, a quantidade de informação que pode alcançá-lo, não excede o número total de recompensas e punições aplicadas. Uma criança que tivesse de aprender a repetir "Casabianca" provavelmente ficaria muito dolorida se o texto só pudesse ser descoberto por uma técnica de "Vinte Perguntas" em que cada "Não" assumisse a forma de uma palmada. Cumpre, por isso, dispor de outros canais de comunicação "não-emotivos". Se estiverem disponíveis, será possível ensinar uma máquina, através de punições e recompensas, a obedecer ordens dadas em alguma linguagem, por exemplo uma linguagem simbólica. Essas ordens deverão ser transmitidas através de canais "não-emocionais". O uso desta linguagem diminuirá enormemente o número de recompensas e punições necessárias.

As opiniões podem variar quanto à complexidade conveniente para a máquina-criança. Pode-se tentar fazê-la tão simples quanto possível, dentro dos princípios gerais. Alternativamente, pode-se optar por um sistema completo de inferência lógica "embutido".(*) Neste caso, a memória seria ocupada em grande parte por definições e proposições. As proposições seriam de vários tipos; por exemplo, fatos bem estabelecidos, conjeturas, teoremas matemáticos demonstrados, enunciados de autoridade, expressões que tenham a forma lógica de proposição mas não de valor-crença. Certas proposições podem ser descritas como imperativas. A máquina deve ser construída de tal forma que tão logo um imperativo seja classificado com "bem estabelecido", a ação apropriada ocorra automaticamente. Para ilustrar isso, suponhamos que o professor diga à máquina: "Faça agora seus deveres de casa". Isto pode resultar em que "O pro-

(*) Ou antes "programado", pois a nossa máquina-criança será programada num computador digital. O sistema lógico, porém, não precisará ser aprendido.

fessor disse: faça agora seus deveres de casa” seja incluído entre os fatos bem estabelecidos. Outro fato pode ser: “Tudo o que o professor diz é verdade”. A combinação dos dois fatos pode levar o imperativo “Faça agora seus deveres de casa” a ser incluído entre os fatos bem estabelecidos, e isso, dada a construção da máquina, significará que os deveres de casa em verdade começaram a ser feitos; o efeito, porém, é deveras satisfatório. O processo de inferência utilizado pela máquina não precisa ser de molde a satisfazer os lógicos mais exigentes. Pode não haver, por exemplo, hierarquia de tipos. Mas isso não tem de significar que falácias de tipo ocorram com frequência assim como não estamos sujeitos a despencar de penhascos sem cerca de proteção. Imperativos adequados (expressos dentro dos sistemas, não fazendo parte das regras do sistema), tais como: “Não use uma classe a menos que seja subclasse de outra mencionada pelo professor”, podem ter efeito similar a: “Não chegue tão perto da borda do penhasco”.

Os imperativos que podem ser obedecidos por uma máquina não dotada de braços e pernas estão limitados a ser de caráter assaz intelectual, como no exemplo acima (fazer os deveres de casa). Importantes entre tais imperativos hão de ser aqueles que regulam a ordem em que as regras do sistema lógico implicado serão aplicadas. Porque, a cada estágio, quando alguém está usando um sistema lógico, há um grande número de escolhas alternativas, qualquer um deles de aplicação possível no que concerne à obediência das regras do sistema lógico. Essas escolhas marcam a diferença entre um argumentador brilhante e outro inepto, mas não a diferença entre um argumentador correto e outro sofismador. Proposições que conduzam a imperativos dessa espécie poderiam ser: “Quando se mencione Sócrates, use o silogismo de Barbara”, ou: “Se um método é mais rápido que outro, não use o método mais vagaroso”. Algumas destas proposições podem ser dadas “por autoridade”, mas outras podem ser produzidas pela própria máquina, isto é, por indução científica.

A idéia de uma máquina que aprende talvez pareça paradoxal a alguns leitores. Como podem as regras de operação da máquina mudar? Elas deveriam descrever completamente como a máquina irá reagir, qualquer que possa ser sua história, quaisquer que sejam as mudanças que sofra. As regras são, pois, de-

veras invariantes no tempo. Isto é bem verdade. A explicação do paradoxo está em que as regras que mudam nos processos de aprendizagem são de caráter menos pretencioso, aspiram apenas a uma validade efêmera. O leitor pode compará-las à Constituição dos Estados Unidos.

Uma característica importante da máquina que aprende é a de que seu professor ignorará amiúde a maior parte do que se está passando no interior da máquina, embora possa até certo ponto predizer o comportamento de sua aluna. Isto se aplicaria muito mais à educação ulterior de uma máquina originada de uma máquina-criança de projeto (ou programa) bem experimentado, e contrasta claramente com o procedimento normal quando se usa uma máquina em operações de computação: o objetivo é então ter um nítido quadro mental do estado da máquina a cada momento da computação. Tal objetivo só pode ser alcançado com luta. A concepção de que “a máquina pode fazer somente aquilo que saibamos como ordenar-lhe que faça” parece estranha em face disso. (*) A maioria dos programas que podemos colocar na máquina dará como resultado ela fazer algo que não consigamos absolutamente entender, ou que consideramos como comportamento de todo aleatório. É de presumir-se que o comportamento inteligente consista num desvio do comportamento inteiramente disciplinado implicado em computação, mas desvio pequeno, que não dê margem a comportamento aleatório ou a voltas repetitivas, sem objetivo. Outro resultado importante de, por meio de um processo de ensino e aprendizagem, prepararmos nossa máquina para seu papel no jogo da imitação é o de que a “falibilidade humana” será provavelmente omitida de maneira natural, isto é, sem “preparação” especial. (O leitor deve conciliar isto com o ponto de vista do argumento número 5, discutido páginas atrás.) Processos aprendidos não propiciam uma certeza de cem por cento quanto ao resultado; se a propiciassem, não poderiam ser de aprendidos.

Será provavelmente sensato incluir um elemento aleatório numa máquina aprendiz. Um elemento aleatório é bastante útil quando estamos buscando a solução de um problema. Suponha-

(*) Compare-se com a afirmação de Lady Lovelace, que não contém a palavra “somente”.

-se, por exemplo, que queiramos encontrar um número entre 50 e 200 que seja igual ao quadrado da soma de seus algarismos; poderíamos começar com 51, depois experimentar 52, e continuar assim até encontrar um número que satisfizesse aquela condição. Alternativamente, poderíamos escolher números ao acaso até achar um que servisse. Este método tem a desvantagem de que se pode experimentar o mesmo número duas vezes. Isso não será grande inconveniente se existirem várias soluções. O método sistemático tem a desvantagem de que pode haver um enorme bloco sem qualquer solução na região que tem de ser investigada primeiramente. O processo de aprendizagem pode ser considerado como busca de uma forma de comportamento que satisfaça ao professor (ou a algum outro critério). Como provavelmente existe um número muito grande de soluções satisfatórias; o método aleatório parece ser melhor que o sistemático. Cumpre notar que ele é usado no processo análogo da evolução. Mas nesta o método sistemático não é possível. Como se poderiam guardar todas as diferentes combinações genéticas tentadas, de modo a evitar sua repetição?

É de esperar que as máquinas acabem por competir com o homem em todos os campos puramente intelectuais. Quais, porém, os melhores para começar? Mesmo esta é uma decisão difícil. Muitas pessoas acham que uma atividade assaz abstrata, como o jogo de xadrez, seria o melhor. Pode-se também sustentar que o mais conveniente é prover as máquinas dos melhores órgãos sensoriais que o dinheiro possa comprar, e ensiná-la a compreender e falar inglês. Tal processo poderia acompanhar o do ensino normal de uma criança. Coisas seriam apontadas e nomeadas, etc. Mais uma vez, não sei qual a resposta certa, mas penso que ambos os enfoques deveriam ser tentados.

Podemos avistar só um pequeno trecho do caminho à nossa frente, mas ali já vemos muito do que precisa ser feito.

[Título original: "Computing Machinery and Intelligence". Extraído de: Edward A. Feigenbaum e Julien Seldman (org.), *Computers and Thought*, Nova York, McGraw-Hill, 1963. Tradução de Marcia Epstein.]

CIBERNÉTICA E AÇÃO

A. MOLES

A cibernética se define como a ciência geral dos sistemas ou organismos, independentemente da natureza física dos elementos ou órgãos que os constituem. Estuda, portanto, estruturas de *relações* entre elementos e não propriamente os elementos em causa, que reduz a certo número de propriedades funcionais, traduzindo essa atitude no denominar esses elementos *átomos de estruturas* — ou "caixas negras", segundo a terminologia adotada pelos anglo-saxões. A cibernética se coloca, pois, entre as disciplinas gerais do espírito, como a lógica, o pensamento filosófico, a metodologia ou a matemática.

(I) História da cibernética

Sabemos que a cibernética recebeu tal nome de Wiener e von Neumann, por volta de 1945, depois da série de trabalhos do primeiro, a propósito dos processos de regulação, e dos trabalhos do segundo, acerca da analogia entre os mecanismos do cérebro humano e os dos primeiros computadores — aos quais von Neumann deu importante contribuição. Essas idéias, entretanto, já haviam sido examinadas de modo mais ou menos preciso por outros autores. Sem remontar a Platão, que caracterizou as atribuições do piloto (*kubernetes*) ou a Ampère, que viu a cibernética em termos de uma ciência do governo, os princípios da regulação foram objeto de importantes trabalhos de matemáticos (Volterra, Van Der Pol, Kolmogoroff) e de engenheiros (Hermann Schmidt e Zuse). A ci-

ISAAC EPSTEIN
(Organizador)

CIBERNÉTICA
E
COMUNICAÇÃO

Tradutores:

ISAAC EPSTEIN, HAROLDO DE CAMPOS, LEONIDAS HEGENBERG,
MARCIA EPSTEIN, MARIA LUCIA CACCIOLA e
OCTANNY SILVEIRA DA MOTA



EDITORA CULTRIX
SÃO PAULO
EDITORA DA UNIVERSIDADE DE SÃO PAULO

1973