

Lógica moderna e ciências cognitivas: uma controversa contribuição da lógica no desenvolvimento das ciências da mente

Monalisa M. Lauro¹, monalisalauro@gmail.com.br; **Carolina Blasio**²

1. Psicóloga e especialista em Filosofia Moderna e Contemporânea pela Universidade Federal de Juiz de Fora (UFJF), MG.
2. Psicóloga e mestranda em Ciência da Religião pela Universidade Federal de Juiz de Fora (UFJF), MG.

RESUMO: Discute-se neste ensaio a relação entre o desenvolvimento da Lógica Moderna e o surgimento das Ciências Cognitivas, incluindo sua ênfase nos trabalhos de John McCarthy, nos anos 50, quando houve o desenvolvimento da Inteligência Artificial Simbólica. Ao final, são discutidas críticas sobre as dificuldades da Inteligência Artificial Simbólica elaboradas por Alan Turing, John Searle, e, sobretudo, por Gödel.

Palavras-chave: Lógica, ciências cognitivas, filosofia da mente.

RESUMEN: Lógica moderna y ciencias cognitivas: una controversial contribución de la lógica en el desenvolvimiento de la Lógica Moderna y el surgimiento de las Ciencias Cognitivas, incluyendo su énfasis en los trabajos de John McCarthy, en los años 50, cuando hubo un desenvolvimiento de la Inteligencia Artificial Simbólica. Al final, son discutidas críticas sobre las dificultades de la Inteligencia Artificial Simbólica elaboradas por Alan Turing, John Searle, y , sobre todo, por Godel.

Palabras llaves: Lógica, ciencias cognitivas, filosofía de la mente.

ABSTRACT: Modern logic and cognition sciences: a controversial contribution of the logical in the development of the sciences of mind. In the essay it's argued the connection between the development of the Modern Logic and the emergence of Cognition Science, including the emphasized of that connection in the John McCarthy' s works during the 1950's , when the Symbolical Artificial Intelligence (AI) has developed. Finally, critics about the difficulties of the Symbolical AI are discussed, presented by Alan Turing, John Searle and especially by Gödel.

Keywords: Logical, cognition sciences, philosophy of mind.

Introdução

A obra Ideografia de Frege (*Begriffsschrift*), de 1879, irá contribuir para uma grande revolução na lógica da qual emergiu, historicamente, boa parte da filosofia do século XX. Em seu trabalho, Frege se empenhou em mostrar que a aritmética pode ser construída exclusivamente a partir de leis lógicas, afastando a lógica da análise gramatical e a aproximando-a da análise matemática.

Apesar dos seus aspectos notáveis, a construção de Frege foi rapidamente minada pela descoberta de uma contradição por Russell, em 1902. Mesmo assim, Russell prosseguindo com a obra de Frege, não apenas querendo fundamentar a aritmética, mas com o sonho de criar uma ciência rigorosa e fundadora de todas as outras ciências, conduziu seu trabalho de fundar a matemática sobre uma base puramente lógica, única suscetível de garantir a sua objetividade. Juntamente com Whitehead, escreve o **Principia Mathematica** (1900-1913), em que elaborou uma teoria em que a matemática é representada como desenvolvimento da lógica, e afastou as contradições que afetavam o trabalho de Frege. Russell também reduziu a análise da linguagem em sua totalidade às estruturas lógicas, o que significa fazer da lógica uma espécie de linguagem ideal (COSTA, 2002). A lógica de Russell, ao passar a constituir uma espécie de linguagem ideal, não ficou mais estritamente ligada às matemáticas, passando a ser um instrumento de análise científica e geral, aplicada também na análise filosófica (HAACK, 1978). Desta forma, Russell levou ao desenvolvimento de estreitos vínculos entre a matemática, a lógica e as ciências empíricas.

No presente trabalho, interessa-nos ressaltar uma nova relação desenvolvida, desde então, entre a lógica e as ciências da mente, cujos primórdios remontam à criação da ciência da computação nos anos 30, a partir dos trabalhos de estudiosos como Alan Turing, e ao subsequente surgimento da Inteligência Artificial (IA), com destaque ao desenvolvimento de uma IA lógica realizada por John McCarthy a partir do final dos anos 50. Ainda, abordamos algumas dificuldades que têm sido apresentadas nesta articulação entre lógica e ciências da mente, especificamente, mencionando objeções derivadas dos trabalhos de Turing, Searle e Gödel.

I – Áureos anos da inteligência artificial

Nos anos 30 do século passado, o matemático inglês Alan Turing propôs um procedimento efetivo (mecânico) na lógica e matemática, formalizando a noção de algoritmo: uma máquina teórica simples, capaz de manipular símbolos de um sistema de regras próprias (TURING, 1936, 1937). Turing estava envolvido com a resolução de um dos problemas matemáticos apresentado por David Hilbert em 1928, que buscava a construção de provas matemáticas consistentes e finitas para a matemática clássica, o Problema da Decisão (*Entscheidungsproblem*), o qual consistia em saber se haveria um procedimento efetivo e absoluto que demonstrasse (deduzisse) todos os enunciados matemáticos verdadeiros (teoremas) de um dado conjunto de premissas (axiomas) (TEIXEIRA, 1998).

Através da elaboração teórica de sua máquina, Turing representou que uma computação seria um processo limitado a um número finito de instruções não ambíguas, expressas por um número finito de símbolos que produza um número limitado de resultados, e que fosse possível de ser executado por qualquer ser humano apenas com papel e lápis (COPELAND, 2002). Portanto, esta máquina é uma representação formal de um procedimento efetivo. Assim, em 1936, simultaneamente a um trabalho análogo de Church, Turing estabelece a tese de que toda computação efetiva pode ser simulada por uma Máquina de Turing (op. cit.).

Embora não especulasse e nem pretendesse construir um computador digital, a Máquina de Turing tornou-se o princípio formal que norteia a construção dos modernos computadores, os quais devem sua real possibilidade de construção ao trabalho de John Von Neumann na década de 40. Como a Máquina de Turing é uma máquina abstrata, definida apenas pela função que realiza e podendo ser construída com qualquer tipo de material, não tardou muito para que a atividade mental humana fosse considerada uma versão biológico-cerebral de uma Máquina de Turing (TEIXEIRA, 2004).

Essa era a idéia que vigorou fortemente nos anos 50 e 60, período otimista de formação da inteligência artificial (IA), em que a analogia entre o sistema computacional e os processos humanos de pensamento foi amplamente usada. Além de funcionar como uma ferramenta valiosa de trabalho, o computador passou a servir como um modelo do pensamento humano (GARDNER, 1985). Pensar seria realizar um procedimento algorítmico

Cabe sublinhar que em 1950, Turing em seu artigo **Computação e Inteligência** (TURING, 1950) questionará sobre a possibilidade de uma máquina pensar, propondo que se pode dizer que uma máquina “pensa”, se for capaz de simular um comportamento idêntico ao humano. Tal procedimento comportamental ficou conhecido como Teste de Turing: se o comportamento de uma máquina for indistinguível do exibido por um ser humano, então pode se atribuir a esta máquina a capacidade de pensar.

Mesmo já existindo algumas questões centrais da IA a respeito da relação entre pensamento humano e máquinas, uma efetiva relação entre lógica, computação e ciências da mente somente se consolidará no verão de 1956, ano em que pesquisadores interessados em teoria dos autômatos, redes neurais e estudos da inteligência se reuniram em Dartmouth, Estados Unidos. Este encontro marca um momento crucial na história das ciências cognitivas em geral e, particularmente na história da IA, uma vez que visava à construção e à descrição de programas computacionais que se “comportassem inteligentemente”, estabelecendo assim as bases para o desenvolvimento de uma nova ciência que estuda os processos mentais à luz do funcionamento computacional (GARDNER, 1985).

Entre os especialistas em Ciência da Computação presentes nesta conferência, é de nosso particular interesse destacar o trabalho de John McCarthy, criador do termo *Artificial Intelligence*, que pode ser considerado o principal responsável pela introdução da lógica como ferramenta para a construção de sistemas lógicos em IA, firmando assim a relação entre pensamento humano, máquinas e lógica. Desde então, as perspectivas da IA baseada na lógica tiveram um avanço significativo em várias partes do mundo, especialmente entre lógicos interessados em provar teoremas matemáticos.

II – A criação de John McCarthy

O trabalho de McCarthy consiste em uma abordagem formal rigorosa, em que atos inteligentes ou bom senso são reduzidos a um conjunto de relações lógicas ou axiomas que podem ser expressos com precisão, em que todos os processos de conhecimento humanos são ou podem ser concebidos em termos puramente lógicos (MCCARTHY, 1999). McCarthy enfatizou o uso da

lógica de primeira ordem na representação do raciocínio em IA. Searle definirá essa abordagem simbólica da IA, em que as máquinas contariam com mentes reais, como 'IA Forte' (SEARLE, 1980).

No final dos anos 50, McCarthy desenvolveu a linguagem de alto nível *Lisp*, que acabou se tornando a linguagem de programação dominante, também neste período escreveu o artigo **Programs with common sense** (1959), em que apresentou um programa hipotético, o Advice Taker, que pode ser considerado o primeiro sistema de IA completo, projetado para usar o conhecimento com a finalidade de solucionar problemas, bem como, incorporar conhecimento geral do mundo. O principal tema de sua abordagem nesse programa era a separação clara entre o conhecimento (na forma de regras) e o componente de raciocínio (RUSSEL; NORVING, 2003).

Seu programa seria capaz de usar operações lógicas aplicadas a conhecimento do mundo representado por meio de sentenças declarativas, equivalentes a sentenças lógicas de primeira ordem. Os métodos de raciocínio lógico seriam empregados para deduzir conseqüências deste conhecimento e, assim, ações seriam executadas pelo programa (NILSON, 1998; MCCARTHY, 1959).

O comportamento da máquina ainda poderia ser aperfeiçoado, os programas aprenderiam a partir da experiência, como efetivamente os homens fazem. Para tanto, seria necessário que o programa fosse autônomo. Somente se o programa for capaz de aperfeiçoar seu comportamento de forma independente, poderíamos afirmar que ele aprende com a experiência (MCCARTHY, 1959).

Ainda com uma brevíssima história, podemos observar a forte influência que a IA sofre da lógica. Tal influência pode ser constatada em outro importante artigo de McCarthy, **Some Philosophical problems from the standpoint of Artificial Intelligence**, de 1969, no qual faz referência a 35 trabalhos de lógica dos 58 que constam no artigo. A motivação em recorrer à formalização lógica na resolução de problemas em IA parece não ter mudado muito ao longo dos anos. Mesmo havendo uma diminuição na referência à literatura lógica, principalmente após o seu desenvolvimento e constituição como ramo específico da Ciência da Computação, os teóricos deste ramo ainda concordam que a lógica tem um importante papel em áreas centrais da IA, seja como ferramenta de análise, como base para a representação do conhecimento ou como linguagem de programação.

Vemos que a IA, com sua proposta de mecanização do raciocínio, proporcionou significativos avanços à lógica. Novos problemas e metodologias emergiram, questões teóricas, até então vagamente mencionadas na literatura lógica, passaram a ser tratadas dentro da IA, gerando *insights* sobre raciocínio de causa e efeito das ações, persistência de estados de interação entre ações e

processos físicos e desenvolvimento de diferentes lógicas. O próprio McCarthy procurou estender o escopo da lógica ao propor a formalização de raciocínios do senso comum, raciocínios pré-científicos usados na resolução de problemas cotidianos, como, por exemplo, ir de casa ao aeroporto (MCCARTHY, 1959). De fato, todo este desenvolvimento que temos presenciado só foi possível pela articulação das teorias lógicas com as aplicações práticas de planejamento de autônomos.

III – Mentas, máquinas e seus problemas

A relação entre mentes e máquinas parece ser marcada por dificuldades e questões como: Será que as atividades mentais humanas podem ser simuladas por sistemas artificiais? Podemos sustentar uma analogia entre mente e máquina ou haverá uma diferença entre elas? Serão todas as formas de processos mentais humanos representadas em algoritmo?

Várias críticas a esta proposta vêm sendo formuladas ao longo dos anos de vida da IA e das ciências da mente em geral e, algumas destas têm por base trabalhos e pensamentos feitos muito tempo antes da criação da IA e das ciências da mente. Estas críticas vêm suscitando propostas ainda não conclusivas a respeito dos processos mentais e de qual a melhor forma de estudá-los. No presente trabalho abordamos, de forma breve, algumas considerações apresentadas por Turing, Searle e Gödel.

Primeiramente podemos mencionar que o próprio Turing, cujo trabalho é a raiz de toda essa problemática, reconhecia uma dissimilaridade entre atos mentais e processo algorítmico (TEIXEIRA, 2004). Como já mencionamos, Turing procurava demonstrar provas absolutas de consistência de um sistema por meio de um procedimento efetivo ou algorítmico, resolvendo assim o Problema da Decisão levantado por Hilbert.

Seu objetivo era criar um procedimento mecânico capaz de demonstrar todos os enunciados matemáticos verdadeiros (teoremas) de um dado conjunto de premissas (axiomas). Mas não bastava demonstrar que quando houvesse um processo algorítmico para resolver um problema, este poderia ser executado em sua máquina teórica, era necessário também saber se este procedimento de demonstração era finito, ou seja, se após um determinado número de passos a máquina parasse. Foi exatamente isto que Turing tentou verificar ao propor a Máquina de Turing Universal – MTU (TURING, 1936, 1937).

A proposta de uma Máquina de Turing Universal trata de considerarmos uma Máquina de Turing como um *input* ou programa a ser rodado por uma outra máquina, a MTU. Diante da questão de se conseguir um procedimento mecânico finito e absoluto, tal como questionado por Hilbert, o Problema da

Parada, Turing concluiu que tal máquina (MTU) é matematicamente inconcebível. Saber se a MTU pára ou não é uma coisa que a própria máquina não pode fazer. Tal decisão requereria uma inteligência não algorítmicamente expressa (TEIXEIRA, 2004; TURING 1936, 1937).

Entretanto, a idéia de uma representação formal de um procedimento efetivo faz surgir um mito, passando pela filosofia, bem como as ciências da mente com efeitos perniciosos (COPELAND, 2002). Turing não mostrou que suas máquinas podem resolver qualquer problema que possa ser resolvido por instruções, regras estabelecidas e explícitas, ou procedimentos, nem provou que a MTU pode computar qualquer função que qualquer computador, qualquer arquitetura possa computar. Essas proposições seriam uma extensão da tese própria do Turing, que trazem uma questão empírica em aberto, pois existem especulações que há processos físicos – assim, potencialmente, operações/máquinas – cujo comportamento conformam-se a funções não computáveis por máquina de Turing (op. cit.).

A literatura da teoria computacional da mente contém numerosos endossos de proposições equivalentes ou similares à tese entendida que são consideradas como sendo a Tese de Turing (ou Church). Essa confusão, aliada a idéia de compatibilidade entre sistemas mecânicos ou biológicos, gera algumas sustentações fundamentais das Ciências Cognitivas, como por exemplo, a idéia de que os processos mentais devem ser expressos nos termos da máquina de Turing, ou seja, que pensar seria realizar um procedimento algorítmico (BODEN, 1988 apud COPELAND, 2002).

Tocando um problema epistemológico relativo a essas afirmações, mesmo se a proposta da IA forte enquanto uma teoria da mente se confirmasse, ou seja, mesmo se fossemos capazes de descrever os processos mentais em termos computacionais e de distinguir quais sistemas seriam genuinamente mentais dos que não seriam, criaríamos um novo problema. Pois, se máquinas simples como termostatos podem ter crenças da mesma forma como os homens, tal como pretendia McCarthy (1979 apud SEARLE, 1980), como se explica o que é especificamente mental sobre o mental, ou seja, o que distinguiria, dentre este “pan-mentalismo”, “mentes de máquinas”, das “mentes de humanos” (op. cit.).

Os computadores por mais que a tecnologia pudesse avançar e eles pudessem fazer tudo o que um ser humano faz, teriam limitações que os impediriam de ter uma vida mental semelhante a nossa, pois nunca poderiam adquirir consciência (intencionalidade) do que estão fazendo (TEIXEIRA, 2003).

Foi seguindo este raciocínio que John Searle desenvolveu seu conhecido argumento do quarto chinês apresentado no artigo **Minds, Brains, and Programs**, de 1980. Searle critica a proposta da IA forte, quando esta afirma

que o computador apropriadamente programado possui estados cognitivos, e não estaria apenas simulando-os, e que os programas desse modo explicariam a cognição humana.

Para Searle, esses programas seriam manipuladores de signos vazios segundo regras sintáticas, sem condições suficientes de entendimento. A atribuição de entendimento e intencionalidade a máquinas se daria de forma metafórica e por analogia. Ou seja, uma máquina não entenderia uma história de dada língua, mas apenas receberia estímulos do texto, manipularia os signos segundo as regras previamente estabelecidas e produziria uma resposta compatível, ou mesmo, idêntica a que um humano que dominasse esta mesma língua produziria.

Voltando ao Problema da Parada da Máquina de Turing, podemos afirmar que este vai ao encontro do trabalho desenvolvido por Gödel, uma vez que este problema é um caso de indecidibilidade (TEIXEIRA, 2004; TURING, 1936, 1937). Em seu **Teorema da Incompletude**, de 1931, o matemático Gödel ataca um problema central da fundamentação da matemática: saber se um dado conjunto de postulados é internamente consistente, ou seja, se podemos deduzir do mesmo todas as fórmulas possíveis e não contraditórias (NAGEL; NEWMAN, 1958).

Gödel estabelece que em um dado sistema lógico existem proposições verdadeiras que não serão passíveis de demonstração dentro do sistema, exceto se o sistema for inconsistente. Haveria um abismo entre verdade e prova, pois alguns enunciados matemáticos verdadeiros não são passíveis de demonstração usando as regras do sistema, sendo preciso estar fora do mesmo para conhecer esta verdade (op.cit.). Assim como uma Máquina de Turing Universal, pois esta é incapaz de determinar se seu próprio programa é finito, ou seja, capaz de parar (TURING, 1936-37).

O artigo de Gödel foi um marco tanto na história da lógica, quanto na da matemática, mas seu trabalho também teve grande repercussão sobre outros campos de conhecimento, dos quais nos interessa particularmente as ciências da mente. Neste campo, as conclusões de Gödel versam sobre a possibilidade de se construir uma máquina comparável à mente humana. Veremos que de acordo com Lucas (1961), o teorema de Gödel parece provar que as mentes não podem ser explicadas por máquinas.

De acordo com esse autor, uma máquina/sistema, se consistente, não produz como verdade uma asserção de sua própria consistência, conseqüentemente a mente sendo realmente uma máquina, não poderia alcançar a conclusão de que seja consistente. E prossegue afirmando que os humanos não estão confinados a fazerem inferências dedutivas, isto é, eles são auto-reflexivos, concebendo seu desempenho e simultaneamente algo externo a ele. A máqui-

na, por outro lado, necessitaria para o relato de seu próprio desempenho do acréscimo de uma nova máquina. Assim, qualquer modelo mecânico, daquilo que se entende por mente, deveria incluir um mecanismo que pudesse enunciar verdades aritméticas, porque isto é uma coisa que mentes podem fazer. A isto, Lucas conclui que nenhuma máquina pode ser um modelo completo ou adequado da mente, pois estas são essencialmente diferentes das máquinas.

Basta sabermos se essa impossibilidade se estende a qualquer tipo de máquina, como Lucas afirma, ou, se isto cabe apenas às máquinas de Turing. E mesmo se existe alguma possibilidade da replicação de processos mentais (ou cerebrais) de forma artificial, ainda não podemos ignorar um outro problema, delicado por não ser formal, e relativo à conceitualização da mente (ou cérebro), bem como dos seus processos.

Referências bibliográficas

COPELAND, B. J. The church-turing thesis. **The Stanford encyclopedia of Philosophy**. Stanford: Edward N. Zalta, 2002. Disponível em: <<http://plato.stanford.edu/archives/fall2002/entries/church-turing/>>. Acesso em: 14 fev. 2006.

COSTA, C. **Filosofia da linguagem**. Rio de Janeiro: Jorge Zahar, 2002.

FREGE, G. Sobre o sentido e a referência. In: **Lógica e filosofia da linguagem**. São Paulo, Cultrix: EDUSP, 1978.

GARDNER, H. **A nova ciência da mente**. São Paulo: EDUSP, 2003.

HAACK, S. **Filosofia das lógicas**. São Paulo: UNESP, 2002.

LUCAS, J. R. Minds, machines and Gödel. **Philosophy**, n. 36, p. 112-127, 1961. Disponível em: <<http://cogprints.org/356/00/lucas.html>>. Acesso em: 12 fev. 2006.

MCCARTHY, J. **Programs with common sense**, 1958. Disponível em: <<http://www-formal.stanford.edu/jmc/>>. Acesso em: 12 fev. 2006.

_____. **Philosophical and scientific presuppositions of logical AI**. 1999. Disponível em: <<http://www-formal.stanford.edu/jmc/>>. Acesso em: 11 fev. 2006.

MCCARTHY, J.; HAYES, P. J. **Some philosophical problems from the standpoint of Artificial Intelligence**. 1969 Disponível em: <<http://www-formal.stanford.edu/jmc/>>. Acesso em: 11 fev. 2006.

NAGEL, E.; NEWMAN, J. R. **A prova de Gödel**. São Paulo: Perspectiva, 2003.

NILSON, N. **Artificial intelligence: a new synthesis**. San Francisco, CA: Morgan Kaufmann Publishers, 1998.

RUSSEL, S.; NORVING, P. **Inteligência artificial**. Rio de Janeiro: Elsevier, 2003.

SEARLE, J. Minds, brains, and programs. **Behavioral and brain sciences**, n. 3, p. 417-457, 1980. Disponível em: <<http://members.aol.com/NeoNoetics/MindsBrainsPrograms.html>>. Acesso em: 30 mar. 2006.

TEIXEIRA, J. F. **Mentes e máquinas: uma introdução à ciência cognitiva**. Porto Alegre: Artes Médicas, 1998.

_____. **Mente, cérebro e cognição**. Petrópolis: Vozes, 2003

_____. **Filosofia e ciência cognitiva**. Petrópolis: Vozes, 2004.

TURING, A. On computable numbers, with an application to the Entscheidungs problem. **Proceedings of the London Mathematical Society**, v. 42, n. 2, p. 230-265, 1936-37. Disponível em: <<http://www.abelard.org/turpap2/tp2-ie.asp>>. Acesso em: 14 fev. 2006.

_____. Computação e inteligência. In: TEIXEIRA, F. **Cérebros, máquinas e consciência: uma introdução à filosofia da mente**. São Carlos: EdUFSCar, 1996.