

Complexidade computacional e medida da informação: caminhos de Turing e Shannon

JOSÉ ROBERTO CASTILHO PIQUEIRA¹

Introdução

DURANTE os anos 1990, havia no Instituto de Estudos Avançados da USP o chamado “Grupo de Ciência Cognitiva”, liderado pelo saudoso professor Henrique Schützer Del Nero. Era uma pequena semente de multidisciplinaridade plantada com sacrifício, convivendo até com certo escárnio de pessoas mais conservadoras.

Vários alunos e professores que passaram pelas reuniões realizadas às quartas-feiras de manhã incorporaram às suas linhas de pesquisa assuntos amplamente discutidos naquela pequena sala da reitoria velha.

Para mim, o primeiro contato com a ideia de complexidade foi ali. Lembro-me de muitas horas de conversa sobre o que é um “fenômeno complexo” e se era possível objetivar essas ideias.

Havia gente originária das Ciências Humanas que falava de Edgar Morin; das Biológicas interessadas em Neurociência; e o pessoal de Exatas querendo descrever o mundo pela Matemática.

Eu trabalhara, nos anos 1980, em empresa brasileira produtora de equipamentos para redes de comunicação de dados, o que exigira conhecimento da chamada “Teoria de Informação”, devida a Claude E. Shannon (Shannon; Weaver, 1963) e, por isso, pensava que complexidade estava diretamente relacionada com o conceito de medida de informação.

Neste novo século, o debate sobre complexidade adquiriu grandes proporções com a criação de grupos de estudo das mais variadas orientações em todo o mundo. A Física Estatística, a Neurociência, as Ciências Sociais, a Economia e a Engenharia estão cada vez mais interessadas em fenômenos emergentes de interações não lineares entre os mais diversos tipos de agentes (Northrop, 2011).

Ao longo desses anos, não perdi de vista que uma medida objetiva de complexidade poderia originar-se na entropia informacional de Shannon (Shannon; Weaver, 1963). Aprendi, também, que outras medidas interessantes de complexidade foram propostas (López-Ruiz; Mancini; Calbet, 1995; Shiner; Davison; Landsberg, 1999), mais compatíveis com os processos naturais.

Associar as definições de teoria da informação a medidas de complexidade,

entretanto, é bastante interessante quando se trata de caracterizar sistemas de computação, levando à ideia de “complexidade computacional”, conceito restrito, mas de grande utilidade (Kolmogorov, 1965).

Este artigo se ocupa dessa discussão: mostrar como os conceitos de medida de informação (Shannon) estão relacionados com complexidade computacional (Turing) (Desurvire, 2009).

Não há, aqui, qualquer conceito original. Trata-se, apenas, de um relato das diversas interpretações de importantes trabalhos de cientistas cujas teorias permitiram a emergência do complexo fenômeno da ubiquidade da informação que parece mudar hábitos e relações pessoais.

Inicia-se uma seção de breves dados biográficos de Shannon e Turing, que permitem entender o contexto dos problemas por eles estudados. A seguir, após resumida apresentação da máquina de Turing, discute-se o conceito de complexidade computacional do ponto de vista algorítmico (Turing) e informacional (Shannon). Finaliza-se com uma descrição sobre as similaridades e diferenças originárias das duas abordagens apresentadas.

Breve contextualização histórica

Nascido nos Estados Unidos da América em 30 de abril de 1916, Claude E. Shannon obteve o título de doutor no MIT, em 1937, com trabalho notável em “Álgebra de Boole”, propondo circuitos elétricos capazes de executar as principais operações da Lógica clássica.

Quatro anos antes (23 de junho de 1912) de seu nascimento, em Londres, nascera Allan M. Turing, que também se interessou por encontrar meios de realizar operações lógicas e aritméticas, fazendo uso de máquinas. Suas ideias resultaram no importante conceito de “Máquina de Turing”, paradigma abstrato para a computação, apresentado durante seus estudos, no King’s College, em Cambridge, no ano de 1936. Entre 1936 e 1938, Turing viveu em Princeton-NJ onde realizou seu doutorado estudando problemas relativos à criptografia.

Assim, Shannon e Turing, de maneira independente, trabalhavam, simultaneamente, em comunicações e computação, dois tópicos que, combinados, hoje proporcionam recursos antes inimagináveis para o mundo moderno das artes, da ciência, da medicina, da tecnologia e das interações sociais.

Contemporaneamente à eclosão da Segunda Guerra Mundial, Shannon e Turing gestavam ideias abstratas sofisticadas, tentando associá-las ao mundo concreto das máquinas que, gradativamente, tornavam-se fatores de melhoria da qualidade de vida das populações.

A Segunda Grande Guerra utilizou-se de tecnologias sofisticadas para a destruição. Os bombardeios aéreos causaram muitas mortes e devastaram cidades. Evitá-los e preveni-los eram questões de vida ou morte e, para tanto, ouvir as comunicações dos inimigos e decifrar seus códigos era uma atividade indispensável.

Os países do eixo tinham desenvolvido sofisticadas técnicas de comunicação criptografada utilizada para planejar ataques inesperados às forças aliadas. Shannon e Turing, então, com seu conhecimento sofisticado da matemática da informação deduziram as regras alemãs de codificação, levando os aliados a salvar muitas de suas posições de ataques nazistas.

Pode-se dizer que uma boa parte da inteligência de guerra dos aliados vinha desses dois cérebros privilegiados.

Finda a guerra, Shannon passou a trabalhar nos laboratórios Bell, propondo a “Teoria da Informação”, em 1948. Com carreira profícua, notável pela longevidade e muitos trabalhos importantes, deixou sua marca nas origens das comunicações digitais. Faleceu aos 85 anos (em 24 de fevereiro de 2001), deixando grande legado intelectual e tecnológico.

Turing, após o término da guerra, ingressou como pesquisador da Universidade de Manchester, sofrendo ampla perseguição por ser homossexual. Mesmo vivendo na avançada Inglaterra, foi condenado à castração química, em 1952. Essa sequência de dissabores levou-o ao suicídio, em 7 de junho de 1954.

Shannon viu sua teoria transformar o mundo, com o nascimento da internet. Turing, entretanto, não viu sua máquina se transformar em lap-tops e tablets que hoje povoam, até, o imaginário infantil.

Máquina de Turing e complexidade computacional

Há na literatura (Desurvire, 2009) e na internet (Schechter, s.d.) um grande número de descrições completas sobre a máquina de Turing que pode ser objeto de estudo mais aprofundado. Aqui faremos uma breve descrição, apenas para poder introduzir a ideia de computador universal e de algoritmo.

Uma máquina de Turing é constituída, de maneira abstrata, por uma fita de comprimento infinito, isto é, uma sucessão de células de memória que podem assumir o valor “0” ou “1”. A fita é munida de uma cabeça de leitura e uma de escrita, podendo ser movida para a esquerda ou para a direita uma célula por vez (Figura 1).

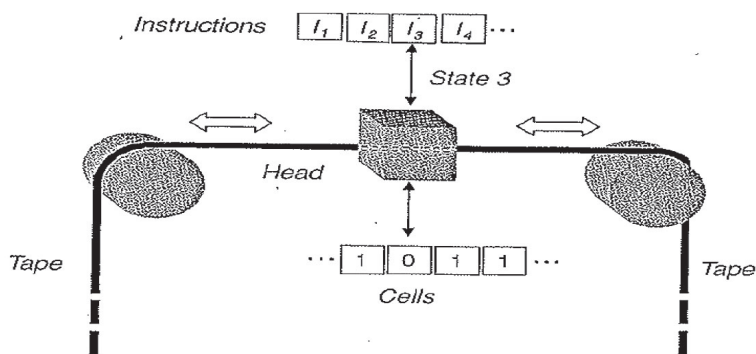


Figura 1 – Máquina de Turing.

As operações da cabeça e da fita são definidas por uma tabela de instruções $\{I_1, I_2, \dots, I_n\}$, chamada tabela de ação.

Vários sites na internet contêm programas dirigidos para a máquina de Turing (disponíveis em www.ams.org e <http://ironphoenix.org/tril/tm>). Lá, é possível baixar programas que permitem o entendimento de como uma máquina de Turing computa.

Com essas ideias em mente, Turing, em seu tempo de doutorado nos Estados Unidos, sob a orientação de Alonzo Church, em 1936, propôs o conceito de Máquina de Turing Universal (UTM), viabilizando a construção do computador programável.

A arquitetura computacional, proposta por John Von Neumann em 1945 e usada pelos computadores de hoje, fundamenta-se no trabalho de Turing de 1936 e sua base pode ser resumida na tese de Church-Turing: “Uma UTM é capaz de resolver todos os problemas solucionáveis por um algoritmo ou por um método computacional efetivo”.

Deve-se observar que, de maneira simplificada, um algoritmo é uma sequência de operações lógicas e aritméticas, com a finalidade de dar a resposta a um problema passível de ser colocado em linguagem matemática.

Chega-se, então, ao conceito de complexidade computacional, entendido como o número de operações necessárias para a execução de um programa, isto é, para a execução de um conjunto de algoritmos (Desurvire, 2009). Como qualquer programa originário de algoritmo computacional pode ser executado por uma UTM, a medida da complexidade computacional de um programa ou sequência, do ponto de vista algorítmico, pode ser entendida como o número de estados da tabela de ação, quando o programa ou a sequência são executados em uma UTM (Kolmogorov, 1965).

Neste ponto, é importante ressaltar que a medida de complexidade computacional, do ponto de vista algorítmico, está relacionada a uma sequência ou a um programa, proporcionando, portanto, uma medida determinista da complexidade computacional.

Shannon e entropia informacional como medida de complexidade

O conceito de entropia informacional, proposto por Shannon (Shannon; Weaver, 1963) tem origem no mundo das comunicações, considerando-se que um sistema proposto com essa finalidade é composto por uma fonte, capaz de emitir sequências de dados, um meio, através do qual a sequência transita, atingindo um receptor, destino da mensagem.

Os trabalhos em Teoria da Informação dividem-se em três grupos: medir a informação emitida pela fonte, medir a capacidade do canal e escolher um código que permita o trânsito dos dados com pequena e arbitrária taxa de erros.

Para a discussão aqui estabelecida, interessa a medida de informação associada a uma sequência, escolhida entre as diversas sequências possíveis de emissão pela fonte.

A abordagem é probabilística, associando-se às diversas sequências, suas probabilidades de ocorrência e definindo-se informação individual como sendo uma variável aleatória que assume valores altos para sequências pouco prováveis e baixos, para as muito prováveis.

Ao valor esperado (média) dessa variável aleatória é dado o nome de entropia informacional e sua unidade é bits por sequência. Assim, a entropia informacional é relativa à fonte como um todo e, diferindo da complexidade algorítmica, não está associada a uma sequência individual, mas a todas as sequências possíveis de serem emitidas pela fonte.

Embora esse conceito não tenha, inicialmente, a proposta de medir complexidade computacional, pode perfeitamente servir para isso. Basta pensar que, dado um programa, é possível criar diferentes sequências para realizá-lo e atribuir a cada uma delas um valor de probabilidade.

A medida de complexidade computacional, do ponto de vista informacional, pode ser dada pela entropia de Shannon, que assume valor máximo caso todas as sequências sejam equiprováveis (Shannon; Weaver, 1963).

Conclusões

Apresentamos, de maneira qualitativa, duas maneiras diferentes de medir complexidade computacional: uma determinística (Turing) e uma probabilística (Shannon), originárias de pensamentos e enfoques diferentes dos dois grandes pensadores que teceram a era moderna da informática.

Como a natureza prega peças e nunca se sabe se Deus joga dados, para sequências longas, que são os casos mais frequentes na prática, pode ser demonstrado que as duas medidas coincidem.

Referências

DESURVIRE, E. *Classical and Quantum Information Theory*. Cambridge: Cambridge University Press, 2009.

KOLMOGOROV, A. N. Three approaches to the definition of the concept “quantity of information”. *Problemy Peredachi Informatsii*, v.1, p.3-11, 1965.

LÓPEZ-RUIZ, R.; MANCINI, H. L.; CALBET, X. A statistical measure of complexity. *Physics Letters A*, v.209, p.321-6, 1995.

NORTHROP, R. B. *Introduction to Complexity and Complex Systems*. Florida: CRC Press, 2011.

SHANNON, C. E.; WEAVER, W. *The Mathematical Theory of Communication*. Urbana; Chicago: Illini Books Edition, 1963.

SHINER, J.; DAVISON, M.; LANDSBERG, P. Simple measure for complexity. *Physical Review E*, v.59, n.2, p.1459-64, 1999.

SCHECHTER, L. M. *A vida e o legado de Alan Turing para a ciência*. s. d. Disponível em: <<http://www.dcc.ufjf.br/~luisms/turing/seminarios.pdf>>.

RESUMO – Este artigo apresenta, qualitativamente, os conceitos de complexidade computacional algorítmica (Turing) e de complexidade computacional informacional (Shannon), enfatizando como pensamentos independentes, de naturezas diferentes, produziram conceitos matemáticos similares e de grande utilidade para a computação moderna.

PALAVRAS-CHAVE: Complexidade, Máquina de Turing, Medidas computacionais.

ABSTRACT – This paper qualitatively presents the concepts of algorithmic computational complexity (Turing) and informational computational complexity (Shannon), emphasizing how independent thinking, of different nature, produced similar mathematical concepts with great usefulness for modern computation..

KEYWORDS: Complexity, Turing machine, Computational measures.

José Roberto Castillo Piqueira é graduado e mestre em Engenharia Elétrica pela Escola de Engenharia de São Carlos da USP e doutor e livre-docente em Engenharia de Sistemas pela Escola Politécnica da USP. Professor titular e diretor da Poli-USP.

@ – piqueira@lac.usp.br

Recebido em 22.4.2015 e aceito em 28.9.2015.

¹ Escola Politécnica, Universidade de São Paulo, São Paulo/São Paulo, Brasil.