



RESUMO

Este estudo tem como objetivo geral apresentar uma exegese da teoria da informação de Claude Elwood Shannon, abordando seus conceitos fundamentais, como entropia, capacidade de canal, redundância e codificação eficiente. A intenção é explicar como a teoria permite quantificar a quantidade de informação transmitida, independentemente de seu significado, e de que forma ela fundamenta o desenvolvimento de sistemas de comunicação mais confiáveis e eficientes. Para isso, o texto destaca os princípios matemáticos da teoria de Shannon, incluindo a relação entre previsibilidade de símbolos e quantidade de informação, além do papel da redundância na resistência da mensagem a ruídos. A metodologia empregada consiste na revisão teórica, fundamentada na leitura dos principais textos de Shannon sobre *A Mathematical Theory of Communication*. Assim, o artigo visa consolidar a compreensão dos conceitos da teoria da informação de maneira acessível, ressaltando a importância da separação entre informação e significado.

Palavras-chave: Shannon. Informação. Comunicação. Entropia. Codificação.

Conceptual aspects of Information Theory according to Claude E. Shannon

ABSTRACT

The general objective of this study is to present an analysis of Claude Elwood Shannon's information theory, addressing its fundamental concepts, such as entropy, channel capacity, redundancy, and efficient coding. The intention is to explain how the theory allows us to quantify the amount of information transmitted, regardless of its meaning, and how it underpins the development of more reliable and efficient communication systems. To this end, the text highlights the mathematical principles of Shannon's theory, including the relationship between symbol predictability and amount of information, as well as the role of redundancy in message resistance to noise. The methodology employed consists of a theoretical review, based on a reading of Shannon's main texts on *A Mathematical Theory of Communication*. Thus, the article aims to consolidate the understanding of information theory concepts in an accessible manner, emphasizing the importance of separating information from meaning.

Keywords: Shannon. Information. Communication. Entropy. Coding.

1 Introdução

A informação é uma daquelas palavras que parecem óbvias até que se tente defini-las. Seu uso cotidiano é tão amplo que seu significado depende fortemente do contexto em que é empregada. Não por acaso, diferentes autores, de matemáticos à filósofos, propuseram definições que refletem os aspectos que desejam destacar. O matemático George Lfrah (1997, p. 805-814), em *História universal dos algarismos*, por exemplo, chegou a reunir mais de duas dezenas de definições distintas de informação ao tentar construir uma formulação que fosse ao mesmo tempo ampla e precisa acerca da complexidade do termo. Em sentido semelhante, Keith Devlin (1995, p. 1-9) argumenta em *Logic and Information* que a insistência ao tentar construir uma única definição seria ingenuidade, uma vez que conceitos científicos frequentemente amadurecem ao longo do tempo, de modo que o que era entendido inicialmente é reformulado à luz de novos contextos e descobertas. Investigar a informação exige mais do que uma definição: exige situá-la historicamente, epistemologicamente e tecnicamente.

Warren Weaver, nessa perspectiva, acrescenta que informação é o grau de liberdade que se tem ao selecionar uma mensagem. Isso significa que, segundo Weaver, informação é uma medida da variedade de escolhas disponíveis quando se vai transmitir algo. Se só existe uma opção, não há liberdade de escolha, logo, nenhuma informação. Mas se há várias possibilidades e escolhemos uma, essa escolha reduz a incerteza, o que é interpretado como informação. A informação não está na mensagem isolada, mas no ato de escolher entre várias alternativas possíveis. É o contexto estatístico que dá sentido técnico ao termo. Na frase “o dado caiu no número 6”, para saber quanta informação a mensagem traz, é preciso conhecer: o dado tinha quantas faces? Era um dado comum de seis lados? Ou um especial com vinte? Esse conjunto de possíveis mensagens, com suas probabilidades associadas, é o que forma o contexto estatístico. A compreensão desse contexto é fundamental para calcular a entropia, que mede a informação na linguagem de Shannon.

Weaver ressalta que embora as mensagens, na prática, frequentemente tenham significado, esse aspecto não é levado em consideração pela teoria¹. E assim, reforça o que já havia dito Shannon, que os componentes semânticos, isto é, o que a mensagem quer dizer, ou a quais ideias e objetos ela se refere, são irrelevantes para os propósitos da engenharia da comunicação². “Frequentemente, as mensagens têm significado; isto é, elas se referem a, ou estão correlacionadas de acordo com algum sistema com certas entidades físicas ou conceituais. Estes aspectos semânticos da comunicação são irrelevantes para o problema de engenharia” (1949, p. 31). A preocupação da teoria está na eficiência e na fidelidade da transmissão, não nos conteúdos simbólicos.

Diante da compreensão de que a quantidade de informação está relacionada à incerteza e à escolha entre múltiplas alternativas, vamos explorar como essa quantificação é realizada e quais são seus limites. Para tanto, a teoria de Shannon apresenta conceitos essenciais, como a entropia, que permite medir essa incerteza de maneira objetiva. Assim, ao entender os fundamentos que sustentam a transmissão de informações, podemos aprofundar a análise sobre a eficiência dos canais de comunicação e os fatores que influenciam sua capacidade de transmitir dados de forma consistente.

2 Fundamentos da teoria de Shannon

A teoria da informação de Shannon se debruça sobre a transmissão eficiente de dados em vez de se preocupar com o significado intrínseco da informação. O valor da teoria reside na quantificação da informação, tratando-a não pelo seu sentido, mas pela sua codificação e pela probabilidade de ocorrência dos símbolos que a compõem. Pouco importa se a mensagem é uma declaração de amor ou uma lista de compras; a intenção é garantir que cada “pedacinho” da mensagem chegue ao destino sem perdas. Ou seja, a informação, para ser tratada e medida, precisava ser desvinculada de sua carga semântica. Precisava virar algo quantificável, como o

¹ “The word information, in this theory, is used in a special sense that must not be confused with its ordinary usage. In particular, information must not be confused with meaning” (Weaver *in* Shannon; Weaver, 1949, p. 8).

² “Frequently the messages have a meaning; that is they refer to or are correlated according to some system with certain physical or conceptual entities. These semantic aspects of communication are irrelevant to the engineering problem”.

bits.

Quanto à subjetividade do significado da informação, o sentido de uma mensagem depende da interpretação do receptor, o que significa que uma mesma sequência de símbolos pode gerar compreensões diversas, embora a informação transmitida em si seja idêntica. “Por exemplo: na mensagem ‘a bolsa caiu’ a quantidade de informação é sempre a mesma” (Pineda, 2006, p. 29). Para os indivíduos essa oração pode ter diversos significados, de uma mala caída no chão a uma queda no mercado financeiro, mas para um sistema de comunicação, a quantidade de informação que ela carrega é a mesma, pois se trata da mesma sequência de *bits*. O sistema não se importa com ambiguidade, ele só quer entregar os *bits*, o contexto e a interpretação ficam a cargo do usuário final. Fatores como o conhecimento prévio, o idioma e o interesse do receptor são elementos que influenciam a interpretação da mensagem, mas não alteram a quantidade de informação transmitida pelo canal³.

Segundo Shannon, o processo geral de comunicação envolve três elementos fundamentais: agentes, recursos e métodos. O emissor e o receptor são dois personagens importantes no processo de comunicação. O emissor é aquele que inicia a comunicação, que tem uma mensagem para compartilhar. O receptor está do outro lado, pronto para receber e compreender essa mensagem. Entre esses personagens existe o canal de comunicação, uma espécie de elo entre o emissor e o receptor. Este canal é o meio físico ou virtual por onde a mensagem transita. Em uma conversa, pode ser o ar o canal de comunicação, um cabo de fibra óptica para a internet, ondas de rádio para transmissão ou até mesmo um papel para uma carta. O canal é o “caminho” que a mensagem percorre. Geralmente, a mensagem não é transmitida em sua forma bruta, sendo necessário o emissor codificá-la. No contexto de Shannon, a codificação é o processo de transformar a mensagem original em um conjunto de símbolos discretos. Esses símbolos são como unidades básicas de informação – podem ser letras do alfabeto, números, *bits* (no caso de computadores), ou até mesmo tons musicais. O emissor, então, pega a sua mensagem e, usando um repertório de símbolos previamente acordados, que tanto

³ Nesse contexto, os conceitos de “sinal” e “ruído” são elementos bastante citados na teoria da informação. O sinal seria a própria mensagem em sua forma “viajante”, aquilo que será transmitido ou enviado. O ruído é tudo aquilo que atrapalha, que distorce, que tenta bagunçar a mensagem durante seu percurso.

ele quanto o receptor conhecem, a transforma. Por exemplo, se a mensagem é “cachorro”, ela pode ser codificada como uma sequência de *bits* que representa cada letra. Uma vez codificada, essa sequência de símbolos é enviada através do canal. Ao chegar ao receptor, o processo se inverte. O receptor recebe os símbolos codificados e precisa decodificá-los. Isso significa traduzir esses símbolos de volta para a mensagem original, usando o mesmo repertório e as mesmas regras que o emissor usou para codificar (Shannon, 1963, p. 33-35).

No artigo *A Mathematical Theory Of Communication*, Shannon propõe que a quantidade de informação de uma mensagem pode ser medida pela entropia, entendida como grau de incerteza e imprevisibilidade de um conjunto de símbolos.

Se uma fonte só puder produzir uma mensagem particular, sua entropia é zero, e nenhum canal é necessário. Por exemplo, uma máquina de computação configurada para calcular os dígitos sucessivos de π produz uma sequência definida sem elementos de aleatoriedade. Nenhum canal é necessário para transmitir isso a outro ponto (Shannon, 1949, p. 18, *tradução nossa*)⁴.

Diferentemente da entropia na termodinâmica, que está relacionada à desordem física, aqui o conceito de entropia diz respeito à quantidade de incerteza contida em uma mensagem. Quanto mais difícil for prever o próximo símbolo, maior é sua entropia. Se ao ler um texto uma letra aleatória aparece fora de qualquer padrão esperado, ela tem alta entropia. Já elementos esperados ou repetidos tendem a reduzir esse valor. Pode parecer contraintuitivo, mas é simples. A entropia é inversamente proporcional à informação que já se possui: quanto mais sabemos, menor a incerteza, e portanto, menor a entropia. Se temos certeza de que a próxima letra será “A”, a incerteza é nula e, logo, a entropia também. Mas se todas as letras possíveis têm igual probabilidade de ocorrer, estamos diante de uma entropia máxima, um cenário de máxima imprevisibilidade.

Shannon também explora a relação entre estrutura linguística e transmissão de informação. Ele observa que idiomas naturais possuem regras gramaticais, sintáticas e semânticas que moldam a forma como as mensagens são construídas, e, por consequência, a forma como devem ser modeladas para fins de transmissão

⁴ “If a source can produce only one particular message, its entropy is zero, and no channel is required. For example, a computing machine set up to calculate the successive digits of π produces a definite sequence with no chance element. No channel is required to ‘transmit’ this to another point”.

informativa. Para modelar essas mensagens é necessário considerar essas regras.

A mensagem pode ser de vários tipos: (a) Uma sequência de letras, como em um sistema de telégrafo ou teletipo; (b) Uma única função do tempo $f(t)$, como na rádio ou telefonia; (c) Uma função do tempo e de outras variáveis, como na televisão em preto e branco — aqui, a mensagem pode ser considerada como uma função $f(x, y, t)$ da intensidade da luz no ponto (x, y) e no tempo t em um tubo receptor; (d) Duas ou mais funções do tempo, por exemplo, $f(t), g(t), h(t)$ — isso ocorre na transmissão de som em “três dimensões” ou se o sistema visa atender vários canais individuais por multiplexagem; (e) Várias funções de várias variáveis — na televisão em cores, a mensagem consiste em três funções $f(x, y, t), g(x, y, t), h(x, y, t)$ definidas em um espaço tridimensional — podemos também pensar nessas três funções como componentes de um campo vetorial definido na região — de modo semelhante, várias fontes de televisão em preto e branco produzirem “mensagens” que consistem em várias funções de três variáveis; (f) Combinações variadas também ocorrem, por exemplo, na televisão com um canal de áudio associado (Shannon, 1948, p. 2, tradução nossa).

Outro aspecto crucial que Shannon destaca é o papel da redundância linguística. Ele argumenta que a repetição ou previsibilidade presente nas línguas naturais atuam como um mecanismo de segurança frente às imperfeições no canal de transmissão:

Uma aproximação do ideal teria a propriedade de que, se o sinal for alterado de maneira razoável pelo ruído, o original ainda poderá ser recuperado. Em outras palavras, a alteração geralmente não o aproximará mais de outro sinal razoável do que o original. Isso é obtido ao custo de uma certa quantidade de redundância na codificação. A redundância deve ser introduzida da maneira adequada para combater a estrutura específica do ruído envolvido. No entanto, qualquer redundância na fonte geralmente será útil se for aproveitada no ponto de recepção. Em particular, se a fonte já possuir certa redundância e não for feita nenhuma tentativa de eliminá-la para adequação ao canal, essa redundância ajudará a combater o ruído (Shannon, 1948, p. 24, tradução nossa).

Nesse sentido, a redundância inserida de forma estratégica desempenha um papel útil na construção da resiliência comunicacional. Quando a mensagem contém elementos previsíveis, seja por repetição de padrões sintáticos, estrutura esperada ou mesmo regras da linguagem, ela se torna mais fácil de recuperar, mesmo que partes sejam distorcidas ou perdidas durante a transmissão. É como se a própria estrutura da mensagem oferecesse ao receptor pistas para preencher as lacunas causadas pelo ruído ou outros tipos de interferências. Quanto maior a redundância,

menor é a entropia, a mensagem se torna mais previsível e, por consequência, mais tolerante a falhas. Isso é especialmente relevante em canais ruidosos, nos quais a integridade da transmissão pode ser comprometida. Por outro lado, uma comunicação com alta entropia, com pouca previsibilidade e mínima repetição, tende a carregar mais informação por símbolo, o que é ideal sob o ponto de vista da eficiência. No entanto, essa riqueza informacional exige maior precisão na transmissão e deixa o sistema mais suscetível a erros. Isso revela que um bom sistema de comunicação precisa equilibrar essas duas coisas, pois não adianta só mandar muita informação se ela for sensível a ruídos, nem mandar algo super resistente se for 90% repetição e 10% conteúdo.

A partir dessa base, Shannon avança para o conceito de entropia relativa, que tem o objetivo de comparar a quantidade de informação real produzida por uma fonte com o máximo de informação que ela poderia gerar, caso todos os símbolos tivessem a mesma chance de ocorrer. A ideia é simples: se certos símbolos aparecem com muita frequência, a mensagem se torna previsível e essa previsibilidade reduz a entropia. Assim, a entropia relativa funciona como uma medida do quanto ainda há de novidade ou surpresa na mensagem, em comparação com seu potencial máximo. Quanto mais próxima de 1 for essa medida, mais imprevisível e informativamente densa é a mensagem. Já valores baixos indicam alta redundância.

Outro ponto crucial da teoria de Shannon é entender que nem todo *bit* transmitido realmente carrega uma nova informação. Isso porque, durante o processo de comunicação, muitos *bits* são usados como sinais de controle ou para organizar a transmissão, por exemplo, *bits* de sincronização ou de confirmação, que não acrescentam conhecimento novo ao receptor. O que realmente importa na transmissão de informação é a quantidade de incerteza que ela consegue reduzir no receptor. Dessa forma, a mensagem consistente não é apenas o número de *bits* transmitidos, mas a quantidade de incerteza eliminada, a redução da dúvida sobre o conteúdo original da mensagem. Essa diferença nem sempre é visível, mas é fundamental para entender os limites de um canal de comunicação. Como Shannon aponta, “a capacidade de um canal é o supremo da taxa na qual a informação pode ser transmitida de forma confiável” (1948, p. 4) e também “a taxa máxima de transmissão de informação quando a probabilidade de erro é arbitrariamente

pequena” (1948, p. 40). Ou seja, não é só uma questão de quantidade de *bits* transmitidos por segundo, mas sim de quanto conteúdo informativo esses *bits* efetivamente carregam.

Essa distinção entre sinalização e informação ajuda a compreender o papel da redundância. Shannon observa que “em uma situação ideal, um *bit* de sinalização carregaria um *bit* de informação, mas isso muitas vezes não ocorre devido à redundância” (1948, p. 7). Essa redundância pode ser útil, por exemplo, para combater ruídos ou perdas, mas quando é excessiva e não aproveitada de forma inteligente, ela compromete o uso eficiente do canal. Para contorná-lo, Shannon propõe o uso de códigos eficientes, que buscam aumentar a entropia dos símbolos transmitidos e, assim, aproveitar melhor a capacidade do canal. Como ele explica, “por meio de uma atribuição adequada das probabilidades de transição, é possível maximizar a entropia dos símbolos no canal até atingir a capacidade do canal” (1948, p. 16).

Toda essa discussão aponta para o verdadeiro objetivo da teoria: garantir que os sinais enviados carreguem o máximo de informação útil possível, dentro dos limites do canal e enfrentando seus ruídos. A engenharia da comunicação, nesse contexto, se torna o esforço de codificar inteligentemente, reduzir desperdícios e equilibrar redundância e imprevisibilidade com precisão.

Até aqui foram discutidas apenas fontes de informação discretas ligadas à transmissão de textos por diversos métodos, porém, para fontes contínuas, a situação é diferente. Shannon justifica essa situação pelo fato de que os valores de uma fonte contínua não são restritos a valores discretos.

Considere o conjunto de mensagens de uma longa duração, digamos, T segundos. A fonte é descrita fornecendo a densidade de probabilidade, no espaço associado, de que a fonte selecionará a mensagem em questão $P(x)$ Como, geralmente, os canais têm uma certa quantidade de ruído e, portanto, uma capacidade finita, a transmissão exata é impossível. No entanto, isso não resolve a questão real. Na prática, não estamos interessados na transmissão exata quando temos uma fonte contínua, mas apenas na transmissão dentro de uma certa tolerância. A questão é: podemos atribuir uma taxa definida a uma fonte contínua quando exigimos apenas uma certa fidelidade na recuperação, medida de uma forma adequada (1948, p. 48, tradução nossa).

Nesse trecho, Shannon argumenta que não é realista esperar uma reprodução perfeita da informação gerada por fontes contínuas, como gravações de

som ou vídeo, pois todo canal possui limitações, sejam ruídos, perdas ou restrições de capacidade do canal de transmissão. O que se busca, na prática, é garantir que a recuperação ocorra dentro de um nível aceitável de qualidade, ou seja, com uma fidelidade suficiente.

Essa questão se conecta ao tratamento digital dos sinais contínuos. Como explica o próprio Shannon:

Muitas fontes de informação, como sinais de áudio ou vídeo em sua forma original, produzem sinais que variam continuamente no tempo e no espaço; ou seja, seus valores não são restritos a conjuntos discretos e finitos. Para transmitir esses sinais por sistemas digitais, eles são tipicamente “quantizados”, ou seja, convertidos em uma sequência de valores discretos (1948, p. 48, tradução nossa).

Grande parte das informações com as quais interagimos, fala, música, imagens em movimento, são naturalmente analógicas e contínuas. Isso significa que, a cada instante, o sinal pode assumir qualquer valor dentro de um intervalo, sem limitação a um conjunto finito de possibilidades. Já os sistemas digitais — como computadores e redes — funcionam exclusivamente com valores discretos. Para tornar esses dois mundos compatíveis, utiliza-se a quantização: um processo que agrupa esses valores contínuos em faixas e os representa por símbolos discretos (como os *bits* do código binário). A quantização, nesse contexto, atua como uma ponte entre o mundo físico e o digital.

A partir dessa diferenciação, Shannon introduz a ideia de fidelidade de reprodução e propõe uma conexão direta entre a qualidade desejada na recuperação da mensagem e a quantidade de informação necessária para alcançá-la. Quanto mais exigente for o padrão de qualidade, maior será a taxa de *bits* por segundo necessária para garantir que o sinal reconstruído se aproxime suficientemente do original.

Essa relação mostra que transmitir uma informação com boa qualidade depende da capacidade do canal disponível, ou seja, do quanto ele consegue “aguentar” de informação por segundo. Se a qualidade desejada for muito alta, é preciso que o canal consiga acompanhar esse fluxo. Caso contrário, mesmo que tentemos enviar todos os dados, o canal não dará conta, e a mensagem final chegará com perdas ou distorções. Ainda assim, isso não quer dizer que tudo está perdido: mesmo em canais limitados ou com interferências, é possível usar formas

inteligentes de codificação e organização dos dados para garantir que a mensagem chegue suficientemente fiel ao original, dentro de um nível de qualidade aceitável.

3 Considerações finais

A Teoria Matemática da Comunicação proposta por Shannon estabelece que a informação contida em uma mensagem pode ser quantificada por meio de uma medida chamada entropia, diretamente relacionada à distribuição de frequência dos símbolos transmitidos. Esses símbolos pertencem a um repertório finito, previamente definido, que dá forma à codificação das mensagens. Um dos aspectos mais inovadores da abordagem de Shannon está na separação entre quantidade de informação e significado: ele afirma que o conteúdo semântico é irrelevante para os propósitos da engenharia da comunicação.

A partir de uma abordagem estatística, Shannon mostra que a entropia se reduz a zero quando há certeza absoluta sobre qual símbolo será transmitido, indicando ausência total de novidade. Por outro lado, a entropia atinge seu valor máximo quando todos os símbolos possuem a mesma probabilidade de ocorrência, ou seja, quando a mensagem é completamente imprevisível. Essa relação entre previsibilidade e informação revela o impacto da teoria: quanto mais incerta for a próxima escolha, maior é a quantidade de informação carregada. E é justamente essa lógica matemática, livre de interpretações subjetivas, que contribui para o desenvolvimento de sistemas de comunicação mais eficientes e consistentes.

Referências

DEVLIN, Keith. **Logic and information**. Cambridge: Cambridge University Press, 1995

IFRAH, Georges. **História universal dos algarismos**: a inteligência dos homens contada pelos números e pelo cálculo. Tradução de Alberto Muñoz e Ana Beatriz Katinsky. Rio de Janeiro: Nova Fronteira, 1997

PINEDA, José Octávio de Carvalho. **A entropia segundo Claude Shannon**: o desenvolvimento do conceito fundamental da Teoria da Informação. 2006. 126 f. Dissertação (Mestrado em História da Ciência) – Pontifícia Universidade Católica de São Paulo, São Paulo, 2006.

SHANNON, Claude Elwood. A mathematical theory of communication. **Bell System Technical Journal**, v. 27, n. 3, p. 379–423; n. 4, p. 623–656, jul./out. 1948.

SHANNON, Claude Elwood; WEAVER, Warren. **The mathematical theory of communication**. Urbana: University of Illinois Press, 1963.

WEAVER, Warren. Some recent contributions to the mathematical theory of communication. *In*: SHANNON, Claude E.; WEAVER, Warren. **The mathematical theory of communication**. Urbana: University of Illinois Press, 1949. p. 8.

Recebido: 07/09/2025
Aprovado: 20/10/2025