

Da comunicação à informação: quando a prática se sobrepõe à teoria

From communication to information: when practice overshadows theory

Aurélio Bianco Pena^{*1}, Cibelle Celestino Silva¹

¹Universidade de São Paulo, Instituto de Física de São Carlos, Grupo de História Teoria e Ensino de Ciências, São Carlos, SP, Brasil.

Recebido em 07 de agosto de 2021. Revisado em 22 de novembro de 2021. Aceito em 22 de novembro de 2021.

Usualmente atribui-se a Claude Elwood Shannon o papel de fundador da área da física hoje conhecida como teoria da informação. Neste artigo, trazemos um estudo histórico dos desenvolvimentos do campo desde de 1922 até os estudos imediatamente posteriores ao grupo de publicações de 1948, procurando entender como o campo se desenvolveu, os papéis dos diversos pesquisadores envolvidos e os motivos que fizeram com que Shannon fosse considerado fundador do campo. Durante o período em questão pudemos observar mudanças nos aspectos priorizados na pesquisa, passando da busca do entendimento do que seria transmitido em uma comunicação para o foco em aplicações tecnológicas em detrimento da compreensão do significado de informação.

Palavras-chave: História da Física, Teoria da Informação, Claude Shannon.

The foundation of the field of physics currently known as information theory is usually attributed to Claude Elwood Shannon. This article provides a historical study of the developments in the field from 1922 to the studies immediately following the 1948 series of publications, seeking to understand how the field developed, the various researchers' roles, and why Shannon is considered the founder of the field. Along the period in focus, we could observe changes in the aspects prioritized in the research. It moved from the search for understanding what is conveyed in communication to focusing on the technological applications to the detriment of understanding the meaning of information.

Keywords: History of physics, Information theory, Claude Shannon.

1. Introdução

Em uma caverna, um homem faz desenhos na pedra, desenhos de outros homens, animais e plantas. Alguns séculos depois, em um círculo ao redor de uma fogueira, uma matriarca conta histórias aos outros membros da tribo. A capacidade de transmitir e receber informação é tão antiga quanto a própria humanidade, a eficiência da transmissão e o desenvolvimento da sociedade estão intimamente relacionados. Um salto no tempo e notamos a invenção de logogramas, hieróglifos e alfabetos; a tradição oral lentamente dá lugar à escrita. Nasceram impérios, os mensageiros romanos e os tambores da África correm e ecoam pelos continentes, carregando consigo notícias e informações. Finalmente, nos aproximamos da era moderna, as cartas cedem lugar para as novas maravilhas tecnológicas da revolução, o telégrafo dos Chappe sinaliza pelos céus de Paris, mensagens passam a se deslocar a velocidades maiores que qualquer homem ou cavalo. O domínio da eletricidade apenas impulsionou o progresso, e em meio a muitos aspirantes a inventores do telégrafo elétrico, surgem Samuel Morse e Alfred Vail nos Estados Unidos, William Cooke e Charles

Wheatstone na Grã-Bretanha, e o padre Roberto Landell de Moura no Brasil, palavras agora viajavam em fios ou pelo ar, e não demorou muito até cruzarem a barreira do Atlântico.

O desenvolvimento de um tratamento matemático para a informação, e conseqüentemente a fundação do que hoje conhecemos como teoria da informação, ocorreu na primeira metade do século XX e muitas vezes é atribuída ao matemático e engenheiro estadunidense Claude Elwood Shannon (1916-2001), que popularmente recebe a alcunha de *pai da teoria de informação* [1-3].

Este trabalho mostrará um panorama geral dos caminhos percorridos por diversos pesquisadores até a célebre publicação de Shannon em 1948, com enfoque na construção gradual e colaborativa dos conceitos e ideias. Além disso, buscamos estudar os motivos que levaram aos desenvolvimentos imediatamente posteriores ao trabalho de Shannon a se distanciarem da discussão de questões ontológicas da nova teoria, entre elas os desentendimentos a respeito da melhor definição de entropia no contexto informacional, com o foco das comunidades se voltando fortemente às aplicações tecnológicas, tais como a construção de “máquinas pensantes”.

Para tal, apresentamos alguns dos desenvolvimentos a partir da década de 1920, em particular os trabalhos

* Endereço de correspondência: aurelio.pena@usp.br

realizados nos laboratórios Bell pelos engenheiros Harry Nyquist (1889–1976) e Ralph Hartley (1888–1970) e na Europa pelo engenheiro eletrônico Karl Küpfmüller (1897–1977) que buscavam responder algumas perguntas importantes para o desenvolvimento da tecnologia da comunicação tais como: qual a velocidade máxima de transmissão de informação¹? Quanta informação uma fonte produz? Como “medir” a “quantidade de informação” produzida? Há um limite na quantidade de informação que pode ser enviada por um canal? Como transmitir de forma confiável mesmo lidando com ruídos?

Em seguida, analisamos a publicação do trabalho “A Mathematical Theory of communication” [4] por Claude Shannon em 1948, considerado o artigo precursor do novo campo de estudos ao procurar matematizar a transmissão de informação. Depois, desafiamos a ideia de um trabalho fundador único pela análise de trabalhos de coetâneos de Shannon que chegaram a conclusões semelhantes, com destaque para o desentendimento acerca da melhor definição de ‘quantidade de informação’ (entropia no contexto informacional) entre Shannon e o renomado matemático estadunidense Norbert Wiener (1894–1964), um dos pensadores mais influentes do período. Finalmente observamos os desenvolvimentos posteriores a 1948, limitando nossa análise à valorização da aplicação tecnológica da teoria em detrimento das investigações sobre a ontologia da informação.

Optamos pelo recorte acima, pois explorar todos os desenvolvimentos dos trabalhos de 1948 seria esboçar toda a história da teoria da informação incluindo suas diversas inclusões na psicologia, biologia, termodinâmica entre outros campos, o que não é o objetivo do presente artigo.

2. Metodologia

À disposição de um historiador da ciência existem diversas abordagens metodológicas possíveis para se estudar o passado. Neste trabalho utilizaremos a diacrônica [5], na qual procuramos observar o passado a luz de seus próprios problemas, métodos e soluções, sem impor a visão moderna como correta ou melhor que as demais concorrentes. Em contraste a essa forma de abordar o passado temos a abordagem anacrônica [5], a qual discrimina entre as teorias concorrentes baseado em fatos desconhecidos para a época e utiliza termos e jargões modernos e sem sentido para o contexto do estudo. Embora essa abordagem esteja superada e não seja mais utilizada por historiadores da ciência profissionais, ela ainda é comum em trabalhos de cunho histórico escritos por cientistas divulgadores e jornalistas. Para escapar do anacronismo, buscamos estudar e entender as fontes primárias – textos dos próprios cientistas da

época – em seu contexto científico, construindo uma narrativa que considere também outros contextos, tais como filosófico, social, cultural e político a depender das questões históricas investigadas. Todavia, é importante apontar que uma historiografia puramente diacrônica e contida no passado é impossível e indesejável do ponto de vista pedagógico, afinal os interlocutores deste trabalho estão no presente [6].

3. Precedentes

Ao longo da década de 1920, as sequelas da Primeira Guerra Mundial assolavam a Europa e os últimos soldados americanos retornavam para os Estados Unidos para encarar uma década de profundas transformações sociais, políticas e econômicas [7]. O êxodo rural se intensificou conforme mais e mais agricultores trocaram suas propriedades por trabalhos fabris; e o fordismo, recém introduzido na indústria automobilística passou a dominar grande parte das cadeias produtivas. Isso gerou uma explosão na capacidade produtiva e consequentemente, conforme a população se tornava assalariada, na capacidade do mercado consumidor de absorver os novos produtos. Carros, máquinas de lavar roupa, torradeiras, rádios, telefones, entre outros produtos, passaram a integrar o dia-a-dia dos estadunidenses. A popularização do rádio e do telefone na década de 1920, impulsionada pelas primeiras redes comerciais de rádio e expansão de centrais e linhas telefônicas, criou uma forte demanda tecnológica na área da comunicação, o que culminou com a união de diversos grupos de pesquisa sob uma única instituição: os Laboratórios Bell [8].

3.1. Os Laboratórios Bell

Formados como um braço de pesquisa da *American Telephone and Telegraph Company* (AT&T), os Laboratórios Bell promoveram um encontro único para a área da comunicação: engenheiros e matemáticos trabalhando juntos. Tal encontro, normalmente tenso² [8], elevou a instituição a um centro de matemática aplicada sem precedentes. No entanto, o instituto se mantinha praticamente invisível para o mundo acadêmico. Seus membros, normalmente, se limitavam a publicar no periódico de circulação interna, o *Bell Labs Technical Journal*, reduzindo o alcance das descobertas.

O periódico começou a ser publicado em 1922 e sintetiza bem as propostas dos Laboratórios Bell. Os volumes iniciais trazem publicações que orbitam em torno da comunicação, de seus aspectos teóricos e principalmente tecnológicos. Nesse cenário, alguns trabalhos se destacaram, em particular os artigos de Harry Nyquist e

¹ Na linguagem jargão da época o termo adequado seria inteligência, como discutiremos a seguir.

² A tensão se dava por um conflito de interesses: engenheiros, preocupados com o problema prático, valorizavam a aplicação dos métodos e não a validade universal da teoria; enquanto que matemáticos buscavam teoremas gerais não vinculados apenas a casos específicos.

Ralph Hartley que foram de grande importância para a formação posterior de uma teoria da comunicação, por isso, discutiremos alguns de seus aspectos em mais detalhe.

3.1.1. Os trabalhos de Harry Nyquist

Nyquist nasceu na Suécia em 1889 e imigrou ainda jovem para os Estados Unidos, obteve seu doutoramento em física por Yale em 1917 passando a integrar a equipe dos Laboratórios Bell no mesmo ano. Em 1924, apresentou a palestra “Certain factors affecting telegraph speed” a qual gerou uma publicação de mesmo título [9]. Nela, Nyquist explora algumas variáveis que pareciam prejudicar a transmissão de inteligência³, além de propor uma fórmula geral para a velocidade máxima de transmissão de inteligência em um sistema com um determinado código.

O primeiro fator abordado foi o formato das ondas transmissoras, Nyquist estuda ondas retangulares, senoidais e uma onda modificada proposta por ele. Segundo o cientista, a maior eficiência das ondas senoidais que estava sendo defendida por muitos engenheiros da época⁴ era falsa [9]; em sistemas operados em condições ótimas (máxima velocidade⁵ de transmissão de inteligência) as ondas retangulares e as ondas modificadas seriam mais eficientes.

Em seguida, Nyquist faz algumas observações importantes. A primeira delas a respeito de ondas previsíveis não carregarem inteligência, isto é, não carregam consigo nenhuma informação nova e, portanto, não podem transmitir inteligência, nas palavras do autor:

“O fato de a componente [senoidal][...] não carregar inteligência [...] torna-se claro quando consideramos que seus valores são previsíveis a qualquer momento e, portanto, a componente pode ser produzida localmente” [9, p. 223].

No trecho, notamos que Nyquist associa a transmissão de inteligência à incerteza na onda transmitida – uma onda perfeitamente previsível poderia ser produzida pelo destinatário a qualquer momento, assim, o remetente não enviou nenhuma inteligência pelo fio.

³ A palavra informação não era utilizada e os cabos telegráficos de 1924 transmitiam o que os engenheiros denominavam por inteligência. Apenas anos depois a palavra informação começou a ser utilizada, passando a designar a *commoditie* que preenchia o interior dos cabos telegráficos.

⁴ Nyquist cita como exemplos os trabalhos: “A practical Transmitter using the Sine Wave for Cable Telegraphy; Squier on an unbroken Alternating current for Cable Telegraphy” [10] e “A method of transmitting the Telegraph Alphabet Applicable for Radio, Land Lines and Submarine cables” [11] e “On an Unbroken Alternating Current for Cable Telegraphy” [12].

⁵ Vale notar que a definição dessa velocidade é dada por ele no artigo como $W = K \log(m)$, onde k é uma constante e m é o número de valores possíveis do código (Se o código tem pontos e barras ou 0 e 1 temos $m = 2$. O alfabeto tem $m = 26$).

A segunda observação, fortemente apoiada nos trabalhos do engenheiro eletricitista John Carson (1886–1940), diz respeito a limitações de banda conforme a velocidade de transmissão [13, 14]. Isto é, para transmitir sinais em uma taxa determinada é preciso consumir uma largura de banda e, caso o sistema não tenha banda suficiente, parte da inteligência não é transmitida. Esse tema foi explorado por Nyquist em sua publicação de abril de 1928 intitulada “Certain topics in telegraph transmission theory” [15].

Os resultados obtidos por Nyquist foram explorados mais profundamente por Ralph Hartley que, em julho de 1928, publicou uma lei mais geral para a interdependência da velocidade de transmissão e a largura de banda.

3.1.2. Os trabalhos de Ralph Hartley

Companheiro de Nyquist nos Laboratórios Bell e preocupado com as mesmas questões de transmissão de inteligência, o estadunidense Ralph Hartley ministrou um seminário no *International Congress of Telegraphy and Telephony* na Itália em 1927, publicada no periódico interno dos Laboratórios Bell no ano seguinte sob o título “Transmission of Information” [16]. Hartley define o termo “informação” e propõe uma medida quantitativa para essa grandeza, defendendo que é preciso tratar as mensagens como uma sequência aleatória de símbolos e desconsiderar a interpretação que o destinatário faz da mensagem do remetente. Ou seja, do ponto de vista da teoria as mensagens não precisam carregar significado; nas palavras do engenheiro isso implica em “desconsiderar os fatores psicológicos da mensagem” [16, p. 536].

O autor de “Transmission of Information” argumenta que para uma medida baseada apenas em considerações físicas, a mensagem não precisa ser relevante. Segundo o artigo, ela pode inclusive ser completamente arbitrária, como ocorre com uma máquina que seleciona letras aleatórias do alfabeto e as envia por um telégrafo⁶.

A segunda proposta do autor é em relação à forma de medir a informação. Para Hartley, a quantidade de informação que é transmitida no sistema é dada por:

$$H = n \log(s) = \log(s)^n$$

onde s é o número de símbolos para cada seleção (em código Morse 2, no alfabeto 26) e n o número de seleções. Por exemplo, 5 letras corresponderiam a 5 seleções no conjunto do alfabeto de 26 símbolos.

Desconsiderando-se aspectos semânticos, em uma sequência, cada símbolo deve carregar a mesma quantidade de informação. Esperamos, portanto, que uma sequência de 4 símbolos carregue duas vezes

⁶ Hartley compara essa máquina a duas pessoas que não falam a mesma língua enviando mensagens por um telégrafo, as palavras do emissor, seriam praticamente uma sequência aleatória de letras para o receptor, desprovidas de conteúdo semântico.

mais informação do que uma de 2 símbolos e 4 vezes mais informação do que uma sequência de 1 símbolo, o mesmo deve valer para letras formando uma palavra, ou palavras formando um texto⁷. Essa estrutura pode ser representada matematicamente por um logaritmo⁸. Exemplificando: podemos tomar o caso do alfabeto. Cada escolha de letra corresponde a 1 em 26 possíveis símbolos. Portanto, em uma escolha temos 26 possibilidades e em duas temos 676 (26^2) possibilidades, entretanto 2 letras não transmitem uma quantidade quadraticamente maior de informação que 1 letra. Na verdade, ignorando os fatores interpretativos (psicológicos) 2 letras deveriam carregar exatamente o dobro da informação.

A escolha do logaritmo torna possível o aumento de uma letra (símbolo) corresponder a um fator multiplicativo na quantidade de informação (H) e não um aumento exponencial do tipo 26^n , com n sendo o número de letras escolhidas:

$$H = n \log(s) = \log(s)^n$$

$$\text{Primeira escolha: } H = (1) \log(26)$$

$$\text{Segunda escolha: } H = (2) \log(26)$$

$$\text{Enésima escolha: } H = (n) \log(26)$$

Hartley também expande essas conclusões para sinais contínuos como os de telefone, defendendo que esses sinais transmitem quantidades finitas de informação e, portanto, podem ser aproximados por pequenos degraus (símbolos possíveis s na telefonia) que formam a onda completa.

A partir disso, Hartley procura entender o limite da transmissão de informação para cada sistema com uma largura de banda definida, estudando os sistemas existentes à luz da nova teoria. Em suas conclusões, o estadunidense defende que dada uma quantidade de informação, um produto específico de largura de banda por tempo é requerido para que o sistema consiga realizar a transmissão.

As lacunas do estudo de Hartley de 1927, residem em dois pontos principais: (1) a definição de informação dada só funciona em sistemas nos quais cada símbolo (s) tem a mesma chance de ser escolhido. Por exemplo, um dado viciado não poderia ser representado; e (2) os resultados são apoiados exclusivamente na aplicação

tecnológica, não há uma tentativa de generalização matemática dos conceitos.

Uma proposta a respeito do significado de informação foi construída nos Laboratórios Bell, principalmente por Nyquist e Hartley que defendem que informação é incerteza. O primeiro faz essa defesa de forma indireta propondo que uma onda senoidal previsível poderia ser produzida pelo destinatário e por isso o remetente não está enviando informação, afirmando indiretamente que o envio de informação depende da indeterminação do sinal. Já Hartley é mais explícito ao definir sua medida de quantidade de informação – para ele informação é sinônimo de indeterminação.

3.2. O velho continente

As décadas de 1920 e 1930 na Europa foram marcadas pela instabilidade política e social, causas da ascensão de regimes cada vez mais totalitários e extremistas que culminariam na Segunda Guerra Mundial. Nesse contexto foram publicados poucos trabalhos sobre comunicação a despeito da (ou justamente pela) relevância militar do tema. Apenas Karl Küpfmüller foi citado nos desenvolvimentos posteriores.

3.2.1. Os trabalhos de Karl Küpfmüller

Küpfmüller foi um engenheiro e professor alemão que trabalhou durante a década de 1920 para a Siemens & Halske em Berlim. Nesse período realizou estudos sobre transmissão telegráfica e chegou a conclusões semelhantes às de Nyquist nos Estados Unidos – principalmente a respeito das limitações na velocidade de transmissão impostas por limitações na largura de banda de sistemas [17]. O alemão também investigou questões de estabilidade de sistemas [18] criticando a aplicabilidade do critério de Barkhausen no caso geral⁹.

3.2.2. Os trabalhos de Leo Szilárd

O físico nuclear Leo Szilárd (1898–1964), diferente dos demais pesquisadores mencionados anteriormente, não estava estudando fenômenos ligados à comunicação quando propôs uma relação entre entropia e memória. O húngaro estava procurando uma forma de resolver o antigo paradoxo termodinâmico do demônio de Maxwell. Em seu artigo de 1929 [19], o autor defende que a ação do demônio em um sistema termodinâmico implicaria em uma medida e uma decisão; esse processo estaria associado a um aumento de entropia da forma: $\Delta S = k \log(2)$. Expressão semelhante à encontrada por Hartley no caso de um código com dois símbolos.

⁷ Podemos imaginar exceções nas quais grande parte da informação de um texto está contido em uma palavra, por exemplo na frase “não venha amanhã”, remover a palavra “não” muda completamente a informação transmitida. Ou poemas, nos quais muita informação é armazenada em pequenas sequências de símbolos. Todavia, esse tipo de análise pressupõe que não estamos interpretando ou significando a mensagem e, portanto, não estamos desconsiderando os “fatores psicológicos”.

⁸ Quanto a base desse logaritmo, Hartley alega que ela pode ser arbitrária e “a seleção de uma base particular fixa o tamanho da unidade de informação” [16, p. 540]. Shannon, posteriormente vai escolher a base 2 e definir a unidade de informação, batizada por John Turkey (1915–2000), como Bit.

⁹ O critério é uma condição matemática para a oscilação em circuitos elétricos, para mais detalhes ver [20]. Posteriormente, em 1928, Nyquist abordou sistemas semelhantes chegando às mesmas conclusões e avançando em alguns pontos que não são relevantes para o presente estudo [15].

A conexão entre termodinâmica e teoria de informação vem do fato de o modelo de Szilárd permitir interpretação e análise termodinâmicas, que ao mesmo tempo implicam em um processo de decisão binária, sendo possível estabelecer uma relação quantitativa entre informação usada pelo demônio e diminuição de entropia do reservatório. Tal relação foi posteriormente comentada por Norbert Wiener em 1948 ao definir um conceito próprio de informação e, mais tarde, Wiener também tentaria em 1961 unir a teoria de informação à termodinâmica.

3.3. Depressão e guerra

A crise econômica de 1929 nos Estados Unidos marcou o início da recessão econômica que acompanharia o país pela década seguinte. Acompanhando a recessão, notamos uma redução considerável da produção científica sobre a teoria da comunicação em solo estadunidense.

Em 1933, o engenheiro soviético Vladimir Kotelnikov (1908–2005) publicou um artigo sobre a capacidade de transmissão do éter¹⁰, chegando às mesmas conclusões de Nyquist e Küpfmüller a respeito da limitação de banda também restringir a velocidade máxima de transmissão [21]. Porém, esse trabalho não parece ter chegado aos Estados Unidos antes das publicações de Shannon, Wiener e Tuller em 1948 e 1949.

Em 1939 eclodiu a Segunda Guerra Mundial e grande parte dos pesquisadores envolvidos na pesquisa em comunicação foram recrutados para os esforços de guerra, provocando um novo período sem publicações [8]. Esse hiato terminou com uma nova onda de trabalhos a partir de 1946 que discutiremos a seguir.

4. Os Trabalhos de Claude Elwood Shannon

Claude Elwood Shannon nasceu em Petoskey, Michigan em 30 de abril de 1916. Frequentou o ensino médio na *Gaylord High School*, escola na qual sua mãe lecionava, e se formou em 1932. No mesmo ano entrou na Universidade de Michigan e em 1936 obteve uma formação dupla em engenharia elétrica e matemática. Após graduar-se Shannon se inscreveu para operar o Analisador Diferencial de Vannevar Bush (1890–1974) no MIT. A máquina era “uma plataforma metálica de cem toneladas cheia de eixos e engrenagens em movimento” [8, p. 180] dedicada a resolver analogicamente equações diferenciais [8]. A formação dupla de Shannon tornou-o extremamente qualificado para o trabalho de traduzir as equações em movimentos mecânicos no analisador.

Sob orientação de Bush, escreveu sua tese de mestrado relacionando os muitos circuitos, interruptores e relés do

Analisador com a álgebra booleana, resultando em sua primeira publicação em 1938 na *I.E.E.E Transactions* que seria laureada em 1940 pelo Prêmio Alfred Noble das sociedades de engenharia dos Estados Unidos. No mesmo ano da premiação, Shannon recebeu seu doutorado em matemática com um trabalho que relacionava genética e álgebra booleana.

Durante esse período, Shannon desenvolveu interesse pela área da comunicação, passando o verão de 1937 nos Laboratórios Bell. Dois anos depois escreveu uma carta a Bush [22] na qual cita os trabalhos de Hartley [16] e Carson [13] além de dizer:

“Estou tentando provar o seguinte teorema: para quaisquer operadores T R o comprimento de uma mensagem arbitrária f_1 multiplicado por seu espectro essencial e dividido pela distorção do sistema é menor que uma certa constante vezes o tempo de transmissão de F multiplicado por sua largura de espectro essencial ou – grosso modo – é impossível reduzir largura de banda vezes tempo de transmissão para uma distorção constante” [22, p. 2].

Como citado pelo próprio Shannon, a ideia é parecida com o que foi proposto por Hartley para sistemas específicos em 1929 [16]. Entretanto, notamos a tentativa de aplicar uma linguagem matemática e encontrar uma prova para quaisquer operadores.

Shannon passou os anos de 1940–1941 entre os Laboratórios Bell e o Instituto de Estudos Avançados em Princeton sob tutela de Hermann Weyl. Em 1941 Shannon foi recrutado para compor o time de pesquisa dos Laboratórios Bell dedicando-se aos *anti-aircraft directors* – sistemas dedicados a observar aeronaves inimigas e calcular a mira dos mísseis em solo.

Em 1945 Shannon publica o artigo “A Mathematical Theory of Cryptography” no qual utiliza pela primeira vez o termo “*information theory*”. Três anos depois, Shannon apresentaria seu artigo de maior impacto, “A Mathematical Theory of Communication”, publicado em duas partes no periódico interno dos Laboratórios Bell, composto por vinte e três teoremas e sete apêndices com provas matemáticas detalhadas, nem todas corretas, que mudariam o rumo da comunicação [8, 24].

4.1. Informação para Shannon: o bit

O artigo de 1948 de Shannon [4] representa a realização de algumas ideias presentes em uma carta a Vannevar Bush de 1939 [22], em particular, a tentativa de provar teoremas gerais para a comunicação independentes das aplicações tecnológicas ou de casos específicos. Na introdução de seu trabalho Shannon cita os de Nyquist [9] e Hartley [16] como importantes bases, e propõe a inclusão de outros fatores como o ruído e a estrutura real da mensagem na teoria, como veremos em mais detalhes logo a seguir.

¹⁰ Devido aos experimentos do final do século XIX e a relatividade restrita, em 1933 a existência do éter já era questionada por muitos [23] o que pode ter contribuído para a pouca atenção recebida pelos estudos de Kotelnikov.

Nyquist afirmara em 1924 que uma onda previsível não poderia carregar consigo nenhuma informação, ideia utilizada por Hartley para propor uma medida de informação que dependeria das escolhas do transmissor frente às possibilidades disponíveis, considerando apenas casos nos quais os símbolos da mensagem tinham a mesma probabilidade de escolha.

O estilo de pensamento, típico dos Laboratórios Bell (que considera a quantidade de informação como sinônimo de incerteza), será seguido por Shannon. Entretanto ele começa o seu artigo de 1948 definindo algo que fora ignorado pelos seus antecessores: uma unidade de medida de informação e uma representação geral para sistemas de comunicação esquematizado na Figura 1.

A unidade proposta por Shannon é o “bit”, nome proposto pelo estatístico e colega nos Laboratórios Bell John Tukey (1915–2000), como uma abreviação para *binary digits*. O bit é tratado matematicamente por um logaritmo, como feito em 1927 por Ralph Hartley, correspondendo à quantidade de informação armazenada em um sistema de duas posições (como um interruptor). N interruptores possuem 2^N estados podendo, portanto, carregar consigo $\log_2 2^N = N$ bits.

Um sistema de comunicação geral é representado por uma fonte (I) e um transmissor (II) que produzem e codificam a mensagem, enviando-a pelo canal (III). Do outro lado do canal há um decodificador e um destinatário que recebem a mensagem enviada.

Shannon generaliza as ideias de Nyquist e Hartley criando um modelo válido para a comunicação humana¹¹. Em uma língua, as escolhas de letras e palavras para compor um texto não são aleatórias e independentes, mas dependem de uma estrutura anterior. Para ilustrar, tomemos o exemplo da língua portuguesa: quando uma palavra apresenta a letra “M” as únicas consoantes que podem se seguir são “P” e “B”, portanto sempre que um “M” aparece estamos limitados a sete seleções (A, E, I, O, U, P e B). Essa estrutura, na qual a escolha seguinte depende da variável atual, é característica de um processo estocástico. Em seu trabalho, Shannon mostra tratar-se de um processo de Markov, um tipo específico do anterior. A estrutura da linguagem – e, portanto, das mensagens na comunicação – segue uma estrutura estatística do tipo de um processo estocástico que pode aumentar de complexidade conforme adicionamos regras a ele¹².

Com isso, Shannon propõe uma série de condições que devem ser satisfeitas para se definir uma medida para “a

quantidade de ‘escolhas’ envolvidas em uma seleção” e deduz uma expressão matemática da forma:

$$H = K \sum_{i=1}^n p_i \log_s(p_i)$$

onde K é uma constante positiva, p_i as probabilidades de cada evento possível e s é o número de símbolos disponíveis (2 em um código Morse, 26 em um alfabeto, etc.).

Nesse ponto o autor ressalta a semelhança entre essa expressão e a da entropia termodinâmica, em particular o teorema-H de Boltzmann [25], denotando H como uma medida de entropia¹³. A ligação é semelhante à feita anteriormente por Szilárd, entretanto Shannon não cita o húngaro em seu trabalho e afirma, anos depois em uma entrevista [24], que não conhecia os trabalhos de Szilárd no período da publicação.

Shannon também aborda o problema relacionado à capacidade máxima de transmissão de um canal durante um intervalo de tempo determinado, partindo da seguinte definição para a capacidade de transmissão:

$$C = \lim_{t \rightarrow \infty} \frac{\log N(T)}{T}$$

onde $N(T)$ é a quantidade de símbolos que podem ser transmitidos durante um período de tempo T . C é o limite superior para a taxa de transmissão de informação por um canal de comunicação específico, transmissões a uma taxa superior a C podem se sujeitar à perda de informação.

Essa definição foi justificada no artigo pelo caso do teletipo, uma espécie de aparelho telegráfico que envia diretamente um texto digitado em um teclado no posto transmissor até o receptor, e pela aplicação em canais discretos com e sem ruído. Para tal, é preciso considerar a interferência do ruído e realizar um tratamento matemático para esse importante fator limitante nas transmissões. Shannon trata o ruído estocasticamente, de forma semelhante à mensagem, e explora suas definições e teoremas nos exemplos práticos, posteriormente expandindo as conclusões para canais de transmissão contínua, apoiando-se na discretização proposta por Hartley de se considerar o sinal contínuo como sucessivos degraus discretos.

Finalmente, Shannon retorna para a linguagem escrita para propor uma forma de reduzir os erros de transmissão de informação (seja em telégrafos ou telefonia). A língua possui uma redundância inerente, que no inglês, segundo o artigo, pode chegar a 50%, isto é, apenas metade das letras em um texto são relevantes para sua compreensão. Entretanto é exatamente essa redundância que reduz os erros na nossa comunicação, pois não precisamos de todas as letras ou palavras para entender um texto. Com isso em mente, Shannon propõe uma

¹¹ Shannon afirma que um modelo estocástico que considera palavras e não letras seria o melhor modelo para as línguas ocidentais e orientais. Utilizaremos o exemplo das letras porque é didático (do ponto de vista das línguas ocidentais), mas no caso geral se considera palavras como unidades básicas da comunicação.

¹² Se adicionamos regras o suficiente, um texto gerado por uma máquina que escolhe palavras aleatoriamente dentro dessas regras será indistinguível de um texto escrito por uma pessoa, ou seja, um processo estocástico suficientemente complexo pode mimetizar perfeitamente a língua.

¹³ Esta conexão será brevemente comentada na seção 5.3.1.

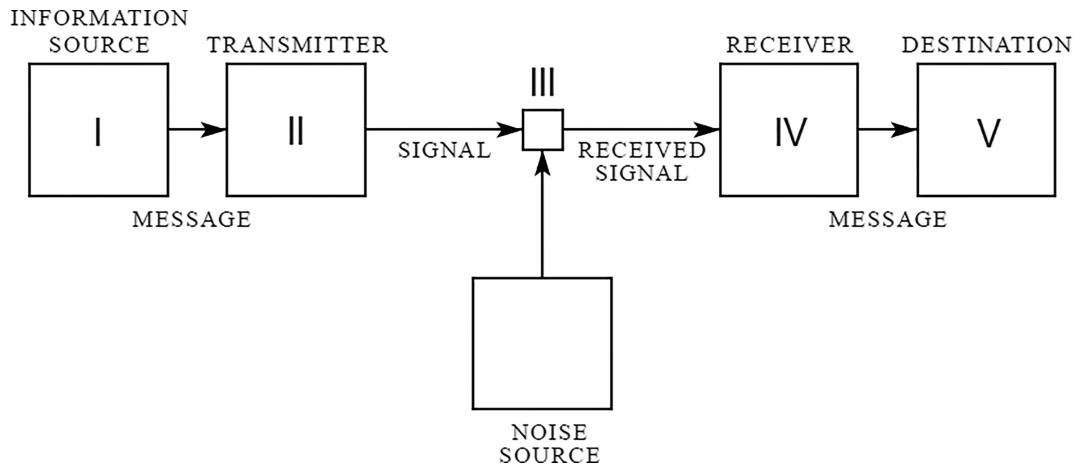


Figura 1: Diagrama proposto por Shannon para um sistema de comunicação geral. Fonte: Adaptada de SHANNON [4].

forma de eliminar o erro pela introdução de redundância no sistema enviando uma mesma mensagem duas ou mais vezes seguidas e comparando o que foi recebido pelo destinatário.

Mas como aumentar a redundância sem perder velocidade de transmissão? Shannon defende que isso pode ser feito pela compressão dos dados, exemplificada na época pelo código Morse. Nesse código, os menores símbolos correspondem às letras mais comuns no inglês, assim a letra “E” que é a mais comum na língua inglesa é apenas um ponto, enquanto que “Z” são dois traços e dois pontos, uma forma eficiente de compressão. Já existiam na época livros de abreviações¹⁴ para utilização em telégrafos e Shannon propôs um sistema de abreviações comum ao transmissor e receptor, de forma que estes codificassem e decodificassem as mensagens enviadas pelo canal de comunicação (Item III da Figura 1).

Nas páginas de “A Mathematical Theory of Communication” notamos algumas ideias facilmente reconhecíveis na computação moderna o que poderia tornar tentador atribuir a autoria das ideias usadas hoje a Shannon. Isso seria cair em na sedutora armadilha do anacronismo, já que conceber conceitos modernos como se eles já estivessem prontos em 1948 é uma deliberada “intervenção” no passado que desconsidera todos os desenvolvimentos posteriores que buscaram adequar, criar e abandonar conceitos e ideias. O historiador da ciência Helge Kragh considera que “atualmente, história anacrônica da ciência raramente é uma estratégia consistente. Pelo contrário, há amplo consenso sobre elogiar

um ideal não anacrônico” [5, p. 89]. Para compreendermos melhor as mudanças nos significados das ideias, precisamos compreender como foram recebidas em seu próprio tempo e contexto.

5. Contemporâneos

Até o momento exploramos o contexto científico na época da publicação do famoso artigo de Shannon, sendo fácil notar que as conclusões e os teoremas propostos em “A Mathematical Theory of Communication” não são fruto de uma mente isolada e brilhante, mas resultado de um processo histórico envolvendo diversos atores que se influenciaram mutuamente. Nesta seção abordamos alguns estudos publicados em um período próximo a obra de Shannon com o objetivo de entender melhor como esta foi recebida e os caminhos que a teoria de informação tomou a partir de 1948.

5.1. Dennis Gabor: os logons de informação

Dennis Gabor (1900–1979) foi um engenheiro elétrico e físico nascido na Hungria, embora seja mais reconhecido por seus trabalhos com holografia – pelos quais foi laureado com o Nobel em 1971 – Gabor publicou o artigo intitulado “Theory of Communication” [26] em 1946 no qual faz deduções a respeito das limitações ligadas à frequência e tempo de transmissão nos sistemas de comunicação além de propor como unidade de medida de informação, nas palavras do autor, o “quanta de informação” denominado “logon”.

O artigo começa reconhecendo os trabalhos de Carson [13], Nyquist [9], Küpfmüller [18] e Hartley [16] mas aponta que o problema da comunicação envolve as bandas de transmissão (limitações de frequência) e o tempo no qual a transmissão acontece (limitações de tempo). Os quatro autores citados estudaram as limitações de frequência; Gabor, por sua vez, incluiu as limitações de tempo em suas pesquisas. Para isso, o húngaro propõe o

¹⁴ Esses livros eram listas de palavras com suas respectivas abreviações, na transmissão telegráfica a cobrança era por letra, assim, um acionista que precisava enviar diariamente tabelas de preços pelos telégrafos podia possuir um livro com abreviações para as palavras mais comuns de seu ofício, se o receptor da mensagem possuir o mesmo livro, um conjunto de 5 letras como “PAAM” pode, naquele contexto representar a frase “O preço da ação na abertura do mercado”. Esses livros eram comuns e existiam diferentes livros para diferentes áreas [8], essa é uma forma de compressão de dados.

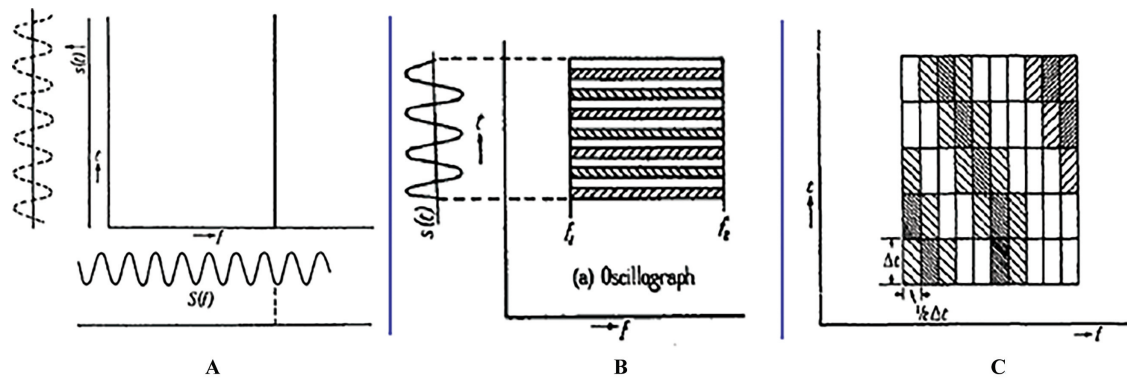


Figura 2: Representação de sinais nos diagramas de informação. Fonte: adaptada de GABOR [26].

uso dos “diagramas de informação” (nome dado por ele) que seriam “representações bidimensionais de sinais, com tempo e frequência como coordenadas” [26, p. 429].

Para entender tais diagramas, vamos analisar a Figura 2(a), onde temos um oscilador harmônico representado no diagrama de informação. Sabendo com exatidão a frequência do oscilador, o tempo de oscilação é completamente indefinido (linha vertical). Um sinal típico possui um intervalo de tempo e conseqüentemente um intervalo de frequências (f_1 a f_2) nos quais ele é transmitido. Representamos esse sinal no diagrama de informação por um retângulo (Figura 2(b)).

Gabor reconhece a similaridade desses diagramas com os gráficos de posição e momento (x e p) da mecânica quântica e deduz a incerteza de tempo (t) e frequência (f). Tomando o aparato matemático da teoria quântica como base e modulando o sinal como uma soma de senos e cossenos complexos, o autor faz um procedimento matematicamente similar ao realizado para se deduzir o princípio de incerteza de Heisenberg para momento e posição, mostrando que a indefinição temporal induz a indefinição na frequência dadas por:

$$\Delta t \Delta f > \frac{1}{2}$$

Em seguida Gabor encontra o sinal que gera o menor valor possível para a desigualdade, ou seja, $\Delta t \Delta f = 1/2$, definindo-o como o “sinal elementar”¹⁵. Com isso qualquer sinal no diagrama de informação poderia ser representado como uma soma de sinais elementares de lados Δt e Δf . Como ilustrado na Figura 2(c) cada um dos retângulos contém um “quanta de informação” e são as unidades mínimas, os logons.

Gabor aplica esses conceitos para alguns casos práticos como modulações de frequência e telefonia. O trabalho traz para a teoria de comunicação o formalismo matemático da mecânica quântica, e uma unidade prática de medida de informação: o logon. Todavia Gabor não discute nenhuma forma de ruído e não fornece aos

engenheiros um método prático de transmitir informação mais eficientemente.

5.2. A tese de William Tuller

William Gordon Tuller (1918–1954) publicou, em abril de 1949, o artigo “Theoretical Limitations on the Rate of Transmission of Information” [27] baseado em sua tese de doutorado num contexto no qual a comunidade científica em polvorosa pelas publicações do ano anterior.

O ponto de partida para a tese de Tuller é a medida de informação proposta por Hartley [16] com a inclusão de uma constante de proporcionalidade (K):

$$H = n(K \log s)$$

novamente n é o número de seleções e s o número de símbolos possíveis.

Com essa medida para quantidade de informação, o autor passa a estudar a transmissão de informação sem ruído. No período, era praticamente unanimidade que canais desprovidos de ruído possuiriam um limite para a transmissão de informação imposto pela geração de um efeito transiente no filtro do emissor, ideia da qual Tuller discordava.

Os demais pesquisadores defendiam que o primeiro pulso enviado criaria um transiente no filtro que adicionaria banda aos demais pulsos do sistema, gerando um efeito acumulativo com cada pulso. Tuller defende que, caso se conheça o filtro com antecedência, é possível adicionar ao sinal recebido uma onda inversa à interferência, anulando o efeito do transiente e permitindo transmissões ilimitadas em sistemas sem ruído. Em seguida, o autor parte para a sistemas com ruído e cria um diagrama para o sistema de comunicação geral, semelhante ao proposto por Shannon (seção 4.1).

A análise de Tuller baseia-se fortemente nas definições de Hartley de 1928 e está limitada às lacunas presentes em seu artigo “Transmission of Information” [16] como a falta de um tratamento para o ruído e a limitação da definição de quantidade de informação aos casos nos quais a probabilidade de escolha dos símbolos é igual. Tuller reconhece tais problemas e procura solucioná-los

¹⁵ O sinal é $\psi(t) = \exp(-\alpha^2(t-t_0)^2) \cos(2\pi f_0 t + \phi)$, onde α , t_0 e f_0 são constantes associadas às características do pulso.

criando uma análise para o ruído e generalizando a quantidade de informação.

A generalização consiste em ampliar a definição de informação como incerteza da escolha, dada por Hartley. Tuller afirma que em um sistema no qual todos os símbolos podem ser selecionados ($s_{av} = s_{max}$)¹⁶, a transmissão da informação é máxima, quando os símbolos são limitados ($s_{av} < s_{max}$) o sistema está transmitindo informação abaixo da capacidade máxima.

Os esforços de Tuller para generalizar a abordagem de Hartley não foram reconhecidos pela comunidade, pois as definições defendidas por ele foram superadas nos trabalhos do ano anterior [4, 28], somados às novas publicações do início de 1949 [29, 30]. Esses fatores mantiveram o artigo fora do radar das grandes discussões.

5.3. A cibernética de Wiener

Formado em matemática aos 14 anos e doutor em lógica matemática em 1913 com apenas 19 anos, o filósofo e matemático estadunidense Norbert Wiener era considerado um dos mais influentes cientistas e matemáticos no período. Durante o esforço de guerra, dedicou-se a problemas relacionados à mira de aeronaves, modelando os desvios do sinal nas miras, ou “ruído”, como um processo estatístico, pois Wiener notara a semelhança desse ruído àquele presente na comunicação. Essa investigação gerou, em 1942, o livro *The extrapolation interpolation and smoothing of stationary time series* [31] que abordava o problema do ruído irreduzível presente em misturas de sinal e ruído.

A obra teve circulação reduzida pelas forças armadas estadunidenses, mas Shannon e Tuller tiveram acesso a ele e o citaram em suas publicações. Todavia o artigo empregava uma linguagem matemática “muito além da capacidade do engenheiro de comunicação comum” [27, p. 2] e acabou esquecido. Após a guerra, a Josiah Macy Foundation passou a promover congressos para estudos multidisciplinares em diversas áreas do conhecimento, com enfoque na interface entre psicologia, fisiologia e engenharia de comunicação [8]. Wiener participou de tais seminários e na mesma época começou a escrever um livro igualmente multidisciplinar sobre cibernética¹⁷ [28].

No mesmo ano da publicação do trabalho de Shannon (1948), Wiener publicou o livro *Cybernetics: or control and communication in the animal and the machine* [28], no qual elabora um tratado extenso discorrendo sobre diversos tópicos, desde mecânica estatística até relações entre informação, linguagem e sociedade.

¹⁶ s_{av} é o nome dado a quantidade de símbolos disponíveis na transmissão.

¹⁷ Algumas fontes atribuem a Wiener a cunhagem do termo cibernética [8], todavia é importante apontar que o termo já era utilizado antes do lançamento do célebre livro de 1948 [28], em particular nas conferências Macy (desde 1946) que definiam o termo como “mecanismos causais e de feedback em sistemas biológicos e sociais” [32], o que Wiener faz é redefinir o termo como “o estudo científico do controle e da comunicação no homem e na máquina” [28].

Ao longo da obra, Wiener define quantidade de informação como uma medida do grau de organização de um sistema. No presente estudo vamos nos limitar a esse aspecto do trabalho, principalmente porque um de seus desdobramentos foi uma disputa pela definição mais adequada para o conceito de quantidade de informação entre Wiener e Shannon. A definição dada por Wiener no *Cybernetics* é:

“A noção de quantidade de informação se liga muito naturalmente a uma noção clássica da mecânica estatística: a entropia. Da mesma forma que a informação é a medida do grau de organização do sistema, a entropia é uma medida do grau de desorganização; e uma é simplesmente o negativo da outra” [33, p. 11].

Da mesma forma que Szilárd partiu da análise do paradoxo do demônio de Maxwell, Wiener considera que a sua definição para quantidade de informação pode ser utilizada em análises do demônio de Maxwell sem citar o húngaro¹⁸, mas não desenvolve a ideia na obra de 1948¹⁹. Contudo, definir quantidade de informação com o grau de organização de um sistema, mostrou-se na contramão dos trabalhos anteriores na teoria da comunicação, pois Hartley e Shannon definem informação como uma medida de incerteza (ou desordem) enquanto Wiener trata informação como certeza (ou ordem).

5.3.1. Dois conceitos, uma informação

Wiener cita, na introdução da segunda edição do livro *Cybernetics* [33], o trabalho de Shannon nos Laboratórios Bell e alega que ambos tiveram a mesma ideia: criar uma medida para a quantidade de informação. Entretanto, uma análise cuidadosa na forma pela qual os autores definem os conceitos mostra que eles são diferentes em seus fundamentos. Informação para Shannon é análoga à entropia de Boltzmann enquanto a informação de Wiener é o negativo dessa entropia.

Vale observar que a decisão de Shannon de pautar a definição de informação na incerteza de uma escolha do sistema é coerente com trabalhos anteriores da equipe dos Laboratórios Bell²⁰. Nyquist, por exemplo, havia discutido a onda senoidal previsível em 1924, concluindo que ela não carregaria nenhuma informação enquanto que Hartley baseou sua noção de informação na indeterminação presente na escolha de um símbolo.

Na década de 1950, ocorreram esforços para explorar os conceitos de entropia de Shannon e Wiener, principalmente para unificar a entropia termodinâmica e a entropia informacional. Entre esses esforços destacaram-se os

¹⁸ Não se sabe se Wiener conhecia ou não o trabalho de Szilárd.

¹⁹ Em 1961, na segunda edição do livro, Wiener faz um tratamento detalhado para o demônio de Maxwell, mas esse desenvolvimento foge ao escopo desse artigo.

²⁰ Essa abordagem foi brevemente exposta na seção 3.1.2.

estudos de Léon Brillouin (1889–1969)²¹ que expandiu o trabalho de Szilárd, procurando no demônio de Maxwell a conexão entre as duas entropias. Wiener, por sua vez, aborda o problema do demônio de Maxwell na segunda edição do *Cybernetics*, publicada em 1961 [33].

Finalmente, é notável a diferença de abordagem dos dois matemáticos. Shannon ataca o problema diretamente, trazendo definições, teoremas e deduções fortemente apoiadas na prática da engenharia. Wiener, por sua vez, traz suas definições em meio a uma obra complexa e multidisciplinar. As diferenças no estilo de apresentação das ideias e nas ênfases dadas à teoria e aplicação influenciaram a escolha da comunidade e, principalmente, contribuíram para a valorização tecnológica da teoria, pois o trabalho de Shannon é direto e aplicado enquanto que o tratado de Wiener é extenso e não ataca os problemas práticos diretamente. A imensa maioria da comunidade utilizou as formulações propostas por Shannon para a prática, “No MIT a ordem era atacar os problemas práticos” [24, p. 52]. Nas palavras de David Forney²², comentando a respeito da sua própria experiência no MIT nesse período: “O conselho era: ‘não trabalhe com a teoria, vá para as aplicações’” [24, p. 52].

6. Comunidade, Sociedade e Informação

A construção da teoria de informação, desde as suas bases na década de 1920 até os trabalhos de 1949, foi um processo que envolveu contribuições de vários pesquisadores de diferentes áreas. Nesta seção vamos explorar os desdobramentos imediatamente posteriores, começando pelas reações à publicação do artigo de Shannon “A Mathematical Theory of communication” procurando entender o caminho que a teoria de informação tomou após 1948.

As recepções ao trabalho no meio acadêmico foram diversas. O matemático estadunidense Joseph Leo Doob (1910–2004) publicou uma revisão da obra de Shannon criticando a abordagem matemática do trabalho, principalmente o tratamento dado à taxa de geração de informação em sistemas. Nas palavras de Doob:

“A discussão é sugestiva, e não matemática, e nem sempre fica claro se as intenções matemáticas do autor são honráveis” [34, p. 2]

Shannon alegava que as “liberdades” tomadas na análise, principalmente ligadas ao limite para o caso contínuo, poderiam ser “justificadas em todos os casos de interesse prático” [4, p. 32]. Ou seja, nos casos reais da engenharia, as demonstrações se sustentavam, entretanto, se sua proposta geral era formular uma

teoria matemática para o campo e não apenas resolver para alguns casos específicos como seus antecessores, as demonstrações precisavam de mais rigor.

Por essa falta de generalidade em algumas análises, matemáticos como Doob [34] criticaram o trabalho e a comunidade de matemáticos rejeitou as demonstrações e os teoremas apresentados, até que alguns de seus pares como Brockway McMillan (1915–2016) [35], Aleksandr Khinchin (1894–1959) [36] e Robert Fano (1917–2016) o último apoiado pelos estudantes do MIT na nova disciplina “*Transmission of Information*” [24], dedicaram-se a formular os teoremas e as provas de forma matematicamente precisa. Assim, somente cerca de apenas vinte anos depois da publicação do artigo, os teoremas de Shannon haviam sido devidamente provados.

Retornando ao período da publicação do artigo de Shannon, 1948, a publicação de outros trabalhos no campo, como o livro *Cybernetics: or control and communication in the animal and the machine* de Wiener, que propunham uma análise interdisciplinar, contribuíram para que o artigo de Shannon (que seria republicado como um livro em 1949) atingisse um público mais amplo, além de engenheiros e matemáticos. A abordagem de Shannon passou a ser conhecida por outras comunidades e a receber atenção de novos grupos de cientistas como biólogos e psicólogos.

Somado a isso, o artigo de Shannon foi recebido nos Estados Unidos em um momento particularmente favorável, no final da Segunda Guerra. Devido à expansão da “ameaça soviética” o público geral passou a valorizar mais a pesquisa científica e tecnológica, principalmente as vinculadas a usos militares, como era o caso das desenvolvidas nos Laboratórios Bell [36]. Shannon também publicou seu artigo em um livro no ano seguinte [30] (1949) em parceria com Warren Weaver (1894–1978). Este escreveu uma revisão do livro em linguagem compreensível voltada ao grande público na *Scientific American* [37] para a obra ampliar ainda mais seu alcance.

Em suma, as reações imediatamente posteriores às publicações do trabalho de Shannon foram negativas por parte dos matemáticos, mas positivas na comunidade de engenheiros. Além disso, o surto de popularidade da obra e a ideia de máquinas pensantes (promovida pelo livro de Wiener) atraiu a atenção da sociedade e dos pesquisadores de outras áreas. Com isso, o conceito de informação se popularizou passando a ser utilizado em muitas disciplinas, como a psicologia, economia e algumas ciências sociais, com significados não necessariamente congruentes ao proposto por Shannon. Conferências passaram a ser organizadas regularmente para explorar a interdisciplinaridade do tópico e suas aplicações. Os conceitos de entropia, informação e redundância pareciam capazes de resolver diversos problemas não diretamente relacionados com a comunicação.

Shannon passou a frequentar as conferências promovidas pela Fundação Josiah Macy Jr [8] com a presença de

²¹ Principalmente no trabalho *La Science et la Théorie de l'information* (Masson, 1959).

²² Forney é professor adjunto de engenharia elétrica no MIT e no período era estudante.

Wiener. As reuniões reuniam biólogos, físicos, psicólogos entre outros pesquisadores para apresentar e discutir problemas relacionados aos seus respectivos campos. A teoria da informação atraiu a atenção de muitos desses cientistas, o que resultou na utilização dos conceitos de forma descuidada em diversas áreas de pesquisa. Em resposta, alguns pesquisadores (como Shannon e Wiener) começaram a criticar o uso da teoria fora do campo da comunicação para o qual ela foi inicialmente proposta. Em 1956 Shannon escreveu [38]:

“A teoria de informação, nos últimos anos, se tornou um ‘efeito manada’²³ científico. Começando como uma ferramenta para o engenheiro da comunicação, ela recebeu uma publicidade extraordinária na imprensa popular e científica. [...] pesquisadores de muitos campos diferentes, atraídos pela festa e pelas novas avenidas abertas para análise científica, estão utilizando as ideias para seus próprios problemas. Aplicações estão sendo feitas na biologia, psicologia, linguística, física básica, economia, teoria da organização, entre outros. [...] Eu pessoalmente acredito que muitos dos conceitos da teoria da informação vão se provar úteis nesses outros campos – e realmente alguns resultados são bastante promissores – mas estabelecer essas aplicações não é um processo trivial de traduzir as palavras para um novo contexto, mas o processo tedioso e lento de criação de hipóteses e verificação experimental” [38, p. 3].

As reações de Shannon às novas aplicações mostram que “A Mathematical Theory of Communication” é um artigo escrito por um engenheiro e matemático para outros engenheiros do campo, que visava primeiramente a aplicação tecnológica na engenharia. Essa interpretação é corroborada pela análise das publicações posteriores do próprio Claude Shannon, de artigos sobre comunicação [29] na presença de ruído e pelas “máquinas pensantes”, em particular uma máquina capaz de jogar xadrez [36] e um “rato” capaz de resolver um labirinto simples [8], o qual poderia ser modificado para tornar o trabalho do “rato” mais complicado.

Notamos, portanto, que as comunidades de engenheiros e cientistas deixaram de se preocupar com as discussões básicas do conceito de informação e entropia para aplicar tais ideias nos mais diversos campos, em particular à engenharia da comunicação que finalmente poderia utilizar equações matemáticas para aprimorar as tecnologias existentes, buscando a eficiência máxima de transmissão de informação.

²³ O original *bandwagon* se refere ao viés cognitivo das pessoas de aderirem a uma ideia unicamente porque muitas pessoas estão fazendo o mesmo.

7. Conclusões

O objetivo deste artigo foi traçar um panorama geral do longo processo que levou à formação de um novo campo de estudos conhecido atualmente como teoria de informação. É notável que foram necessários diversos anos e a atuação de pesquisadores dedicados e não apenas um surto de genialidade de um grande cientista. A caracterização de Claude Shannon como “o pai” da teoria de informação é ingênua e desconsidera desenvolvimentos anteriores, coevos e posteriores à publicação de seu trabalho. Propomos alguns motivos para a valorização da aplicação tecnológica da nova teoria frente ao aprofundamento dos conceitos nos desenvolvimentos posteriores: (1) o artigo de Shannon, que se popularizou entre os engenheiros e pesquisadores, era focado nos aspectos práticos do problema da comunicação e não na discussão dos significados dos conceitos; (2) o desencontro entre Shannon e Wiener acerca da natureza ontológica da informação não teve resultados imediatos – décadas se passaram até a entropia termodinâmica e a entropia informacional se conversarem; e (3) os engenheiros e cientistas envolvidos no campo passaram a priorizar a construção de ‘máquinas pensantes’, um problema fundamentalmente tecnológico, utilizando a abordagem proposta por Shannon.

Portanto fica evidente o processo colaborativo de construção de teorias na ciência, o conceito de informação proposto por Shannon foi inspirado diretamente pelos trabalhos de Nyquist e Hartley e sedimentado na prolongada controvérsia com Wiener e dezenas de outros pesquisadores nas conferências da Fundação Macy. O processo de construção do conhecimento não é linear nem representa os avanços em busca do consenso universal; ele pressupõe a discordância, é por meio dela que o conhecimento evolui.

Agradecimentos

Agradecemos ao professor Dr. Diogo de Oliveira Soares Pinto pelas conversas, ideias e comentários que impulsionaram este trabalho e pelo CNPq pela bolsa de iniciação científica (#111039/2020-7) e pela de bolsa de produtividade em pesquisa (#312748/2018-3).

Referências

- [1] M. Waldrop, *Claude Shannon: reluctant father of the digital age*, disponível em: <https://www.technologyreview.com/2001/07/01/235669/%20claudeshannonreluctantfatherofthedigitalage>, acessado em 22/03/2021.
- [2] <https://news.mit.edu/2001/%20shannon>, acessado em 22/03/2021.
- [3] G. Johnson, *Claude Shannon, mathematician, dies at 84*, disponível em: <https://www.nytimes.com/2001/02/27/nyregion/claudeshannon-mathematician-dies-at-84.html>, acessado em 25/11/2021.

- [4] C.E. Shannon, *The Bell System Technical Journal* **27**, 3 (1948).
- [5] H. Kragh, *An Introduction to the Historiography of Science* (Cambridge University Press, Cambridge, 1987).
- [6] E. Harrison, *Nature* **329**, 6136 (1987).
- [7] B. King e R. Biggs, *Spearhead of logistics: a history of the united states army transportation corps* (Department of the Army, Washington D.C., 2016).
- [8] J. Gleick e A. Calil, *A Informação: uma História, uma Teoria, uma Enxurrada* (Companhia das Letras, São Paulo, 2013).
- [9] H. Nyquist, *Transactions of the American Institute of Electrical Engineers* **XLIII**, 1 (1924).
- [10] A.C. Crehore e G.O. Squier, em: *Transactions of the American Institute of Electrical Engineers* (Philadelphia, 1900).
- [11] G.O. Squier, *Journal of the Franklin Institute* **195**, 5 (1923).
- [12] G.O. Squier, *Proceedings of the Physical Society of London* **27**, 540 (1915).
- [13] J.R. Carson, em: *Proceedings of the American Institute of Electrical Engineers* (New York, 1919).
- [14] J.R. Carson, *Proceedings of the Institute of Radio Engineers* **10**, 1 (1922).
- [15] H. Nyquist, em: *Transactions of the American Institute of Electrical Engineers* (New York, 1928).
- [16] R.V. Hartley, em: *The Bell System Technical Journal*, editado por Nokia Bell Labs (Bell System Technical Journal, Como, 1928), p. 535.
- [17] E.C. Cherry, *Proceedings of the IEE* **98**, 55 (1951).
- [18] C. Bissell, *IEEE control systems magazine* **26**, 3 (2006).
- [19] L. Szilard, *Behavioral Science* **9**, 4 (1964).
- [20] E. Lindberg, em: *IEEE Workshop on Nonlinear Dynamics of Electronic Systems* (Dresden, 2010).
- [21] V.A. Kotelnikov, *Physics-Uspekhi* **49**, 7 (2006).
- [22] N.J.A. Sloane e A.D. Wyner, *Claude E. Shannon: collected papers* (Wiley IEEE Press, New York, 1993).
- [23] R.D.A. Martins, *Revista Brasileira de Ensino de Física* **27**, 11 (2008).
- [24] M.E. Guizzo, *The Essential Message: Claude Shannon and the Making of Information Theory*. Dissertação de Mestrado, Massachusetts Institute of Technology (2003).
- [25] H. Leff e A.F. Rex, *Maxwell's demon 2 entropy, classical and quantum information, computing* (CRC Press, Florida, 2002).
- [26] D. Gabor, *Journal of the Institution of Electrical Engineers-part III: radio and communication engineering* **93**, 26 (1946).
- [27] W.G. Tuller, *Proceedings of the IRE* **37**, 5 (1949).
- [28] N. Wiener, *Cybernetics: or Control and Communication in the Animal and the Machine* (MIT press, Cambridge, 1948).
- [29] C.E. Shannon, *Proceedings of the IRE* **37**, 1 (1949).
- [30] C.E. Shannon e W. Weaver, *The Mathematical Theory of Communication* (The University of Illinois Press, Illinois, 1949).
- [31] N. Wiener, *The Extrapolation, Interpolation and Smoothing of Stationary Time Series* (National Defense Research Council, Washington, 1942).
- [32] H. Von Foerster, *Cybernetics; Circular Causal and Feedback Mechanisms in Biological and Social Systems* (Josiah Macy, Jr. Foundation, New York, 1952).
- [33] N. Wiener, *Cybernetics: Control and Communication in the Animal and the Machine* (MIT press, Cambridge, Massachusetts, 1961), 2^o ed.
- [34] J. Doob, *The Bell System Technical Journal* **27**, 1 (1948).
- [35] B. Mcmillan, *The Annals of mathematical statistics* **24**, 2 (1953).
- [36] J. Pierce, *IEEE Transactions on Information Theory* **19**, 1 (1973).
- [37] W. Weaver, *Scientific American* **181**, 1 (1949).
- [38] C.E. Shannon, *Scientific American* **2**, 1 (1956).