JOSÉ OCTÁVIO DE CARVALHO PINEDA

A ENTROPIA SEGUNDO CLAUDE SHANNON: O DESENVOLVIMENTO DO CONCEITO FUNDAMENTAL DA TEORIA DA INFORMAÇÃO

São Paulo – 2006

JOSÉ OCTÁVIO DE CARVALHO PINEDA

A ENTROPIA SEGUNDO CLAUDE SHANNON: O DESENVOLVIMENTO DO CONCEITO FUNDAMENTAL DA TEORIA DA INFORMAÇÃO

Dissertação apresentada à Banca Examinadora da Pontifícia Universidade Católica de São Paulo, como exigência parcial para obtenção do título de mestre em História da Ciência sob a orientação do Prof. Dr. José Luiz Goldfarb.

Pontifícia Universidade Católica
São Paulo – 2006

Aprovação da Banca Examinadora:

____

____

____

A Beatriz.

Agradecimentos

Ao Prof. Dr. José Luiz Goldfarb, pela orientação e incentivo;

ao Prof. Dr. Luiz Carlos Petry, pelas referências valiosas;

aos Professores do Programa, pela iniciação;

à ERM Brasil, seus colaboradores e gestores, pelas oportunidades;

a Terezinha e Christian, pela leitura atenta e paciente;

a minha família, pelo apoio e compreensão.

Resumo

Esta dissertação tem por objetivo investigar as origens do conceito de Entropia formulado por Claude Shannon no desenvolvimento da Teoria da Informação, bem como as influências que este e outros conceitos da mesma teoria tiveram em outras ciências, em especial a Física.

Partindo de sua origem na Mecânica Estatística, o conceito de Entropia foi transformado numa medida de quantidade de informação por Shannon. Desde então, a abordagem proposta pela Teoria da Informação influenciou outras áreas do conhecimento, e ocorreram várias tentativas de integrá-la às teorias físicas. A análise das obras dos principais formuladores da Teoria da Informação, colocadas em seu contexto histórico, aliada à análise das propostas de integração desta teoria com a Física permitirá demonstrar que a interação atual entre as áreas ainda se dá ao nível de abordagem dos problemas físicos, e não numa forma mais fundamental como era a expectativa de alguns cientistas.

Palavras-chave: Entropia, Teoria da Informação, Mecânica Estatística, Física, Ciência do Século XX, Segunda Guerra Mundial, Telecomunicações, Criptografia, Claude Shannon.

Abstract

This dissertation's objective is to investigate the origins of the concept of Entropy as defined by Claude Shannon in the development of the Information Theory, as well as the influences that this concept and other ones from the same theory had over other sciences, especially in Physics.

Starting from its origin in Mechanical Statistics, the concept of entropy was transformed into a measure of amount of information by Shannon. Since then the approach proposed by Information Theory has influenced other areas of knowledge and there were many attempts of integrating it with physical theories. The analysis on Information Theory main authors' works viewed under a historical outlook, added to the analysis of proposals for its integration with Physics will allow to demonstrate that the integration is currently at the level of approach to physical problems and not at a more fundamental level as it was some scientists expectation.

Keywords: Entropy, Information Theory, Statistical Mechanics, Physics, 20th Century Science, World War II, Telecommunications, Cryptography, Claude Shannon.

Sumário

Ilustrações

Introdução

As tecnologias digitais permeiam nosso cotidiano, estando presentes de formas ora evidentes, ora sutis. Convivemos com suas bênçãos e maldições, freqüentemente ignorando os fundamentos científicos que explicam e viabilizam o funcionamento de todos os dispositivos a nos rodear e servir. A *Teoria da Informação*, desenvolvida a partir da década de 1940, é um dos pilares da assim chamada Era Digital. A compreensão de seus conceitos favoreceu o desenvolvimento de aplicações jamais pensadas até então, e elevando ao nível de ciência a atividade de engenheiros e técnicos daquela época. Ainda, demonstrou formalmente como "quantizar" informação obtida de fontes contínuas, fornecendo as bases teóricas para a convergência digital, que no início do século XXI começa a se materializar. A Teoria da Informação também é importante pela influência que tem causado no pensamento científico desde então.

Qualquer atividade científica que passa a ter ampla aceitação corre os riscos da notoriedade. Suas proposições podem ser simplificadas em excesso para consumo popular, seus pensadores mais influentes podem ser elevados ao status de "precursores" ou "pais" da ciência, e suas conclusões podem ser generalizadas de maneira imprópria noutras áreas do conhecimento. Talvez o exemplo acabado deste fenômeno seja a Mecânica Quântica: atualmente ela é invocada para explicar não apenas fenômenos subatômicos, mas por que a lua está no lugar onde esperamos que ela esteja, e também por que o universo conspira a nosso favor na literatura dita "auto-ajuda". Outros exemplos, porém limitados à esfera acadêmica, nos são dados por Sokal em *Imposturas Intelectuais*, dentre os quais destaca o uso indevido da Matemática, da Teoria do Caos e da Prova de Gödel por pensadores de grande reputação.

Num grau menor que a Mecânica Quântica, é o que acontece com a Teoria da Informação: inicialmente destinada a resolver problemas de engenharia de telecomunicação, ganhou aceitação quase que imediata das comunidades tecnológica e acadêmica onde militava Claude Shannon, seu autor mais influente. Desde então, parece ter se tornado um modismo científico, tendo influenciado outras áreas como a Física, Química, Biologia e Economia. A situação se propagou de tal forma que provocou a intervenção de Shannon, poucos anos após a publicação de sua obra que causou tamanho furor, a *Teoria Matemática da Comunicação*. Num editorial publicado no principal periódico de engenharia elétrica da época, Shannon declarou-se surpreso com as repercussões da Teoria, que estaria gerando contribuições significativas em outras áreas, mas alertou para o uso inconseqüente dos conceitos e descobertas que poderiam advir do modismo ("bandwagon"). A advertência, como se sabe, não foi completamente eficaz, o que parece ter contribuído para o afastamento de Shannon da Teoria. Obviamente, a interação da Teoria da Informação com outros saberes não é nem imprópria nem indesejável. De fato, tal interação tem produzido trabalhos interdisciplinares interessantes, inclusive no Brasil[1].

Talvez a Teoria da Informação trilhe o caminho da popularização entre o público leigo: bastaria a adesão de um punhado de celebridades científicas e algum apoio dos veículos de comunicação (através do jornalismo científico e dos periódicos de divulgação científica). Vemos sinais desta tendência ao encontrarmos artigos de cientistas da estatura de Stephen Hawking e Leonard Susskind sobre a possibilidade de um buraco negro consumir informação. De

---

[1] Isaac Epstein, (Universidade Metodista de São Paulo) e Norbert Fenzl (Universidade Federal do Pará) produziram obras interessantes que serão citadas no devido tempo.

qualquer forma, a Informação pode vir a ser um novo paradigma da ciência, uma vez que pensadores como Mário Schenberg[2] reconhecem sua importância na Física e antecipam uma possível transformação.

A Teoria da Informação e sua história são relativamente desconhecidas no Brasil. Seja pelos interesses das elites dirigentes em formar mão-de-obra tecnicamente capacitada ao invés de pensadores e cientistas, ou pelas deficiências estruturais da Universidade, ou quaisquer outros motivos (os motivos são importantes, mas não serão abordados neste trabalho), o fato é que a Teoria da Informação e seus principais autores são desconhecidos pelos graduados em Ciências da Computação e pela elite profissional de Tecnologia da Informação. Transcender o caráter instrumental do saber nesta e em qualquer área será sempre oportuno e desejável, a menos que se pretenda dar continuidade à formação de técnicos e tecnólogos que por vezes parece ser a prioridade da formação universitária nos dias de hoje.

Numa obra de História da Ciência é natural que as fronteiras entre historiografia e ciência sejam elásticas. Como ensina Roberto Martins[3], entre os diversos níveis discursivos que uma obra de História da Ciência aceita, o conteúdo científico em si é válido e relevante. Indo um pouco além, este trabalho incluirá alguns conteúdos da ciência em questão, demonstrando que algumas previsões da Teoria da Informação são facilmente verificáveis, e para elucidar alguns aspectos controvertidos da teoria, especialmente da aparente tensão que existe entre os pensadores da Comunicação e os Matemáticos a

---

2 Schenberg expressa esta opinião em *Pensando a Física*.

3 Martins discorre sobre o tema em "Ciência *versus* historiografia: os diferentes níveis discursivos nas obras sobre história da ciência", *in* A. M. Alfonso-Goldfarb & M. H. R. Beltran, orgs., *Escrevendo a história da ciência: tendências, propostas e discussões historiográficas*, pp. 115-45.

respeito da irrelevância do conteúdo na determinação da quantidade de informação de uma mensagem. Estas idéias serão apresentadas já no capítulo 1, que tratará de alguns aspectos relevantes da vida, da personalidade e dos trabalhos de Shannon; dos fundamentos da Teoria da Informação e suas aplicações; e da origem e significado do conceito de Entropia utilizado para designar a quantidade de informação de uma mensagem, que conta com versões contraditórias atribuídas ao próprio Shannon.

O capítulo 2 se iniciará com a história da primeira metade do século XX, com ênfase aos desenvolvimentos científicos e tecnológicos que concorreram para a elaboração da Teoria da Informação, colocados em função do contexto político e econômico. A ênfase será no desenvolvimento do rádio e sua aplicações militares, as quais empregavam maciçamente técnicas de criptografia e criptologia, que têm muitos pontos em comum com a Teoria da Informação. Alan Turing, o cientista cuja contribuição individual à criptologia é reconhecida como mais relevante, teve uma breve interação com Shannon durante o curso da Segunda Guerra Mundial, quando descobriram as similaridades de suas idéias e trabalhos. A seguir, tratará em maior profundidade a gênese do conceito de Entropia empregado na Teoria da Informação: a Teoria dos Gases, de Boltzmann, que define Entropia como uma medida estatística de desordem molecular, e a Entropia na Teoria Matemática da Comunicação, de Shannon. O conceito de Informação na Cibernética de Wiener também será considerado, apesar de estar relacionado à influência de Szilard, o que leva a posições irreconciliáveis com a Teoria da Informação conforme proposta por Shannon. Ainda neste capítulo, serão abordadas as tentativas de generalização dos conceitos da Teoria da Informação na Física do século XX, em especial por

Brillouin (que se baseia em Szilard e Wiener), e algumas de suas contradições, dificuldades e críticas. Também se abordará a liberdade com que se usam os conceitos da Teoria da Informação na Física contemporânea nos trabalhos de Hawking, Preskill e Susskind, que parece fornecer uma nova alternativa de abordagem de certas questões ao invés de uma estrutura formal consolidada.

Encerrando este trabalho, o capítulo 3 tratará de novas perspectivas da Teoria da Informação, iniciadas com o trabalho de integração de Epstein para, em seguida, indicar possíveis antecedentes conceituais na Economia baseados em alguns conceitos formulados por Adam Smith e Alfred Marshall. A seguir, partindo da proposição de Vilfredo Pareto de que a Economia é regida pelos fenômenos psicológicos, serão analisados alguns aspectos da Neuropsicologia de Alexander Luria, cujas relações com alguns fenômenos estudados pela Teoria da Informação aparentemente não foram identificadas até hoje. Por fim, proporá uma abordagem mais ampla para o conceito de Entropia, antes da conclusão em si, que resumirá as principais idéias apresentadas ao longo da dissertação.

As obras primárias desta dissertação são: *The Mathematical Theory of Communication*, de Claude Shannon; *Lectures on Gas Theory*, de Ludwig Boltzmann; *Cybernetics*, de Norbert Wiener; *Science and Information Theory*, de Leon Brillouin. Utiliza-se o sistema de notas numeradas para as referências (conforme definido no Manual de Estilo de Chicago), que é costumeiramente usado em Humanidades e Ciências Sociais.

1. Fundamentos da Teoria da Informação

1.1. Aspectos Biográficos e do Pensamento de Shannon

Muito se tem escrito a respeito da vida e da obra de Claude Shannon. Como geralmente acontece, mais destaque é dado aos aspectos pitorescos de seu comportamento de que a suas idéias e motivações, e suas obras são mais celebradas do que entendidas[4]. Os aspectos de sua imagem pública que aderem ao estereótipo do cientista excêntrico serão contextualizados em função de seus pensamentos e interesses. Ao abordar aspectos de sua biografia, pretende-se destacar os aspectos relevantes de suas idéias e motivações, o que possivelmente contribuirá para uma melhor compreensão de sua obra e sua trajetória acadêmica. Como Shannon não registrou em suas obras suas idéias filosóficas, a inferência a partir de seus dados biográficos e das poucas entrevistas que concedeu parece constituir a única forma de penetrar em suas convicções.

Claude Elwood Shannon nasceu em 30 de abril de 1916 em Petoskey, no estado de Michigan, Estados Unidos, e faleceu em 24 de fevereiro de 2001 de causas naturais. Passou seus primeiros 16 anos na pequena cidade de Gaylord, Michigan, e freqüentou a escola pública onde sua mãe foi professora de línguas e diretora. Seu interesse por objetos mecânicos se manifestou desde a infância, e foram

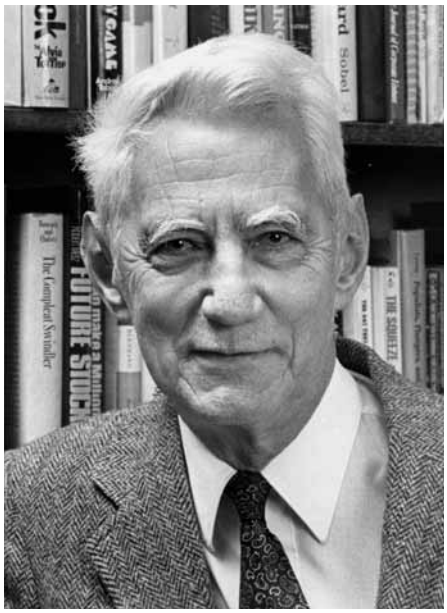

Fig. 1 - Claude E. Shannon, 1995.

[4] Este fenômeno pode ser atribuído à popularidade de algumas de suas obras, em especial a *Teoria Matemática a Comunicação*. De fato, a Teoria da Informação teve grande repercussão na mídia em geral, chegando a ser considerada mais importante que a Teoria da Relatividade na edição de Dezembro de 1953 da revista *Fortune*. Para detalhes, consultar: O. Aftab *et alii*, "Information Theory after Shannon's 1948 Work", pp. 6-7.

inúmeras as engenhocas que construiu para sua própria diversão, como um barco controlado por rádio e um telégrafo entre sua casa e a de um vizinho usando um par de fios de arame farpado que rodeava um pasto. Tinha gosto por jogos mentais, como a resolução de criptogramas, adquirido pela leitura de "O escaravelho dourado", de Edgar Allan Poe, e similares. Ele admirava o inventor Edison, e os cientistas Newton, Darwin, Einstein e Von Neumann[5]. Na universidade, obteve dupla graduação em Matemática e Engenharia, e tornando-se Mestre em Engenharia e Doutor em Matemática[6]. Profissional da pesquisa científica e tecnológica, ele trabalhou como pesquisador nos Laboratórios Bell durante cerca de 15 anos a partir da década de 1940, até se dedicar totalmente a sua atividade de docente no MIT. Sobre os tempos nos Laboratórios Bell, recorda sua satisfação com as condições de trabalho que lhe eram oferecidas:

> *Trabalhei nos Laboratórios Bell por quinze anos, e depois disso tornei-me consultor deles. Eles me davam uma grande liberdade. Para começar, podia-se trabalhar com o que se quisesse, em sua próprias idéias; eles não vinham e diziam "trabalhe nisto!" Pelo menos, não para mim*[7].

Entretanto, em entrevista concedida a Robert Price, Shannon se contradiz a respeito da pretensa liberdade que gozava nos Laboratórios Bell. Perguntado sobre sua *Teoria da Comunicação dos Sistemas de Sigilo*, um trabalho científico sobe criptografia com evidentes finalidades militares (e que contém muito do conteúdo da posterior *Teoria Matemática da Comunicação*), ele afirma:

---

5 [s.a.], "Biography of Claude Elwood Shannon", *in* C. E. Shannon, *Collected Papers*, p. xi.

6 *Ibid.*, pp. xi-xii;

7 A. Liversidge, "Profile of Claude Shannon", *in* C. E. Shannon, *Collected Papers*, pp. xxiii. Citação no original: "I worked at Bell Labs for fifteen years, and after that I was a consultant there. They gave you great freedom. To begin with you could work on what you wanted, your own ideas; they didn't come and say, "work on this!" At least, not to me".

> *Minha primeira motivação foi a teoria da informação, e eu usei a criptografia como um meio de legitimar o trabalho*[8].

Em entrevista concedida a Anthony Liversidge, Shannon revela seus interesses e motivações. Segundo ele mesmo, desde menino ele apreciava jogos, não só os que exigiam destreza física, mas também os que exigiam raciocínio: ele possuía estes atributos em graus elevados. Isto explica o gosto pelo malabarismo e pelas invenções "inúteis" (segundo seu próprio critério[9]) que desenvolveu durante toda a vida. Shannon afirma ser movido pela curiosidade e pelo prazer de resolver problemas e aumentar seus conhecimentos; prêmios e retorno financeiro nunca foram suas intenções. Ele admite que sempre se envolveu com atividades que considerava "divertidas", o que inclui sua dissertação de mestrado em engenharia elétrica (*A Symbolic Analysis of Relay and Switching Circuits*), no qual estabelece a relação entre o mecanismo de chaveamento elétrico de circuitos (usado nos computadores de então) e a álgebra de Boole.

> *Resolver aquilo [relacionar a topologia dos circuitos de chaveamento com a lógica booleana] foi muito divertido. Acho que me diverti mais fazendo isso do que em qualquer outra coisa em minha vida, falando em termos criativos*[10].

É importante salientar que sua dissertação é considerada um marco na Ciência da Computação por ter provido os fundamentos necessários para a construção de computadores cada vez mais sofisticados[11]. Até então, as máquinas de cálculo mais poderosas eram analógicas, como o Analisador

---

8 R. Price, "A conversation with Claude Shannon: one man's approach to problem solving", p. 124. Citação no original: "My first getting at that was information theory, and I used cryptography as a way of ligitimizing the work."

9 *Ibid.*, pp. xxiii.

10 *Ibid.*, pp. xxii-xxvi. Citação no original: "Working that out was a lot of fun. I think I had more fun doing that than anything else in my life, creatively speaking".

11 Esta opinião é compartilhada por alguns especialistas. Para mais detalhes, consultar: M. Waldrop, *Reluctant Father of the Digital Age,* pp. 66-7; consultar também: UCSD-TV, *Claude Shannon, Father of the Information Age,* depoimento de Andrew Viterbi.

Diferencial de Vannevar Bush, de quem Shannon foi assistente antes de obter seu grau de Mestre em Engenharia Elétrica. Entretanto, os dispositivos analógicos eram projetados e resolver problemas determinados (como resolver parcial ou totalmente equações diferenciais) e tinham limitações de precisão, pois se valiam de medições de propriedades físicas macroscópicas. A transformação dos processos de cálculo da forma analógica para a digital (contável) através de formulação algébrica possibilitou fazer do computador uma máquina programável, ao mesmo tempo em que aumentava a precisão dos resultados[12].

Na mesma entrevista, Shannon rejeita as tentativas de transformar a Teoria da Informação numa espécie de crença religiosa[13], e declara suas verdadeiras crenças. Ele se declara ateu, e afirma crer na teoria de Darwin e na concepção de que os seres vivos nada mais são do que máquinas complexas. Quanto às máquinas construídas pelo homem, ele pensa que um dia elas poderão ser aperfeiçoadas a ponto de superar os seres humanos em qualquer atividade física ou intelectual[14]. Numa pergunta anterior, ele já havia declarado sua crença na segunda lei da termodinâmica, no Big-Bang e no aumento da entropia em longo prazo (mesmo que, de forma localizada, ela possa decrescer por algum tempo)[15]. É interessante observar os pontos de contato de suas crenças pessoais com o pensamento positivista em voga na época, e sua total

---

12 Para uma análise contextualizada sobre a longa transição da computação analógica para a digital, consultar: P. Edwards, *The Closed World: computers and the politics of disclosure in Cold War America*, pp. 66-70.

13 A este respeito, existem relatos de que na década de 1950 a Teoria da Informação era levada tão a sério a ponto de ser quase que como uma religião no M.I.T e em outras universidades norte-americanas. Os adeptos da TI julgavam estar vivenciando um momento de transformação no qual diversos problemas científicos seriam resolvidos pela nova abordagem oferecida pela Teoria da Informação. Para mais detalhes, consultar J. Gleick, *Caos, a Criação de Uma Nova Ciência*, pp. 245-7.

14 A. Liversidge, *in* C. E. Shannon, *Collected Papers*, p. xxx.

15 *Ibid.*, p. xxviii.

concordância com os princípios desenvolvidos por Boltzmann a respeito da entropia e o universo[16].

Os aspectos materialistas de seu pensamento podem ajudar a explicar seu afastamento em relação à Teoria da Informação. Para alguns, o afastamento teria se dado pelo descontentamento com os rumos que a Teoria havia tomado: a distorção de seus conceitos quando aplicados noutras disciplinas o incomodava[17]. Soma-se a isto a aversão a linhas de pensamento místico-religiosas, que era a reação que o conhecimento da Teoria da Informação provocava em muitos. O resultado pode ter causado em Shannon o desejo de não mais participar de tais derivações e digressões realizadas a partir de suas idéias. Sobre este assunto, respondendo a Liversidge se seu afastamento era decorrente da exaustão da teoria[18], ele responde:

> *...eu apenas desenvolvi interesses diferentes*[19].

Entretanto, os motivos declarados por Shannon podem não ser os únicos. Parecem haver indícios de que estava em curso uma disputa dentro da sociedade de estudos da TMC, o PGIT[20]. O primeiro grupo era formado pelos "puristas", que defendiam o uso da TMC em assuntos tecnológicos, enquanto que no segundo grupo estavam os pesquisadores que defendiam as extensões da teoria. Entre os "puristas" estavam Shannon, Wiener e Fano, que eram os pesquisadores mais influentes do PGIT, e sendo que os dois primeiros eram tidos como os criadores da área. O descontentamento dos "puristas" parece não

---

16 O pensamento de Boltzmann a este respeito será abordado mais adiante neste capítulo, e com maior profundidade no capítulo 2.

17 M. Waldrop, *op. cit.*, p. 71.

18 A. Liversidge, *in* C. E. Shannon, *Collected Papers*, p. xxviii. Segundo o entrevistador, Marvin Minsky teria dito que todos os teoremas importantes da Teoria da Informação já tinham sido provados.

19 *Ibid.*, p.71. Citação no original: "...I just developed different interests."

20 O Grupo Profissional da Teoria da Informação (PGIT) foi criado pelo Instituto de Engenheiros de Rádio (IRE) para estimular a difusão, o debate e o desenvolvimento da Teoria da Informação.

estar apenas relacionado ao uso inadequado da TMC, mas também pela manutenção da "seriedade" da área, que contava com maciço suporte financeiro dos militares norte-americanos: alguns pesquisadores de outras áreas estariam modificando seus projetos para tentar inserí-los na Teoria da Informação com o propósito de obter verbas de pesquisa mais facilmente. O posicionamento de Shannon, expresso no editorial *The Bandwagon*[21] serviu para coibir os abusos, mas pode ter abortado muitas tentativas sérias de aplicação da TMC em outras áreas, uma vez que o próprio foro de debates passou a ser controlado pelo grupo dominante[22].

Debates sobre a Teoria da Informação e suas extensões também ocorriam nas conferências sobre cibernética patrocinadas pela Fundação Macy. As "Confererências Macy" foram em número de dez, e ocorreram de 1946 a 1953 sob a coordenação de Norbert Wiener e John von Neumann, que reuniam um seleto grupo multidisciplinar de pesquisadores para as apresentações e discussões. Os coordenadores estavam interessados em explorar as analogias entre seres vivos e máquinas para, principalmente, possíveis aplicações militares (como, por exemplo, a criação de sistemas de orientação para mísseis teleguiados)[23]. Shannon participou algumas vezes como convidado, e numa de suas apresentações ocorreu um intenso debate a respeito de um tipo particular de extensão: a aplicação da TMC no nível semântico da comunicação[24]. Noutra

---

21 O termo "bandwagon" é empregado no sentido de "modismo" (ou, para usar um termo mais antigo e direto, de "coqueluche"), ou seja, uma atividade à qual indivíduos se lançam simplesmente porque outros também o fazem. O editorial, publicado no *IRE IT Newsletter* de Dezembro de 1953, pedia que os assuntos fossem abordados com mais rigor matemático e preferencialmente em aplicações de engenharia, embora reconhecesse que a Teoria da Informação pudesse contribuir noutras áreas de conhecimento. O estilo do texto é elegante e conciliatório, mas deixa clara a posição de Shannon em favor da limitação da abrangência da teoria.

22 O. Aftab *et alii*, "Information Theory after Shannon's 1948 Work", pp. 8-12.

23 P. Edwards, *op. cit.*, p. 189.

24 *Ibid.*, pp. 203-4.

conferência, Shannon apresentou Theseus, um “camundongo eletrônico” que conseguia encontrar a saída de um labirinto. Foi uma das poucas vezes em que uma máquina em funcionamento foi apresentada nestas conferências[25]. Com relação ao Theseus e suas engenhocas em geral, é interessante notar que Shannon parece subestimar a importância delas em seu trabalho, referindo-se a elas como divertidas ou inúteis. No entanto, algumas delas constituíram demonstrações da viabilidade de suas idéias (“proof of concept”), como é caso do Theseus e da máquina que jogava xadrez. Se Shannon age assim por convicção, por modéstia ou por qualquer outro motivo é uma questão para futuras investigações.

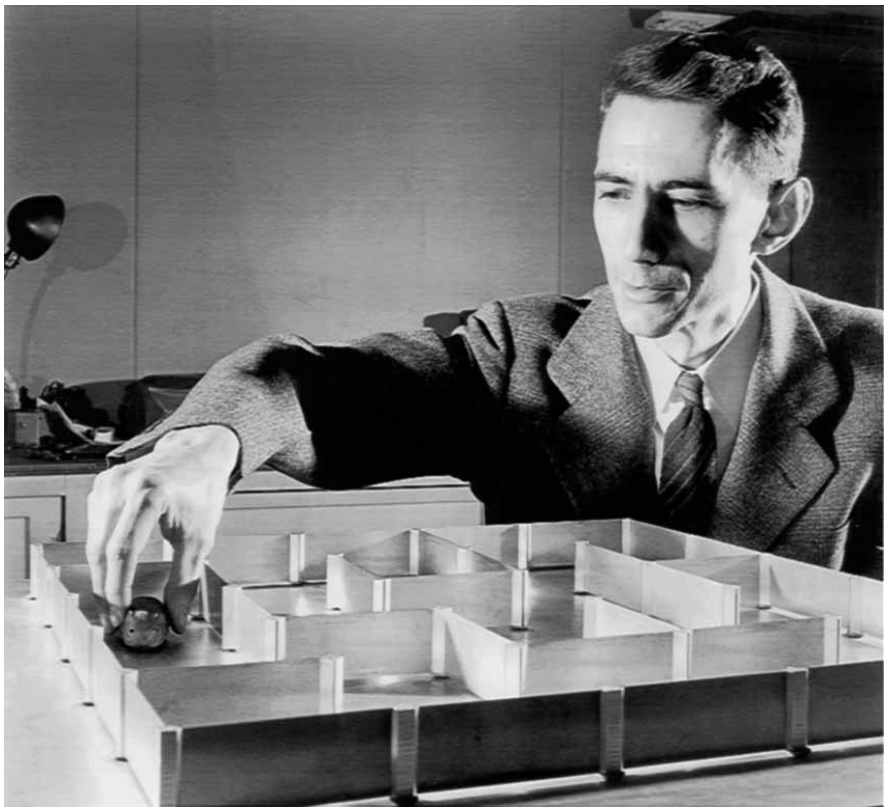

Fig. 2 - Claude Shannon e Theseus, década de 1950.

[25] *Ibid.*, p. 191.

É interessante mencionar a breve interação Shannon com Alan Turing durante a Segunda Guerra (que será abordada no capítulo 2), bem como sua participação no *Project X*, que foi um projeto secreto conduzido pelos Laboratórios Bell para a criptografia da voz humana. Até então, eram utilizados métodos analógicos de distorção de ondas sonoras, o que era pouco eficaz por requerer apenas um ouvinte atento e paciente para entender o que se transmitia. O Sistema X, ao contrário, trabalhava com a quantização das ondas sonoras, sendo que ao resultado deste processo seriam adicionados valores arbitrários determinados por uma "chave", que deveria ser de conhecimento apenas do transmissor e do receptor da mensagem[26]. Este método foi proposto e descrito matematicamente por Shannon em sua *Teoria da Comunicação dos Sistemas de Sigilo* (finalizada em 1945, mas mantida secreta até 1948), e posteriormente na própria TMC sem o componente criptográfico. O desenvolvimento do Sistema X propiciou a colaboração de Shannon no trabalho de lingüística de Chomsky:

> *Shannon posteriormente trabalhou com Noam Chomsky, para quem a Lingüística se tornou o estudo do humano como um Sistema X, executando transformações complexas sobre "núcleos" de informação, decodificados por transformação reversa pelo ouvinte*[27].

Embora não tenhamos detalhes sobre este trabalho conjunto, é certo que Shannon exerceu alguma influência no pensamento de Chomsky, que chegou a publicar um artigo no *IEEE Transactions on Information Theory*[28]. Neste

---

26 *Ibid.*, pp. 199-201.

27 *Ibid.*, p. 202. Citação no original: "Shannon later worked with Noam Chomsky, for whom linguistics became the study of the human as X system, performing complex transformations on "kernels" of information, decoded by reverse transformations by the listener."

28 O *IEEE Transactions on Information* Theory foi criado após a fusão do IRE (Instituto dos Engenheiros de Rádio) e do AIEE (Instituto Norte-Americano de Engenheiros Elétricos), gerando o IEEE (Instituto de Engenheiros Elétricos e Eletrônicos). Para mais detalhes, consultar: http://www.ieee.org/portal/pages/about/whatis/index.html.

trabalho, que foi previamente apresentado num simpósio de cibernética em 1956, Chomsky questiona algumas premissas de Shannon na TMC no que diz respeito a linguagens naturais[29], ao mesmo tempo em que aplica conceitos derivados do Sistema X, no que talvez possa ser considerada uma versão preliminar de sua Gramática Gerativa Transformacional[30].

Concluindo: embora Shannon tenha declarado que tinha total liberdade para escolher o tema de seu trabalho nos Laboratórios Bell, a escolha de trabalhar com sistemas de sigilo talvez não tenha sido simples coincidência entre suas vontades e os interesses de seu empregador (determinados pelo contexto da Segunda Guerra Mundial), caso contrário não lhe teria sido preciso mudar a abordagem de seu trabalho em teoria da comunicação para uma teoria sobre criptografia. Parece lícito questionar se a "liberdade" não seria melhor entendida como a possibilidade de trabalhar em qualquer um dos projetos em curso nos Laboratórios Bell, somada ao ambiente informal e à abundância de recursos tecnológicos que eram oferecidas a seus pesquisadores mais importantes.

Sobre o temperamento e a personalidade de Shannon, é possível supor que sua criatividade parece estar bastante ligada à busca de divertimento e à percepção de que muito de seu trabalho era diversão. Seu interesse pelo lúdico e seu temperamento jovial coexistiam com aspectos mais "sérios" de sua

---

29 Segundo observa Verdú, Chomsky discorda que as estatísticas das linguagens naturais podem ser aproximadas tão bem quanto se queira ao modelo das Cadeias de Markov. Para mais detalhes sobre este comentário, consultar: S. Verdú, "Fifty Years of Shannon Theory", *in* S. Verdú & S. McLaughlin, orgs., *Information Theory – 50 Years of Discovery*, pp. 26-7. Para mais detalhes sobre a argumentação e as propostas de Chomsky em questão, consultar: N. Chomsky, "Three models for the description of language", pp. 113-24.

30 P. Edwards, *op. cit.*, p. 229. Miller sugere que Chomsky tenha se aproveitado da notoriedade da Teoria da Informação para divulgar sua Gramática Gerativa Transformacional. Para detalhes sobre esta opinião no contexto da história das Ciências Cognitivas, consultar: G. Miller, "The cognitive revolution: a historical perspective", p. 142.

personalidade, que se manifestam no formalismo de seus trabalhos, na dedicação com que se empenhava em cada projeto, e na tenacidade com que enfrentava cada desafio intelectual. Todos estes aspectos, exercidos em graus elevados, somados a uma formação acadêmica sólida e à oportunidade de trabalhar com ampla liberdade, parecem ter sido decisivos para o pleno exercício de suas potencialidades[31]. Não menos importante é notar a rede de relações e de influências mútuas estabelecidas entre Shannon e os mais destacados cientistas de sua época, em especial com Wiener, von Neumann e Turing: cada um deles tinha objetivos específicos, mas todos atuavam na solução de problemas militares, em conformidade com o contexto histórico no qual estavam inseridos.

[31] As considerações sobre características da personalidade de Shannon são análises de senso comum baseadas nos dados disponíveis, e não devem ser tomadas como uma avaliação psicológica profissional.

1.2. Origem do Conceito de Entropia usado por Shannon

O conceito físico de Entropia está relacionado tanto a um estado como a uma tendência: no primeiro caso, ao grau de desorganização da matéria; no segundo, à tendência de desorganização de toda matéria. A Termodinâmica afirma que a entropia nunca diminui num sistema fechado, ou seja, seu grau de desorganização pode aumentar, mas jamais diminuir. Disto decorre que a entropia é um estado dinâmico que varia em função do estado inicial de organização da matéria e do tempo, caracterizando um processo irreversível. As inúmeras interpretações e conseqüências físicas e cosmológicas são abrangentes e profundas; entretanto, este trabalho impõe que apenas os aspectos que mais se relacionam com a Teoria da Informação sejam considerados, o que impõe abordar a entropia conforme pensada por Ludwig Boltzmann.

Boltzmann define a entropia em termos estatísticos dentro de um contexto mecânico. Como se sabe, a estatística estuda as propriedades de uma amostragem ou população, ou seja, de um conjunto finito de objetos. De fato, é impossível estudar as propriedades macroscópicas da matéria pelas características individuais de cada molécula. Assim sendo, a entropia que Boltzmann define é uma estatística sobre uma quantidade de matéria, ou seja, um número que descreve as moléculas coletivamente[32]. As dificuldades de se chegar a uma fórmula satisfatória são imensas: 1) as moléculas têm várias propriedades físicas, que são expressas por grandezas fundamentais, como massa e extensão, e derivadas, como velocidade e momento (e, ainda, o momento é uma grandeza vetorial e a posição espacial é expressa por 3

[32] L. Brillouin, *Science and Information Theory*, p. 119.

distâncias); 2) o domínio das grandezas derivadas das moléculas é contínuo, ou seja, cada molécula pode apresentar infinitas velocidades e momentos[33]; 3) a quantidade de moléculas num sistema observável é muito grande para um mapeamento individual[34].

A abordagem de Boltzmann baseou-se numa proposição de Maxwell, conhecida como *Lei da Distribuição de Velocidades*. A descrição do que aconteceria dentro de um recipiente repleto de gás fornece a explicação: as moléculas, partindo de um estado inicial organizado (supondo que fossem introduzidas no recipiente com a mesma velocidade), começariam pouco a pouco exibir variações de suas velocidades individuais em decorrência das colisões entre elas (no modelo, o choque com as paredes do recipiente foi considerado como perfeitamente elástico, o que não mudaria a velocidade da molécula que colidisse contra ela), até que atingisse um estado de desordem máxima[35]. As velocidades que uma molécula pode assumir são infinitas, mesmo que a variação entre a menor e a maior delas seja finita. Boltzmann propõe que o estado mais molecularmente desorganizado seria aquele em que a distribuição das velocidades fosse homogênea. É conveniente notar que é a distribuição das velocidades, e não as velocidades em si, que são homogêneas: velocidades homogenias provocariam, ao contrário, uma distribuição desigual de velocidades.

A proposição de Boltzmann resulta em equações integrais cuja resolução é impraticável (uma alternativa melhor surgiria posteriormente com Max Planck).

---

33 Como explica Penrose, para mapear posições e *momenta* é necessário um espaço de fase de dimensão 6n, onde n é o número de partículas. Para mais detalhes, consultar: R. Penrose, *The road to reality: a complete guide to the laws of the universe*, pp. 217-21, 690-1.

34 *Ibid.*, pp. 119-20.

35 L. Boltzmann, *Lectures on Gas Theory*, p. 36.

Segundo a teoria quântica, átomos e moléculas não se encontram em qualquer estado, mas somente em estados estáveis discretos, sendo que a transição de um estado para outro envolve absorção ou emissão de energia. A contagem destes estados quantizados dá a medida da entropia do sistema[36]. Ainda assim, a determinação precisa da entropia continua impraticável, o que levou ao desenvolvimento de métodos de cálculo aproximados, através da criação de amostras discretas de velocidade, ou seja, criando faixas de velocidades[37]. É interessante notar que Shannon empregou princípios semelhantes na *Teoria Matemática da Comunicação* quando formulou o método de transformação de sinais contínuos em discretos. Ele propõe a divisão do *continuum* de sinais num número grande, mas finito, de pequenas regiões às quais se aplicariam os mesmos cálculos dos sinais contínuos[38].

---

36 L. Brillouin, *op. cit.*, p. 120.

37 O método é conhecido como "coarse graining". Para mais detalhes, consultar: R. Penrose, *op.cit.*, pp. 690-2.

38 C. Shannon, "The Mathematical Theory of Communication", *in* C. Shannon & W. Weaver, *The Mathematical Theory of Communication*, p. 81. Uma análise mais aprofundada será apresentada no capítulo 3.

1.3. Fundamentos de Teoria da Informação

A palavra "informação" tem vários significados, e seu entendimento pode ser tão vago a ponto de se precisar conhecer o contexto em que é empregada para entendê-la. Os principais autores adotam definições diferentes, baseadas nas características que pretendem destacar. Georges Ifrah apresenta 26 diferentes definições de "informação" para tentar esboçar uma nova definição que contemplasse todos os aspectos importantes[39]. A falta de consenso é tal que Devlin propõe em *Logic and Information* que não se dê muita importância à definição de informação, argumentando que tal definição seria prematura e que ela seria possível a seu tempo, à semelhança de outros conceitos científicos cujo entendimento original diferia de maneira significativa do entendimento atual[40].

Segundo Weaver, informação é o grau de liberdade que se tem ao selecionar uma mensagem, mas não a uma mensagem em si, mas considerando-se todo o processo de seleção das possíveis mensagens[41]. Ele, entretanto, adverte que "informação" tem um sentido específico na Teoria, e que não se deve confundir "informação" com "significado" [42], reforçando o que Shannon já havia afirmado:

> *Freqüentemente, 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*

---

39 G. Ifrah, *História universal dos algarismos*, T. 2, pp. 805-14.

40 K. Devlin, *Logic and Information*, pp. 1-2. Para conhecer o pensamento de Devlin sobre "informação", é interessante assistir ao vídeo de sua apresentação "Does information really exist?" no Simpósio de Filosofia da Informação (Instituto de Lógica, Linguagem e Computação da Universidade de Amsterdã, 2004) em: http://live.izi-services.nl/poi_kdevlin.

41 *Ibid.*, p. 9.

42 W. Weaver, "Some Recent Contributions to the Mathematical Theory of Communication", *in* C. Shannon & W. Weaver, *op. cit.*, p. 8. Citação no original: "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".

> *comunicação são irrelevantes para o problema de engenharia. (grifo do original)* [43]

Tratar a informação independente do sentido que ela possa ter é uma redução ao nível básico de qualquer linguagem, uma vez que não será tratada a informação em si, mas sua codificação: isto garante que a informação possa ser quantificada. Quanto ao significado da informação, parece lícito afirmar que ele tenha algum grau de subjetividade: o significado depende de interpretação, o que implica que uma mesma mensagem pode ter significados diferentes, embora a informação transmitida seja a mesma. Por exemplo: na mensagem "a bolsa caiu" a quantidade de informação é sempre a mesma quer ela signifique que "uma valise foi acidentalmente jogada ao chão" ou "houve uma variação negativa no índice da bolsa de valores"[44]. Além das ambigüidades, a interpretação é um processo influenciado pelas características do receptor, como, por exemplo: o conhecimento da língua em que a mensagem está codificada, o conhecimento do tema ao qual a mensagem se refere, o interesse que o receptor tem pelo assunto ou à relevância que ele atribui à mensagem[45]. Neste contexto, os conceitos de "sinal" e "ruído" deixam de ser precisos:

> *Se uma pessoa recebe algo (uma mensagem) por telefone, sendo que ela considera uma parte (da mensagem) útil e outra não, e se alguém quiser chamar a parte útil de sinal, isto dificilmente será um problema matemático*[46].

---

43 *Ibid.*, p. 31. Citação no original: "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".

44 Extrair o significado de uma mensagem é uma tarefa complexa que ainda não tem uma solução adequada: uma conseqüência disto é o desempenho insatisfatório dos *softwares* de tradução de textos, e mesmo dos de correção ortográfica.

45 A respeito das limitações da linguagem escrita, consultar D. Olson, *O mundo no papel: as implicações conceituais e cognitivas da leitura e da escrita*, pp. 107-30.

46 C. Shannon, *apud* P. Edwards, *op. cit.*, p. 204. Citação no original: "if a person receives something over a telephone, part of which is useful to him and part of which is not, and you want to call the useful part the signal, that is hardly a mathematical problem."

Uma definição que parece aderir melhor à essência da Teoria da Informação é a de Isaac Epstein, que sintetiza o pensamento de Shannon de forma objetiva e clara:

> *Informação é uma redução de incerteza, oferecida quando se obtém resposta a uma pergunta*[47].

Também é esclarecedora a explicação oferecida por Fenzl e Hofkirchner, propondo uma de teoria unificada da informação, na qual "informação" se relaciona com a identificação da diversidade:

> *O conceito de* ***informação*** *está intimamente relacionado com a idéia de transformação,* ***emergência da novidade****. (grifos do original)*[48]

Antes de se receber uma mensagem, é suposto que exista um conjunto finito de mensagens possíveis de serem recebidas, e que cada uma destas mensagens tenha probabilidade de ocorrência. Assim sendo, a incerteza quanto à informação a ser recebida diminui quando a mensagem é recebida. O quanto de incerteza que a mensagem recebida reduz está relacionado à probabilidade de sua ocorrência: o recebimento de uma mensagem mais provável tem menos informação que uma mensagem menos provável. Exemplificando: num informe meteorológico que prevê as condições climáticas na cidade de Maceió, "ensolarado" tem menos informação do que "nevasca". Enquanto que o recebimento da primeira mensagem não seja novidade para qualquer um que esteja familiarizado com o clima do Nordeste brasileiro, o recebimento da segunda mensagem será genuinamente uma ocorrência notável.

---

47 I. Epstein, *Teoria da Informação*, p. 35.

48 N. Fenzl & W. Hofkirchner, "Emmergence and Interaction of natural systems: the role of information, energy and matter in the perspective of a Unified Theory of Information", p. 6. Citação no original: "The concept of **information** is closely related to the idea of transformation, **emergence of novelty**."

Para Shannon, a descrição de processo genérico de comunicação inclui agentes, recursos e métodos. Numa análise simplificada, temos: como agentes, o Emissor e o Receptor; como recurso, o canal de comunicação e a codificação; como método, a codificação das mensagens em símbolos. O emissor, valendo-se de um canal de comunicação, envia ao receptor uma mensagem codificada por um processo conhecido por ambos os agentes. O processo de codificação envolve a tradução da informação em símbolos discretos, retirados de um repertório de símbolos previamente acordado[49]. Shannon propõe que a quantidade de informação deve ser entendida como a entropia da mecânica estatística, e observa que a entropia tem características interessantes: à medida que a ocorrência de um grupo de símbolos se torna mais provável que a dos outros sinais do repertório, a entropia decresce. A entropia máxima só é atingida quando a ocorrência de todos os símbolos é equiprovável (ou seja, não existe tendência de concentração de probabilidades em algum grupo de símbolos). Quando existe certeza sobre qual símbolo vai ser transmitido, a entropia é zero[50]. Utilizando os conceitos previamente expostos, é lícito entender que a quantidade de informação decresce quando ocorre uma tendência de concentração das probabilidades de ocorrência de certos símbolos, ou seja, que o conjunto apresenta menor diversidade. No limite, quando não houver diversidade no conjunto, a quantidade de informação será nenhuma.

Shannon observa também que as mensagens usualmente transmitidas são codificadas numa linguagem natural. Analisando a estrutura sintática do idioma

---

49 C. Shannon, "The Mathematical Theory of Communication", *in* C. Shannon & W. Weaver, *op. cit.*,pp. 33-5.

50 *Ibid.*, pp. 50-1.

inglês, ele conclui que existe um alto grau de redundância[51]. *Redundância* é tudo aquilo que é não é fundamental para o entendimento de uma mensagem, e pode ser entendida como uma medida complementar à entropia, ou seja, é a quantidade de entropia que seria necessária para que a mensagem tivesse entropia máxima. Neste sentido, após definir que a entropia de uma mensagem deve considerar as probabilidades de ocorrência de todos os sinais da mensagem, Shannon estabelece o conceito de *entropia relativa*, que é a relação entre a entropia da mensagem e a máxima entropia possível[52]. A necessidade de "fazer sentido" faz com que a língua imponha construções (por exemplo: regências, concordâncias, radicais) que fazem com que algumas palavras e letras sejam mais prováveis do que outras[53]. Obviamente, nem tudo que se pode escrever faz sentido, resultando em que a eqüiprobabilidade dos símbolos seja impossível: a entropia de um texto que tem significado jamais será a maior possível. Curiosamente, a premissa de que o significado nada tem a ver com informação passa a ter nova dimensão após esta constatação: o significado impede que a mensagem tenha informação codificada da forma mais eficiente. Parece ser bastante improvável que uma mensagem com significado possa atingir a entropia máxima, a menos que seja relativamente curta (como um provérbio, um "hai-kai" ou um "limerick") a ponto de não permitir a repetição de sinais: neste caso, a redundância é devida à codificação, não ao significado. Por outro lado, parece ser igualmente improvável que qualquer mensagem com entropia máxima possa ter significado: esta possibilidade só seria concebível se fosse possível tomar qualquer mensagem deste tipo, rearranjar seus

---

51 *Ibid.*, p. 56.

52 *Ibid.*

53 *Ibid.*, pp. 39-40.

componentes numa ordem qualquer, e a mensagem ainda assim fizesse sentido (qualquer que ele fosse)[54].

A partir da definição de entropia relativa, Shannon formula o conceito de Capacidade de Canal: num canal de transmissão qualquer, a velocidade nominal do canal equivale à entropia máxima que ele pode transmitir por unidade de tempo. Isto ocorrerá quando a mensagem tiver redundância máxima. Neste caso, é necessário distinguir as unidades de sinalização das unidades de informação: ambas são medidas em bits, porém um bit de sinalização costuma conter menos de um bit de informação, ou seja: um bit de sinalização só transportará um bit de informação quando sua entropia relativa for máxima[55].

Como a motivação principal da Teoria Matemática da Comunicação era tecnológica, Shannon aborda o problema fundamental da transmissão de informações: garantir que a mensagem enviada seja recebida sem erros. Erros, segundo Shannon, são ocorrências fortuitas que causam interferências no canal de comunicação, podendo fazer com que os sinais sejam corrompidos[56]. Ele propõe que o risco de corrupção de sinais numa mensagem seja mitigado pela inserção intencional de redundância: deveriam ser adicionados sinais que permitissem identificar se a mensagem foi corrompida. Tais sinais deveriam ser calculados em função dos sinais previamente enviados, sendo que o receptor deveria realizar o mesmo cálculo e, se a mensagem estiver íntegra, chegar ao

---

54 *Ibid.*, pp. 40-4. Shannon propõe um processo de geração de mensagens baseado em probabilidades de ocorrências de letras e de palavras que demonstra a transição de um texto sem sentido para um texto com algum sentido.

55 *Ibid.*, pp. 59-61.

56 *Ibid.*, pp. 65-6.

mesmo resultado[57]. Shannon também propõe que seria possível elaborar um algoritmo de recuperação de erros, no que chamou de "Codificação Eficiente"[58].

Shannon, até este ponto, tratou apenas as fontes de informação discretas, que geralmente são formas de transmissão de textos por diferentes métodos. Para fontes contínuas, a situação é diferente, pois a própria codificação da informação introduz ruído na mensagem. Shannon justifica esta situação pelo fato de que uma fonte contínua necessitaria de um repertório infinito de símbolos, que por sua vez exigiria uma capacidade de transmissão infinita, o que é impossível por na existir um canal livre de ruído. Desta forma, a codificação de sinais de fontes contínuas é, por definição, inexata. Com este problema identificado, é introduzido o conceito de *fidelidade de reprodução*, que é o grau de similaridade entre a informação obtida na fonte contínua com sua representação em valores discretos[59].

Em síntese, Shannon inicia a *Teoria Matemática da Comunicação* afirmando que a informação de uma mensagem pode ser mensurada por quantidade que ele chamou de "entropia", e que é relacionada com a freqüência dos símbolos transmitidos (os símbolos fazem parte de um repertório finito, que define a forma de codificação das mensagens). Ele adverte que a quantidade de informação de uma mensagem é independente de seu significado. Como conseqüência da abordagem estatística adotada, demonstra que a entropia zero é obtida quando existe a certeza da transmissão de um único símbolo e, no caso oposto, a entropia máxima é obtida quando a freqüência dos símbolos é equiprovável. Tendo resumido os principais conceitos da teoria, torna-se viável

---

57 *Ibid.*, pp. 66-70.

58 *Ibid.*, p. 80.

59 *Ibid.*, pp. 108-9.

citar as principais aplicações decorrentes dela e conduzir uma breve verificação experimental.

1.4. Aplicações e Verificação Experimental

Uma aplicação que foi impulsionada pela *Teoria Matemática da Comunicação* é a inclusão de redundância controlada nas transmissões e armazenamento de informação para mitigar o risco de corrupção por ruídos. Como o próprio Shannon reconhece, outros cientistas já haviam se dedicado ao problema da detecção e correção de erros e ele não gasta mais de que uma página para abordar o assunto e colocá-lo no contexto de sua teoria. Devido à repercussão que a Teoria causou, é de se supor que esta breve menção tenha contribuído para despertar a atenção de outros para a questão[60]. Atualmente, todos os protocolos de transmissão de dados contêm algoritmos de verificação e correção de erros, em geral uma variante do CRC[61]. No armazenamento de informação digital, existem circuitos de verificação de erros em memória volátil e sistemas RAID[62] de tolerância a erros para dados armazenados em baterias ("arrays") de discos rígidos.

O conhecimento da relação que existe entre entropia e freqüência proporcionou a criação de algoritmos de compressão de dados. É sabido que as mensagens digitais transmitidas ou armazenadas são codificadas usando-se um repertório comum, que é a tabela de caracteres ASCII[63], que contem 256 valores

---

60 Para mais detalhes sobre a influência da Teoria da Informação nas técnicas de detecção e correção de erros, consultar: http://en.wikipedia.org/wiki/Error_correcting_code.

61 CRC é o acrônimo de "cyclic redundant code" e designa uma classe de algoritmos que geralmente empregam equações polinomiais na detecção e correção de erros. Para mais detalhes, consultar: http://en.wikipedia.org/wiki/CRC-32.

62 RAID é acrônimo de "redundant array of independent disks" e designa uma classe de sistemas de tolerância a falhas, que usa um conjunto de discos operando de forma sincronizada, e que permite a continuidade do funcionamento em caso de falha de um dos discos. Para mais detalhes, consultar: http://en.wikipedia.org/wiki/Redundant_Array_of_Independent_Disks.

63 Na verdade, usa-se uma variante ou extensão do padrão ASCII. Existem outras formas de codificação, como o EBCDIC, ainda empregada em redes privadas que usam computadores ditos "de grande porte" ("mainframe"). Os padrões ASCII e EBCDIC codificam o alfabeto latino e sinais gráficos comuns; outras línguas necessitam de conjuntos diferentes. A despeito das origens independentes e

discretos, cada um correspondendo a uma letra, algarismo ou símbolo. A forma original de qualquer conteúdo digital não considera a entropia existente em seu conteúdo, o que significa que, em geral, a quantidade de bits gerados seja maior que a entropia. A compressão de dados pode ser entendida como um processo de eliminação de redundância, ou de adequação do tamanho da mensagem à sua entropia.

Para que as predições da Teoria da Informação sejam verificadas empiricamente, foi realizado o seguinte experimento: foram gerados 8 arquivos de dados com extensão (tamanho) em bits idênticos (1 Megabyte, ou 1.048.576 octetos), mas com conteúdos diferentes; estes arquivos serão submetidos a 3 processos de compactação diferentes, e suas extensões resultantes foram comparadas com o tamanho original de forma a verificar se o conteúdo de cada arquivo tem influência na eficiência da compactação[64]. Detalhes da metodologia encontram-se descritos no anexo 1.

A Teoria da Informação afirma que: 1) a entropia é relacionada com a freqüência dos símbolos transmitidos (ou armazenados); 2) a entropia zero é obtida quando existe a certeza da transmissão de um único símbolo; 3) a

---

desenvolvimentos próprios, atualmente todas as formas de codificação são consideradas subconjuntos do padrão Unicode. Para mais detalhes, consultar: http://www.unicode.org/ e http://en.wikipedia.org/wiki/Unicode.

64 A respeito dos programas compactadores, é necessário notar que: 1) o conteúdo adicionado pelos compactadores a título de controle (como, por exemplo, o novo repertório que codifica informação) deve ser subtraído; 2) a eficiência de cada algoritmo depende, em última análise, de sua capacidade de reconhecer padrões dentro do arquivo e de codificar o conteúdo buscando reduzir o espaço ocupado. A implementação dos conceitos da teoria na forma de programa de computador produz resultados compatíveis com os valores calculados pela teoria, mas não significa que a extensão do arquivo gerado em bits corresponda à entropia do arquivo original, nem mesmo nos casos onde a compactação produz um arquivo idêntico ao arquivo original (quando, teoricamente, a entropia máxima é alcançada já na origem). Isto se dá porque o algoritmo pode ter falhado em detectar um padrão suficientemente complexo. Para uma análise mais aprofundada sobre a importância do reconhecimento de padrões na Teoria da Informação, consultar A. D. Wyner, J. Ziv, & A.J. Wyner, "On the Role of Pattern Matching in Information Theory", *in* S. Verdú & S. McLaughlin, orgs, *op. cit.*, pp. 1-12.

entropia máxima é obtida quando a freqüência dos símbolos é equiprovável; 4) a entropia não está relacionada ao significado ou a questões subjetivas. Estas são as conseqüências da teoria que serão postas à prova. Os arquivos gerados tiveram seu conteúdo atribuído de forma a verificar se os resultados são consistentes com as previsões teóricas: Texto1.txt contém apenas espaços em branco; Testo2.txt contém a letra "H" repetida por toda a extensão do arquivo; Texto3.txt contém uma seqüência de 16 caracteres distintos, repetida ao longo do arquivo; Texto4 contém uma seqüência de 32 caracteres repetida ao longo do arquivo; Texto5.txt contém uma seqüência aleatória de algarismos; Texto6.txt contém uma seqüência aleatória de letras maiúsculas; Texto7.txt contém uma seqüência aleatória de todos os caracteres da tabela ASCII; Texto8.txt contém a transcrição da Bíblia, truncada em 1 MB[65].

Os resultados obtidos, sumarizados na tabela que se segue, comprovam todas as previsões, validando os princípios e as conclusões da Teoria da Informação:

Arquivo	Extensão Compactada (octetos)			Variação Média (%)	Conteúdo	Consequências Verificadas
	Windows XP	WinZip	WinRAR			
Texto1.txt	1152	1152	602	-99,91	espaços em branco	1, 2
Texto2.txt	1152	1152	602	-99,91	repetição de "H"	1, 2
Texto3.txt	2185	2185	623	-99,84	repetição de seqüência de 16 caracteres	1, 2
Texto4.txt	2708	2708	639	-99,81	repetição de seqüência de 32 caracteres	1, 2
Texto5.txt	492.641	492.553	485.997	-53,23	algarismos aletórios	1, 3
Texto6.txt	665.995	665.571	667.227	-36,46	letras maiúsculas aleatórias	1, 3
Texto7.txt	1.052.672	1.048.854	1.048.650	0,14	caracteres aleatórios	1, 3
Texto8.txt	292.291	291.249	246.019	-73,63	Bíblia em inglês, versão King James	1, 4

Os arquivos Texto1.txt e Texto2.txt resultaram em compactação idêntica apesar de diferirem totalmente no conteúdo (um é todo em branco, e o outro é

---

65 É importante evidenciar que a eficiência dos programas de compactação *não* é o objetivo do teste, embora os resultados possam indicar que algum programa produza melhores resultados. As situações reais são muito mais complexas, e os programas de compactação empregam algoritmos adaptados a cada tipo de arquivo. Considerando-se que a aderência a uma necessidade específica costuma ser mais relevante que um "benchmark" genérico, não é esperado que os resultados obtidos sejam conclusivos a respeito da pretensa "superioridade" de um produto sobre outro. O que se pretende com o teste é avaliar a consistência dos resultados frente às previsões da TMC.

inteiramente preenchido com a letra “H”). A característica comum entre eles é que só há um símbolo possível em toda a mensagem, ou seja, existe a certeza com relação a qual símbolo será usado: a teoria diz que a entropia é zero. De fato, a compactação destes arquivos foi de 99,91%[66].

Os arquivos Texto3.txt e Texto4.txt contêm várias vezes a mesma seqüência de caracteres. O resultado deixa claro que existe pouca entropia na mensagem: é, basicamente, a entropia contida na primeira ocorrência do texto-base, que se dilui progressivamente a cada repetição.

Os arquivos Texto5.txt, Texto6.txt e Texto7.txt contêm seqüências aleatórias. Entretanto, nos dois primeiros o repertório é limitado (10 e 26 símbolos, respectivamente), enquanto que no último o repertório coincide com todo o conjunto de caracteres disponíveis (256 símbolos). A quantidade de símbolos do repertório influencia no cálculo da entropia: em se tratando de símbolos cujas ocorrências são equiprováveis, quanto maior for o repertório, menos freqüente será a ocorrência de cada símbolo individualmente[67]. Os resultados da compactação destes arquivos confirmam estes argumentos: a eficiência da compactação é progressivamente menor, ou seja, as entropias progressivamente aumentam[68].

Finalmente, a compactação do arquivo Texto8.txt, que contem um texto com significado, foi maior que a obtida pela compactação da seqüência aleatória

---

66 Este é um dos objetivos do Algoritmo de Codificação de Entropia, de Huffman (aluno de Fano no MIT). Para mais detalhes, consultar http://en.wikipedia.org/wiki/Huffman_coding.

67 W. Weaver, “Recent Contributions to the Mathematical Theory of Communications”, *in* C. Shannon & W. Weaver, *op. cit.*, p. 16.

68 Como se pode perceber, a compactação do Texto7.txt resultou num arquivo maior que o original. Este fenômeno se explica pela necessidade de se gravar dados de controle no arquivo compactado em todos os processos de compressão: tais dados orientam a descompactação do arquivo, uma vez que o conteúdo foi recodificado segundo um novo repertório. Para mais detalhes sobre compressão de dados, consultar http://en.wikipedia.org/wiki/Data_compression.

de algarismos (Texto5.txt). Este resultado comprova a existência de significativo grau de redundância nas linguagens humanas, o que diminui a entropia das mensagens, além de fornecer a demonstração definitiva de que o significado não tem qualquer influência na entropia[69].

Estas conclusões se aplicam às fontes discretas, onde a compactação fidedigna ("lossless") é possível. Para fontes contínuas, a própria codificação da informação é intrinsecamente inexata. Aliando-se os princípios da teoria sobre fontes discretas ao conhecimento da fisiologia dos órgãos sensoriais humanos[70], é possível modificar o conteúdo dos arquivos de forma a produzir compactações mais eficientes. Neste caso, a compactação de informações captadas de fontes contínuas pode também não ser fidedigna à codificação originalmente gerada. A compactação não-fidedigna, dita "com perdas" ("lossy") consegue taxas de compressão significativamente mais altas, o que é especialmente útil devido às extensões maiores dos arquivos multimídia[71].

No caso de arquivos de som, é possível eliminar as freqüências fora espectro audível, o que comumente se faz na própria digitalização dos sons. Também é possível, dentro de determinados limites, modificar a taxa de amostragem do som (quantidade de vezes em que o som é captado e

---

69 É importante notar que o algoritmo que gerou o arquivo Texto8.txt não garante que a freqüência de todos os símbolos do repertório é equiprovável em termos estritos: para obter-se tal condição, seria necessário garantir que a contagem de cada símbolo do repertório resultasse num mesmo número. Nestes termos, as freqüências dos símbolos gerados são muito próximas, mas não necessariamente iguais. Um algoritmo que garante que a contagem de cada símbolo seja igual pode ser encontrada no anexo 2.

70 *Ibid*., p.111. Shannon fornece como exemplo dados experimentais sobre a sensibilidade da audição humana em relação às características das ondas sonoras: com relação à fase, é praticamente insensível; com relação à amplitude e à freqüência, é aproximadamente logarítmica.

71 A compactação com perdas é aceitável nestes tipos de conteúdo porque o objetivo final deles é a apreciação humana. A capacidade cognitiva dos seres humanos faz com que consigamos lidar com informações imprecisas e parciais (o que, como se sabe, é um dos fundamentos da Psicologia Gestalt).

reproduzido), sem que haja perda significativa na percepção que ele provoca no ouvinte quanto executado.

Em se tratando de imagens, outras técnicas são possíveis, como, por exemplo: 1) reduzir a quantidade de cores, o que corresponde a uma redução do repertório; 2) usar algoritmos de suavização na imagem inteira ou em parte dela, que aproximam a cor de cada ponto da imagem em função das cores dos pontos adjacentes, o que faz com que as freqüências de ocorrência das cores aumentem pela convergência imposta pelo algoritmo, resultando numa entropia menor, ao mesmo tempo em que reduz o repertório pela eliminação de cores não mais usadas.

2. Entropia e Teoria da Informação

2.1. Contexto Histórico

O período que será analisado se inicia cerca de 50 anos antes da publicação da *Teoria Matemática da Comunicação*. Dentre as inúmeras possibilidades de abordagem de um período tão estudado e complexo como o Século XX, o recorte temático impõe que se dê ênfase ao desenvolvimento das telecomunicações, que foi instrumento de transformação geopolítica em tempos de guerra antes dos impactos sócio-culturais em tempos de paz. A trajetória narrativa abordará: 1) a invenção do rádio e sua importância estratégica nas comunicações diplomáticas e militares; 2) o uso generalizado de transmissões telegráficas criptografadas de complexidade crescente a partir da Primeira Guerra Mundial; 3) os esforços dos serviços de inteligência inglês, francês e polonês em decifrar as mensagens alemãs; 4) o trabalho de Alan Turing em criptologia durante a Segunda Guerra Mundial. O objetivo desta abordagem é evidenciar as relações existentes entre a nova tecnologia de comunicações sem fio com os interesses militares decorrentes das duas Grandes Guerras, a importância dos sistemas de criptografia como estratégia defensiva, culminando com o trabalho de Alan Turing, que tem pontos de contato interessantes com os trabalhos de Shannon em sistemas de sigilo (que, como foi mostrado no capítulo I, contém os elementos básicos da TMC).

2.1.1. O Rádio

Nos cinqüenta anos anteriores ao aparecimento da rádio-difusão, surgiu e se disseminou a transmissão pública de mensagens por telégrafo. Uma rede

mundial, que utilizava cabos de transmissão terrestres e submarinos, chegou a ser estabelecida pelo do império britânico:

> *Foi a Inglaterra que primeiro construiu um sistema de telégrafo submarino global e que se manteve a (apenas) um passo adiante de seus competidores em todas as formas de telecomunicações civis e militares até 1945*[72].

O emprego de comunicação sem fio ("wireless"), ou seja, via rádio, surgiu no início do século XX. Entre os vários pesquisadores que trabalhavam no desenvolvimento do rádio na mesma época[73], destaca-se Marconi, que, por sua atuação empreendedora, esteve próximo de construir um monopólio mundial de telecomunicações. O monopólio não foi efetivamente implantado devido à progressiva percepção de que a nova tecnologia era fundamentalmente importante para os interesses estratégicos das potências nacionais de então, que conseguiram desenvolver formas de impedir que Marconi obtivesse o controle global das comunicações:

> *Desde o princípio, Marconi entendeu a importância do rádio para comunicação internacional. Ao contrário de outros inventores, ele não poupou esforços para controlar patentes, tanto suas como de outros. Seus competidores individuais jamais conseguiram controlar tantas patentes como ele conseguiu. Na luta geopolítica pelo rádio internacional, três técnicas foram desenvolvidas para evitar que Marconi construísse um monopólio efetivo: o acordo internacional, a manipulação da política doméstica, e a combinação de patentes*[74].

---

72 P. Hugill, *Global Communications since 1844: Geopolitics and Technology*, p. 2. Citação no original: "It was Britain that first constructed a global submarine telegraph system and Britain kept (just) one step ahead of its competitors in all forms of civil and military telecommunications until 1945.

73 A invenção do rádio, entendido como o uso de ondas eletromagnéticas a transmissão de informações sem a utilização de cabos, é disputada por Nikola Tesla, Guglielmo Marconi e Alexander Popov, entre outros. Neste trabalho, que não tem como objetivo discutir a invenção do rádio, não nos posicionaremos a favor de Marconi apesar de ele ser reconhecido como o inventor do rádio pelas obras de referência pesquisadas. Entretanto, ele será amplamente citado devido a seus bem-sucedidos esforços em estabelecer uma indústria a partir de suas invenções.

74 *Ibid.*, p. 92. Citação no original: "From de beginning Marconi understood the importance of radio for international communication. Unlike other inventors, he therefore went to considerable lengths to control patents, both his owns and those originated by others. His individual competitors were never able to control as many patents as he did. In the geopolitical struggle for international radio three techniques were developed to block Marconi from developing an effective monopoly: the international agreement; manipulation of domestic politics; and patent pooling".

No final do século XIX, Marconi investigava um fenômeno peculiar que acontecia aos circuitos elétricos, que era a indução de corrente elétrica à distância, produzido por um circuito em outro, sob certas circunstâncias. Após sucessivas experimentações e modificações nos circuitos, Marconi conseguia transmitir pulsos de informação à distância de 2,5 km. O telégrafo, forma de comunicação à distância predominante durante as cinco décadas anteriores, exigia linhas de transmissão ponto-a-ponto e o conseqüente controle territorial por onde passassem os cabos[75], o que não seria necessário com o sistema de Marconi. Tendo emigrado da Itália para a Inglaterra pouco depois, dedicou-se ao melhoramento do sistema após obter a patente da invenção, e conseguiu enviar mensagens através do Canal da Mancha. Seu invento ganhou notoriedade durante a cobertura da competição de iatismo mais importante da época, o America's Cup de 1899: os jornalistas que cobriam o evento mandavam relatórios diários da corrida através do rádio para suas editorias em Nova Iorque a tempo de serem incluídas na edição do dia seguinte[76].

Porém, o grande desafio de Marconi era provar que a transmissão seria possível entre pontos muito distantes. Desafiando os críticos de sua invenção, que alegavam ser um sistema de alcance limitado já que o sinal não poderia vencer a curvatura da Terra[77], Marconi conseguiu enviar sinais telegráficos da Inglaterra ao Canadá, a uma distância de 3.500 km. Surgiram várias teorias

---

75 *Ibid.*, p. 83.

76 S. Singh, *The Code Book: the Science of Secrecy from Ancient Egypt to Quantum Cryptography*, pp. 101-2.

77 A idéia de que era necessário haver mútua visibilidade entre transmissor e receptor não é absurda. De fato, grande parte das telecomunicações necessita de "visada", como é o caso de algumas transmissões por satélites, como nas transmissões de TV e de dados: a antena receptora precisa estar apontada para o satélite para que a reflexão das ondas no disco parabólico as faça convergir para um ponto focal, ampliando a intensidade do sinal, grandemente diminuída pela longa distância percorrida, uma vez que tais satélites geoestacionários orbitam a Terra a cerca de 40.000 km de altura. Para mais detalhes, consultar: http://en.wikipedia.org/wiki/Satellite_television.

para explicar o fenômeno, entre elas a que afirmava que as ondas eletromagnéticas acompanhavam a superfície da Terra, e a que previa a existência de uma alta camada atmosférica refletora. O fenômeno permaneceu sem uma explicação definitiva por mais de duas décadas, quando foi descoberta a ionosfera e houve a verificação experimental de que as ondas de rádio de determinadas freqüências ricocheteiam entre a ionosfera e a terra, fazendo com uma transmissão possa dar várias voltas no planeta se tiver potência suficiente[78]. Hoje é sabido que tanto os críticos iniciais como os teóricos posteriores tinham sua parte de razão: dependendo da freqüência, a onda eletromagnética pode se curvar à superfície do planeta, ser refletida na ionosfera ou seguir em direção ao espaço, conforme representa a ilustração a seguir:

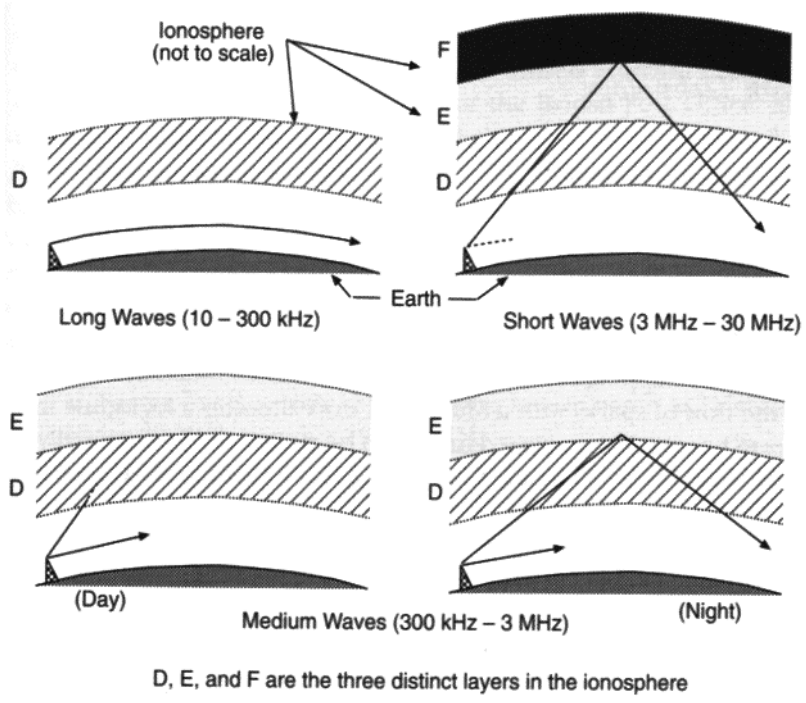

Fig. 3 – A reflexão das ondas eletromagnéticas em função de suas freqüências

[78] A história da primeira transmissão transatlântica por Marconi foi questionada recentemente, embora não haja dúvida de seu sucesso numa segunda transmissão realizada um ano depois. Para mais detalhes, consultar: http://en.wikipedia.org/wiki/Ionosphere .

A ilustração mostra as trajetórias das ondas eletromagnéticas em função de suas freqüências e da insolação: as ondas longas acompanham a curvatura da Terra; as ondas curtas se propagam em linha reta e são refletidas na ionosfera; as ondas médias se propagam em linha reta e são refletidas pela ionosfera somente à noite.

A realização de Marconi enseja uma breve digressão a respeito dos caminhos nem sempre concordantes da Ciência e das Tecnologias. É interessante ver um feito tecnológico pondo em dúvida as crenças científicas de seu tempo. Mário Schenberg cita outro exemplo deste fenômeno: a máquina a vapor, que transformava calor em trabalho mecânico, foi amplamente usada pela indústria num tempo em que a conversibilidade entre as várias formas de energia não era bem compreendida, uma vez que a teoria do calor predominante era a do fluido calórico[79]. Avaliando as últimas três décadas do século XX segundo os termos de Khun, assistimos à crescente interdependência da Ciência e da Tecnologia, possivelmente decorrente do longo período de ciência normal que se seguiu às revoluções científicas do início do século XX: a Relatividade e a Mecânica Quântica. Estas duas teorias, embora ainda sejam incompatíveis, não rivalizam uma com a outra, e ambas transformaram o modo de vida da humanidade, e é possível argumentar que estamos aprimorando e ajustando o que foi introduzido por elas, enquanto aguardamos uma nova revolução que subverta o saber consensual. A intimidade entre Ciência e Tecnologia poderá fazer com que uma precise da outra para a próxima revolução. Mais importante, é a atitude de inconformismo com a situação estabelecida, e o reconhecimento de que o que sabemos não é definitivo, como adverte Schenberg:

---

[79] M. Schenberg, *op. cit.*, p. 142.

> *Há um conservadorismo que nos deixa presos a muitas idéias complicadas*[80].

A invenção de Marconi despertou o interesse do militares. Até então, só era possível ter comunicação rápida à distância através de cabos telegráficos, de custo elevado. Além disso, embarcações ficavam incomunicáveis por longos períodos, e era impossível fazê-las participar de uma súbita mudança de estratégia. Entretanto, a grande desvantagem do rádio era a impossibilidade de garantir sigilo nas transmissões, uma vez que as mensagens são irradiadas para todos os lados e trafegam pelo ar, o que facilita o trabalho de interceptá-las[81].

Uma das formas de introduzir sigilo neste tipo de comunicação é o emprego de técnicas de criptografia. Criptografia é a arte de tornar a mensagem inteligível apenas para seu receptor, na qual são usadas técnicas de transposição e de substituição de sinais. A reversibilidade é um pressuposto básico para qualquer processo criptográfico, e o elemento secreto que garante sua correta reversão é chamado de "chave"[82]. Para a criptografia ser eficiente, ela deve evitar as características que tornam o texto decifrável, sendo que a principal delas é a análise estatística dos sinais e das palavras. Esta é a área de contato entre a criptografia e a criptologia com a Teoria da Informação, como será mostrado no decorrer deste capítulo.

As primeiras aplicações militares de criptografia por rádio consistiam em duas etapas: na primeira, ocorria a transformação da mensagem original numa mensagem cifrada, que era realizada por um especialista; na segunda, a

---

80 *Ibid.*, p. 204.

81 S. Singh, *op.cit.*, p. 102.

82 Existem outras técnicas de sigilo, como a Esteganografia, que busca esconder a existência de uma mensagem secreta dentro uma mensagem aparentemente "inocente". Para mais detalhes sobre terminologia e técnicas básicas de criptografia, consultar D. Kahn, *The Codebreakers: the story of secret writing*, pp. xv-xvii.

mensagem cifrada era passada a um telegrafista, que transmitia a mensagem tal qual a recebeu. Tal procedimento era repetido na ordem inversa pelo receptor, que assim obtinha a mensagem original. Para que o processo fosse bem-sucedido, a chave de cifragem deveria mudar a cada mensagem, caso contrário o algoritmo poderia facilmente ser descoberto por análise estatística sobre as mensagens interceptadas. Isto impunha um procedimento rigoroso de criação e uso de chaves, que eram registradas em livros. A mesma sistemática era usada pelas grandes potências militares do início do século XX.

Em 1914, durante a Primeira Guerra Mundial, um livro de códigos criptográficos alemão foi recuperado do naufrágio do Magdeburg, afundado pela marinha russa no Golfo da Finlândia. Os Russos, após comprovarem a eficácia dos códigos, cederam-nos ao Almirantado Inglês. Esta informação foi mantida em segredo até 1923, quando foi tornada pública por Winston Churchill[83], na época o comandante do almirantado[84]. Além deste, outros livros de códigos foram recuperados pelos ingleses: o primeiro de um "destroyer" alemão na batalha de Heligoland Bight, ainda em 1914 [85]; e outro em 1916 do Zeppelin Z-32, abatido em Billericay[86]. Estes livros auxiliaram o grupo de criptólogos britânicos (conhecidos por "Room 40", um dos endereços que ocuparam nas instalações da marinha) a decifrarem muitas das mensagens interceptadas. A decodificação das mensagens poderia ser feita sem eles, mas a informação contida nos livros facilitava o processo. A inteligência francesa não usava os livros, o que não a impediu de realizar decodificações importantes,

---

83 *Ibid.*, p. 141.
84 D. Kahn, *op. cit.*, p. 268.
85 *Ibid.*, p. 269.
86 *Ibid.*, p. 273.

mesmo que parciais e fragmentadas. A cessão deste tipo de informação aos franceses teria sido considerada um risco para manutenção do domínio naval britânico[87].

O trabalho dos serviços de inteligência, obviamente, não se resume a decifrar mensagens: para garantir superioridade estratégica, é desejável que o inimigo ignore que seus códigos foram violados, caso contrário ele tomará medidas para mudá-los ou reforçá-los. Um exemplo deste estratagema foi a decifração do conhecido Telegrama Zimmermann. Como preparativo para a guerra, o governo alemão planejava manter os Estados Unidos ocupados em seu próprio território para que não se envolvesse no conflito europeu. O plano era financiar o governo mexicano numa incursão ao território norte-americano para a retomada de algumas regiões anexadas aos Estados Unidos, como o Novo México, Texas e Arizona. Entretanto, a mensagem do chanceler Zimmermann ao presidente do México foi interceptada e decifrada pelo Room 40, e em seguida repassada ao governo norte-americano. Este ato de guerra aos Estados Unidos foi retaliado a sangue-frio: foi montada uma sofisticada operação para ocultar a capacidade de decifração dos aliados, que envolveu o acionamento de uma agente secreto americano no México para que invadisse a central telegráfica mexicana e roubasse a mensagem já traduzida. A mensagem roubada foi revelada à opinião pública e justificou o ingresso dos Estados Unidos na Primeira Guerra Mundial. Ainda como manobra de diversionismo, a inteligência inglesa "plantou" na imprensa de seu país uma crítica fictícia do comandante da inteligência inglesa a sua própria equipe por terem sido incapazes de antever o plano alemão. O inquérito alemão concluiu que houve traição de um

---

87 *Ibid.*, p. 277.

colaborador no México, e eles continuaram usando códigos similares durante toda a guerra[88].

2.1.2. A Máquina Enigma

A revelação de que os ingleses tinham livros de códigos alemães deu impulso a uma máquina de criptografia inventada em 1918 na Alemanha, mas que não tinha sido adotada em larga-escala devido a seu alto custo unitário: a máquina Enigma. Este dispositivo foi inventado por A. Scherbius, que obteve a patente de sua invenção em 1928 nos Estados Unidos, conforme mostra a reprodução que se segue:

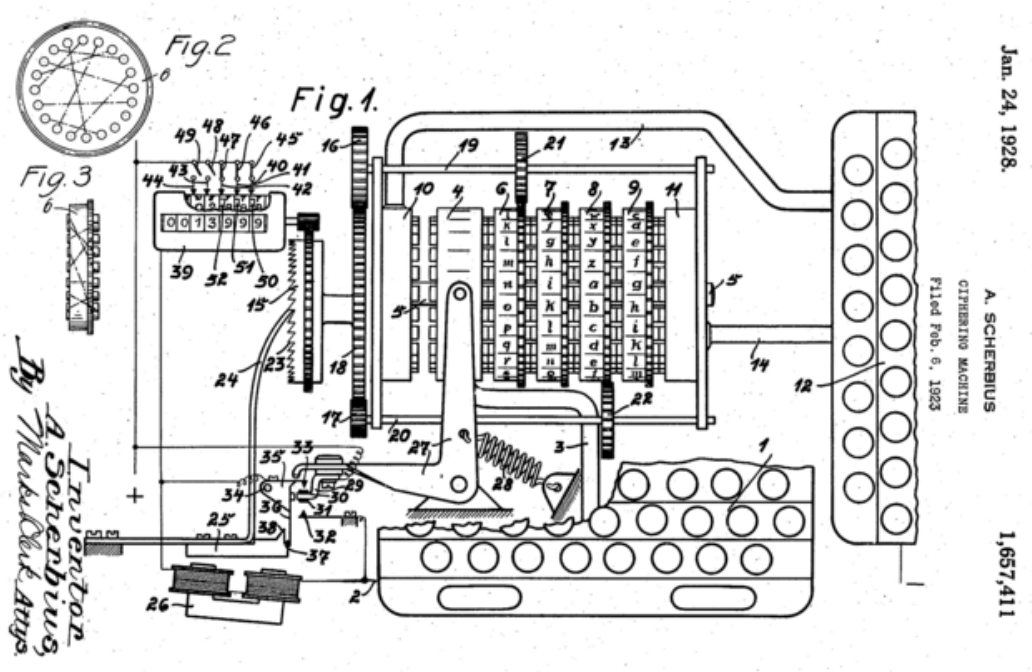

Fig. 4 – Patente norte-americana da Máquina Enigma, concedida em 1928.

A representação esquemática exibida na patente revela parcialmente sua complexidade, que pode ser melhor entendida pela observação de sua aparência real, exibida a seguir:

---

88 S. Singh, *op.cit.*, pp. 107-15.

Fig. 5 - Um exemplar de Máquina Enigma

A máquina Enigma era composta por engrenagens de seleção de chaves, um painel iluminado e um teclado que cobriam um intrincado mecanismo: com a regulagem correta de seu mecanismo, cada letra teclada acendia outra letra no painel, sendo que os ajustes da máquina se alteravam a cada toque. O total de possibilidades de ajuste era de cerca de 10^{16}. Inicialmente rejeitada pelas grandes empresas privadas, também não foi bem recebida pelos militares alemães:

> *Os militares alemães foram igualmente indiferentes porque ignoravam o dano causado por suas cifras inseguras durante a Grande Guerra*[89].

Esta visão mudou com a revelação de Churchill, reforçada pela história da guerra segundo a Marinha Real Britânica, na qual era atribuída grande

---

89 *Ibid.*, p.138. Citação no original: "The German military were equally unenthusiastic, because they were oblivious to the damage caused by their insecure ciphers during the Great War".

importância ao serviço do Room 40 para a vitória dos aliados. Os militares alemães, desejosos de não repetir o recém-descoberto fiasco, acharam na máquina Enigma a solução de seus problemas[90].

Os alemães começaram a utilizar a Enigma a partir de 1926. As mensagens criptografadas pela máquina Enigma eram tão complexas que provocou a desistência dos serviços de inteligência ingleses, americanos e franceses em tentar decifra-las. Os aliados, tendo vencido a guerra, não temiam mais a Alemanha, o que pode ter causado uma falsa sensação de segurança. Entretanto, o serviço de inteligência polonês não compartilhava deste sentimento[91]. A Polônia, tornada independente após a guerra, estava encravada entre dois inimigos poderosos: a leste, a Rússia, ávida pela expansão do estado comunista, e a oeste a Alemanha, desejosa de recuperar o território perdido. Seu serviço de inteligência, apesar dos fracassos em decifrar as mensagens da Enigma, perseverou, coletando e organizando, por anos a fio, todas as mensagens alemãs interceptadas. Ajudado pela traição de um alemão, o manual de operação da máquina Enigma foi obtido pelos franceses, que os repassaram aos poloneses[92]. Os poloneses continuaram a estudar as mensagens e descobriram alguns pontos vulneráveis decorrentes do uso de palavras e expressões da língua alemã, mas ainda não conseguiam decifrar mensagens inteiras. Em 1939, com a iminência da invasão da Polônia pelo exército nazista, o serviço de inteligência polonês enviou o resultado de suas pesquisas aos ingleses e aos franceses. Duas semanas depois, a Polônia foi invadida[93].

---

90 *Ibid.*, pp. 141-2.
91 *Ibid.*, pp. 143-4.
92 *Ibid.*, pp. 144-6.
93 *Ibid.*, pp. 158-60.

2.1.3. Alan Turing

As descobertas polonesas deram novo ânimo aos ingleses, que decidiram investir num novo serviço de criptologia. Foi criado um centro de treinamento em criptologia, o GC&CS ("Government Code & Cipher School") em Bletchley Park, Buckinghamshire, sendo que matemáticos e cientistas eram os alvos do recrutamento, ao invés de lingüistas e eruditos como no Room 40. Em Bletchley Park, durante o outono de 1939, os matemáticos e cientistas recrutados aprenderam sobre o funcionamento da máquina Enigma e as técnicas polonesas[94]. Em seguida, começaram a desenvolver suas próprias técnicas de exploração das mensagens alemãs, basicamente pela mesma falha que ajudou os poloneses: as repetições de palavras e de chaves, que de fato não eram um problema da tecnologia Enigma, mas uma questão de uso inadequado dela[95]. Dentre os muitos colaboradores, o que mais se destacou foi Alan Turing.

Alan Turing ingressou no King's College em Cambridge em 1931, num período de intenso debate sobre matemática e lógica deflagrado por Göedel e do qual participavam Russel, Whitehead e Wittgenstein[96]. Influenciado por estas idéias e também pelas de von Neumann, publicou em 1937 o trabalho *On Computable Numbers*[97], onde propõe formas de encontrar problemas indecidíveis, que acabou por fornecer o modelo conceitual do computador moderno[98]. Ele foi convidado para ingressar em Bletchley Park em 1939, no dia seguinte à declaração de guerra da Inglaterra à Alemanha. Além de ajudar na

---

94 *Ibid.*, pp. 158-60.

95 *Ibid.*, pp. 164-5.

96 *Ibid.*, p. 166.

97 Para mais detalhes as influências de Turing no desenvolvimento do *On Computable Numbers*, consultar A. Hodges, *Alan Turing: the Enigma*, pp.79-110.

98 S. Singh, *op.cit.*, pp. 168-9.

decifração de mensagens, Turing atuava no núcleo consultivo ("think tank") do grupo, onde seu trabalho era desenvolver uma nova técnica de ataque às mensagens da Enigma antes que os alemães se dessem conta dos erros que cometiam deste o 1926 e decidissem corrigí-los[99]. Baseando-se em suas idéias expressas em *On Computable Numbers*, Turing projetou uma máquina que acoplava diversas máquinas Enigma trabalhando para achar chaves, a partir de ajustes previamente selecionados por um criptólogo experiente[100]. O fato é que o aumento de eficiência decorrente da mecanização da geração de mensagens criptografadas (pelo uso da Enigma) só teve uma contrapartida no processo de decifração com a invenção de Turing (chamada de "bombe"). No fim da Segunda Guerra Mundial, 50 destas máquinas, estavam em operação, e abasteciam o almirantado com informações vitais sobre as táticas inimigas[101].

Em 1943, durante o curso da guerra, Turing viajou aos Estados por conta do acordo de cooperação militar entre ingleses e norte-americanos. Entre os lugares que visitou, estavam os Laboratórios Bell: lá, encontrava-se diariamente com Claude Shannon na hora do chá, o que resultou numa interação significante para ambos[102]. Segundo Andrew Hodges, Turing teria se fascinado pelo tipo de trabalho que Shannon fazia e com a liberdade que os Laboratórios Bell lhe oferecia: este modelo de trabalho não existia na Inglaterra, supostamente por não ser bem-visto pelas empresas britânicas[103]. Turing teria se identificado com Shannon por sua bagagem acadêmica e por ter, como ele, abordado a criptologia por métodos científicos. Comparando os aspectos de

---

99 *Ibid.*, pp. 169-70.
100 *Ibid.*, pp. 174-6.
101 *Ibid.*, pp. 185-7.
102 A. Hodges, *op. cit.*, pp. 249-50.
103 *Ibid.*, p. 250.

identificação de Turing com Friedman, outro pesquisador da Bell, e Shannon, Hodges afirma:

> *Em profundidade intelectual, Shannon é que era o correspondente de Alan, e eles descobriram muito em comum*[104].

De acordo com Hodges, as definições de "máquina" e "comunicação", dois conceitos existentes desde o início da civilização humana, só foram formuladas de forma precisa e matemática, respectivamente, por Turing em *On Computable Numbers*, e Shannon em *Mathematical Theory of Secrecy Systems*. Além do paralelo entre os trabalhos de ambos, haveria uma espécie de reciprocidade, pois Turing tinha proposto uma medida de informação, o "deciban", semelhante ao "bit" de Shannon (a diferença seria a base do logaritmo usada no cálculo):

> *Um **ban** de peso de evidência fazia algo dez vezes mais provável; um dígito binário ou **bit** fazia algo duas vezes mais certo. (grifos do original)* [105]

Apesar da conexão fundamental entre as duas teorias, eles não estavam autorizados a discuti-las abertamente. Shannon também tinha pensado, de forma independente, a respeito de máquinas lógicas, o que o levou a sua famosa dissertação de mestrado. Turing teria deixado Shannon impressionado ao lhe apresentar seu *On Computable Numbers*, sendo que seus principais pressupostos de ambas as obras seriam de aceitação recíproca. Ambos também estariam interessados em reproduzir o funcionamento do cérebro numa máquina, mas com uma amplitude (ou pretensão) maior por parte de Shannon, uma vez que Turing teria comentado a outros, num almoço:

---

104 *Ibid.*, p. 250. Citação no original: "In intellectual depth it was Shannon who was Alan's opposite number, and they found a good deal in common".

105 *Ibid.*, p. 250. Citação no original: "A **ban** of weight of evidence made something ten times as likely; a binary digit or **bit** made something twice as definite".

> *Shannon quer alimentar o Cérebro não apenas com **dados**, mas com coisas **culturais**! Ele quer tocar **música** para ele! (grifos do original)*[106]

Infelizmente para Turing, as semelhanças entre sua vida e a de Shannon não foram muitas. Enquanto Shannon obteve condições únicas para desenvolver seus interesses e conquistou reconhecimento imediato por suas realizações, o trabalho de Turing foi mantido secreto pelo governo britânico ainda por muitos anos além do término da Segunda Guerra. O governo britânico distribuiu milhares de máquinas Enigma tomadas dos alemães a suas antigas colônias, que acreditavam que as mensagens produzidas por elas eram impenetráveis: assim, os ingleses tiveram acesso irrestrito às comunicações de suas ex-colônias ainda por muitos anos[107]. Turing cometeu suicídio em 1954, motivado pela perseguição que sofreu depois que sua homossexualidade foi descoberta. As informações sobre as atividades desenvolvidas em Bletchley Park, conhecidas pelo codinome de "Ultra", foram tornadas públicas somente na década de 1970[108].

2.1.4. Telecomunicações e Geopolítica

A respeito da importância geopolítica do domínio das telecomunicações, Hugill oferece uma síntese:

> *Se informação é poder, quem quer que domine os sistemas mundiais de telecomunicação comanda o mundo*[109].

---

106 *Ibid.*, p. 250. Citação no original: "Shannon wants to feed not just *data* to the Brain, but *cultural* things! He wants to play *music* to it!".

107 S. Singh, *op.cit.*,p. 187.

108 *Ibid.*, p. 189.

109 P. Hugill, *op.cit.*, p. 2. Citação no original: "If information is power, whoever rules the world's telecommunications system commands the world."

A empresa de Marconi conseguiu ligar por rádio todos os territórios do império britânico em 1927, e foi forçada em 1929 a se fundir com a companhia telegráfica (via cabos) inglesa, formando a "Imperial and International Communications", mais tarde renomeada para "Cable & Wireless". Insatisfeito com os termos da fusão, que não lhe garantia o controle da empresa apesar de sua participação acionária majoritária, Marconi retirou-se do negócio e a Cable & Wireless passou a servir exclusivamente aos interesses hegemônicos ingleses[110], cujas intenções podem ser inferidas do mapa-múndi peculiar reproduzido a seguir:

Fig. 6 – C&W "Great Circle" Map, cortesia de Cable & Wireless Archive, Porthcurno.

110 *Ibid.*, p. 49.

O mapa representa as linhas de telecomunicações britânicas em 1945, e posiciona Londres no centro do mundo, que é para onde convergem as linhas de comunicação internacionais. Na parte superior, lê-se: "Bretanha, o Centro do Mundo", e os países da comunidade britânica são destacados em vermelho. A julgar pelo mapa, parece ser evidente que a pretensão inglesa era de permanecer no controle das telecomunicações mundiais.

Os britânicos só começaram a fazer concessões nesta área devido às necessidades da guerra, mas apenas para os Estados Unidos, cujos canais de comunicação transatlânticos passavam necessariamente por Londres até então:

> *Uma medida de retenção do poder britânico é que só no final de 1943 o governo americano teve acesso a um cabo transatlântico de telégrafo que não passasse pela Inglaterra*[111].

Da mesma forma, a capacidade de decifração de mensagens dos ingleses foi mantida em segredo enquanto foi possível:

> *Os ingleses atribuíam mais importância ao Ultra do que a qualquer outro aspecto da Segunda Guerra Mundial. Embora tivessem compartilhado as tecnologias do radar centimétrico e do motor a jato com a América durante a Missão Tizard de 1940, bem antes da entrada oficial dos americanos na guerra, não antes de maio de 1943 eles compartilhariam plenamente o Ultra*[112].

Em síntese, o rádio ofereceu como vantagens a mobilidade de comunicação e o barateamento das telecomunicações. A falta de sigilo que lhe é inerente não impediu sua disseminação, tendo ainda propiciado o desenvolvimento de técnicas de sigilo manuais e mecanizadas. Parece apropriado considerar a criptografia uma aplicação da Teoria da Informação,

---

111 *Ibid.*, p. 50. Citação no original: "One measure of the retention of British power is that not until late 1943 did the American government have access to a transatlantic teleph cable that did not pass through Britain."

112 *Ibid.*, p. 144. Citação no original: "The British attached more importance to Ultra than to any other aspect of World War II. Although they shared the technologies of centimetric radar and the jet engine with America during the 1940 Tizard Mission, well before official American entry into war, it was May 1943 before they fully shared Ultra.

apesar da primeira ter precedido a segunda na obra de Shannon. A entropia, entendida como uma grandeza que mede quantidade de informação por processos estatísticos, é o conceito central da TMC e que também permeia os trabalhos de criptografia de Shannon e Turing.

Com relação às telecomunicações, parece acertado inferir sua importância geopolítica pela estratégia da Inglaterra de manter o controle sobre a maior parte das linhas de comunicação internacionais até a década de 1940, somente vindo a fazer concessões aos Estados Unidos em função da fragilidade de sua posição durante a Segunda Guerra Mundial. Considerando-se ainda que os trabalhos de Turing permaneceram secretos até a década de 1970, é possível inferir o grau de importância atribuído a estes aspectos no contexto geopolítico mundial[113].

[113] É interessante notar que as mesmas questões sobre poder e dominação ocorrem atualmente nos debates a respeito da internacionalização do controle da internet (que está sob o controle de um órgão do governo norte-americano).

2.2. Boltzmann e a Entropia na Teoria do Gás

O cientista alemão Ludwig Boltzmann, no século XIX e início do XX, dedicou-se ao estudo da termodinâmica estatística, a qual descrevia a pressão e temperatura dos gases como função do movimento de suas moléculas. Numa época em que a existência dos átomos ainda era questionada, seu trabalho sofreu críticas contundentes que contribuíram para levá-lo ao isolamento e ao suicídio[114]. Não obstante, os princípios que defendia demonstraram-se adequados para descrever diversos aspectos da natureza dos gases, e seu conceito de Entropia influenciou pesquisadores do século XX no desenvolvimento da Teoria da Informação. Sua obra fundamental é o *Lectures on Gas Theory* (tradução em língua inglesa do *Vorlesungen über Gastheorie*, editada em dois volumes em 1896 e 1898, respectivamente); dela serão retirados o conceito e as implicações da entropia na Teoria dos Gases.

Os principais pressupostos da teoria são antecipados já no título da Parte I do *Lectures on Gas Theory*: "Teoria dos gases com moléculas mono-atômicas cujas dimensões são desprezíveis comparadas ao livre percurso médio". Nele se pode distinguir a visão mecanicista da matéria, entendida como composta por unidades discretas de tamanho diminuto. Baseado nos conceitos da Termodinâmica de Clausius, que definiu o Calor como movimento molecular, Boltzmann estabelece os limites de sua teoria: as questões epistemológicas sobre as relações entre modelos com a realidade, motivo de discussões

---

[114] Boltzmann tinha saúde frágil e sofria de neurastenia. Seu temperamento causou-lhe diversas inimizades no meio acadêmico e mesmo o afastamento de amigos, como Ostwald. Sofreu oposição continuada de Mach e seus partidários, que alegavam que a existência de átomos seria "metafísica" por ser um fato não observável e, portanto, incompatível com temas "científicos". Mais informação a respeito da vida, obra e polêmicas de Boltzmann pode ser obtida em D. Lindley, *Boltzmann's Atom: the great debate that launched a revoltuion in physics*.

infindáveis, seriam evitadas. Propõe-se a desenvolver sua teoria livre de posições "dogmáticas", sejam elas tanto atomistas como anti-atomistas, descrevendo sua Teoria dos Gases inicialmente como "analogia mecânica"[115], embora afirme acreditar na descontinuidade real da matéria[116] e se propunha a provar que a analogia mecânica não é apenas aparência superficial[117].

Tendo definido o calor como movimento das moléculas, Boltzmann explica que o calor é uma forma de energia cinética, mas que não é visível porque somos capazes de percebê-la apenas quando ela atua igualmente em todas as moléculas de um corpo: nos demais casos, o movimento molecular seria percebido como calor. Ele procura justificar o engano de se considerar o calor como sendo uma "substância" (uma clara referência à teoria do fluido calórico), que ocorreria pelo fato da energia cinética se transmitir sempre de um corpo cujas moléculas têm mais energia para outro cujas moléculas têm menos energia, aliado ao fato de que a energia cinética não se destrói. Continuando, afirma que a natureza das forças de coesão molecular não é ocupação da presente teoria, embora reconheça que sua incompreensão até aquele tempo pudesse causar imprecisão nas definições de estado físico[118].

A relação entre o movimento molecular e estado físico é caracterizada por Boltzmann de forma simplificada, o que se justificaria, segundo o autor, pelo desconhecimento da natureza das forças de atração e repulsão em nível molecular até então. Assim sendo, define o estado sólido como aquele em que cada molécula tem uma posição de repouso, na qual é mantida pela repulsão das

---

115 L. Boltzmann, *op. cit.*, p. 26.
116 *Ibid.*, pp. 27-8.
117 *Ibid.*, p. 28.
118 *Ibid.*, pp. 28-9.

moléculas adjacentes quando delas se aproxima e pela atração quando delas se afasta. Desta forma, o movimento térmico colocaria a molécula em oscilação pendular ao redor de sua posição de descanso, e se todas as moléculas tiverem o mesmo tipo de movimento, suas posições relativas seriam mantidas e o corpo teria uma forma fixa. A única conseqüência visível do movimento molecular seria a dilatação do corpo à medida que a amplitude do movimento pendular de cada molécula se ampliasse. O estado líquido, por sua vez, foi caracterizado pelo movimento molecular de intensidade suficiente a permitir que uma molécula se imiscuísse entre moléculas vizinhas, sobrepujando as forças de atração e repulsão moleculares: o corpo sólido, neste caso, se derreteria. Além de um limite definido, o movimento térmico seria de intensidade tal que as forças de atração e repulsão seriam desprezíveis, e as moléculas se movimentariam livremente pelo espaço: o corpo, então, se evaporaria[119].

Partindo destes princípios, Boltzmann expõe seu conceito de gás: moléculas que se movimentam livremente num espaço fechado suficientemente grande. Na ausência de forças externas, o movimento das moléculas seria retilíneo com velocidade constante, assemelhando-se a projéteis disparados por uma arma de fogo; somente se desviariam caso passassem muito perto de outras moléculas, ou quando se chocassem contra as paredes do recipiente que as contivesse. Neste sentido, a pressão do gás seria interpretada como a ação das moléculas contra as paredes do recipiente[120]. Boltzmann prossegue com o cálculo da pressão de um gás a partir de fundamentos exclusivamente mecânicos, como a conservação da energia cinética e o movimento dos centros

---

[119] *Ibid.*, p.30.

[120] *Ibid.*, p. 30.

de gravidade das moléculas. O modelo se inicia de forma muito simples: considera que as moléculas são esferas totalmente elásticas e que sua deformação é desprezível ao se chocarem com as paredes do recipiente que contém o gás, consideradas como superfícies completamente lisas e elásticas. Desta forma, elabora equações que relacionam a força dos choques das moléculas com a força de reação da parede do recipiente, concluindo que a pressão do gás é o valor médio da soma de todas as pequenas pressões que as moléculas transmitem ao se chocarem com o recipiente. Boltzmann prossegue com as demonstrações e prepara as equações básicas para os capítulos que se seguem[121].

O Capítulo 1 da Parte 1 antecipa em seu título os principais pressupostos do Teorema-H: as moléculas são esferas elásticas e não existe movimento macroscópico aparente no gás. Boltzmann inicia com a prova da Lei da Distribuição de Velocidades de Maxwell[122]. Como pressuposto inicial, considera que o recipiente contém apenas um gás, composto por moléculas idênticas que se comportam como esferas completamente elásticas ao colidirem entre si. Neste cenário, mesmo que todas as moléculas tenham velocidades iguais num momento inicial, em algum tempo ocorreriam colisões que determinariam que as velocidades das moléculas se alterassem (aumentando ou diminuindo), após cada colisão. No decorrer do tempo, existindo um número suficientemente grande de moléculas no recipiente, todas as possíveis velocidades ocorreriam, e cada molécula assumiria uma velocidade entre zero e uma velocidade que seria

---

121 *Ibid.*, pp. 30-5.

122 Segundo Penrose, a distribuição Maxwelliana é normal (gaussiana). Para mais detalhes, consultar: R. Penrose, *op. cit.*, pp. 694-6. Mais especificamente, é uma "distribuição qui" (do tipo exponencial) com três graus de liberdade. Para mais detalhes, consultar:
http://en.wikipedia.org/wiki/Maxwell-Boltzmann_distribution#Distribution_of_speeds .

muito superior à velocidade inicial comum a todas as moléculas. Neste estado final, poder-se-ia calcular a lei de distribuição de velocidades entre as moléculas. Num caso mais geral, poder-se-ia considerar a existência de dois gases no recipiente que naturalmente teriam moléculas de massas distintas, chamando-as de moléculas-*m* e o de moléculas-*m1*. Boltzmann propõe que se considere a inexistência de forças externas, e que as paredes do recipiente atuem apenas como "refletoras" das moléculas que se chocassem a elas. Nestas circunstâncias simplificadas, a velocidade de cada molécula só seria alterada por colisões com outras moléculas[123].

Na seqüência, Boltzmann procura demonstrar a validade de se aplicar a Lei da Distribuição de Velocidades de Maxwell ao modelo que propõe. Esta lei demonstraria que um gás, partindo de um estado molecular ordenado, tende a um estado molecular desordenado à medida que o tempo passa, atingindo um estado final "estável" no qual seriam observadas moléculas se deslocando em todas as velocidades possíveis. Os conceitos de ordem e desordem molecular, embora não formalmente definidos, podem ser inferidos pelos exemplos e situações hipotéticas descritas no capítulo: a ordem seria o estado advindo de propriedades dinâmicas comuns entre as moléculas do gás, principalmente velocidade e direção de movimento[124]. A primeira etapa da argumentação de Boltzmann consiste na tentativa de provar que tal estado desordenado e irreversível é um estado possível no sistema em questão. A segunda etapa consiste em demonstrar que este estado "Maxwelliano" é o único possível: este é o objetivo do Teorema-H.

---

123 L. Boltzmann, o*p.cit.*, p. 36.

124 *Ibid.*, pp. 40-1.

2.2.1. O Teorema-H

O Teorema-H se inicia com a declaração de seu objetivo: provar que um gás que tenha atingido a distribuição de velocidades de suas moléculas de acordo com a lei de Maxwell permaneceria neste estado indefinidamente, sendo que as colisões entre as moléculas não mais alterariam a distribuição de velocidades. Boltzmann parece tê-lo chamado de Teorema do Mínimo quando de sua publicação, mas o nome pelo qual a demonstração passou a ser referenciada e conhecida leva o nome da grandeza que mede a desordem molecular de um gás (nas condições já apresentadas). O autor utiliza a letra *H* para nomear uma grandeza que soma um aspecto particular de todas as moléculas do gás:

> *Devemos agora calcular a soma H de todos os valores das funções logarítmicas correspondentes num dado momento a todas as moléculas-m e moléculas-m1 contidas num elemento de volume*[125].

Boltzmann emprega as equações formuladas até então para somar as variações em *H* decorrentes dos 3 tipos de colisões possíveis: o das moléculas-*m* entre si (que ocorre em ¼ dos casos), o das moléculas-*m1* entre si (que também ocorre em ¼ dos casos), e entre as moléculas-*m* e moléculas-*m1* (que ocorre em 1/2 dos casos). Ele verifica que a fórmula resultante é composta por elementos que têm sempre o mesmo sinal e que sempre aumentam quando seus argumentos aumentam. Desta forma, a quantidade *H*, por somar 3 termos de sinal negativo que sempre aumentam, só pode diminuir. Esta grandeza poderia permanecer constante ao longo do tempo caso o gás estivesse em estado estacionário. Conclui, finalmente, a partir da assunção de que o estado de

---

[125] *Ibid.*, p. 50.

distribuição de velocidades molecularmente desordenado não se altera uma vez atingido, que *H* só pode decrescer e que a distribuição de velocidades no gás deve convergir para aquela preconizada por Maxwell[126].

Boltzmann faz, a seguir, algumas considerações sobre os significados matemático e físico da quantidade *H*, incluindo algumas questões epistemológicas. Conclui, após extensa argumentação, exemplificação e analogias, que a fato de *H* somente decresce é uma condição probabilística: o aumento da ordem molecular de um gás não seria impossível, mas altamente improvável. Assim sendo, o gás, partindo de um estado menos provável, tenderia a lentamente assumir um estado mais provável. Quanto às simplificações que teria usado em seus modelos, procura justificá-las como necessárias para se conseguir o entendimento dos processos, diferenciando-as de meras omissões que introduziriam erros nos cálculos. Admite que o uso de equações diferenciais sejam fórmulas de aproximação (não seriam corretas porque as quantidades envolvidas não seriam infinitas, embora muito grandes), o que seria compensado pela vantagem de se obter uma melhor percepção. Quanto ao significado físico de *H*, Boltzmann associa-o ao negativo da entropia, relacionado com a segunda lei da Termodinâmica, à qual atribui caráter probabilístico[127].

Na seção final do *Lectures*, Boltzmann tece extensas considerações a respeito das leis físicas deduzidas ao longo da obra, das quais selecionaremos algumas que mais se relacionam com o tema deste trabalho. Boltzmann reafirma de forma sintética e articulada as premissas, demonstrações e

---

126 *Ibid.*, pp. 50-5.

127 *Ibid.*, pp. 74-5.

conclusões previamente expostas, acrescentando considerações de ordens físicas e filosóficas. O autor parece acreditar que seu modelo mecânico pode ser extrapolado do laboratório ao universo, como se pode observar na seguinte passagem:

> *Para explicar o fato de que os cálculos baseadas nestes pressupostos correspondem a processos realmente observáveis, deve-se assumir que um sistema mecânico enormemente complicado representa uma boa imagem do mundo, e que todas as partes, ou pelo menos a maioria delas, ao nosso redor estão inicialmente num estado muito ordenado e, consequentemente, muito improvável. Quando este é o caso, e sempre que duas ou mais partes dele interagem, o sistema formado por estas partes é inicialmente um estado ordenado, que por si mesmo vai rapidamente proceder para o estado desordenado mais provável*[128].

O autor reafirma que a transição de um estado desordenado para ordenado é apenas extremamente improvável: num sistema fechado e com quantidade finita de moléculas, o sistema poderia voltar a um estado ordenado após um intervalo de tempo inconcebivelmente longo – o que deve ser tomado não como uma refutação, mas como uma confirmação da teoria. Segundo seus cálculos, gases misturados num recipiente de volume 100 ml, só manifestariam alguma ordem perceptível após um intervalo de aproximadamente $10^{10000000000}$ anos, o que em termos práticos significaria “nunca”. Observa que, num espaço de tempo tão longo, eventos muito improváveis poderiam acontecer e se repetir, como, por exemplo, todos os edifícios do mundo se incendiando num mesmo dia, ou todos os habitantes de um país morrendo por causas acidentais num mesmo dia: comenta, de forma bem-humorada, que as seguradoras sempre tiveram bons resultados ignorando estes riscos[129].

---

[128] *Ibid.*, p. 443.

[129] *Ibid.*, p. 443-4.

2.2.2. Crítica à Fenomenologia

Sobre a irreversibilidade dos processos naturais, Boltzmann expõe suas reservas ao analisar os procedimentos da Termodinâmica Geral: considera que a irreversibilidade é inferida da observação, e adverte que o "dogma" da auto-suficiência da fenomenologia deve ser evitado, uma vez que as possibilidades de observação são extremamente restritas (inclusive em termos de tempo). Esta posição, segundo o autor, deve ser entendida com os devidos cuidados: não se trataria de ignorar os resultados da observação, mas de reconhecer suas limitações[130]. A auto-suficiência da fenomenologia que Boltzmann combatia era sua defesa contra as críticas de Ernst Mach, com quem manteve um tenso relacionamento acadêmico. Mach defendia que se deveria embasar qualquer conhecimento em evidências experimentais, e rejeitava a teoria cinética dos gases por ela se basear na existência de átomos, cuja existência não poderia ser comprovada pela observação; o segundo argumentava que a rejeição da natureza atômica em favor da natureza contínua da matéria era uma questão de escolha entre duas hipóteses que deveria ser decidida em prol daquela que oferecesse os melhores resultados[131]. É interessante notar o claro engajamento de ambos aos ideais positivistas, embora o fizessem com interpretações divergentes.

Ao longo do *Lectures*, Boltzmann desenvolve uma síntese do pensamento científico de seus contemporâneos sobre a Teoria do Gás, à qual acrescenta, desenvolve e demonstra idéias e conceitos próprios. Sua posição científico-filosófica é sempre colocada de forma clara: ele busca uma explicação mecânica

---

130 *Ibid.*, p. 445-6.

131 J. Uffink, "Boltzmann's Work in Statistical Physics".

para os fenômenos termodinâmicos baseada na concepção de que a natureza da matéria é corpuscular. Sua abordagem da Termodinâmica é estatística, baseado no fato de que é impossível mapear cada partícula individualmente devido à imensa quantidade delas. Tenta generalizar suas teorias de forma a explicar o funcionamento do universo, e propõe uma revisão criteriosa do que considera "dogmático". Embora intensamente criticado em sua época, os princípios propostos por Boltzmann ainda se fazem presentes na Mecânica Quântica e tiveram influência importante numa nova área de conhecimento criada no século XX: a Teoria da Informação. De fato, sua explicação de que "tudo no universo tende a se transformar no sentido do estado menos provável para o mais provável" é ao mesmo tempo simples, intuitiva e profunda, e tem sido amplamente utilizada para explicar fenômenos da Física Quântica e na Cosmologia.

2.2.3. A Transição para o Estado Mais Provável

Utilizando uma linha argumentativa semelhante à que ele próprio usa, poderíamos tecer considerações epistemológicas sobre a questão da "transição do estado menos provável para o mais provável": esta proposição parece conter certa obviedade que (talvez por isto mesmo) escapa à percepção comum. Sua capacidade preditiva é de amplitude máxima, embora exija que a determinação do estado mais provável de fenômenos específicos seja resultado do processo de verificação de modelos teóricos através de experimentações. Esta visão de mundo desencadeou no século XX um intenso debate entre os partidários da Mecânica Quântica e os Físicos Relativistas: enquanto os primeiros aceitam que os processos estatísticos como inerentes ao universo, os segundos rejeitam esta

idéia, numa posição que pode ser sintetizada na conhecida frase de Einstein manifestando sua crença de que "Deus não joga dados"[132]. De qualquer forma, a ciência contemporânea parece cada vez mais ligada à incerteza inerente dos modelos probabilísticos, o que se pode verificar pela crescente popularização do conceito em textos científicos, inclusive de divulgação[133]. Atualmente, a Entropia é calculada pela fórmula de Boltzmann-Planck:

$$S = k \ln P$$

onde k é igual a 1,38 x 10^{-16} erg/°C (chamada de constante de Boltzmann) e P é a quantidade de configurações discretas possíveis[134].

A nova abordagem introduzida por Boltzmann no estudo dos gases foi uma importante influência na Física do século XX pela introdução da descontinuidade da matéria e pela formulação estatística[135]. A Teoria da Informação também foi diretamente influenciada, uma vez que ela emprega tratamento estatístico em sua formulação e usa a entropia para definir quantidade de informação. Para alguns autores, entretanto, as relações entre a Física e a Teoria de Informação não se limitam ao compartilhamento de princípios e fórmulas. De fato, cientistas e pensadores como Wiener, Brillouin e

---

132 A origem da controvérsia decorre da interpretação formulada por Niels Bohr e Werner Heisenberg, conhecida como Interpretação de Copenhagen. Mais detalhes (inclusive sobre a retórica utilizada no debate, com a resposta de Bohr a Einstein para que ele "não dissesse a Deus o que fazer"), consultar: http://en.wikipedia.org/wiki/Copenhagen_interpretation. Para uma breve introdução ilustrada aos fenômenos em discussão, consultar: http://en.wikipedia.org/wiki/Bohr-Einstein_debates. Para conhecera opinião de Schenberg sobre alguns aspectos paradoxais da interpretação estatísca, consulta: M. Schenberg, *op.cit.*, pp. 161-2.

133 Como se sabe, a física quântica e a cosmologia lançam mão de recursos como o "colapso da função de onda" para exprimir probabilidades em todos os níveis, de partículas subatômicas ao universo. A popularização do conceito em obras de divulgação pode ser constatada em S. Hawking, *O universo numa casca de noz*, pp. 106-29, em M. Kaku, *Hyperspace*, pp. 254-65 e em J. D. Barrow, *Theories of Everything*, pp. 86-92.

134 L. Brillouin, *op. cit.*, pp. 119-20.

135 M. Schenberg, *op.cit.*, p. 146. Schenberg lembra que Boltzmann escreveu a Planck aconselhando-o a considerar a descontinuidade na solução de problema do corpo negro, sem a qual seria impossível resolve-lo. Aceitando o conselho, Planck chegou à solução do problema, e mais tarde utilizou o mesmo princípio na criação da mecânica quântica.

Ifrah propõem que existe um nível de interação mais profundo unindo a informação às dimensões e propriedades já conhecidas da matéria e da energia. Entretanto, antes de explorar tais possibilidades, faz-se necessário analisar com maior profundidade a própria Teoria da Informação nas visões de Shannon e Wiener.

2.3. Shannon e a Teoria Matemática da Comunicação

Para Shannon, a questão fundamental da comunicação é a capacidade de se reproduzir remotamente uma mensagem, com o grau de fidelidade desejado. Neste sentido, ele ignora o aspecto semântico da mensagem em favor de uma abordagem quantitativa. Não interessa à teoria qual o significado da mensagem: o que importa é que cada mensagem é o resultado de uma escolha entre um conjunto de mensagens possíveis, ou seja, de um conjunto de mensagens possíveis, apenas uma foi selecionada[136]. Convém notar que, apesar desta abordagem ser melhor entendida e aplicada a canais discretos, seus princípios também são aplicáveis aos canais contínuos[137]. Antes de prosseguir, é necessário caracterizar e distinguir os dois tipos.

Entende-se por *Canais Discretos* os sistemas de comunicação que operam com um conjunto finito de símbolos, que mantém correspondência bi-unívoca com um conjunto arbitrário de representação. O telégrafo, o teletipo, o telex e todas as formas de transmissão digital são exemplos de canais discretos. No caso do telégrafo, os símbolos são as letras do alfabeto latino e os algarismos indo-arábicos conforme definidos pelo Código Morse, e são formados por pontos (sinais curtos) e traços (sinais longos), e separados por pausas ("espaços") entre letras e entre palavras. Desta forma, cada combinação de pontos e traços entre duas pausas deve corresponder à codificação de um símbolo no Código Morse, ou não será um símbolo válido. No caso da transmissão digital por modem, sabe-se que é usual empregar dígitos binários

---

136 C. Shannon, "The Mathematical Theory of Communication", *in* C. Shannon & W. Weaver, *op. cit.*, p. 31.

137 *Ibid.*, p. 81.

agrupados em octetos ou bytes (oito dígitos binários), cujas combinações correspondem a símbolos definidos conforme a Tabela ASCII.

Canais Contínuos, por outro lado, são sistemas de comunicação que prescindem de um conjunto de símbolos, podendo operar com qualquer valor ou intensidade dentro de uma faixa de valores ou intensidades determinadas. A ausência de um repertório finito de símbolos implica que existem infinitos valores possíveis dentro dos limites pré-determinados; esta é a característica principal das transmissões analógicas, ou seja, aquelas cuja interpretação depende do estabelecimento de uma analogia entre diferentes escalas contínuas, a real e a de representação, através da intensidade do sinal transmitido[138]. A reprodução de uma música num disco de vinil é um exemplo de transmissão analógica, uma vez que inexiste um repertório finito e arbitrário de códigos na gravação. A teoria de Shannon promove a "discretização"[139] do sinal contínuo através da divisão do *continuum* de mensagens e sinais em numerosos grupos de pequenas regiões, que, à medida que a amplitude das regiões diminui e o número de grupos aumenta, os resultados obtidos se aproximam dos resultados da continuidade[140]. Também existem Canais Mistos, que combinam processos discretos e contínuos, como o PCM[141] e Transmissões de Televisão com *Closed-Captioning*[142].

---

138 R. Tenório, *Cérebros e Computadores*, p. 69.

139 *Ibid*., pp. 69-70.

140 C. Shannon, "The Mathematical Theory of Communication", *in* C. Shannon & W. Weaver, *op. cit*., p. 81.

141 Pulse-Code Modulation é uma técnica de digitalização de sons usada nas em transmissões e gravações digitais. Para detalhes sobre PCM, consultar: http://www.webopedia.com/TERM/P/PCM.html e http://en.wikipedia.org/wiki/PCM.

142 As legendas são codificadas digitalmente e transmitidas durante o intervalo de reposicionamento vertical do sinal analógico, chamado de VBI (Vertical Blanking Interval). Para detalhes sobre *Closed-Captioning*, consultar http://en.wikipedia.org/wiki/Closed_caption, e para detalhes sobre VBI consultar: http://en.wikipedia.org/wiki/Vertical_blanking_interval.

2.3.1. O Problema de Engenharia

A partir do conhecimento das diferenças entre o analógico e o digital, seria lícito perguntar o porquê da digitalização, uma vez que o analógico pode parecer mais vantajoso que o digital uma vez que contempla infinitas possibilidades, e, portanto, que poderia representar qualquer coisa mais fielmente. A resposta não se encontra na TMC, mas nos é oferecida por Viterbi: no analógico não há como se distinguir entre o sinal e o ruído, pois o ruído encontra-se necessariamente dentro do *continuum* de valores válidos para um sinal[143].

Os sinais elétricos, ao serem transmitidos num meio físico, são susceptíveis à ocorrência fortuita de ruído, além de sofrer atenuação, ou seja, a diminuição de sua intensidade devido à resistência do meio. Amplificar o sinal ao longo do meio de transmissão significa ampliar o ruído agregado a ele até então, sendo que após ser amplificado o sinal continua susceptível a ruídos no resto da trajetória. A cada amplificação intermediária do sinal analógico, a relação sinal/ruído tende a se degradar. Também não se pode ampliar a potência do transmissor indefinidamente[144]. A digitalização resolve o problema da amplificação do ruído: na transmissão binária só se reconhecem dois níveis de intensidade, e nem todo ruído será intenso o suficiente para fazer com que um nível de intensidade se transforme no outro, uma vez que os valores admitem certo grau de tolerância. Ao receber um sinal com ruído, o receptor deverá traduzi-lo num dos 2 valores válidos. Se este valor tiver sido traduzido

---

143 UCSD-TV, *Claude Shannon, Father of the Information Age*, depoimentos de Jack Keil Wolf, índices de tempo 08:20 a 09:03 e 12:20 a 13:00.

144 *Ibid.*

erroneamente em conseqüência do ruído, ele deverá ser tratado pelos algoritmos de detecção e correção de erros.

Quanto à medida da quantidade de informação e grandezas associadas, Shannon optou pela escala logarítmica. Entre as vantagens mencionadas na justificativa desta escolha, destaca-se a conveniência de se empregar uma escala que freqüentemente surge nos teoremas da Teoria da Informação, pois está relacionada com a quantidade de sinais necessários para se codificar determinado repertório de símbolos. Sendo o *n* a quantidade de símbolos do repertório, e *m* a quantidade de sinais distintos, a quantidade de sinais necessária para codificar um símbolo é $\log_m n$. Na base binária, cuja unidade é o *bit* (*b*inary dig*it*), a quantidade de sinais necessária para codificar um repertório de *n* símbolos é $\log_2 n$. Exemplificando: para se codificar os algarismos indo-arábicos, cujo repertório é de 10 símbolos discretos, são necessários $\log_2 10$ bits, ou seja, 3,3 bits. De fato, com 3 bits é possível representar 8 símbolos ($2^3=8$), e com 4 bits é possível representar 16 símbolos ($2^4=16$). Por analogia, para se codificar o alfabeto latino, que é um repertório de 26 símbolos discretos, são necessários $\log_2 26 = 4{,}7$ bits ou $\log_{10} 26 = 1{,}4$ algarismos. Outra vantagem citada por Shannon se relaciona às propriedades matemáticas dos logaritmos, cuja mudança de base é feita por $\log_a x = \log_b x / \log_b$[145].

2.3.2. Modelo de um Sistema de Comunicação

Shannon propõe um modelo de comunicação composto por uma Fonte de Informação (uma pessoa ou uma máquina), que envia mensagens ao

---

145 C. Shannon, "The Mathematical Theory of Communication", *in* C. Shannon & W. Weaver, *op. cit.*, p. 31.

Transmissor, que por sua vez transforma as mensagens em sinais através de um processo de codificação para, em seguida, enviá-las por um Canal (um meio físico qualquer). O sinal é recebido por um Receptor (uma pessoa ou uma máquina), que decodifica os sinais em mensagens, enviando-os ao Destino. O processo de transmissão está sujeito a interferências de diversas naturezas, e que podem corromper os sinais enviados: estas interferências são chamadas de Ruídos.

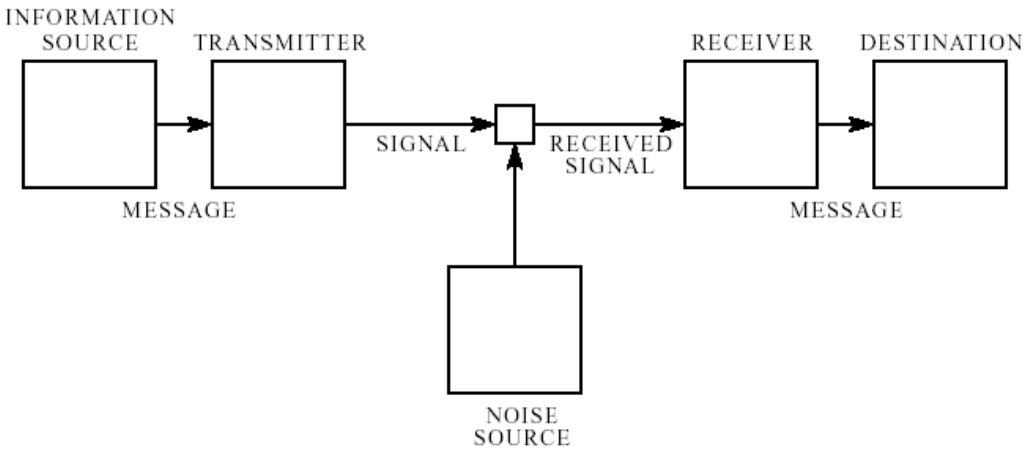

Fig. 7 - Diagrama Esquemático de um Sistema de Comunicação

Partindo deste modelo, Shannon inicia pelo Canal Discreto Sem Ruído por ser o caso geral que servirá de base para o estudo dos demais casos[146].

Shannon define *Capacidade* de transmissão de um canal como a quantidade de sinais transmitidos durante um intervalo de tempo. Ora, a quantidade de sinais depende da base de representação e da quantidade de símbolos do repertório, o que resulta em

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

[146] *Ibid.*, p. 35.

onde N(T) é a quantidade de símbolos permitidos durante o período de amostragem T. A capacidade é usualmente medida em bits por segundo[147]. A seguir, é necessário analisar as características da Fonte de Informação Discreta.

2.3.3. Entropia de Fontes Discretas

A experiência demonstra, afirma Shannon, que o processo de geração de informação é um processo não-determinístico, uma vez que é geralmente impossível saber de antemão o que o canal vai transmitir. A partir deste ponto, Weaver resume o arrazoado de Shannon a este respeito, conforme segue:

> *Um sistema que produz uma seqüência de símbolos (que podem ser letras ou notas musicais ao invés de palavras) de acordo com certas probabilidades é chamado de* ***processo estocástico****, e um tipo especial de processo estocástico no qual as probabilidades dependem dos eventos anteriores é chamado de* ***processo de Markoff*** *ou corrente de Markoff. Dos processos de Markoff que podem de alguma forma gerar mensagens, existe uma classe especial que é de importância primária para a teoria da comunicação, sendo estes chamados de* ***processos ergódicos****. (grifos do original)*[148]

Shannon prossegue, explicando que um processo ergódico deve ser analisado de maneira probabilística, ou seja, considerando em conjunto as probabilidades de ocorrência dos eventos possíveis. Assim sendo, a quantidade de informação gerada por um processo ergódico deve somar todas as probabilidades *p* de ocorrência de evento, multiplicado pelo seu próprio logaritmo. Esta medida é expressa matematicamente por:

---

147 *Ibid.*, pp. 35-6.

148 W. Weaver, "Recent Contribuions to The Mathematical Theory of Comunication", *in* C. Shannon & W. Weaver, *op. cit.*, p.6. Citação no original: "A system which produces a sequence of symbols (which may, of course, be letters or musical notes, say, rather than words) according to certain probabilities is called a stochastic process, and the special case of a stochastic process in which the probabilities depend on the previous events, is called a Markoff process or a Markoff chain. Of the Markoff processes which might conceivably generate messages, there is a special class which is of primary importance for communication theory, these being what are called ergodic processes".

$$H = -K \sum_{i=1}^{n} p_i \log p_i$$

onde *K* é uma constante arbitrária que expressar a unidade de medida. A demonstração é fornecida no apêndice do livro, e decorre dos pressupostos estatísticos assumidos. Esta medida calcula a quantidade média de informação por sinal de um repertório, o que pressupõe que a soma de todas as probabilidades p_i é igual a 1. Epstein oferece uma explicação complementar: esta medida nada mais é do que a média da a*uto-informação* de cada símbolo do repertório (log p_i) ponderada por sua freqüência em determinada língua (p_i)[149].

Shannon nota que expressões desta forma (soma do produto de probabilidades por seus logaritmos) têm papel fundamental Teoria da Informação, e que esta forma da quantidade *H* seria reconhecida como a da entropia da mecânica estatística, como no Teorema-H de Boltzmann. Assim, o termo entropia passa a designar a quantidade de informação, que é representada por *H* na TMC[150].

O caso mais simples que se pode conceber é a ocorrência de um evento com duas possibilidades, cujas probabilidades são *p* e *q* = 1 - *p*. A entropia deste sistema (repertório) é

$$\boldsymbol{H = -\,(\,p \log_2 p \;+\; q \log_2 q\,).}$$

---

149 Epstein, *op. cit.*, pp. 45-9.

150 C. Shannon, "The Mathematical Theory of Communication", *in* C. Shannon & W. Weaver, *op. cit.*, pp. 48-50. Contudo, é importante notar que o Teorema-H não trata da entropia, e sim da prova da lei da distribuição de velocidades de Maxwell, onde Boltzmann relaciona *H* à entropia, sem, contundo, estabelecer qualquer igualdade. A fonte indicada por Shannon para justificar esta afirmação não vem de Boltzmann, mas de uma obra de 1938 de R. Tolman. Por similaridade, pode-se dizer que a entropia de Shannon deriva da fórmula de um dos conceitos "análogos de entropia" de Gibbs. Para mais detalhes, consultar: K. Denbigh, "How subjective in entropy?", pp. 110-1. Ainda, alguns autores se referem à entropia com "Entropia de Boltzmann-Gibbs-Shannon", como é o caso de Pereira & Stern em *Inferência Indutiva com Dados Discretos: Uma Visão Genuinamente Bayesiana*, p. 62.

Este caso pode ser melhor entendido se representado graficamente, conforme segue:

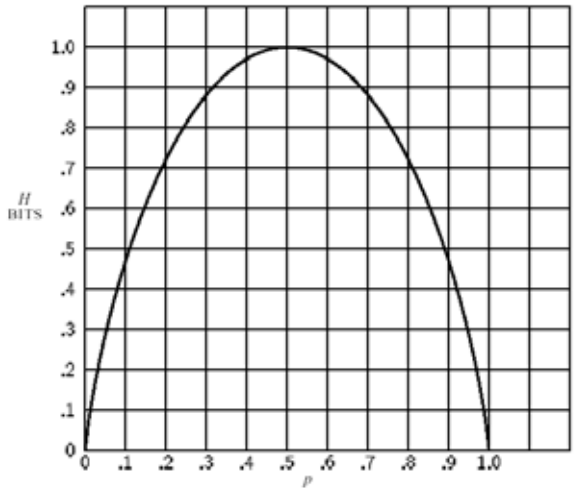

Fig. 8 - Variação da Entropia no caso de duas possibilidades com probabilidades *p* e (1-*p*)

Shannon observa que a entropia máxima ocorre quando as possibilidades são equiprováveis: quanto maior for a diferença entre as probabilidades, menor será a entropia, e este é a situação de maior incerteza. No caso particular de uma das probabilidades ser 100%, a entropia será zero[151]. Weaver nota, em complemento a esta idéia, que a entropia também aumenta quanto maior for da quantidade de possibilidades, em outras, ou seja, tomando-se dois conjuntos de possibilidades equiprováveis, a entropia será maior no conjunto de possibilidades mais numerosas[152].

Ele prossegue, afirmando que os eventos de uma Fonte de Informação Discreta são os símbolos que ela gera, e define a *Entropia de uma Fonte Discreta* por

$$H' = \sum_i f_i H_i$$

---

151 *Ibid.*, pp. 50-1.

152 W. Weaver, "Recent Contributions to The Mathematical Theory of Communication", p.15-6.

onde f_i é a freqüência média do estado *i*, e relacionando-a com a Entropia pela equação

$$H' = mH$$

onde *m* é o número médio de símbolos produzidos por segundo. A seguir, o autor deduz a fórmula da entropia para longas seqüências de símbolos independentes, que é válida para qualquer fonte: toma-se uma seqüência composta de *N* símbolos, e p_i a probabilidade do símbolo *i*, com entropia $-\Sigma\, p_i \log p_i$. Esta seqüência terá alta probabilidade de conter ocorrências de cada símbolo proporcionais às suas probabilidades individuais, ou seja, p_iN. Assim sendo, a probabilidade desta seqüência em particular é

$$p = p_1N \times p_2N \times \ldots \times p_nN.$$

Daí, tem-se

$$\log p \doteq N \sum_i p_i \log p_i$$
$$\log p \doteq -NH$$
$$H \doteq \frac{\log 1/p}{N}.$$

A entropia da fonte discreta é, aproximadamente, o logaritmo do inverso da probabilidade de uma longa seqüência típica de símbolos, dividido pela quantidade de símbolos que contém[153].

Adiante, Shannon conceitua *Entropia Relativa* como sendo a razão entre a entropia de uma fonte com o valor máximo que ela poderia apresentar estando restrita ao mesmo repertório de símbolos. Este valor corresponde ao máximo de Compressão possível na codificação do repertório, entendendo-se por

[153] C. Shannon, *op. cit.*, pp. 53-4.

compressão a técnica de codificar o repertório de maneira usar o canal de forma mais eficiente. A *Redundância* é o complemento da entropia Relativa, e como tal é calculada subtraindo-se a entropia relativa de um. A redundância corresponde aos símbolos mais freqüentes do repertório, ou seja, aqueles de maior probabilidade de ocorrência e, conseqüentemente, de menor entropia.

2.3.4. Codificação Eficiente, Redundância e Compressão

Neste ponto, Shannon está pronto para demonstrar o Teorema Fundamental do Canal Sem Ruído. Ele prova que, com a codificação mais eficiente possível, a entropia determina a capacidade do canal. Por este motivo, é impossível transmitir a uma taxa média superior a **C/H**, o que pode ser traduzido matematicamente como $\boldsymbol{H'/H \leq C/H}$. Para o entendimento adequado desta inequação, é necessário lembrar que a capacidade do canal é a quantidade de sinais transmitidos por segundo, e não a quantidade de informação transmitida por segundo (que é a própria definição de entropia por segundo): o canal transmite tanto entropia como redundância na forma de sinais. O Shannon prova é que, em condições especiais, é possível transmitir informação à taxa da capacidade do canal, o que ocorrerá quando a mensagem não contiver redundância[154].

A forma de reduzir a redundância das mensagens é a *Compressão*. Shannon sugere alguns algoritmos, sendo que o de entendimento mais intuitivo é o seguinte: arranjar as mensagens de extensão *N* em ordem decrescente de probabilidade, e atribuir a elas um número binário crescente. As mensagens mais prováveis serão representadas por números binários mais curtos enquanto

[154] *Ibid.*, p. 58-61.

que as menos prováveis serão codificadas por números mais longos, o que provocará o encurtamento da transmissão. A expansão será realizada pelo Receptor, que realizará o processo inverso. Obviamente, é necessário que tanto o Transmissor como o Receptor tenham capacidade de armazenamento de sinais ("memória") para que a compressão e a expansão possam ser realizadas[155]. Shannon nota que a codificação dos sinais em função distribuição de freqüências de suas ocorrências numa determinada língua já foi empregada de forma limitada quando da criação do Código Morse: a letra *E*, que é a mais freqüente na língua inglesa, é codificada por apenas um ponto, enquanto que letras menos freqüentes como Q, X e Z são representadas por seqüências mais longas de pontos e traços. O autor prossegue, lembrando que a idéia de economizar tempo e capacidade de canal (bem como pagar menos pelas transmissões telegráficas) levou ao uso de abreviações de palavras ou frases mais comumente transmitidas[156]. O assunto será retomado adiante, ao aprofundarmos a análise das abreviações e compressões sintáticas.

Shannon afirma que é impossível a reconstrução fidedigna das mensagens transmitidas num canal ruidoso, uma vez que a ocorrência de ruídos é aleatória e não há como saber quais partes do sinal foram afetadas. Ele formula, na seqüência, o conceito de *Equivocação* (ou Entropia condicional), definindo-a como a ambigüidade média do sinal. A única forma de eliminar a equivocação é a introdução intencional de redundância[157]. Ele descreve um algoritmo criado por outro pesquisador para detecção e correção de erros, embora reconheça suas limitações, uma vez que este método introduz de 3 bits de redundância

---

155 *Ibid.*, pp. 60-1.

156 *Ibid.*, p. 39.

157 *Ibid.*, p. 65-70.

para 4 bits de sinal.[158] Posteriormente, outros pesquisadores criaram técnicas de detecção e correção de erros que não oneram tanto o canal, e que conquanto não sejam infalíveis, trabalham com graus de acerto estatisticamente aceitáveis. Um exemplo deste tipo de algoritmo é o CRC[159], que é amplamente usado em protocolos de redes de computadores.

A redundância está presente em todos os meios de comunicação. A comunicação oral, cujo canal de transmissão é o mecanismo combinado fala e cuja codificação é a linguagem, embute redundâncias semânticas bidirecionais para garantir uma comunicação eficaz. A escrita herda parte da redundância semântica da fala, e ainda carrega a redundância sintática imposta pela gramática. Na escrita, observa Shannon, a redundância da língua inglesa é da ordem de 50%, o que significa que até 50% de seus sinais podem ser removidos sem que a mensagem se torne ininteligível[160]. Este fenômeno pode ser verificado diversos tipos de comunicação escrita, uma vez que parece existir uma tendência e abreviar as expressões mais comumente usadas.

A abreviação é uma forma não-sistemática de eliminar redundância do sinal (a escrita, no caso), e é aceita em maior ou menor grau em todos os tipos de escrita (comercial, acadêmica, informal e técnica, entre outros). Os exemplos são inúmeros, e serão citados apenas alguns poucos e significativos acrônimos e abreviações, com suas respectivas "expansões": 3D (tridimensional), ASAP ("as soon as possible"), CD ("compact disk"), CQD (como queríamos demonstrar),

---

158 *Ibid.*, p. 80.

159 *Cyclic Redundancy Check* é um algoritmo baseado em divisão binária polinomial. Para mais detalhes, consultar http://en.wikipedia.org/wiki/Cyclic_redundancy_check.

160 Este número foi obtido por Shannon, que alega ter chegado ao mesmo resultado que outros pesquisadores obtiveram por diferentes métodos empíricos, como pode ser verificado em C. Shannon, "The Mathematical Theory of Communication", *in* C. Shannon & W. Weaver, *op. cit.*, p. 55-7. Entretanto, Kahn calcula que a entropia da língua inglesa é de cerca de 75%. Para mais detalhes, consultar: D. Kahn, *op. cit.*, pp. 759-62.

CEP (código de endereçamento postal), FAQ ("frequently asked questions"), GLS (gays, lésbicas e simpatizantes), IPI (imposto sobre produtos industrializados), PS ("post scriptum"), Radar ("radio detection and ranging"), RSVP ("répondez s'il vous plait"), TV (televisão) e USB ("universal serial bus"). A existência de abreviações em latim indica que o fenômeno não é recente. De fato, línguas antigas como a hebraica permitem a compressão sintática através da eliminação das vogais e manutenção das consoantes[161], como Olson descreve:

> *As línguas semíticas têm a interessante propriedade de veicular as identidades léxicas através daquilo que chamamos de consoantes. O que conhecemos como vogais é usado apenas para inflexões*[162].

e, a seguir, explica:

> *Como as vogais só contêm informação gramatical, e não léxica ou morfêmica, alguns sistemas de escrita semíticos nunca desenvolveram qualquer recurso para representá-las*[163].

Ainda, lembrando que Umberto Eco ensina que "a linguagem da tese é uma *meta-linguagem*"[164], é possível introduzir uma auto-referência neste próprio texto para indicar alguns casos de compressão: nas notas de rodapé, por exemplo, cada "Ibid." é uma abreviação "Ibidem", que já é uma forma de compressão, uma vez que evita que a referência a uma obra recém-citada seja repetida.

Ao que parece, a compressão sintática é foi um fenômeno duradouro e abrangente cuja explicação matemática demorou séculos para surgir. A Teoria da Informação, embora seja indiferente ao significado da comunicação, oferece

---

161 Neste caso, as ambigüidades são eliminadas pela análise do contexto das frases. Uma breve descrição do sistema de escrita da língua hebráica pode ser encontrada em: http://www.jewishvirtuallibrary.org/jsource/Judaism/alephbet.html.

162 D. Olson, *op. cit.*, p. 99.

163 *Ibid*.

164 U. Eco, *Como se faz uma tese*, p. 166.

modelos para o entendimento em níveis além de suas pretensões. Sobre o significado da comunicação, Weaver sugeriu em 1949 que seria possível a aplicação dos princípios da Teoria da Informação num nível semântico[165], mas esta hipótese ainda não foi satisfatoriamente comprovada.

Shannon utilizou os conceitos de Mecânica Estatística desenvolvidos por Boltzmann e aplicou na análise dos sistemas de comunicação. Medir o grau de aleatoriedade de um sistema talvez seja mais bem compreendido como determinar o grau de liberdade do mesmo, ou seja, a forma com que mensagens se distanciam dos padrões encontrados. Notadamente nos estudos onde a quantidade de dados é muito grande, a aplicação dos conceitos da Teoria da Informação facilita a identificação de informação genuína dos padrões estruturais.

A Ciência tem por objetivo a compreensão da Natureza, o que é conseguido através da observação e da aquisição de dados para que se possa formular e testar predições. Neste aspecto, o reconhecimento de padrões é de suma importância para transformar os dados coletados em informações inteligíveis, caso contrário não haveria sentido em elaborar listas de dados sem sentido. O reconhecimento de padrões permite que se substitua vasta quantidade de dados por *fórmulas* que permitam representar coletivamente todas as medições efetuadas, com maior ou menor grau de precisão. O princípio que permite que isto aconteça é a *Compressibilidade Algorítmica*. Uma seqüência de medições de um fenômeno que não seja completamente aleatória deve ter uma forma de representação que seja menor que a própria seqüência de dados, ou seja, deve

---

[165] W. Weaver, "Recent Contributions to the Mathematical Theory of Communications", *in* C. Shannon & W. Weaver, *op. cit.*, pp. 24-8.

ser algoritmicamente compressíveis[166]. Não parece ser casual que a Teoria da Informação tenha desenvolvido ferramentas para a aplicação sistemática deste princípio no que tange à informação: seu objetivo é o mesmo da Ciência, apesar de possuírem métodos diferentes.

A Teoria da Informação parece ter inaugurado uma nova abordagem para a compreensão de fenômenos das mais diversas áreas. De fato, Schenberg assinala que talvez a informação seja uma propriedade física fundamental ainda não completamente entendida, com potencial para provocar uma revolução na Ciência[167]. Isto pode ser verificado pela influência que a Teoria da Informação já exerce em Física, Genética, Química e na própria Matemática e Estatística de onde se originou[168].

Para facilitar o entendimento das idéias de Boltzmann e Shannon a respeito da entropia, o que é importante para a análise das idéias de Wiener (e de físicos que acreditam na existência de uma relação íntima entre informação e entropia), encontra-se a seguir um quadro comparativo que resume conceitos e conseqüências segundo o pensamento de cada autor.

---

166 J. Barrow, *op. cit.*, p. 14-5.

167 M. Schenberg, *op. cit.*, pp. 99-100.

168 S. Verdú, "Fifty Years of Shannon Theory", *in* S. Verdú & S. McLaughlin, orgs., *op. cit.*, pp. 26-7.

ENTROPIA		
Autor	Boltzmann	Shannon
Objeto	Gases	Informação
Conceito	Desordem Molecular	Quantidade de Informação
Medida	Distribuição de Velocidades das Moléculas	Probabilidade de Ocorrência dos Símbolos
Variação	Aumenta à medida que o movimento das moléculas se torna aleatório.	Aumenta à medida que a ocorrência dos símbolos se torna aleatória (a distribuição se aproxima à eqüiprobabilidade).
Tendência	Num sistema fechado, aumenta até atingir a situação de equilíbrio (distribuição de Maxwell-Boltzmann)[169].	Numa língua natural, a freqüência de símbolos tende à distribuição de Gibbs[170].

[169] Devem-se considerar as restrições do modelo teórico adotado por Boltzmann, bem como as limitações do método de cálculo.

[170] O resultado esperado deve considerar as restrições do modelo adotado por Shannon, que assume que os textos produzidos na língua inglesa são adequadamente descritos como processos ergódicos. É importante notar que, na prática, os sistemas de comunicação não transmitem apenas texto na língua inglesa (também transmite, por exemplo, por exemplo, números codificados e dados analógicos digitalizados) o que tende a desviar os resultados reais da distribuição teórica.

2.4. Wiener e a Entropia na Cibernética

Em seu livro *Cybernetics*, Wiener aborda uma vasta gama de assuntos ligados ao controle e à comunicação. Como já foi mencionado, ele liderava um grupo multidisciplinar que tinha como objetivo desenvolver aplicações militares, sendo que um dos pressupostos era de que se poderia usar o comportamento dos animais para modelar o funcionamento de máquinas. *Cybernetics* talvez possa ser mais apropriadamente descrito como um amplo painel multidisciplinar, ao qual Wiener acrescenta formulações matemáticas e reflexões sobre filosofia e ciência. Ao contrário do que fez Shannon na TMC, Wiener não adota o modelo de demonstração de teoremas.

Na introdução do livro, Wiener declara seu interesse pelo processo de comunicação, o qual acredita estar baseado não em técnicas de engenharia, mas numa noção abrangente de "mensagem" que independe do meio de transmissão, que poderia tanto ser elétrico, mecânico, ou nervoso.

> *A mensagem é uma seqüência discreta ou contínua de eventos mensuráveis no tempo – exatamente o que os estatísticos chamam de série temporal*[171].

Seguindo sua argumentação, descobre-se que ele acredita poder transformar o projeto de engenharia de comunicação numa ciência estatística derivada da mecânica estatística. Ele considera que a mecânica estatística está entranhada em todos os ramos da ciência há mais de um século, e acredita que o domínio desta disciplina na Física poderá ter importância capital na

---

[171] N. Wiener, *Cybernetics, or control and communication in the animal and the machine*, pp. 8-9. Citação no original: "The message is a discrete or continuous sequence of measurable events distributed in time – precisely what is called a time series by the statisticians".

interpretação da natureza do tempo[172]. Com relação à comunicação, Wiener declara que Shannon e R. Fischer, além dele próprio, teriam chegado de forma independente e quase simultânea à ideía da quantificação da informação[173]. Wiener declara que a noção de quantidade de informação se associa naturalmente à noção estatística de entropia de Boltzmann. Entretanto, ao defini-la incorre numa dissensão fundamental com a definição de Shannon:

> *Assim como a quantidade de informação num sistema é uma medida de seu grau de organização, a entropia de um sistema é uma medida de seu grau de desorganização; e uma é simplesmente o negativo da outra*[174].

Como foi anteriormente exposto neste capítulo, a entropia de Shannon e a entropia de Boltzmann apresentam semelhanças notáveis quanto a seus conceitos e decorrências. Entropia é o nome atribuído por Shannon à grandeza que mede quantidade de informação, o que é conflitante com a proposta de Wiener. Foram apresentadas evidências de que ambas medem o grau de aleatoriedade de certos fenômenos. Esta aleatoriedade pode ser interpretada como a tendência ao à desordem molecular decorrente da lei da distribuição das velocidades, ou como a quantificação da aleatoriedade dos símbolos. Claramente, o conceito de "informação" de Wiener não é o mesmo que o de Shannon.

A explicação para o conceito diferenciado de "informação" de Wiener se encontra adiante, quando ele descreve um mecanismo hipotético que seria capaz de diminuir a entropia de um sistema isolado, conhecido como "demônio

---

172 *Ibid.*, p. 10.

173 *Ibid.*, pp. 10-1.

174 *Ibid.*, p. 11. Citação no original: "Just as the amount of information in a system is a measure of its degree of organization, so the entropy of a system is a measure of its degree of disorganization: and the one is simply the negative of the other."

de Maxwell". Na experiência imaginada por Maxwell, um ser inteligente diminuto dentro de um recipiente bipartido repleto de um gás perfeito, poderia permitir a passagem de moléculas de um lado para outro baseado na percepção de suas velocidades: ele operaria uma abertura na parede divisória, de forma a permitir que as moléculas mais energéticas ficassem de um lado e as menos energéticas do outro. Esta construção introduzira ordem ao sistema, uma vez que induziria a uma tendência de concentração na distribuição de velocidades das moléculas em função do lado no qual seriam confinadas dentro do recipiente. Esta hipotética violação das leis da termodinâmica encontrou vários "vingadores", entre os quais Leó Szilard: ele argumentou que a variação negativa da entropia era devida à aquisição de informação, sendo por isto considerado o primeiro a estabelecer uma conexão entre a entropia termodinâmica e a informação[175]. Para tanto, Szilard propôs um sistema de uma única molécula, e substituiu o demônio por um dispositivo mecânico para registrar as variações da entropia.

Apesar de não se encontrar menção a Szilard na argumentação de Wiener, sua influência pode ser inferida pela associação entre entropia mecânica e informação e também pela descrição que Wiener oferece sobre o experimento que permitiria a redução da entropia de um sistema: a abertura poderia ser controlada pelo demônio ou por um equivalente mecânico:

> *...[a abertura seria] operada por um porteiro, [que poderia ser] ou um demônio antropomórfico ou um mecanismo diminuto*[176].

---

[175] H. Leff & A. Rex , *Maxwell's Demon: Entropy, Information, Computing*, p. 16.

[176] N. Wiener, *op.cit.*, p. 57. Citação no original: "... [the opening would be] operated by a gatekeeper, either an anthropomorphic demon or a minute mechanism."

Wiener retoma o assunto adiante, ao tratar de séries temporais, informação e comunicação. A princípio, procura associar informação a conceitos probabilísticos:

> *O que é informação, e como ela é mensurada? Uma das formas mais simples e unitárias de informação é o registro de uma escolha entre duas alternativas simples e igualmente prováveis, quando uma ou outra deverá ocorrer*[177].

Esta noção sucinta serve apenas para embasar os argumentos de Wiener a respeito do processo de aquisição de informação[178] e à demonstração de que "a quantidade de informação é essencialmente uma entropia negativa[179]". Se não nos parece correto supor que a entropia da informação mede sua organização enquanto que a entropia de um sistema mede sua desorganização, tampouco parece acertado afirmar que uma seja oposta à outra. Ora, a entropia de Shannon é uma grandeza adimensional positiva, enquanto a entropia de Boltzmann é uma grandeza positiva expressa por Energia dividida por Temperatura (no Sistema Internacional, J/K). Entretanto, ao fazer tal afirmação, Wiener contribuiu para alimentar as expectativas de integração de alguns físicos, como Brillouin: este cita a definição de Wiener e o capítulo do *Cybernetics* onde ela foi formulada num artigo de 1949 sobre as relações existentes entre a vida, a termodinâmica e a cibernética[180]. Poucos anos mais tarde, Brillouin escreveu *Science and Information Theory*, onde retoma estas idéias e as elabora matematicamente.

---

177 *Ibid*., p. 61. Citação no original: "What is information, and how is it measured? One of the simplest, most unitary forms of information is the recording of a choice between two equally probable simple alternatives, one or the other of which is bound to happen."

178 *Ibid*., pp. 61-2.

179 *Ibid*., p. 64. Citação no original: "...[the] amount of information... is essentially a negative entropy."

180 L. Brillouin, "Life, Thermodynamics, and Cybernetics", *in* H. Leff & A. Rex, orgs., *op. cit.*, pp. 101-2.

Tais divergências entre os conceitos fundamentais da Teoria da Informação existem desde sua origem, e são oriundas daqueles que são considerados os fundadores da disciplina. Shannon e Wiener, que se posicionaram contra as extensões da Teoria da Informação em outras disciplinas, certamente não concordavam entre si a respeito dos próprios fundamentos da teoria. Apesar da evidente existência de pontos de contato, a busca de extensões e correlações da Física para a Teoria da Informação (ou vice-versa) não deixa de ser problemática. O conhecimento destas dificuldades e divergências a partir da origem da TI é necessário para que o conflito entre as duas áreas seja adequadamente analisado a seguir.

2.5. O Diálogo Conflituoso entre a TI e a Física

O tipo de conexão existente entre os conceitos da entropia termodinâmica e da informação suscita debates intensos entre os físicos e os teóricos da informação. A origem destes conflitos origina-se numa proposição de Maxwell para tentar reverter a entropia de um sistema isolado. Numa experiência de pensamento, Maxwell propõe o seguinte: 1) um recipiente bipartido com a mesma quantidade de um mesmo gás em ambos os lados; 2) um ser inteligente de tamanho diminuto, controlando uma abertura na parede divisória; 3) este ser teria a capacidade de identificar as trajetórias e o momento das moléculas que fossem colidir com a abertura na parede divisória, e permitiria apenas a passagem das moléculas mais energéticas para um dos lados, e a das menos energéticas para o outro. A proposição de Maxwell, aparentemente, visava apenas evidenciar contradições na suposta irreversibilidade dos fenômenos físicos, sendo que Maxwell não antecipou as conseqüências que a reversão provocaria. Costuma-se argumentar que, com a entropia do sistema diminuindo, o gás de um lado se tornaria progressivamente quente enquanto que o do outro se tornaria cada vez mais frio; também ocorreriam variações de pressão correspondentes em cada um dos lados[181].

Desde então, diversos físicos se dedicaram ao tema, e dentre eles se destacou Leó Szilard, que estabeleceu qual seria a variação da entropia por um processo de aquisição de informação, em função das probabilidades de ocorrência de dois eventos:

---

[181] H. Leff & A. Rex , orgs., *op. cit.*, pp. 3-13.

> *...o valor médio da quantidade de entropia produzida por uma medição (obviamente, neste caso especial independente das freqüências w1, w2 dos dois eventos): S = k log 2*[182].

As idéias de Szilard provocaram entusiasmo nuns, e ceticismo noutros. Entre aqueles que acreditavam na conexão inequívoca da informação com a entropia estava Brillouin. Influenciado pela argumentação de Szilard e também pela cibernética de Wiener, ele desenvolveu um trabalho minucioso para demonstrar seus pontos de vista. Em *Science and Information Theory*, de 1956, ele tenta conciliar as idéias de Shannon com a termodinâmica, o que o leva a formular o Princípio da Negentropia (entropia negativa), que seria o equivalente termodinâmico da informação, que é uma influência direta das idéias de Wiener[183]. Ainda, ele afirma que Shannon teria definido a entropia com um sinal oposto à definição termodinâmica de entropia, o que não corresponde ao que de fato ocorreu: o sinal negativo da fórmula da entropia de Shannon é decorrência da demonstração algébrica do teorema da entropia e que serve para tornar seu valor positivo[184]. É interessante notar que para conciliar a Termodinâmica à Teoria da Informação, Brillouin propõe uma reforma da Termodinâmica (pelo acréscimo do princípio da negentropia) e apenas uma adequação algébrica na quantidade de informação definida pela Teoria da Informação.

---

182 L. Szilard, "On the decrease of entropy in a thermodynamic system by the intervention of intelligent beings", in H. S. Left & A. F. Rex, orgs., *op. cit.*, p. 127. Citação no original: "...the mean value of the quantity of entropy produced by a measurement is (of course, in this special case independent of the frequencies w1, w2 of the two events): $S = k \log 2$."

183 L. Brillouin, *op. cit.*, pp. 152-61.

184 Na fórmula extrai-se o logaritmo de uma probabilidade (que, por definição, é um número fracionário entre 0 e 1). Sabe-se que o logaritmo de um número fracionário é negativo, e que a soma de vários números negativos resulta num número negativo. Desta forma, o sinal negativo da entropia de Shannon apenas transforma o resultado num número positivo (a equação da entropia de Boltzmann-Planck resulta em números positivos, enquanto *H* resulta num número negativo). Para mais detalhes, consultar a demonstração do teorema da entropia em C. Shannon, "The Mathematical Theory of Communication", *in* C. Shannon & W. Weaver, *op. cit.*, pp. 116-8.

As obras de Brillouin e seus partidários tiveram grande aceitação entre os físicos, mas também atraíram críticas, principalmente pela re-interpretação das idéias de Szilard e pela suposta subjetividade com que tratou o conceito de informação. Do lado dos críticos encontrava-se Karl Popper, que criticou o trabalho de Szilard[185] por considerar a conexão entre conhecimento e entropia como "espúria"[186]: a Termodinâmica seria suficiente para oferecer as respostas procuradas por Szilard[187]. Posteriormente, em 1982, Popper reconhece alguns aspectos positivos na busca do "exorcismo" do demônio de Maxwell:

> *Sem dúvida, algumas interessantes analogias foram descobertas, mas a solidez da obra, e especialmente o papel desempenhado pelo demônio de Maxwell, parecem-me altamente questionáveis*[188].

Em contrapartida, Costa de Beauregard & Tribus propuseram que a Termodinâmica poderia ser uma derivação da Teoria da Informação (numa associação de sentido oposto àquela proposta por Wiener):

> *...a entropia da termodinâmica não passa de um caso especial da entropia da teoria da informação*[189].

e concluem:

> *Acreditamos que, embora seja possível, e muito informativo, deduzir a Termodinâmica de uma teoria geral da Informação, o contrário não é possível*[190].

Talvez a crítica mais clara e fundamentada seja a de Denbigh, que retrocede às origens do conceito de entropia para explicar porque a entropia da informação é diferente da entropia física. Segundo Denbigh, a entropia não

---

185 H. Leff & A. Rex, orgs., *op.cit.*, p. 16.

186 K. Popper, *apud* H. Leff & A. Rex, orgs., *op.cit.*, p. 20.

187 H. Leff & A. Rex , *op.cit.*, pp. 311-2.

188 K. Popper, *apud* H. Leff & A. Rex, orgs., *op.cit.*, p. 317. Citação no original: "No doubt some interesting analogies have been unearthed, but the soundness of the edifice, and especially the part played by Maxwell's demon, seems to me highly questionable."

189 O. Costa de Beauregard & M. Tribus, "Information Theory and Thermodynamics", *in* H. Leff & A. Rex, orgs., *op. cit.*, p. 179.

190 *Ibid.*, p. 181.

pode ser corretamente calculada com as fórmulas existentes, pois elas são métodos aproximativos e é possível que embutam certo grau de subjetividade[191]. A seguir, apóia a posição de Gibbs de interpretar a entropia calculada desta forma como "analogias de entropia", o que inclui não apenas fórmula de Gibbs (que é idêntica á formula da entropia da informação), mas também a fórmula de Boltzmann-Planck[192]. Sobre a entropia da informação, Denbigh afirma que a similaridade formal dela com uma "analogia de entropia" não é evidência de que elas signifiquem ou representem o mesmo conceito[193]. Denbigh também critica Brillouin por seu "princípio da negentropia", onde a entropia representaria o grau de ignorância sobre a organização de um sistema[194].

Ao que parece, a relação entre a Teoria da Informação e a Física ainda não está solidamente estabelecida depois de mais de 70 anos de debate. Um aspecto contraditório daqueles que defendem a unidade das duas disciplinas é a tentativa de conciliar conceitos diferentes, sendo que nenhum deles parece disposto a romper com (ou criticar) as idéias de Shannon ou as de Szilard, Wiener e Brillouin.

De qualquer forma, é inegável que as idéias e o jargão da Teoria de Informação influenciam o pensamento dos físicos contemporâneos. Prova desta influência é a complexa formulação de mecânica quântica proposta em 1976 por Stephen Hawking sobre a entropia de um buraco negro, e que formula um problema que se tornou conhecido como o "paradoxo da informação". Nesta

---

191 Penrose acredita que existe subjetividade uma vez que o método aproximativo usualmente empregado no cálculo, conhecido como "coarse graining", propõe a divisão do espaço de fase em pequenas áreas contendo partículas cujos estados macroscópicos sejam indistinguíveis, e um estado macroscópico indistinguível para um observador pode não o ser para outro. Para mais detalhes sobre esta argumentação, consultar: R. Penrose, *op. cit.*, p. 691.

192 Denbigh, K. "How subjective is entropy?", *in* H. Leff & A. Rex, orgs., *op. cit.*, pp. 109-11.

193 *Ibid.*, p. 113.

194 *Ibid.*, p. 112-5.

proposição, a entropia do buraco negro é igual à área de seu horizonte de eventos:

> *Há um bit de informação sobre o estado interno do buraco negro para cada unidade fundamental da área de superfície do horizonte*[195].

Em 2004, Hawking fez uma declaração pública afirmando que havia reconsiderado sua posição a respeito da perda da informação absorvida por um buraco negro[196]. Ao que parece, poucos entenderam do que se tratava, uma vez que poucos se arriscaram a comentar o assunto. Na declaração, Hawking também admitiu que sua reconsideração o fez perder uma aposta feita contra John Preskill, cujo ponto de vista a respeito do problema é:

> *De acordo com a mecânica quântica, embora processos físicos possam transformar a informação codificada num sistema físico numa forma que seja inacessível na prática, em teoria a informação pode ser sempre recuperada*[197].

Leonard Susskind propõe uma nova abordagem um pouco mais técnica para explicar o "paradoxo da informação" proposto por Hawking. Ele emprega argumentos da relatividade e da teoria-M (teoria modificada das cordas), e sugere que a explicação deste fenômeno poderia ajudar na formulação de uma teoria quântica da gravitação[198]. Susskind oferece uma nova noção de informação, aparente alinhada com a Teoria da Informação (a entropia termodinâmica associada à entropia da informação), porém associada à presença de um átomo num determinado local, que seria representada por 1 bit

---

195 S. Hawking, *op. cit.*, p. 63. Para uma descrição mais detalhada do problema em termos acessíveis ao público não-técnico, consultar: *Ibid*., pp. 121-3.

196 Para mais detalhes, consultar: http://www.newscientist.com/article.ns?id=dn6151.

197 J. Preskill, "On Hawking's Concession". Citação no original: "According to quantum mechanics, although physical processes can transform the information that is encoded in a physical system into a form that is inaccessible in practice, in principle the information can always be recovered."

198 L. Susskind, "Buracos negros e o paradoxo da informação", pp. 18-23.

de informação[199]. Ainda, Susskind sugere uma nova analogia, desta vez com a teoria das cordas, que faria com que a superfície do horizonte de eventos de um buraco negro se assemelhasse a uma gigantesca trama de cordas expandidas:

> *Cada segmento de corda, medindo 10^{-33} cm de comprimento, funciona como um bit. Assim, as cordas permitem que a superfície do buraco negro preserve a imensa quantidade de informação que ele engoliu ao longo do tempo*[200].

Aparentemente, cada autor emprega os conceitos de Teoria da Informação de acordo com suas necessidades. Neste debate convivem interpretações diferentes de um mesmo fenômeno, fenômenos distintos com a mesma nomenclatura, e correlações da unidade de informação com diversos tipos de partículas físicas: não parece haver sinais de que um consenso está sendo construído. Por outro lado, pode ser um indício de que a TI fornece uma abordagem flexível para abordar diversos tipos de fenômenos, sendo que suas associações com fenômenos físicos seriam na forma de analogias ou metáforas, e não como igualdades ou equivalências. De fato, Schenberg acredita que a ainda não se chegou a um consenso:

> *Mas pode ser que esteja faltando algum elemento essencial na nossa visão total do Universo. Muita gente acha que o Universo não seja só a matéria. Acreditam que exista outro elemento, e que seria a informação, ou alguma coisa obtida da informação. Mas, pelo menos até o momento atual, a Física não conseguiu esclarecer esta questão*[201].

Uma vez que as principais semelhanças entre as entropias de Boltzmann e de Shannon já foram apontadas, e tendo introduzido os principais argumentos

---

199 *Ibid.*, p. 21.

200 *Ibid.*, p. 23. O tempo se dilataria progressivamente até a parada total ao se atingir o horizonte de eventos: os segmentos de cordas, que são a unidade elementar de matéria na teoria-M, deixariam de vibrar e se dilatariam até atingir dimensões da ordem do comprimento de Planck (aproximadamente $1{,}6 \times 10^{-35}$ m). Preskill julga desnecessária a invocação de outra teoria para explicar fenômenos da mecânica quântica (à semelhança de Denbigh a respeito da termodinâmica com relação ao demônio de Maxwell). Uma análise mais detalhada a respeito de alternativas de solução do paradoxo da informação pode ser obtida em: J. Preskill, "Do black holes destroy information?".

201 M. Schenberg, *op.cit.*, p. 154.

empregados no debate sobre as “entropias”, torna-se necessário apontar suas principais diferenças. O primeiro aspecto a ser considerado é a natureza dos sistemas em análise: um gás é um sistema no qual não é possível ter acesso individual às moléculas que o compõe, enquanto que uma mensagem é um sistema que pressupõe o acesso individual aos sinais de que é composto. O segundo aspecto é a acurácia do cálculo: o cálculo da entropia termodinâmica é aproximado por ser baseado em semelhanças arbitrárias entre moléculas, enquanto que o da informação é preciso por ser auto-referente (baseia-se numa estatística sobre a codificação de cada mensagem). O terceiro aspecto é a ontologia do objeto de estudo: as moléculas possuem materialidade[202], enquanto que os bits são representações.

Concluindo, existem indícios suficientes de que o papel da informação na Física ainda não está devidamente esclarecido. Tendo explicitado algumas das influências, contradições e polêmicas que a TI provoca dentro da comunidade científica da qual emprestou seu principal conceito, pretendeu-se dar uma visão abrangente dos principais argumentos em discussão, bem como das dificuldades de integração entre as duas disciplinas, antes de se avançar para a parte final deste trabalho e indicar possíveis novas abordagens.

---

[202] Mesmo em se adotando uma postura filosófica idealista que postule que a matéria é uma construção da mente, poder-se-ia argumentar que mesmo assim suas naturezas difeririam uma vez que os símbolos codificados numa mensagem seriam meta-construções.

3. Novas Perspectivas e Conclusões

3.1. Antecedentes Conceituais e Outras Correlações

Nos capítulos anteriores, evidenciou-se a apropriação do conceito de entropia da Mecânica Estatística na *Teoria Matemática da Comunicação* de Shannon, e a transformação deste conceito numa grandeza capaz de medir a quantidade de informação numa mensagem. Também foram analisados diferentes conceitos de quantidade de informação, principalmente nas tentativas de se introduzir a informação como uma nova dimensão física ou grandeza inerente a sistemas físicos. Em todos os casos, sempre se associou a quantidade de informação a conceitos matemáticos decorrentes de probabilidades de ocorrência de certos fenômenos. De fato, o que se tem observado é a inter-relação entre Matemática, Física e Teoria da Informação em diferentes graus, bem como da influência da Teoria da Informação sozinha em outras áreas do conhecimento. Entretanto, a Matemática e a Física podem não ser as únicas influências da Teoria da Informação conforme formulada por Shannon. Com efeito, o objetivo agora é procurar indícios de influências mais sutis, cujo efeito ainda não foi detectado. A partir deste ponto, Teoria da Informação e Entropia serão tratadas segundo os conceitos de Shannon, a menos que explicitamente declarado em contrário.

A abordagem para detectar estas possíveis influências será a retomada de alguns conceitos de quantidade de informação já apresentados, e sua associação com fenômenos observados em outras áreas da atividade humana. O ponto de partida é a observação de Fenzl & Hofkirchnerb de que a quantidade de

informação está ligada à *emergência da novidade*[203], ou seja, de que a quantidade de informação é maior para fenômenos infreqüentes. De fato, a ocorrência de um evento raro desperta mais atenção do que eventos corriqueiros. Epstein faz uma análise abrangente deste tipo de fenômeno considerando diversos aspectos[204]. Quanto aos atributos sintáticos, afirma que:

> *Um fato pode se transformar em notícia quando é raro, e, portanto, inesperado e este atributo da notícia é comum ao conceito de "quantidade de informação" oriundo da* ***Teoria da Informação****. Trata-se de um atributo sintático tanto da notícia como da quantidade de informação (Teoria da Informação) porque concerne a sua freqüência relativa e não ao seu significado. (grifo do original)* [205]

Epstein prossegue, analisando o "valor notícia" de um fato científico sob a ótica do pensamento de Kunh a respeito dos períodos de "ciência normal" e "extraordinária":

> *O fato científico inesperado, no tempo de ciência normal firmemente estabelecida é marginalizado como anomalia à espera de uma explicação dentro do paradigma vigente. Em tempo de ciência extraordinária pode eventualmente adquirir grande importância como confirmador de um novo paradigma*[206].

Sob a ótica do falsifacionismo de Popper, Epstein afirma que fenômenos usuais têm pouco poder explanatório, uma vez que apenas se confirma o conhecimento tido como verdadeiro:

> *Um enunciado com alta probabilidade de ocorrência é cientificamente desinteressante, porque diz pouco e não tem poder explanatório*[207].

---

203 N. Fenzl & W. Hofkirchner, "Emmergence and Interaction of natural systems: the role of information, energy and matter in the perspective of a Unified Theory of Information", p. 6.

204 A fim de se manter consistência conceitual, procurar-se-á estabelecer relações utilizando-se critérios objetivos, como já foi feito nos capítulos anteriores. Assim sendo, "preço" (medida objetiva verificável) terá precedência sobre "valor" (conceito subjetivo). Quando isto não ocorrer, será por respeito a citações originais dos autores, ou para efeito de exemplificação.

205 I. Epstein, "Quando um fato se transforma em notícia no jornalismo e na ciência".

206 *Ibid*.

207 K. Popper, *apud* I. Epstein, "Quando um fato se transforma em notícia no jornalismo e na ciência".

Segundo o exposto, parece existir uma relação entre “raridade” e “quantidade de informação”, ou seja: quanto mais comum ou esperada é a mensagem, menos informação ela contém; por outro lado, quanto mais incomum ou inesperada é a mensagem, mais informação ela contém. Embora estes aspectos já tenham sido suficientemente analisados nos capítulos anteriores, é possível procurar indícios de fenômenos semelhantes já detectados em outras áreas do conhecimento. De fato, existem fenômenos onde se verificam relações análogas na Economia e na Psicologia, os quais serão abordados a seguir.

3.1.1. Antecedentes no Pensamento Econômico

À semelhança do que ocorre com a entropia na Teoria da Informação, existe um fenômeno análogo anterior no qual se verificam variações de um indicador ou grandeza em função de sua raridade. De fato, num sistema de mercado livre, quando outros fatores econômicos (objetivos e subjetivos) que poderiam influir no preço de uma mercadoria permanecem inalterados, a quantidade da mercadoria ofertada faz o preço dela variar da mesma forma: quando é escassa (rara), o preço aumenta; quando é abundante (comum), o preço diminui. A seguir, será verificado o que propuseram Adam Smith e Alfred Marshall a respeito destas relações, bem como a forma com que a Economia e a Psicologia se relacionam através do pensamento de Vilfredo Pareto.

Adam Smith publicou *Wealth of Nations* em 1776. Nesta obra, ele descreve a formação do preço de uma mercadoria como a soma dos valores do arrendamento dos meios de produção, do trabalho empregado e da margem de

lucro: este seria o *preço natural* da mercadoria[208]. O preço de mercado de uma mercadoria, entretanto, pode sofrer variações em função do desequilíbrio entre a necessidade de obtenção desta mercadoria pelos compradores e a quantidade desta mercadoria tornada disponível pelos produtores:

> *O preço de mercado de cada mercadoria é regulado pela proporção entre a quantidade que é realmente trazida ao mercado, e a demanda daqueles que querem pagar o preço natural da mercadoria... Quando a quantidade de qualquer mercadoria que é trazida ao mercado é inferior à demanda efetiva, ... o preço de mercado aumentará mais ou menos acima do preço natural...*[209]

Smith pressupõe que existe um preço fixo para cada mercadoria, e que neste preço a demanda por esta mercadoria é constante. Esta situação de equilíbrio se alteraria em função da quantidade ofertada desta mercadoria: se insuficiente para atender a demanda, seu preço aumenta; se, ao contrário, a quantidade ofertada é superior à demanda, seu preço diminui. Assim sendo, parece lícito estabelecer uma analogia entre a variação do preço de uma mercadoria com o conceito de auto-informação, que é baseado na freqüência de determinado símbolo: com oferta abundante, o preço de mercadoria diminui; com ocorrência abundante, a auto-informação de um símbolo diminui.

Mais de um século adiante, em 1890, Alfred Marshall publicou o tratado *Principles of Economics*, com o qual pretendia compilar e melhorar antigas doutrinas econômicas. Diferentemente de seus antecessores no pensamento econômico (incluindo o próprio Adam Smith), Marshall desenvolve esta obra posicionando a vontade do consumidor como o fator determinante de toda a cadeia de produção, distribuição e comércio. O preço de uma mercadoria, sob sua ótica, também é decorrência dos desejos dos consumidores, e toda

---

208 A. Smith, *Wealth of Nations*, pp. 50-7.

209 *Ibid.*, p. 59.

correlação entre *oferta*, *procura* e *preço* que ele formula é colocada em função da *procura*. Uma de suas proposições baseadas nesta premissa é a *Lei das Vontades Saciáveis* ou *da Utilidade Decrescente,* a qual declara que o benefício obtido pela aquisição de um bem se dilui à medida em que se obtém mais que o necessário:

> *... o benefício adicional que uma pessoa obtém de um dado aumento em seu estoque de um coisa [uma mercadoria], diminui a cada aumento do estoque [desta mercadoria] que ela já tem*[210].

Marshall reconhece as limitações de uma lei com esta generalidade, entre as quais estão: fatores que variam no longo prazo, como a mudança dos gostos do consumidor ou de sua psicologia; e nos casos em que o consumidor não obtém quantidade suficiente de uma mercadoria para satisfazer suas necessidades[211]. Neste cenário ideal, a diminuição da utilidade da mercadoria em função de sua abundância parece ter relação direta com a diminuição da auto-informação dos símbolo à medida que suas freqüências aumentam numa mensagem. De fato, quanto maior o excesso de estoque de determinada mercadoria, menor será sua utilidade média; da mesma forma, quanto mais um símbolo ocorrer em textos de determinada língua, menor será sua auto-informação.

De formas diferentes, tanto Smith como Marshall expressam relações econômicas que têm similaridades com conceitos da Teoria da Informação. A abundância de uma mercadoria determinaria a diminuição de seu preço no mercado ou de sua utilidade média daquele que a possui: estas proposições guardam interessantes semelhanças com os conceitos da Teoria da Informação,

---

210 A. Marshall, *Principles of Economics*, p. 93.

211 *Ibid.*, p; 94.

à medida que a quantidade de informação contida num símbolo ou numa mensagem diminui à medida que suas ocorrências se tornam mais freqüentes ou prováveis. Não existem, todavia, elementos que evidenciem a influência dos conceitos econômicos nos pensamentos de Shannon e de outros que contribuíram para aclarar os conceitos e implicações da Teoria da Informação. Entretanto, dada a universalidade da chamada "lei da oferta e da procura", não parece desprovida de fundamento a idéia de que o pensamento econômico possa ter causado alguma influência na formulação da Teoria da Informação, mesmo que em nível inconsciente. O pensamento de Epstein, que explicita as relações entre a TI com fenômenos observados noutras áreas, parece corroborar esta conjectura.

Vilfredo Pareto publicou em 1906 o *Manual de Economia Política*. A leitura desta obra evidencia a intenção do autor em reescrever os princípios econômicos numa forma mais objetiva a matematizada. Nela, não poupa críticas a seus antecessores no pensamento econômico, mesmo àqueles que, segundo seu critério, formularam proposições corretas. De fato, esta obra é pródiga em gráficos e equações, mas o ponto que se deseja destacar é que Pareto considera que a Economia é uma conseqüência da Psicologia:

> *A Psicologia é, evidentemente, o fundamento da Economia Política e, de modo geral, de todas as Ciências Sociais. Talvez chegue o dia em que possamos deduzir dos princípios da Psicologia as leis da Ciência Social, da mesma maneira que, um dia talvez, os princípios da construção da matéria nos dêem, por dedução, todas as leis da Física e da Química*[212].

Obviamente, esta correlação pode ser inferida das proposições de Marshall uma vez que ele posiciona os desejos e vontades do consumidor como foco de

---

[212] V. Pareto, *Manual de Economia Política*, Vol. I, p. 29.

toda atividade econômica. O que Marshall deixa implícito, Pareto evidencia de forma inequívoca. Uma vez que se aceite a algum grau de subordinação da Economia aos princípios da Psicologia, parece lícito supor que se possa encontrar alguma evidência que relacione alguns conceitos da Teoria da Informação com a Psicologia.

3.1.2. Relações com a Psicologia

Um possível exemplo de área de contato entre a Teoria da Informação e a Psicologia se encontra no *Curso de Psicologia Geral* de Alexander Luria. Nesta obra, o autor discorre sobre diversos aspectos de psicologia e cognição, e é de especial interesse seu pensamento sobre a *Atenção*:

> *O homem recebe um imenso número de estímulos, mas entre eles seleciona os mais importantes e ignora os restantes. Potencialmente ele pode fazer um grande número de possíveis movimentos, mas destaca poucos movimentos racionais que integram as suas habilidades e inibe outras. Surge-lhe grande número de associações, mas ele conserva apenas algumas, essenciais para a sua atividade, e abstrai as outras que dificultam o seu processo racional de pensamento*[213].

Para Luria, a atenção envolve os processo de seleção de informação e o controle de programas seletivos de ação. A seletividade da atividade consciente também se manifesta noutros processos:

> *O caráter seletivo da atividade consciente, que é função da atenção, manifesta-se igualmente na nossa percepção, nos processos motores e no pensamento*[214].

Qualquer atividade seria impossível sem esta seletividade, uma vez que quantidade de informação a ser tratada seria muito grande. Ainda, sem a

---

213 A. Luria, *Curso de Psicologia Geral*, Vol. III, p. 1.

214 *Ibid.*

inibição das associações, qualquer pensamento organizado seria impossível[215]. Luria prossegue, analisando o papel da importância biológica dos estímulos dos animais, que estariam ligados aos instintos, e os compara com o que ocorre no ser humano:

> *Tudo isso [as necessidades determinadas pelos instintos] se refere igualmente ao homem, com a única diferença de que as necessidades e interesses que o caracterizam não têm, em sua grande maioria, caráter de instintos e inclinações biológicos, mas caráter de fatores motivacionais complexos, que se formaram no processo da história social*[216].

Segundo Luria, o nível atenção pode ser determinado em função do grau de automatização da atividade desempenhada. A automatização de uma atividade é obtida através da repetição, que levaria ao sucessivo aprimoramento de sua execução. À medida que o executor se sente seguro de realizar certa atividade, menos atenção de sua parte é necessária:

> *O processo de automatização da atividade leva a que certas ações, que chamavam a atenção, se convertam em operações automáticas e a atenção do homem comece a deslocar-se para objetivos finais, deixando de ser atraída por operações costumeiras bem consolidadas*[217].

Dados estes princípios, é possível estabelecer algumas relações importantes entre a Psicologia e a Teoria da Informação. A primeira é cognitiva: a mente humana é capaz de desenvolver formas de classificar os estímulos e informações recebidas. Assim sendo, o processo de percepção parece interpretar alguns poucos estímulos como “informação” enquanto que a maioria deles é tratada como “redundância” ou como “ruído” (se tiverem intensidade suficiente

---

215 *Ibid.*, pp. 1-2.

216 *Ibid.*, pp. 4-5.

217 *Ibid.*, p. 6.

para tornar ineficaz a percepção daquilo em que se quer dirigir a atenção)[218]. Uma segunda relação possível é com a diminuição da atenção com ações costumeiras, que é análoga à diminuição da quantidade de informação quando se obtém mensagens ou sinais costumeiros. Os decréscimos da quantidade de informação de uma mensagem, do preço de uma mercadoria e da atenção sobre uma atividade estão relacionados ao aumento das freqüências de ocorrência da mensagem, da oferta da mercadoria e do grau de familiaridade da ação, respectivamente.

É importante explicitar que apesar das relações recém-apresentadas não serem quantitativas, elas não deixam de ser objetivas: “preço” não implica em “valor” e “atenção” não implica em “importância”[219]. Tendo apresentado todas estas relações, é possível avançar para a conclusão deste trabalho.

---

[218] A distinção entre informação e redundância já foi abordada no capítulo 1, ao se apresentar os fundamentos da TI.

[219] De fato, as atividades biológicas automáticas não deixam de ser importantes pela pouca atenção consciente que dedicamos a elas, como, por exemplo, a respiração e a circulação sanguínea. O sucesso do processo respiratório é tão costumeiro que raramente nos damos conta dele: este seria o caso mais provável, portanto de menor informação. Entretanto, não conseguir respirar é uma condição muito menos provável (aos indivíduos saudáveis, ao menos) e que desperta nossa atenção a ela quando ocorre. De forma análoga, enquanto os diamantes têm preço elevado em comparação ao ar que respiramos (que ainda é de graça), não se pode negar que o ar é muito mais valioso para nossa sobrevivência do que os diamantes: é concebível que se passe toda a vida sem possuir ou precisar de diamantes, enquanto que não se sobrevive sem ar além de alguns poucos minutos.

3.2. Conclusões

A definição de entropia, formulada na termodinâmica clássica, está relacionada com o grau de desordem de um sistema físico. Desde então, diversas interpretações matemáticas e físicas foram propostas para mensurar esta grandeza e descrever adequadamente suas decorrências. Boltzmann empregou uma abordagem estatística na tentativa de expressar os conceitos de ordem e desordem de um sistema físico: a distribuição das velocidades das moléculas de um gás contido num recipiente é o que determina seu estado, sendo que o estado ordenado seria aquele em que houvesse concentração das velocidades ao redor de determinados valores, enquanto que o estado desordenado seria aquele no qual não houvesse tais concentrações. Boltzmann adotou a premissa de que a matéria seria composta por unidades discretas (no que foi duramente criticado, pois esta não era a concepção dominante na época), introduzindo uma visão "descontínua" dos fenômenos físicos que seria o fundamento da Mecânica Quântica do século XX.

A abordagem discreta de Boltzmann, entretanto, não favoreceu o desenvolvimento de uma fórmula para a entropia que fosse efetivamente calculável, uma vez que ele assumiu que as velocidades das moléculas seriam continuamente variáveis. Planck propôs que o cálculo da entropia considerasse um número finito de intervalos de velocidades, transformando-o num processo contável. Gibbs, todavia, considerava que as fórmulas propostas (inclusive sua própria) seriam "analogias de entropia": a entropia seria um princípio físico definido na Termodinâmica, e a Mecânica Estatística proveria interpretações

matemáticas que descreveriam alguns de seus aspectos de maneira mais ou menos precisa.

Claude Shannon começou a se dedicar ao problema de quantificar a informação de uma mensagem na década de 1940. Antes de ter uma teoria acabada, aproveitou suas principais idéias numa teoria sobre criptografia para os Laboratórios Bell. Devido às necessidades bélicas norte-americanas durante a Segunda Guerra Mundial, a criptografia era um tema estratégico que necessitava ser entendido e dominado. De fato, o sucesso dos órgãos de inteligência americano e britânico em decifrar as mensagens militares alemãs foi decisivo para a vitória dos Aliados. As idéias de Shannon a respeito da quantificação da informação deram origem à *Teoria Matemática da Comunicação*, publicada em 1948 num jornal interno dos Laboratórios Bell e mais tarde republicada em forma de livro. Nesta obra, Shannon associa a Quantidade de Informação ao conceito de Entropia conforme interpretado pela Mecânica Estatística. Neste contexto, a quantidade de informação de uma mensagem não se relaciona ao significado do que se pretende transmitir, e sim à freqüência dos símbolos empregados para codificá-la.

Apesar de existirem analogias marcantes entre os conceitos de entropia da Mecânica Estatística e da Teoria da Informação, inclusive com a semelhança formal entre a entropia de Shannon com uma analogia de entropia de Gibbs, a equivalência dos conceitos é um tema controverso e ainda não totalmente entendido. No entanto, é flagrante a influência da Teoria da Informação em diversas áreas de investigação, ao menos como novo método de abordagem dos problemas existentes. Esta influência tem sido proclamada e celebrada por seus teóricos, que são, em sua maioria, ligados ao desenvolvimento de tecnologias

digitais. Em geral, a influência é apregoada como unidirecional, mas é concebível que existam outras influências que não as originalmente creditadas (que são exclusivamente oriundas da Mecânica Estatística).

De fato, existem indícios de que a Economia já empregava alguns conceitos semelhantes aos formulados por Shannon. A relação existente entre a variação dos preços de uma mercadoria em função da escassez ou abundância é análoga à quantidade de informação em função da freqüência de um símbolo. A relevância do incomum frente ao ordinário é um fenômeno reconhecido também pela Psicologia, o que sugere que sua origem possa ser tão antiga quanto a Humanidade. Assim sendo, como estas ciências reconhecem a influência contemporânea da Teoria da Informação, esta talvez possa retribuir a deferência e reconhecer influências mútuas e antecedentes.

Bibliografia

AFTAB, O., P. Cheung, A. Kim, S. Thakkar, & N. Yeddanapudi. “Information Theory after Shannon's 1948 Work”. *Project History, Massachusetts Institute of Technology*, 2001; http://mit.edu/6.933/www/Fall2001/Shannon2.pdf, agosto de 2005.

ALFONSO-GOLDFARB, A. M. *O que é história da ciência*. São Paulo, Brasiliense, 1994 (Col. Primeiros Passos, Vol. 286).

ALFONSO-GOLDFARB, A. M. & M. H. R. Beltran, orgs. *Escrevendo a história da ciência: tendências, propostas e discussões historiográficas.* São Paulo, Educ/Editora Livraria da Física/FAPESP, 2005.

_____., orgs. *O laboratório, a oficina e o ateliê: a arte de fazer o artificial.* São Paulo, Educ/FAPESP, 2002.

BARROW, J. D. *Theories of everything.* Nova Iorque, Fawcett Columbine, 1991.

BOLTZMANN, L. *Lectures on gas theory*. Trad. inglesa de G. S. Brush. Nova Iorque, Dover Publications, 1995.

BRILLOUIN, L. *Science and Information Theory*. Nova Iorque, Academic Press Inc, 1956.

CHOMSKY, N. “Three models for the description of language”. *IEEE Transactions on Information Theory*, IT-2 (Sept. 1956): 113-24.

CSISZÁR, I. “Applications of Information Theory in Probability and Statistics”. *EEE Information Theory Society Newsletter,* Special Golden Jubilee Issue (Summer 1998): 9-11.

DAY, R. *The modern invention of information: discourse, history and power*. Carbondalle, Southern Illinois Press, 2001.

DEVLIN, K. *Logic and Information*. Cambridge, Cambridge University Press, 1991.

ECO, U. *Como se faz uma tese*. São Paulo, Perspectiva, 1997.

_____. *The language of Mathematics: making the invisible visible*. Nova Iorque, W. H. Freeman/Owl, 1998.

EDWARDS, P. *The closed world: computers and the politics of discourse in Cold War America*. Cambridge, The MIT Press, 1996.

_____. “Virtual Machines, Virtual Infrastructures: the New Historiography of Information Technology”. *Isis*, 89 (1998): 93-9; http://www.si.umich.edu/~pne/PDF/isis_review.pdf, junho de 2004.

EPHREMIDES, A. “The Historian’s Column”. *IEEE Information Theory Society Newsletter*, Special Golden Jubilee Issue (Summer 1998): 5-6.

EPHREMIDES, A. & J. Massey. “1948-1998 Information Theory: The First Fifty Years”. *IEEE Information Theory Society Newsletter*, Special Golden Jubilee Issue (Summer 1998): cover page.

EPSTEIN, I. *Teoria da Informação*. São Paulo, Editora Ática, 1986.

_____. “Quando um Fato se Transforma em Notícia no Jornalismo e na Ciência”. *Boletín ALAIC Comunicación para Latinoamérica*, Año IV, 16, Mayo 2004; http://www.eca.usp.br/alaic/boletin16/Texto-Darcilia-Quando um fato se transforma em notícia no jornalismo e na ciência.htm, junho de 2004.

FENZL, N. & W. Hofkirchnerb. “Emmergence and Interaction of natural systems: the role of information, energy and matter in the perspective of a Unified Theory of Information”. Grupo de Pesquisa Amazônia 21, Universidade Federal do Pará; http://www.gpa21.org/en/pdf/emergence.pdf, novembro de 2005.

FRANK, M. “The Physical Limits of Computing”. *Computing in Science & Engineering*, May/June 2002; http://www.cise.ufl.edu/research/revcomp/physlim/PhysLim-CiSE/c3fra.pdf, julho de 2004.

GEARHART, C. “Planck, the Quantum, and the Historians”. *Physics in Perspective*, 4 (2002): 170–215; http://employees.csbsju.edu/cgearhart/Planck/PQH.pdf, outubro de 2005.

GLEICK, J. *Caos, a Criação de Uma Nova Ciência*. Rio de Janeiro, Campus, 1991.

GOLDFARB, J. L. *Voar também é com os homens: o pensamento de Mário Schenberg*. São Paulo, EDUSP, 1994.

GOLOMB, S., E. Berlekamp, T. Cover, R. Gallager, J. Massey & A. Viterbi. “Claude Elwood Shannon (1916–2001)”. *Notices of the American Mathematical Society*, 49 (1, January 2002): 8-16; http://www.ams.org/notices/200201/fea-shannon.pdf, julho de 2004.

HAGENAUER, J. “The Impact of Information Theory on Communications”. *IEEE Information Theory Society Newsletter*, Special Golden Jubilee Issue (Summer 1998): 6-8.

HAWKING, S. *O universo numa casca de noz*. São Paulo, Mandarim, 2001.

HODGES, A. *Alan Turing: the Enigma*. Londres, Vintage, 1992.

HUGILL, P. *Global Communications since 1844: Geopolitics and Technology*. Baltimore, Johns Hopkins University Press, 1999.

IFRAH, G. *História universal dos algarismos*. Rio de Janeiro, Nova Fronteira, 1997.

KAHN, D. *The Codebreakers: the story of secret writing*. Nova Iorque, Scribner, 1996.

KAKU, M. *Hyperspace*. Nova Iorque, Anchor Books Doubleday, 1995.

KLIR, G. *Fuzzy sets, uncertainty and information*. Nova Iorque, Prentice-Hall, 1988.

KOLMOGOROV, A. "A Great Mathematician's View of Shannon (foreword to Papers in Information Theory and Cybernetics by C. E. Shannon published in Russian in 1963)". *IEEE Information Theory Society Newsletter,* Special Golden Jubilee Issue (Summer 1998): 22.

KUHN, T. *A estrutura das revoluções científicas*. São Paulo, Editora Perspectiva, 2003.

LEFF, H. & A. Rex, org. *Maxwell's Demon: Entropy, Information, Computing*. Bristol, Adam Hilger, 1990.

LINDLEY, D. *Boltzmann's Atom: the great debate that launched a revolution in physics*. Nova Iorque, The Free Press, 2001.

LURIA, A. *Curso de Psicologia Geral*. Rio de Janeiro, Civilização Brasileira, 1979.

MACHADO, N. & M. Cunha, orgs. *Linguagem, conhecimento, ação: ensaios de epistemologia e didática*. São Paulo, Escrituras Editora, 2003 (Col. Ensaios Transversais, Vol. 23).

MARSHALL, A. *Principles of Economics*. Amherst, Prometheus Books, 1997 (Great Minds Series).

MASSEY, J. "Shannon and Cryptography". *IEEE Information Theory Society Newsletter,* Special Golden Jubilee Issue (Summer 1998): 12-3.

MILLER, G. "The cognitive revolution: a historical perspective". *Trends in Cognitive Sciences*, 7 (3, March 2003): 141-4; http://www.cogsci. princeton.edu/~geo/Miller.pdf, novembro de 2005.

ROSSI, P. *Naufrágios sem espectador: a idéia de progresso*. São Paulo, Editora UNESP, 2000.

OLSON, D. *O mundo no papel: as implicações conceituais e cognitivas da leitura e da escrita*. São Paulo, Ática, 1997.

PARETO, V. *Manual de Economia Política*. São Paulo, Nova Cultural, 1988 (Col. Os Economistas).

PELEGRINI, C. H. "Claude Elwood Shannon e a Revolução Digital". Dissertação de Mestrado. São Paulo, Pontifícia Universidade de São Paulo, 2005.

PENROSE, R. *The road to reality: a complete guide to the laws of the universe*. Nova Iorque, Knopf, 2005.

PEREIRA, C. A. B. & J. M. Stern. *Inferência Indutiva com Dados Discretos: Uma Visão Genuinamente Bayesiana*; http://www.ime.usp.br/~jstern/infestcomp/livs.pdf, março de 2006.

PRESKILL, J. *Do black holes destroy information? Cornell University*, 1992; http://xxx.lanl.gov/PS_cache/hep-th/pdf/9209/9209058.pdf, janeiro de 2006.

_____. *On Hawking's Concession*. *California Institute of Technology*, 2004; http://www.theory.caltech.edu/~preskill/jp_24jul04.html, janeiro de 2006.

PRICE, R. "A conversation with Claude Shannon: one man's approach to problem solving". *IEEE Communications Magazine*, 5 (22, May 1984): 123-6.

SCHENBERG, M. *Pensando a Física*. 5ª ed. São Paulo, Landy, 2001.

SHANNON, C. *Collected Papers*. Piscataway, IEEE Press, 1992

_____. *Communication Theory of Secrecy Systems*. Urbana/Chicago, Illinois State University Press, 1948.

_____. "The Bandwagon (Editorial)". *IRE Information Theory Society Newsletter*, IT-2 (March 1953): 3.

SHANNON, C. & W. Weaver. *The Mathematical theory of communication*. 5ª ed. Urbana/Chicago, Illinois State University Press, 1963.

SINGH, S. *The Code Book: the Science of Secrecy from Ancient Egypt to Quantum Cryptography*. Nova Iorque, Anchor Books, 1999.

SMITH, A. *Wealth of Nations*. Amherst, Prometheus Books, 1991 (Great Minds Series).

SOKAL, A. & J. Bricmont. *Imposturas intelectuais: o abuso da ciência pelos filósofos pós-modernos*. Rio de Janeiro/São Paulo, Record, 1999.

SUSSKIND, L. “Buracos negros e o paradoxo da informação”. *Scientific American Brasil*, edição especial (8, 2005): 18-23.

SZILARD, L. “On the Decrease of Entropy in a Thermodynamic System by the Intervention of Intelligent Beings”. *Z. Phys.* 53 (1929): 840.

TENÓRIO, R. *Cérebros e computadores: a complementaridade analógico-digital na informática e na educação*. São Paulo, Escrituras Editora, 2003 (Col. Ensaios Transversais, Vol. 2).

TRIBUS, M & E. McIrvine. “Energy and Information”. *Scientific American Magazine*, 224 (September 1971): 179-190.

UNIVERSITY OF CALIFORNIA TELEVISION (UCSD-TV). *Claude Shannon - Father of the Information Age*. San Diego, 2002; http://webcast.ucsd.edu:8080/ramgen/UCSD_TV/6090_Shannon.mpg.rm

UFFINK, J. "Boltzmann's Work in Statistical Physics". *The Stanford Encyclopedia of Philosophy* (Summer 2005 Edition); http://plato.stanford.edu/entries/statphys-Boltzmann/, junho de 2005.

VERDÚ, S. & S. McLaughlin, orgs. *Information Theory – 50 Years of Discovery*. Nova Iorque, Wiley-IEEE Computer Society, 1999.

WALDROP, M. “Reluctant Father of the Digital Age”. *Technology Review*, (July-August 2001): 64-71.

WIENER, N. *Cybernetics or the Control and Communication in the Animal and the Machine*. 2ª ed. Cambridge, The MIT Press, 1965.

_____. “What is Information Theory? (Editorial)”. *IRE Information Theory Society Newsletter*, IT-2 (June 1956): 42.

WIKIMEDIA FOUNDATION, INC. *Wikipedia*. http://en.wikipedia.org, junho de 2004.

Iconografia

Fig. 1 - Claude Shannon, 1995.
Cortesia de Lucent Technologies.

Fig. 2 - Claude Shannon e Theseus.
Cortesia de Lucent Technologies.

Fig. 3 - A Reflexão das ondas eletromagnéticas em função de suas freqüências.
Fonte: P. Hugill, *Global Communications since 1844: Geopolitics and Technology*, p. 87.

Fig. 4 - Patente Norte-Americana da Máquina Enigma, concedida em 1928.
Fonte: D. Kahn, *The Codebreakers: the story of secret writing*, p. 423.

Fig. 5 - Um exemplar da Máquina Enigma.
Autor: Declan McCullagh
"Copyright Declan McCullagh/mccullagh.org, all rights reserved."

Fig. 6 - Cable&Wireless "Great Circle" Map.
Autor: MacDonald Gill, 1945.
"Courtesy of Cable & Wireless Archive, Portchurno."

Fig. 7 - Diagrama Esquemático de um Sistema de Comunicação
Fonte: C. Shannon, "The Mathematical theory of communication", p. 34.

Fig. 8 - Variação da Entropia no caso de duas possibilidades com probabilidades p e (1-p)
Fonte: C. Shannon, "The Mathematical theory of communication", p. 50.

Anexos

1. Metodologia da Verificação Experimental

O programa de geração dos arquivos foi codificado na linguagem de programação *VBA* (*Visual Basic for Applications*) embutida no programa *Microsoft Excel 2003*, operando na plataforma *Microsoft Windows XP SP2* com processador *AMD Athlon XP 3000+*.

O arquivo contendo a Bíblia em formato textual foi obtido em http://www.o-bible.org/download/kjv.txt.

```
Sub GerarArquivos()

'  Verificação Experimental do Conceito de Entropia
'  através da compactação de arquivos.

'  Este programa gera arquivos com diferentes conteúdos
'  para posterior compactação.


'----- Arquivo contendo só espaços em branco
Call GravarSequencia("C:\Entropia\Texto1.txt", 1048576, " ")

'----- Arquivo contendo só a letra H
Call GravarSequencia("C:\Entropia\Texto2.txt", 1048576, "H")

'----- Arquivo contendo repetição de 16 carcateres distintos
Call GravarSequencia("C:\Entropia\Texto3.txt", 1048576, "0123456789ABCDEF")

'----- Arquivo contendo repetição de 32 carcateres
Texto$ = "LUDWIG BOLTZMANN CLAUDE SHANNON "
Call GravarSequencia("C:\Entropia\Texto4.txt", 1048576, Texto$)

'----- Arquivo contendo sequencia aleatória de algarismos
Call GravarAleatorio("C:\Entropia\Texto5.txt", 1048576, "0", "9")

'----- Arquivo contendo sequencia aleatória de letras maiúsculas
Call GravarAleatorio("C:\Entropia\Texto6.txt", 1048576, "A", "Z")

'----- Arquivo contendo sequencia aleatória de todos os caracteres
Call GravarAleatorio("C:\Entropia\Texto7.txt", 1048576, Chr$(0), Chr$(255))

'----- Arquivo contendo a Biblia em inglês, limitado a 1 MB
Call Truncar("C:\Entropia\kjv.txt", "C:\Entropia\Texto8.txt", 1048576)

End Sub
```

```
Sub GravarSequencia(ByVal Arquivo$, ByVal Tamanho, ByVal Texto$)

    '   Parâmetros:
    '       Arquivo$ = nome do arquivo a ser gerado
    '       Tamanho  = quantidade de caracteres a gravar
    '       Texto$   = texto a ser reptido dentro do arquivo

    TamTex = Len(Texto$)    ' tamanho do texto
    Gravados = 0            ' quantidade gravada

    Open Arquivo$ For Output As #1

    Do Until Gravados >= Tamanho
        Print #1, Texto$;
        Gravados = Gravados + TamTex
    Loop

    Close #1

End Sub


Sub GravarAleatorio(ByVal Arquivo$, ByVal Tamanho, ByVal Texto1$, ByVal Texto2$)

    '   Parâmetros:
    '       Arquivo$ = nome do arquivo a ser gerado
    '       Tamanho  = quantidade de caracteres a gravar
    '       Texto1$  = início da faixa de texto
    '       Texto2$  = fim da faixa de texto

    FaixaIni = Asc(Texto1$)             ' início da faixa
    FaixaExt = Asc(Texto2$) - FaixaIni  ' fim da faixa
    Gravados = 0                        ' quantidade gravada

    Open Arquivo$ For Output As #1

    Do Until Gravados >= Tamanho
        Print #1, Chr$(FaixaIni + (Rnd() * FaixaExt));
        Gravados = Gravados + 1
    Loop

    Close #1

End Sub


Sub Truncar(ByVal Origem$, Destino$, Tamanho)

    '   Parâmetros:
    '       Origem$  = nome do arquivo a ser lido
    '       Destino$ = nome do arquivo a ser gerado
    '       Tamanho  = quantidade de caracteres a gravar

    Gravados = 0        ' quantidade gravada

    Open Origem$ For Binary As #1
    Open Destino For Output As #2

    Do Until Gravado >= Tamanho
        Texto$ = Input(1024, #1)
        Print #2, Texto$;
        Gravado = Gravado + 1024
    Loop

    Close #1, #2

End Sub
```

2. Algoritmo de Geração de Arquivos com Distribuição Equiprovável de Símbolos

São válidas as mesmas condições técnicas mencionadas no anexo anterior.

```
Sub Equiprobabilidade()

'--- geração da base de todos os caracteres

Todos$ = ""

For i% = 0 To 255
    Todos$ = Todos$ + Chr$(i)
Next

Open "C:\Entropia\Equiprob.txt" For Output As #1

For i% = 1 To 4096

    '--- introdução de permutações aleatórias:
    '--- para cada posição, sorteia-se outra para que
    '--- os caracteres contidos nelas sejam permutados.

    For j% = 1 To 256

        '--- j% = posição corrente
        '--- k% = posição que permutará conteúdo com a posição j%

        k% = 1 + Int(Rnd() * 255)

        '--- execução da permuta

        TrocaJ$ = Mid$(Todos$, j%, 1)
        TrocaK$ = Mid$(Todos$, k%, 1)
        Mid$(Todos$, k%, 1) = TrocaJ$
        Mid$(Todos$, j%, 1) = TrocaK$
    Next

    Print #1, Todos$;

Next

Close #1

End Sub
```

3. Conteúdo do CD-ROM

- Dissertação em Formato PDF

- Instruções e Arquivos para Reprodução da Verificação Experimental

- Claude Shannon - *A Mathematical Theory of Communication*
  Fonte: Lucent Technologies (Bell Labs)

- Warren Weaver - *Recent Contributions to The Mathematical Theory of Communication*
  Fonte: Evergreen State College

- Sergio Verdú – "Fifty Years of Shannon Theory"
  Fonte: Princeton University

- UCSD-TV - *Claude Shannon - Father of the Information Age* (video on-line)
  Fonte: University of California, San Diego

- Isaac Epstein – "Quando um fato se transforma em notícia no jornalismo e na ciência" (página on-line)
  Fonte: Associación Latinoamericana de Investigadores de la Comunicación

- Programas Acessórios para visualização do conteúdo e execução da Verificação Experimental.