

SOBRE A ATUALIDADE DE PROPOSIÇÕES DE LUDWIG BOLTZMANN

ROBERTO CINTRA MARTINS

RESUMO: Partindo da tese de que Boltzmann seria um cientista incompreendido, uma vez que ele se manteve ligado, até o final de sua vida, ao declinante paradigma newtoniano, procuramos mostrar que, mesmo assim, importantes tópicos de seu pensamento científico permanecem atuais, sendo empregados, por exemplo, na teoria da informação. Ao mostrarmos a atualidade de algumas idéias de Boltzmann, estamos contribuindo para homenageá-lo neste ano em que comemoramos o sesquicentário de seu nascimento.

ABSTRACT: We believe that Boltzmann can not be completely by well understood if one retains attached to our contemporary point of view. Boltzmann never gave up his strong belief on the Newtonian paradigm. In spite of his commitment to this paradigm we show that some of his most significant scientific ideas are present in the science of our days. For example, we can mention information theory. Finally, the present work is to be understood as a homage to the one of most important physicist of all times.

Considerações Iniciais

Com a iniciativa de um encontro voltado para a História da Física no século XIX, o NEHCT da UFRJ busca apresentar (trazer e vivenciar no presente) um momento decisivo da História da Ciência. O ensejo de datas significativas como os 150 anos do nascimento de Ludwig Boltzmann leva-nos a revisitá-lo, pesquisar e discutir a obra de cientistas do passado, como estudantes que buscam uma oportunidade de reencontro com seus mestres, para ao mesmo tempo, como é típico dos estudantes, reverenciar, criticar e aprender. Neste sentido, o NEHCT vem mais uma vez exercer seu papel de nuclear e difundir um conhecimento histórico vivo e vivenciado, reafirmando assim seu compromisso com a pesquisa de fronteira e o vínculo com a tradição histórico-científica e filosófica.

Este vínculo entre o novo e o tradicional em ciência, indispensável à própria formação da consciência e da identidade da comunidade científica, tem na segunda metade do século XIX uma referência privilegiada, uma vez que é aí que se dá a última grande ruptura de paradigma na Física, ao mesmo tempo em que se encontram os últimos elos que norteavam os cientistas da época no sentido da “passagem” a um novo paradigma, vinculada ao compromisso e à lembrança do paradigma anterior. Portanto, uma ruptura de paradigma referenciada a uma memória e a um compromisso histórico¹.

Assim, por maiores que sejam as ressalvas da ciência contemporânea à visão de mundo dita mecanicista, típica da chamada Revolução Científica dos séculos XVI e XVII, devemos ainda lembrar que a ruptura-símbolo de todas as transformações ocorridas na ciência na última virada de século dar-se-á na Física e ainda receberá o nome de Mecânica.

Costuma-se colocar a Termodinâmica, o Eletromagnetismo e a Ótica como os “territórios” onde teriam-se evidenciado as mais importantes “rachaduras” no edifício da Física na segunda

¹ Cf. Os conceitos de “ciência normal”, “paradigma”, “anomalias” e “revoluções científicas” em T. S. Kuhn: “The Structure of Scientific Revolutions”, University of Chicago, 2ª edição, 1970, em especial os capítulos II a VII.

metadedoséculoXIX,ensejadasporcontradiçõesentreteoriaeexperimento,insolúveisnosmoldesdo paradigmaentãovigente. O estudo do espectro de radiação do corpo negro, a partir do qual M. Planck chega à primeira formulação conhecida da Mecânica Quântica, é emblemático como experimento-síntese desses três territórios². (Planck consegue inicialmente descrever o fenômeno mediante uma expressão analítica. Já a explicação teórica, que é a primeira formulação conhecida da Mecânica Quântica, somente lhe foi acessível, em suas próprias palavras, por um seu “ato de desespero”³.)

Hoje é difícil e desafiante a tarefa que nos colocamos de buscar avaliar as contribuições desse período, levando em conta o contexto cultural, o “em torno” à ciência da segunda metade do século XIX, pois estamos diante de homens que tiveram que optar entre o novo e o tradicional, em uma perspectiva que não lhes permitiu, como nos é facultado hoje, conciliar as duas visões. Planck, por exemplo, já foi caracterizado como um “revolucionário contra a vontade”⁴, enquanto Einstein, como quem mais resolutamente optou pelo novo.

Quando nos debruçamos sobre a História da Física em torno à virada do século, a par de reconhecer um momento de ruptura, é também parte de nosso dever reconsiderar a contribuição dos que mais resolutamente tenham abraçado a tradição e dedicado o melhor de suas energias a salvá-la. Naquele momento, abraçar uma tradição até às últimas consequências poderia equivaler a acompanhá-la em seu processo de superação, a ser superado juntamente com ela: não apenas uma questão de “preferência”, portanto, mas uma questão “vital”. Assim, sempre que busquemos avaliar este momento, como ora o fazemos, estaremos trazendo para o presente os nomes daqueles que procuraram “desesperadamente” salvar a Mecânica. Ao lado de Planck, estaremos reverenciando a memória daqueles que buscaram reduzir a Termodinâmica a uma mecânica, no caso, uma mecânica estatística. Aqui, ao lado de Maxwell e Gibbs, estaremos falando, em especial, de Ludwig Boltzmann.

Boltzmann coloca nossa tarefa diante de um desafio maior, mesmo porque vivemos imersos em uma cultura científica condicionada pela busca de sucesso, entendido não apenas como o alcançar de uma interpretação verdadeira sobre o real, mas sobretudo como o alcançar de um controle e domínio sobre a natureza e a sociedade. Esta imersão em uma “cultura científica” é reguladora de nosso dia a dia enquanto cientistas, na medida em que o sucesso nos é outorgado ou recusado por nossos “pares”, o que nos dá bem a medida do quanto estamos caminhando “par i passo” com o “novo”, com o “espírito da época”. Mesmo com relação a muitos dos mais renomeados “espíritos rebeldes” da ciência, entre os quais o nome de Galileu é emblemático, nossa “cultura científica” sabe reconhecer no incompreendido um gênio, quando não em vida, ao menos postumamente.

Esta “cultura científica”, que virtualmente confunde fazer ciência com fazer carreira, pode mesmo ser suficiente para dar conta de inúmeros espíritos rebeldes e gênios incompreendidos, ao reconhecê-los, mais cedo ou mais tarde, como “aqueles que tinham razão”. Entretanto, ao nos debruçarmos sobre a vida e a obra de Boltzmann, passamos a reconhecer um destino que não se enquadra facilmente nos limites desta interpretação, e passamos a ter que admitir a existência de vidas que assumiram um compromisso indissolúvel com tradições em declínio e, ao tentar salvá-las, tiveram que acompanhá-las até às últimas consequências.

O fim da Mecânica Newtoniana e o destino de Boltzmann são, nesta perspectiva, indissociáveis. Nesse sentido, relembrar sua vida e obra é para nós um desafio maior, pois apesar de imersos em uma “cultura científica do sucesso”, temos que admitir que, se a Mecânica Newtoniana é uma visão parcial e

2 Para um detalhamento, ver R. C. Martins: “Planck e o Nascimento da Mecânica Quântica: Sugestões para o Estudo de Condicionantes Históricos Recentes”, in: F. R. R. Évora (ed.): “Século XIX: O Nascimento da Ciência Contemporânea”, Coleção CLE/UNICAMP, CAMPINAS, 1992, p.287-308.

Para um relato do próprio Planck, cf. M. Planck: “Zur Geschichte der Auffindung des physikalischen Wirkungsquantums”, in: “Die Naturwissenschaften”, n.14/15, p. 153-159, 1943.)

3 Cf. M. Planck, em entrevista a R. W. Wood, em 7 de outubro de 1931, in: A. Hermann: “The Genesis of Quantum Theory”, MIT Press, 1971.

4 Cf. E. Segrè: “Dos Raios X aos Quarks”, UNB Brasília, 1980, em especial cap. IV.

incompleta se comparada à Mecânica Quântica, então Boltzmann não tinha razão ao preferir a tradição ao novo. Neste sentido, nem mesmo hoje poderíamos enquadrá-lo, tal como Galileu, entre os “gênios incompreendidos”. Em sua época, já surgia um novo paradigma, que se afirmou como superior para nossa “cultura científica”. Boltzmann viveu em um contexto que lhe permitia ter acesso e optar por esse novo, mas apegou-se a um corpo teórico agonizante, até o fim. Como interpretá-lo e reverenciá-lo hoje, se estamos condicionados a valorizar em ciência aquilo que é reconhecido como “bem sucedido” por nossos “pares” e pela “comunidade”, e não aquilo que foi superado?

Breves Considerações Iniciais sobre três Contribuições de Boltzmann

Iniciaremos nossa consideração a respeito da obra de Boltzmann destacando de forma sucinta suas mais conhecidas contribuições à Teoria Cinética dos Gases e à Mecânica Estatística.

De acordo com a conhecida lei de Stefan-Boltzmann, a potência por unidade de área irradiada por um corpo negro é dada por

$$R = \sigma T^4$$

onde R é a potência por unidade de área, T é a temperatura absoluta e σ é a constante de Stefan.

Aqui, um fato crucial que irá motivar Planck a se aprofundar no estudo do espectro de radiação do corpo negro é que R depende somente da temperatura, o que outorga à lei de Stefan-Boltzmann a características, muito atraente para Planck, de “lei universal”.

Considerando um sistema de partículas em um espaço de coordenadas x, y, z e momentos p_x , p_y , p_z , sendo a energia E função destas seis coordenadas, a distribuição de Maxwell-Boltzmann dá a probabilidade de uma partícula se encontrar no volume $dv=dx dy dz dp_x dp_y dp_z$:

$$f(x, y, z, p_x, p_y, p_z) dv = C e^{-E/KT} dv$$

onde T é a temperatura absoluta, K é a constante de Boltzmann e C é uma constante dada pela condição de normalização

$$\int_{-\infty}^{\infty} C e^{-E/KT} dv = 1$$

A distribuição de Stefan-Boltzmann será objeto do item 7 deste trabalho, onde apresentaremos analogias que inspiraram um procedimento recente em programação matemática.

Finalmente, cabe destacar a contribuição de Boltzmann sobre a qual nos deteremos com maior ênfase neste trabalho: sua conhecida formulação probabilística do conceito de entropia.

Talvez não haja em toda a Física conceito mais polêmico ou mais rico em desdobramentos filosóficos que o de entropia. Na conceituação originária de Clausius entropia pode ser entendida como uma medida da qualidade da energia, relacionando-se diretamente com a parte da energia dissipada nos processos de transmissão de calor (processos termodinâmicos) e não reaproveitável na forma de trabalho mecânico.

Para produzir trabalho mecânico, o calor deve necessariamente transferir-se de um sistema para outro, estando o primeiro a uma temperatura mais elevada que o segundo. Sendo T a temperatura na qual este processo se dá e DQ a quantidade de calor trocada no processo, Clausius define entropia como a quantidade

$$S = \frac{\Delta Q}{T}$$

Em 1865, Clausius formula os dois princípios da Termodinâmica em palavras de alcance cosmológico:
A energia do Universo é contante.

A entropia do Universo tende a um máximo.

O segundo princípio pode ser interpretado como um reconhecimento da tendência à equalização do Universo em termos de temperatura, o que equivaleria a um processo de homogeneização ou de “esquecimento” (perda) de simetria original. A “morte entrópica” do Universo expressaria uma tendência inexorável em direção a um estado onde nenhum trabalho mecânico seria possível, pois não haveria mais dois sistemas a temperaturas distintas, que pudessem servir como “fonte quente” e “fonte fria” na produção de trabalho mecânico. Esta tendência à indistingtibilidade em termos de temperatura leva à conceituação da entropia como uma medida de estado de qualquer sistema macoscópico, que cresce com a perda de simetria do sistema.

Tendo como referência um sistema de partículas, o conceito de entropia ganha com Boltzmann uma nova conotação. A entropia passa a ser entendida como uma medida da distribuição das partículas em termos de posição espacial e quantidade de movimento. Aqui, máxima entropia passa a significar distribuição homogênea ou máxima desordem, quando a probabilidade de uma certa partícula se encontrar em uma determinada posição com uma certa quantidade de movimento é idêntica à probabilidade de qualquer outra partícula específica se encontrar na mesma situação.

Sendo P_i a probabilidade de ocorrência de uma determinada configuração macroscópica i de um sistema, então a entropia do sistema nessa configuração é dada segundo Boltzmann por:

$$S_i = -k \log_2 P_i$$

Supondo-se ainda que N configurações macroscópicas distintas sejam possíveis para o sistema, então sua entropia poderá assumir N valores S_1, S_2, \dots, S_n e a entropia média do sistema é dada por

$$S = \sum_{i=1}^N P_i S_i = -K \sum_{i=1}^N P_i \log_2 P_i$$

onde S é máximo quando $p_i = 1/N$ ($i = 1, 2, \dots, N$).

O conceito de Entropia em Boltzmann e o Nascimento da Mecânica Quântica

É conhecida a relutância de M. Planck em aceitar a concepção probabilística de entropia de Boltzmann. Entretanto, esta concepção veio a se tornar componente indispensável da formulação quântica da energia, estabelecida por Planck de forma pioneira em 1900 e que está no cerne do nascimento da Mecânica Quântica.

Assim, é de textos originais de autoria de Planck⁵ que se depreende que a expressão quântica para a energia (e) em função da frequência (f)

5 Cf. M. Planck, op. Cit. na nota 2. Neste item, pode-se supor sem perda de generalidade que todos os logaritmos são tomados na base 2. Isto é evidente na formulação de entropia em Boltzmann. Embora não seja explicado por Planck, também pode ser assumido em suas expressões e equações envolvendo entropia, a menos de constantes multiplicativas.

$$e = hf$$

é derivada a partir da equação diferencial parcial relacionando entropia (S), energia (U) e frequências (f) de radiação:

$$\frac{\delta S}{\delta U} = \frac{1}{af} \log \left(1 + \frac{aCf}{U} \right)$$

que, integrada, resulta

$$S = C \left[\left(\frac{U}{hf} + 1 \right) \log \left(\frac{U}{hf} + 1 \right) - \frac{U}{hf} \log \frac{U}{hf} \right]$$

onde $h = aC$ é aqui ainda apenas uma denominação escolhida por Planck, sem qualquer atribuição de significado físico.

Neste ponto crucial, Planck recorre à concepção probabilística de entropia de Boltzmann, ao mesmo tempo em que postula a hipótese decisiva da discretização da energia total de radiação U_{total} como composta de P "quanta" elementares de energia distribuídos entre N osciladores.

$S = K \log p$ (entropia segundo Boltzmann)

$U_{total} = NU - Pe$ (postulado da discretização da energia segundo Planck)

onde

U = energia média por oscilador

P = número inteiro, supostamente muito grande

e = "quantum" de energia, supostamente muito pequeno

Para a entropia total dos N osciladores, decorre:

$$S_{total} = NS = k \log p = k \log \frac{(N+P-1)!}{P!(N-1)!} = \tilde{k} \log \frac{(P+N)!}{P!N!}$$

onde a segunda passagem faz uso da formulação de entropia de Boltzmann e a terceira decorre do postulado de Planck e expressa o número de maneiras em que P "quanta" de energia podem se distribuir entre N osciladores.

Pela aplicação da fórmula de Stirling:

$$NS \cong k[(P+N) \log(P+N) - \log P - N \log N]$$

$$S \cong k \left[\left(\frac{P}{N} + 1 \right) \log \left(\frac{P}{N} + 1 \right) - \frac{P}{N} \log \frac{P}{N} \right]$$

Finalmente, da simples comparação entre as duas expressões obtidas para S, e considerando que $NU = Pe$, Planck conclui:

$$K = C$$

$$e = hf$$

inaugurando assim a Mecânica Quântica ao determinar a expressão do “quantum” de energia de radiação como proporcional à frequência.

O movimento assim iniciado na virada do século veio a superar a Física Clássica, na mais radical ruptura ocorrida na Física desde o século XVII. Neste contexto, a Mecânica Estatística pode ser vista como uma sucessão de tentativas de preservação do antigo paradigma:

“...probability was introduced into thermodynamics precisely for saving the mechanistic representation of nature. ... The history of statistical mechanics is simple: as soon as one version was challenged or refuted, another on a still more complex level was offered instead. The successive contributions of Boltzmann illustrate this point amply as well as instructively”⁶.

Como fundador da Mecânica Estatística, Boltzmann propôs sucessivas “explicações” para o fenômeno do calor, baseadas em uma abordagem probabilística mediante a qual ele tentou vincular os fenômenos microscópicos às variáveis de estado do sistema macroscópico, preservando a validade das leis da Física Clássica ao primeiro nível. Boltzmann foi neste sentido um conservador pertinaz, assim como foi Planck. Não deixa de ser paradoxal esta sua contribuição involuntária ao surgimento de um novo paradigma na Ciência, que estaria destinado a superar a visão de mundo da Física Clássica, da qual ele foi um dos mais contumazes e vigorosos defensores.

O Conceito de Entropia em Boltzmann e a Ciência da Informação

Um dos desdobramentos mais ricos e polêmicos do conceito probabilístico de entropia em Boltzmann foi sua extensão ao campo da Ciência da Informação.

Com base na formulação original de Clausius e na conceituação de Boltzmann, a entropia tem sido tomada como uma medida:

- do nível de qualidade da energia, entendida como o grau de seu aproveitamento para realização de trabalho mecânico

- do grau de desorganização de um sistema físico, ou mais especificamente, do grau de não distinção entre seus componentes

- da perda de simetria (ou de distinguibilidade) nas características de um sistema físico (posição, quantidade de movimento, energia, temperatura)

A partir dessas considerações, C. Shannon e W. Weaver⁷ propõem uma conceituação para a quantidade de informação de uma mensagem, a ser medida pelo número de decisões binárias necessárias para identificar um objeto, ou seja, pela quantidade de “desinformação” (desordem, entropia) que é removida ou superada pela mensagem.

Para se identificar um entre dois objetos distintos equiprováveis, é necessária uma decisão binária. Entre quatro objetos, duas decisões binárias. Entre oito, três decisões. Assim, a quantidade de informação (isto é, decisões binárias) contida na escolha de um entre $N=2^n$ objetos equiprováveis é dada por

$$I = \log_2 N = \log_2 2^n = n$$

ou

$$I = -\log_2 p$$

6 Cf. N. Georgescu - Roegen: “The Entropy Law and the Economic Process”, Harvard University Press, 1976, p. 148-149.

7 Cf. C. Shannon e W. Weaver: “The Mathematical Theory of Communication”, The University of Illinois Press, Illinois, 1949.

onde

$p = 1/N = 1/2^n$ é a probabilidade de escolha de um objeto.

Em geral, a quantidade de informação associada à identificação de um objeto ou evento de probabilidade de ocorrência p é dada por

$$I = -\log_2 p$$

Notando-se que esta proposição é coerente com a noção intuitiva de que a ocorrência de eventos pouco prováveis implica em maior informação do que a ocorrência de eventos mais prováveis.

Para o caso de N eventos não equiprováveis (distribuição não uniforme), a quantidade de informação média por evento é dada por

$$I = \sum_{i=1}^N p_i \log_2 p_i$$

onde p_i é a probabilidade de ocorrência do i -ésimo evento ($i = 1, 2, \dots, N$)

Aqui, I pode ser interpretada como a quantidade de entropia que é subtraída de um sistema quando se dispõe de uma mensagem (informação) a seu respeito.

Assim, um sistema que inicialmente tenha entropia S_0 passará a ter uma entropia menor (menos desordem), devido ao ganho de informação $I = -p \log_2 p$, associado à ocorrência de um evento de probabilidade p . A nova entropia do sistema será dada por:

$$S = S_0 - I = S_0 + p \log_2 p < S_0$$

Assim, a ocorrência de eventos, enquanto significa ganho de informação sobre o sistema, tem também o significado de redução de entropia, conforme sugere de forma pioneira N. Wiener, ao conceituar informação como anti-entropia: “

O processo de ganho de informação transcorre segundo Wiener em sentido oposto (ao do aumento da entropia). Ele é um processo de redução de entropia ou de “anti-entropia”, isto é, um processo de ordenamento crescente”.⁸

O autor agradece a indicação desta referência, ainda atual, a Roberto dos Santos Bartholo Jr.)

Depreende-se destas considerações que o conceito de informação em Shannon, Weaver e Wiener é herdeiro do legado de Boltzmann. Entretanto, a “migração” do conceito de entropia da Mecânica Estatística para o campo da Ciência da Informação não se deu isenta de críticas e polêmicas.⁹

Entropia e Ganho de Conhecimento

Na linha de pensamento de N. Wiener, Pfeiffer¹⁰ sugere mais uma analogia interessante entre informação enquanto anti-entropia e o conhecimento na ciência moderna. A sugestão de Pfeiffer é portanto herdeira indireta do legado de Boltzmann.

Partindo da concepção de K. Popper¹¹ segundo a qual as hipóteses ou “conjecturas” científicas devem satisfazer à condição de refutabilidade, Pfeiffer constrói modelos simples do processo de refutação e escolha de hipóteses na pesquisa científica.

8 Cf. W. Pfeiffer: “Allgemeine Theorie der technischen Entwicklung als Grundlage einer Planung und Prognose des technischen Fortschritts”, van den Hoeck & Ruprecht, Göttingen, 1971, p.40.

9 Cf. N. Gergescu-Roegen, op. Cit., especialmente o epêndice B (“Ignorance, Information and Entropy”), p. 388-406.

10 Cf. W. Pfeiffer, op. Cit. Na nota 7, p. 40-43

11 Para um detalhamento, cf. K. Popper, “Conjectures and Refutations: The Growth of Scientific Knowledge”, London, 1963.

O processo inicia-se na vigência de uma determinada "teoria científica", isto é, de um conjunto de proposições hipotéticas coerentemente interligadas por relações de derivabilidade. Nestas condições, dada a coerência lógica interna deste conjunto, somente poderão portar informação nova aqueles fenômenos observáveis que contradizem este "sistema teórico ordenado", isto é, fenômenos que não se deixam ordenar dentro desse sistema e portanto são vistos como improváveis ou surpreendentes.

A superação da contradição entre sistema teórico e a realidade somente é possível mediante a reestruturação do sistema teórico ou "modelo" da realidade, isto é, mediante a construção de um ordenamento teórico "superior", que possa englobar o novo fenômeno, não mais enquanto improvável ou inesperado, mas como provável e esperado. Desta forma, a informação potencial (entropia) contida no fenômeno é transformada em informação adquirida.

A quantidade de informação potencial "contida" no novo fenômeno observado é igual à quantidade de desorganização (entropia) que pode ser eliminada (transformada em anti-entropia) pela reestruturação do sistema teórico, em face do fenômeno. Esta medida depende do número de hipóteses (conjecturas) explicativas novas suscitadas pela ocorrência inesperada do fenômeno em pauta, bem como de suas respectivas probabilidades.

Suponhamos que, diante de um fenômeno inesperado nos quadros da teoria vigente, se coloquem n hipóteses H_1, H_2, \dots, H_n , como conjecturas iniciais a serem avaliadas de acordo com sua capacidade de explicar o fenômeno. Tratando-se de um fenômeno totalmente inesperado, pode-se supor inicialmente que cada uma das n hipóteses tenha igual chance de vir a ser aceita como "a hipótese explicativa".

$$P(\text{aceitação de } H_i) = p_i = 1/n \quad (i = 1, 2, \dots, n)$$

O método científico seria então o instrumento mediante o qual se pode proceder à escolha de uma destas hipóteses para ser aceita como "a explicação" do fenômeno, enquanto as demais hipóteses seriam refutadas. Assim, ao final do processo, teríamos:

$$P(\text{aceitação de } H_k) = p_k = 1$$

$$P(\text{aceitação de } H_i) = p_i = 0 \quad (i \neq k)$$

onde se supõe que a hipótese H_k foi a aceita.

Neste contexto, o ganho de conhecimento científico corresponde a um processo de reestruturação de teorias, que se inicia em um estado de entropia máxima S_0 onde n conjecturas equiprováveis "convivem" indistintamente, e termina em um estado de entropia mínima S , onde somente "sobrevive" uma hipótese, valendo as relações.

$$\begin{aligned} S_0 &= -\sum_{i=1}^n \frac{1}{n} \log_2 \frac{1}{n} = \\ &= -\frac{1}{n} n \log_2 \frac{1}{n} = \log_2 n \\ S &= 1 \log_2 1 = 0 \end{aligned}$$

Em outros termos, o processo inicia-se em um estado de máxima informação "em potencial", a ser desvendada ($I_{pot} = -\log_2 n$), e termina em um estado em que não há mais informação a ser desvendada ($I_{pot} = 0$).

Pfeiffer oferece um exemplo ilustrativo que parte de quatro hipóteses equiprováveis, e evolui em um processo de ganho de informação que abrange quatro "sistemas teóricos" conforme a tabela a seguir.

Sistema	Probabilidades associadas às hipóteses h_1, h_2, h_3, h_4	Entropia = informação potencial $S = I_{pot}$	Anti-Entropia = Ganho de Informação = I
a	1/4, 1/4, 1/4, 1/4	2 bits	
			0,25 bits
b	1/2, 1/4, 1/8, 1/8	1.75 bits	
			0.69bits
c	3/4, 1/8, 1/8, 0	1,06 bits	
			1.06
d	1,0,0,0	0	
			$I_{total} = 2.00$

Gonzales¹² nos oferece um exemplo de ganho de conhecimento nos mesmos moldes do aqui apresentado, referindo-se a uma situação onde existem oito presos em uma cela, sendo que apenas um deles poderá escapar, a ser escolhido através de lançamentos de uma moeda. inicialmente, cada preso tem probabilidade 1/8 de vir a ser escolhido e portanto temos

$$S_0 = - \sum_{j=1}^8 \frac{1}{8} \log_2 \frac{1}{8} = -8 \frac{1}{8} \log_2 \frac{1}{8} = 3$$

Já ao final do processo, um único preso foi escolhido para fugir e tem-se

$$S = 1 \log_2 1 = 0$$

A incerteza acerca de quem seria o escolhido para fuga foi reduzida de 8 possibilidades para uma única, e a quantidade de informação associada a esta redução é de 3 bits. Não por acaso, esta última quantidade coincide com o número de decisões binárias equiprováveis (lançamentos de moeda) necessárias para se reduzir a incerteza de 8 para 1; após um primeiro lançamento, excluem-se 4 presos da possibilidade de fuga, após o segundo lançamento, excluem-se dois presos do grupo remanescente e, finalmente, após o terceiro lançamento, estará identificado o preso que poderá fugir.

Este exemplo nos sugere que, para cada distribuição de probabilidades discreta, deve-se associar uma medida de entropia ou informação potencial. Aqui, no caso de uma distribuição discreta e uniforme, tem-se

12 Cf. M. E. Gonzales: "O Nascimento da Ciência Cognitiva e suas Raízes na Física do Século XIX", in: F. R. R. Évora (ed.): "Século XIX: O Nascimento da Ciência Contemporânea", Coleção CLE/UNICAMP, CAMPINAS, 1992, p.413-433).

$$S_0 = I_{pot} = \log_2 n$$

onde n é o número de valores distintos da variável aleatória em questão.

No próximo item, veremos alguns casos específicos que corroboram essa sugestão e tentaremos generalizá-la para distribuições contínuas, seguindo sugestão de J. von Neumann e N. Wiener.

Entropia e Métodos Estatísticos

O conceito de entropia em Boltzmann guarda ainda relações interessantes com conceitos e métodos estatísticos.

Para distribuições de probabilidade discretas, uma medida de entropia decorre diretamente da sugestão proposta ao final do item anterior. Assim sendo X uma variável aleatória discreta que pode assumir n valores distintos x_1, x_2, \dots, x_n e sendo p_i a probabilidade de X assumir o valor x_i , temos:

$$S_0 = I_{pot} = -\sum_{i=1}^n p_i \log_2 p_i$$

onde S_0 e I_{pot} são a entropia e a informação potencial associadas à distribuição, respectivamente.

Referindo-se a uma sugestão de J. Von Neumann, e abstraindo o sinal negativo, N. Wiener propõe uma extensão do conceito para distribuições contínuas¹³:

$$S_0 = I_{pot} = \int_{-\infty}^{\infty} [\log_2 f(x)] f(x) dx$$

Fazendo-se uso da conhecida “lei do estatístico inconsciente”, I pode ser considerado idêntico ao valor esperado (média) da variável aleatória Y dada por

$$Y = \log_2 f(X)$$

$$\text{ou } I_{pot} = E(\log_2 f(X))$$

Como exemplos, pode-se calcular facilmente a informação (I) e a entropia (S) associadas a algumas distribuições discretas e contínuas, tais como:

¹³ Cf. N. Wiener. “Cybernetics or Control and Communication in the Animal and the Machine”, The MIT Press, Cambridge, Massachusetts, 2a edição, 1961, p.62.

Distribuição	Parâmetro	$I_{pot} = S_0$	Condição para I_{pot} máximo (S_0 máximo)	Condição para S_0 máximo (I_{pot} mínimo)
de Bernoulli	p	$p \log_2 p + (1-p) \log_2 (1-p)$	$p=1$ ou $p=0$	$p = 1/2$
Geométrica	p	$(p \log_2 p + (1-p) \log_2 (1-p))/p$	$p=1$	$p = 0$
Exponencial	λ	$\log_2 (1/e)$	$\lambda=\infty$	$\lambda = 0$
Normal	ν, σ	$\log_2 (\varepsilon \pi \lambda^2)^{-1/2}$	$\sigma=0$	$\sigma = \infty$

Outra analogia interessante¹⁴ surge quando consideramos o conhecido método de máxima verossimilhança para estimação de parâmetros. Dada a função densidade de probabilidades $f(x_i)$ de uma variável aleatória contínua X , onde θ é um parâmetro a ser estimado, e uma amostra aleatória x_1, x_2, \dots, x_n dessa variável aleatória, a função de verossimilhança é dada por:

$$L(\theta) = g(x_1, x_2, \dots, x_n, \theta)$$

O método consiste em obter um estimador que maximize a função L , dados os valores x_1, x_2, \dots, x_n . Em geral, sendo x_1, x_2, \dots, x_n uma amostra aleatória, tem-se:

$$L(\theta) = g(x_1, x_2, \dots, x_n, \theta) = \prod_{i=1}^n f(x_i, \theta)$$

Por razões de cálculo, em geral é conveniente proceder à maximização equivalente da função

$$\log L(\theta) = \sum_{i=1}^n \log f(x_i, \theta)$$

No caso de distribuições discretas, o procedimento é análogo e leva à maximização da função

$$\log L(\theta) = \sum_{i=1}^n \log P(X = x_i, \theta)$$

14 O autor agradece a percepção desta analogia a Fernando Yassuo Chiyoshi. Para um detalhamento do método de máxima verossimilhança para estimação de parâmetros, cf. A. M. Mood e F. A. Graybill: "Introduction to the Theory of Statistics", Mc-Graw-Hill Book Company, Inc., New York, 1963, p. 178-185.

Em ambas as expressões anteriores, é comum adotar logaritmos de base neperiana, mas pode-se supor base 2, pois a maximização da função de verossimilhança equivale à maximização de seu logaritmo em qualquer base maior que 1. Desta forma, $\text{Log } L()$ pode ser interpretado como a informação contida nos valores amostrais conhecidos x_1, x_2, \dots, x_n .

Assim, o método de máxima verossimilhança equivale a pesquisar o valor que maximize a informação contida nos valores amostrais conhecidos x_1, x_2, \dots, x_n ou, em outros termos, que minimize a entropia da distribuição em face da ocorrência desses valores. Trata-se portanto de um método que busca transformar a maior quantidade possível de informação potencial (entropia) contida na amostra em informação efetiva ou adquirida (cf. item 5).

Aplicação de Conceitos da Termodinâmica em problemas de Otimização.

Em programação matemática, foi proposto recentemente¹⁵ um novo método para pesquisa do ponto ótimo de funções, inspirado em conceitos da Termodinâmica, que vem apresentando vantagens com relação aos métodos tradicionais. Por este método, a busca pode “escapar” de ótimos locais, o que é especialmente vantajoso no caso de funções objetivo com muitos ótimos locais. A função objetivo não precisa ser quadrática nem diferencial. Além de ser aplicável a funções que não podem ser maximizadas por outros métodos, ele ainda permite um reconhecimento mais detalhado da função, tanto de seus máximos locais, reconhecidos como tal, como também de seus “vales” e “planícies”.

Trata-se do método de “recozimento simulado” ou “simulated annealing”, assim chamado porque o procedimento adotado é análogo ao de resfriamento de metais fundidos. Após um resfriamento (“annealing”), o metal chega a um estado de baixa energia, que é um estado desejável do ponto de vista da Engenharia Metalúrgica.

Eventuais flutuações aleatórias de energia, inerentes ao processo, permitem que o sistema “escape” de estados de energia mínima local e atinja o mínimo global. Se o metal fosse resfriado rapidamente (“quenching”), ele poderia não escapar de estados de energia mínima local e, ao terminar o processo de resfriamento, ele poderia conter mais energia que o metal resfriado lentamente. O resfriamento lento e cuidadoso leva o material para um estado cristalino altamente ordenado, de baixa energia e baixa entropia. Um resfriamento rápido, ao contrário, leva a defeitos e incrustações vítreas no material.

O método de “recozimento simulado” em programação matemática trata a função objetivo tal como a energia é tratada em sistemas físicos submetidos a “annealing”: tenta minimizá-la, admitindo temporariamente eventuais incrementos, para encontrar finalmente o mínimo global.

Nos limites deste trabalho não cabe apresentar um detalhamento do método, mas apenas suas profundas relações com o legado da Termodinâmica. Assim, em grande parte dos métodos de minimização tradicionais, procede-se a uma busca iterativa que somente aceita novos pontos com menor valor da função, o que é análogo a resfriar rapidamente (“quenching”) um sistema físico. O resultado mais provável é ficar “cravado” em um local, sem possibilidades de avançar em direção ao mínimo global. Ao contrário, pelo método de recozimento simulado, são permitidos aumentos temporários na função objetivo (o que equivale a “aquecimentos”), sob controle de um parâmetro análogo à temperatura absoluta. A altas temperaturas, apenas o comportamento “macroscópico” da função objetivo é relevante para a pesquisa do ótimo. Conforme a temperatura diminui, adota-se um procedimento de busca mais “fino”, onde o “material” (a função) é pesquisado com mais detalhe, para se obter um ponto final ótimo.

15 Para um detalhamento, cf. W. L. Goffe, G. D. Ferrier e J. Rogers: “Global Optimization of Statistical Functions with Simulated Annealing”, *Journal of Econometrics* 60 (1994), p. 65-99, North Holland e A. Corana, M. Marchesi, C. Martini e S. Ridella: “Minimizing Multimodal Functions of Continuous Variables with the “Simulated Annealing” Algorithm”, *ACM Transactions on Mathematical Software*, Vol. 113, n. 3, September 1987, p. 262-290.

O autor agradece o acesso a esta literatura a Ana Maria dos Santos Vianna, Fernando Yassuo Chiyoshi e José Paulo de Araújo Mascarenhas.

Um aspecto interessante é o papel do parâmetro “temperatura” no controle do processo de otimização ou “resfriamento”. A cada ponto alcançado X, procede-se à busca de um novo ponto X', gerado aleatoriamente, através de distribuição uniforme n-dimensional, no interior de um paralelepípedo em torno do ponto anterior. Os valores da função objetivo (as “energias”) nos dois pontos são então comparados. Se $f(x')$ for menor que $f(x)$, o ponto x' é aceito como novo ponto de partida.

Caso $f(x')$ seja maior que $f(x)$, estaria ocorrendo um “aquecimento”, que somente terá alta chance de ser aceito se for pequeno tendo em vista o parâmetro “temperatura”. O critério de aceitação utilizado, denominado “Metropolis”, consiste em calcular $f(x') - f(x)$ (valor numérico do “aquecimento”) e aceitar o novo ponto com probabilidade.

$$P(\text{aceitação de } x') = e^{-f(x')-f(x)/T}$$

Note-se aqui que dois fatores contribuem para a aceitação de um eventual “aquecimento”:

- que este “aquecimento” seja pequeno
- que a temperatura seja alta

Em resumo, quando ocorre um “aquecimento”, este será admitido, com alta probabilidade, se $f(x) - f(x') < T$, o que implica em

$$P(\text{aceitação de } x') = e^{-f(x)-f(x')/T} \approx 1$$

A analogia com a Termodinâmica estende-se aqui à Física Moderna.

É sabido que a distribuição de Maxwell-Boltzmann encontra sua extensão, a nível de partículas sub-atômicas, nas distribuições de Fermi-Dirac e de Bose-Einstein. É interessante lembrar aqui a condição para que estas duas últimas distribuições recaiam na de Maxwell-Boltzmann¹⁶.

$$e^{-E/kT} \ll 1 \text{ ou } E \gg kT$$

Aqui notamos mais uma vez um legado da Termodinâmica. A partir das relações $f(x') - f(x) < T$ e $E > kT$, pode-se notar que o critério de aceitação de um eventual “aquecimento” em “simulated annealing” guarda relação com o critério de aceitação da aproximação das distribuições de Fermi-Dirac e Bose-Einstein à de Maxwell-Boltzmann. No primeiro caso, a aceitação de um “aquecimento” ocorre preferencialmente quando este é pequeno se comparado com o parâmetro “temperatura”. Já no segundo, a aproximação é admitida para sistemas de alta energia ($E \gg kT$).

Na literatura sobre “simulated annealing”, a função objetivo é associada à energia E. Tomando-a entretanto como análoga a E/k (energia normalizada pela constante de Boltzmann), o critério de aceitação de um “aquecimento” torna-se

$$f(x') - f(x) = E'/k - E/k \ll T \text{ ou } E' - E \ll kT$$

O que equivale ao critério de aceitação das “novas” distribuições (Fermi-Dirac ou Bose-Einstein) como preferíveis à de Maxwell-Boltzmann, em sistemas com baixa energia.

16 Cf. P. A. Tipler: “Física Moderna”, Editora Guanabara Dois, Rio de Janeiro, 1981, p. 304.

Com este exemplo, cremos ter apresentado um dos mais recentes desdobramentos da Termodinâmica e da Mecânica Estatística, que estão "migrando" e encontrando aplicações em outras áreas de conhecimento.

Conclusão

A oportunidade de revisitar a obra de L. Boltzmann permite-nos aprender com a dificuldade de trazê-lo para o presente, onde predomina, no meio científico, uma "cultura do sucesso" (cf. Item I.).

Neste contexto, é difícil avaliar a contribuição de um eminente cientista cujas teorias "não deram certo" e que, ainda em vida, viu suas proposições serem rejeitadas por um paradigma científico emergente.

Uma tentativa de contornar esta dificuldade, inerente a estudos em História da Ciência em nosso tempo, é pesquisar e apresentar exemplos de implicações e desdobramentos, mesmo que involuntários, da obra que "não deu certo". De preferência, exemplos que "deram certo", como é o caso da Mecânica Quântica, da Ciência da Informação e de recentes métodos em Programação Matemática. Assim, estar-se-ia tentando "esclarecer" o leitor, no sentido de que, na obra em questão, "nem tudo teria dado errado".

O fenômeno da "migração" de conceitos e de metodologias entre áreas de conhecimento distintas é já bastante reconhecido e tem sido explorado na formação de novas áreas de pesquisa interdisciplinar. Neste contexto, o legado de Boltzmann pode ser considerado particularmente rico, pois impressiona a gama de conceitos provenientes de sua obra, atualmente presentes em diferentes áreas da pesquisa.

Mesmo reconhecendo os limites da tentativa empreendida neste artigo, pois não rompe com a "cultura do sucesso" e não ousa dizer que "não dar certo" em ciência é também uma opção, o autor consideraria seu esforço recompensado se sua leitura puder trazer alguma sensação de nossa dívida com relação à obra de Boltzmann.

ROBERTO CINTRA MARTINS é Professor do Instituto de Matemática e COPPE da Universidade Federal do Rio de Janeiro/UFRJ

Endereço: COPPE/UFRJ - Caixa Postal 68.507 - CEP 21.945-970 - Rio de Janeiro - RJ - Brasil

Tel/Fax:(021)590-4144

Revista da SBHC, n. 13, p. 81-94, 1995