Universidade de São Paulo/Faculdade de Educação novembro/2012
SEMA - Seminários de Ensino de Matemática
Coordenador: Nílson José Machado

A Matemática da Informação

Marisa Ortegoza da Cunha
marisa.ortegoza@gmail.com

Em torno de 1950, houve o desenvolvimento simultâneo- e independente – de três áreas de estudo, que logo se entrelaçaram:

- Cibernética – ciência geral de controle de sistemas;
- Informática – teoria matemática da informação, que define uma medida da quantidade de informação contida numa mensagem;
- Computação – realização automatizada de cálculos, pela decomposição de tarefas complexas numa sequência de tarefas simples encadeadas.

Podemos considerar que:
Comunicação é transmissão de uma informação; informação é a novidade contida numa notícia ou numa mensagem, e notícia é uma ordenação de signos (ou, nas palavras de H.Zemaneck, "uma espécie de mistura de surpresa e recapitulação".)

A informação de uma notícia é:

- Do ponto de vista do emissor, uma medida do grau de liberdade na escolha dos elementos da mensagem, e
- Do ponto de vista do receptor, uma medida do grau de imprevisibilidade da notícia.

Para que uma comunicação se processe, os elementos essenciais são:
Emissor da mensagem – [canal – meio] – receptor da mensagem

Os problemas na transmissão de uma informação podem ser classificados em três tipos ou níveis:
Problemas sintáticos – Com que precisão os símbolos utilizados na comunicação podem ser transmitidos?
Problemas semânticos – Com que precisão os símbolos transmitidos conduzem o significado desejado?
Problemas pragmáticos – Quão efetivamente o significado recebido afeta a conduta do receptor, segundo a intenção do emissor da mensagem?

Shannon e Weaver, em 1948, desenvolveram uma teoria da medida da quantidade de informação de uma mensagem, abordando exclusivamente o aspecto sintático dessa mensagem.
Uma teoria da informação possui os seguintes componentes:

- Um repertório fixo de n signos elementares;
- Uma coleção de regras para encadear os signos.

Uma mensagem é um elemento do repertório ou um encadeamento de tais elementos. A quantidade de informação de uma mensagem depende do repertório de mensagens existentes.

Vamos considerar, inicialmente, que as mensagens tenham, todas, a mesma probabilidade de ocorrer (isto é, um repertório equiprovável).

- n = 1

Repertório = {A}
Neste caso, há somente uma mensagem possível. Logo, não há liberdade de escolha para o emissor nem imprevisibilidade para o receptor. Não há novidade. A quantidade de informação é zero.

- n =2

Repertório = {A, B}
Neste caso, há uma liberdade de escolha; uma decisão entre duas. Para identificar uma mensagem desse repertório basta uma pergunta do tipo SIM/NÃO, ou seja, basta uma decisão binária.
A teoria da informação adota a decisão binária (bit, de **bi**nary digi**t**), como unidade da quantidade de informação (que indicaremos por I). Podemos, então, dizer que, neste caso,

$$I(A) = I(B) = 1 \text{ bit}$$

- n = 4

Repertório = {A, B, C, D}
Para identificar uma mensagem escolhida nesse repertório, precisaremos de duas perguntas do tipo SIM/NÃO, de duas decisões binárias.
Logo, I(A) = I(B) = I(C) = I(D) = 2 bits.
Observe que dizer que uma mensagem tem 2 bits de informação é dizer que ela tem 1 chance em 4, de ocorrer.

Procedendo dessa forma, podemos construir a seguinte tabela:

n	1	2	4	8	16	...	n
I	0	1	2	3	4	...	k

Notemos que, para n igual a uma potência de 2, a quantidade de informação **I** de cada mensagem é igual a um número k tal que

$$n = 2^k$$

Temos agora a seguinte questão: como determinar a quantidade de informação de uma mensagem escolhida num repertório cuja cardinalidade não seja uma potência de 2?

Por exemplo, seja <u>n = 3.</u>
Repertório = {A, B, C}
Podemos identificar o elemento escolhido por meio de 1 pergunta ou necessitar de 2 perguntas para isso. A quantidade de informação será, então, um número entre 1 e 2 – uma "média" desses valores.

Por analogia ao que ocorre quando n é potência de 2, o problema se resume a:

dado n, obter k tal que $n = 2^k$

Assim, dado um repertório Rep = $\{A_1, A_2, \ldots, A_n\}$, equiprovável, a quantidade de informação da mensagem A_i é dada por:

$$I(A_i) = \log_2 n$$

Até aqui, como todas as mensagens possuíam a mesma probabilidade de ocorrência ($p = 1/n$), possuíam também a mesma quantidade de informação ($\log_2 n$). Assim, a quantidade de informação está diretamente relacionada com a variedade oferecida pelo repertório: repertórios com um maior número de elementos fornecem maior quantidade de informação.

O problema que surge, a seguir é: Como medir a quantidade de informação de mensagens de um repertório **não** equiprovável?

Para Shannon e Weaver, quanto mais rara uma mensagem, mais informação ela contém, ou seja, a informação de uma mensagem é inversamente proporcional à sua chance de ocorrer.
Por exemplo, dado o repertório Rep = $\{A_1, A_2, \ldots, A_n\}$, com a distribuição de probabilidades $p(A_i) = p_i$, por extensão do que vimos para repertórios equiprováveis, a quantidade de informação da mensagem A_i é dada por

$$I(A_i) = \log_2 \frac{1}{p_i}$$

Trata-se, realmente, de uma generalização do caso equiprovável, pois se todas as mensagens têm probabilidade de ocorrência igual a 1/n,

$$I(A_i) = \log_2 \frac{1}{p} = \log_2 1/(1/n) = \log_2 n,$$

como vimos anteriormente.

Uma palavra sobre os Infons

Os matemáticos Jon Barwise, John Perry e Keith Devlin propuseram, por volta de 1990, uma unidade para quantificar a informação de uma mensagem, levando em consideração o seu conteúdo (ou seja, sua semântica).

De forma simplificada, essa unidade - o infon – pode ser descrita como formada por dois elementos: 1 bit, que traduziria simplesmente a ocorrência (ou não) de determinado fato referente a certos objetos; e uma relação, ou um feixe de relações entre os elementos constitutivos do contexto em que tais objetos se inserem. Assim, o infon é descrito como uma (n+1)-upla:

$$(b;\ c_1, c_2, c_3, \ldots, c_n)$$

Em que b assume o valor zero ou o valor 1 e cada parâmetro c_i ($i = 1, \ldots, n$) assume um valor compatível com o contexto em que o fato em questão ocorre (caso em que $b = 1$) ou não (caso em que $b = 0$).

O infon representaria uma informação a respeito de uma situação, considerando-se um certo contexto. Infelizmente, não há notícias de que a pesquisa a respeito da aplicabilidade dos infons tenha se desenvolvido, nos últimos anos.

Bibliografia

ASHBY, W. Ross. *Introducción a la cibernética*. Buenos Aires: Ediciones Nueva Visión, 1976 (1956).

BARWISE, Jon & Perry, John. *Situations and Attitudes*. Cambridge: MIT Press , 1983.

DEVLIN, Keith. *Logic and Information.* New York: Cambridge University Press, 1991.

MACHADO, Nílson J. e CUNHA, Marisa O. - Bits e infons: a matemática da informação*. in Linguagem, Conhecimento, Ação – ensaios de epistemologia e didática*. (Machado, N.J. e Cunha, M. O. org.). Coleção Ensaios Transversais, vol. 23. São Paulo: Escrituras, 2003.

MASER, Siegfried. *Fundamentos de teoria geral da comunicação*. São Paulo: E.P.U., EDUSP, 1975.

SHANNON, Claude E. & WEAVER, Warren. *The mathematical theory of communication*. University of Illinois Press, 1969.

SINGH, Jagjit. Teoria de la información, del lenguaje y de la cibernética. Marid: Alianza, 1972.