

# Informação e redundância. Os Quadros da Incerteza\*

António Fidalgo  
Universidade da Beira Interior

## Índice

1 Informação e certeza	1
2 Os bits de informação	2
3 A redundância necessária	4
4 A redundância desejável	6
5 Os quadros variáveis da redundância	8

## 1 Informação e certeza

Os dois conceitos fundamentais da Teoria Matemática da Comunicação, proposta por Shannon e Weaver, a saber, informação e redundância, são definidos mediante um conceito filosófico de larga tradição, o conceito de certeza. É sabido que a teoria matemática da comunicação é fundamentalmente uma teoria sobre a quantidade e a medição da informação veiculada por um canal. Ora a grande intelecção desta teoria é que a informação dada é inversamente proporcional à sua probabilidade, ou seja, que a informação é uma propriedade estatística de um signo ou de uma mensagem. Quanto mais provável for um signo, menor a sua informação. Informação é incerteza e redundância é certeza.

---

\*Publicado em José Manuel Santos e João Carlos Correia, orgs, *Teorias da Comunicação*, Covilhã: Universidade da Beira Interior, 2004, p. 15-28.

O modelo cartesiano de conhecimento também assenta na noção de certeza, visto que só conhecemos verdadeiramente aquilo de que temos a certeza absoluta, de que de todo não podemos duvidar. A intelecção fundamental de Descartes é justamente a de que o “cogito” é a base sólida da ciência enquanto edifício de conhecimento verdadeiro e de que essa primeira certeza é modelo, critério e pedra de toque de todos os outros conhecimentos posteriores. O que caracteriza a certeza e verdade do “cogito” são a clareza e a distinção da percepção em que é dado, pelo que todas as percepções claras e distintas podem e devem ser acrescentadas ao núcleo das primeiras certezas.

Poder-se-á dizer que em Descartes sabemos aquilo de que estamos certos e que, portanto, a tarefa do cientista ou do investigador é a de alargar a esfera de certezas cujo núcleo é o “cogito”, num movimento contínuo e infinito de tornar conhecido o que é desconhecido, ou seja, de tornar certo o que é incerto.

À primeira vista o cartesianismo e a teoria matemática da comunicação estariam em campos completamente opostos na medida em que enquanto o primeiro associa o conhecimento à certeza o segundo identifica informação com incerteza. Tal entendimento, todavia, não teria em conta a diversidade de como os termos “informação” e “conheci-

mento” são aqui utilizados. Cometer-se-ia o erro de reduzir o cartesianismo a uma concepção patrimonial de conhecimento, o que sabemos é aquilo de que já estamos certos, e de o contrapor a uma concepção processual de informação, do que ainda não sabemos. São, contudo, coisas distintas e não podem ser contrapostas sem mais. Aliás, nada impede de considerar que o processo cartesiano de adquirir certezas é um processo informativo, pois que a incerteza é condição de novos conhecimentos. Conhecer não seria repisar o já sabido, mas em tornar o incerto e o desconhecido em certo e conhecido.

Uma e outra posição estariam erradas, ou seja, tanto a contraposição como a compaginação dos dois modelos descuram ou ignoram a especificidade radical da noção de informação na teoria matemática da comunicação, especificidade essa que impede de estabelecer quaisquer associações apressadas entre os dois modelos. É que a certeza cartesiana é substantiva, isto é, tem-se a certeza do conteúdo de uma percepção, ao passo que a certeza da teoria matemática da comunicação é de cariz meramente probabilístico, não tendo minimamente em conta o conteúdo ou o sentido da informação.

Weaver chama explicitamente a atenção para o facto de o termo “informação” ser usado num sentido muito especial e não dever ser confundido com o seu sentido habitual.<sup>1</sup> De um ponto de vista probabilístico, “duas mensagens, uma das quais seja repleta

<sup>1</sup> Claude E. Shannon and Warren Weaver, *The Mathematical Theory of Communication*, Urbana and Chicago, University of Illinois Press, (1949) 1963. “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.”, pg. 8

de significado e outra não tenha qualquer sentido, podem ser exactamente equivalentes.”<sup>2</sup> Ou, dito de uma forma mais exacta, o “termo informação na teoria da comunicação não respeita tanto ao que de facto se diz, mas como ao que poderia ser dito. Isto é, informação é uma medida da liberdade de escolha quando se selecciona uma mensagem.”<sup>3</sup>

É claro agora que, perante o sentido probabilístico de informação, não é possível contrapor a certeza cartesiana do conhecimento à incerteza da informação, nem tão pouco associá-las. O conhecimento cartesiano é um conhecimento substantivo, ou seja, um conhecimento certo de algo bem determinado. As percepções claras e distintas que constituem o conhecimento cartesiano são unidades repletas de significado.

Ora, e esta é a tese que me proponho defender aqui, a incerteza informativa e probabilística tem como condição a certeza substantiva enquanto quadro de incerteza. Só tem sentido falar de incerteza probabilística se houver um quadro não infinito de possibilidades, quadro esse que é escolhido por razões já não probabilísticas, mas de conteúdo ou de estratégia. Veremos que é esse o sentido essencial de redundância.

## 2 Os bits de informação

Sendo a informação unicamente probabilística, uma unidade de informação não é uma mensagem, a percepção ou a intelecção de

<sup>2</sup> *ibidem*.

<sup>3</sup> - *ibidem*, p.8. No ensaio original, Shannon escreve que: “These semantic aspects of communication are irrelevant to the engineering problem. The significant aspect is that the actual message is one selected from a set of possible messages”. *Ibidem*, p. 31.

um facto, de uma ocorrência ou de um estado de coisas, mas sim a possibilidade de escolha mais simples, isto é, de escolha entre duas mensagens, seja o conteúdo destas qual for. Como diz Weaver: “As duas mensagens entre as quais há que escolher, numa tal selecção, podem ser o que quisermos. Uma poderia ser o texto da Bíblia, na versão de King James, e a outra poderia ser ‘Sim’.”<sup>4</sup>

É esta unidade de escolha simples, alternativa, entre duas mensagens que pode ser representada por “zero” e “um”, sendo “zero” o sinal para a primeira escolha da primeira mensagem e “um” o sinal para a escolha da segunda mensagem. Um circuito eléctrico aberto ou fechado, uma lâmpada apagada ou acesa, podem então funcionar como sinais alternativos de não, correspondendo ao “zero”, ou sim, correspondendo ao “um”. Se houvesse apenas uma mensagem então não haveria qualquer informação na medida em que não haveria qualquer incerteza. À partida saber-se-ia que era aquela e não outra. O mínimo da informação, a unidade, é a escolha entre duas mensagens.

É justamente daqui que nasce a ideia de medir a informação pelo logaritmo das escolhas possíveis. Reduzindo a informação às unidades mínimas, a escolhas simples entre duas alternativas, verifica-se que é proporcional ao logaritmo de 2 de base 2, ou seja à unidade. Com efeito 1, a unidade, é o logaritmo de 2 com base 2, cuja expressão matemática é a seguinte:  $\log_2 2 = 1$ . É a esta unidade de informação que se passou a chamar um “bit”, abreviação do termo inglês “binary digit”, isto é, de dígito binário. Enquanto o sistema decimal utiliza dez dígitos, do 0 ao 9, o sistema binário tem apenas dois dí-

gitos, 0 e 1. E porque estes dois números podem representar quaisquer escolhas alternativas, então o “bit” é a unidade de informação ao representar a situação de escolha simples entre duas mensagens.

Definir a informação como o logaritmo do número de escolhas possíveis tem desde logo a seguinte vantagem: se a um único circuito aberto ou fechado, e representado por 0 ou por 1, corresponde apenas uma unidade de informação, a dois circuitos correspondem duas unidades de informação ou dois bits. Ora isso corresponde a 4 escolhas possíveis: 00, 01, 10, 11. No caso de 00 os dois circuitos encontram-se desligados, no segundo caso o primeiro circuito está desligado e o segundo ligado, no terceiro caso o primeiro está ligado e o segundo fechado, e no quarto caso os dois estão ligados. Se tivermos 3 circuitos teremos então 3 bits ou 8 possibilidades, com 4 bits 16 possibilidades, e assim por diante. À sucessão, um a um, de bits corresponde um aumento exponencial de possibilidades. Duplicando o tempo, o número de possibilidades é elevado ao quadrado, ou seja, duplica-se o algoritmo; o que significa duplicar a informação medida logaritmicamente.

Medir a informação em bits significa, portanto, medir o número de escolhas possíveis. Um exemplo ilustrativo será o de encontrar uma palavra num dicionário através do número de escolhas alternativas em divisões sucessivas do dicionário: a palavra encontra-se ou na primeira ou na segunda parte do dicionário; depois na primeira ou na segunda parte da divisão certa; e assim sucessivamente, até chegar à palavra. A quantidade de informação será o número de divisões e escolhas necessárias até chegar à palavra pretendida.

<sup>4</sup> - *ibidem*, p. 9.

A teoria matemática da comunicação só ganha, porém, verdadeira pertinência quando analisa as possibilidades a partir da sua probabilidade. É que num conjunto possível de escolhas há umas que são mais prováveis que outras. Se tomarmos como exemplo a língua enquanto sistema de múltiplas e variadas combinações de signos linguísticos, então será fácil verificar que a possibilidade de a um adjectivo se seguir um substantivo é muito superior à de se lhe seguir um advérbio. Quer isto dizer que a língua é um sistema em que certas combinações de signos são mais prováveis que outras e que a tarefa da teoria matemática é a de medir essa probabilidade. É óbvio que o grau de probabilidade de que, numa frase bem construída em português, às palavras “O homem que ontem ...” se siga um verbo, como por exemplo “esteve”, “falou”, é altíssima, enquanto a probabilidade de se lhe seguir um substantivo é baixa, embora não igual a zero. Seria possível, com efeito, continuar com “sábado”. Mas a probabilidade de se lhe seguir uma forma verbal no futuro “estará”, “falará”, essa sim seria igual a zero.

Numa sequência de signos o grau de liberdade de escolha é condicionado e limitado pelas escolhas prévias. Há casos em que a determinação é total, como no caso em que no português escrito à letra “q” se segue sempre um “u”. Neste caso não há qualquer liberdade de escolha e, portanto, a informação é nula.

A razão por que podemos falar de entropia na comunicação é precisamente porque a selecção dos signos discretos de que se compõe uma mensagem é comandada por probabilidades. Se houver uma grande liberdade de escolha, então a entropia é grande e podemos dizer que há muita informação. Se a or-

ganização for elevada, e não houver grande margem para escolher, então a informação é baixa.

### 3 A redundância necessária

Muitas vezes o conceito de redundância é contraposto ao conceito de ruído, entendido este como toda e qualquer perturbação que interfira no processo comunicacional. Entendido apenas neste aspecto, o conceito de redundância será unicamente uma reduplicação da mensagem, ou de partes da mensagem, em ordem a confirmar a mensagem, e a obviar ao ruído. Contudo, o conceito de redundância é mais abrangente do que o de correlato de ruído, e muito mais importante quando relacionado com a noção de entropia.

Se considerarmos todas as possibilidades como tendo a mesma probabilidade, então temos entropia máxima. Neste caso, a falta de organização é total e a liberdade de escolha é completa. A este estado limite contrapõe-se a entropia efectiva de um dado estado de coisas ou de uma certa fonte de informação, onde existe condicionamento de selecção. A relação da entropia efectiva com a entropia máxima é a entropia relativa da fonte.<sup>5</sup> Assim, por exemplo, se a entropia relativa de uma determinada fonte de informação for de 0.8, isso significa que a liberdade de escolha dos signos para constituir a mensagem é de 80% relativamente aos 100% de liberdade de selecção que a entropia máxima permitiria. A redundância é justamente a diferença que existe entre a entropia máxima e a entropia relativa.<sup>6</sup>

<sup>5</sup> - *ibidem*, p. 13.

<sup>6</sup> - “One minus the relative entropy is called the redundancy.” *ibidem*.

Weaver acrescenta que a redundância “é a fracção da estrutura da mensagem que não é determinada pela livre escolha do emissor, mas antes pelas regras estatísticas aceites que regem o uso dos signos em jogo.”<sup>7</sup> Por exemplo, as concordâncias de número, singular ou plural, de género, masculino ou feminino, de tempo, passado ou presente ou futuro, são claramente redundâncias que regem a constituição de frases em português. Se alguém quiser definir com o artigo o substantivo “homens” tem de o fazer respeitando o género e o número, ou seja, tem de ser “os homens” e não pode ser nem “o homens”, caso em que apenas respeitaria o género, nem “as homens”, em que respeitaria o número, mas não o género. Porquê o termo redundância para designar esta parte da mensagem que restringe a liberdade de escolha? Weaver diz que é porque essa parte da mensagem é desnecessária no sentido de que se faltasse a mensagem continuaria a estar essencialmente completa.<sup>8</sup>

De facto, as partes redundantes da mensagem constituem algo que não traz novidade e, portanto, serão desnecessárias nesse sentido. A não necessidade da citada redundância do artigo definido em português em género e número torna-se clara quando comparado com o artigo definido inglês “the” que não conhece nem género nem número. Tendo isto em conta, e ainda e sobretudo a personalização das formas verbais que exis-

<sup>7</sup> - *ibidem*.

<sup>8</sup> - “It is sensibly called redundancy, for this fraction of the message is in fact redundant in something close to the ordinary sense; that is to say, this fraction of the message is unnecessary (and hence repetitive or redundant) in the sense that if it were missing the message would still be essentially complete, or at least could be completed.” *ibidem*.

tem na língua portuguesa em muito maior grau que na língua inglesa (às seis formas pessoais dos verbos portugueses – de *eu amo* a eles *amam* - correspondem no geral apenas duas nos verbos ingleses – *I, you, we, they love, he loves*), é quase certo que a percentagem de redundância em português é superior aos 50% de redundância que Shannon e Weaver atribuem à língua inglesa, em que apenas metade das letras ou das palavras que escrevemos ou dizemos são de livre escolha de quem fala, e que a outra metade é ditada pela estrutura estatística da língua.<sup>9</sup>

Na parte redundante de uma mensagem, ou seja, na parte que escapa à livre escolha da fonte de informação, há que distinguir entre elementos imprescindíveis, estruturantes da própria mensagem, e elementos prescindíveis. Weaver parece, com efeito, significar que toda a redundância, pelo facto de o ser, é desnecessária, mas esse não é o caso, nem o pode ser. Tomemos de novo uma língua natural, o português, como exemplo, e tentemos numa mensagem, género telegrama, eliminar todas as redundâncias possíveis. Rapidamente verificamos que, sob pena de incompreensão, há regras que têm de ser necessariamente observadas. Essas regras não são da livre escolha da fonte de informação, antes lhe são impostas pela estrutura da língua e, como tal, redundantes.

Quando Shannon escreve que a informação não se reporta a uma mensagem, mas sim à escolha dentro de um conjunto (*set*) de mensagens possíveis, isso significa que

<sup>9</sup> - “It is most interesting to note that the redundancy of English is just about 50 per cent, so that about half of the letters or words we choose in writing or speaking are under our free choice, and about half are really controlled by the statistical structure of the language.” *ibidem*.

esse conjunto tem de estar identificado no processo de comunicação entre emissor e receptor. Esse conjunto é ele próprio de cariz redundante, na medida em que não é objecto de escolha, mas se encontra dado à partida. Voltemos ao caso dos circuitos. Se tivermos 3 circuitos sabemos que existem 8 combinações possíveis na medida em que 3 bits são  $2^3$ . Mesmo tomando como entropia máxima a liberdade de usar essas 8 combinações com o mesmo grau de probabilidade, teremos necessariamente de ter em conta, isto é, como um dado prévio, que há apenas 3 e não mais circuitos, que não pode aparecer um quarto circuito em jogo. Ou seja, a total liberdade de escolha tem de ser feito a partir de um determinado conjunto de possibilidades, e é justamente esse determinado conjunto prévio que constitui o quadro necessário da informação.

É certo que há sistemas em que o grau de redundância é muito superior ao de outros. Uma língua natural, para ser compreensível, tem de obedecer a regras de estrutura e de sentido, regras que ao serem conhecidas de antemão por emissor e receptor representam informação partilhada por ambos. São obviamente elementos redundantes em qualquer mensagem construída nessa língua. No totoloto, porém, o grau de incerteza é muito maior. Mas mesmo aqui há um elemento redundante, a saber, o número de elementos que poderão ser seleccionados. Se a selecção de seis números fosse feita a partir de 99 números em vez de 49, então a incerteza seria muito maior. Mas mesmo o maior número possível de elementos não eliminaria toda a incerteza. A definição dos elementos passíveis de serem seleccionados é condição de selecção e, portanto, um elemento de redundância.

Seguindo o próprio modelo comunicacional de Shannon – fonte de informação, transmissor, sinal emitido, canal, sinal recebido, receptor, destinatário –, é condição fundamental de uma comunicação efectiva que a codificação da mensagem seja bem feita. O código constitui aqui um elemento essencial que deve ser partilhado por ambos os lados do processo comunicativo. Ora é justamente essa partilha, esse ponto comum, que constitui a redundância necessária subjacente à comunicação. Ou seja, apesar da redundância não representar qualquer informação, ela é fundamental para a exactidão da mensagem e mesmo para a sua ocorrência.<sup>10</sup>

#### 4 A redundância desejável

Atendendo a que informação ou incerteza e redundância ou certeza são imprescindíveis numa mensagem, coloca-se a questão da justa medida ou proporção entre informação e redundância numa mensagem. Se por um lado, há tendência para que uma mensagem contenha a maior informação possível, por outro lado também se pretende que a mensagem seja o mais rigorosa possível. São tendências que, de algum modo, se opõem, e daí que se coloque a questão sobre a redundância desejável de uma mensagem. Por outro lado, a adequação dos códigos utilizados é extremamente pertinente neste campo. Determinado código pode ser muito mais eficaz a

<sup>10</sup> - "Shannon e Weaver mostram como a redundância facilita a exactidão da decodificação e fornece um teste que permite identificar erros. Só me é possível identificar um erro ortográfico devido à redundância da linguagem. Numa língua não redundante, mudar uma letra significaria mudar a palavra." John Fiske, *Introdução ao Estudo da Comunicação*, Lisboa: Edições Asa, 1993, p. 25.

codificar uma mensagem na medida em que aumenta o grau de informação, sem prejuízo do rigor. As linguagens especializadas são um exemplo dessa adequação e eficácia.<sup>11</sup>

Há códigos ou quadros mais adequados do que outros para a formação de mensagens. Existem quadros que dão menos origem a entropia, quadros que permitem mais criatividade (escolha), quadros mais rígidos, quadros mais maleáveis. É que podemos subir uma escala mais e também escolher os quadros dentro dos quais podemos formar uma mensagem.

Provavelmente aqui poderíamos alargar o sistema comunicacional de Shannon, dando também um significado de redundância ao canal, e não apenas ao código. A mesma mensagem pode ser enviada por diferentes canais, e existe a liberdade de escolha desses canais. Posso telefonar a alguém para lhe dizer uma coisa, posso enviar-lhe um email, posso enviar-lhe uma carta ou um fax, ou posso mandar um recado por outrém. Mas mesmo estas escolhas são feitas dentro de um certo quadro de redundância e por isso temos de verificar as diferentes probabilidades.

Se uma relação adequada entre informação e redundância é crucial para o sucesso comunicativo ao nível técnico, ou seja, ao nível da exactidão com que os signos da comunicação podem ser transmitidos, sê-lo-á ainda mais relativamente ao que Weaver chama os níveis B e C do processo comunicativo, a saber, o nível semântico, relativo à precisão com que os signos transmitidos

convêm ao significado desejado, e o nível da eficácia, relativo à eficácia com que o significado da mensagem afecta da maneira desejada a conduta do destinatário.

Só os diferentes contextos, as razões, as causas e os fins de uma mensagem, podem estabelecer qual a melhor relação entre informação e rigor que a mensagem deve conter. Aliás, a optimização dessa relação pode ser diferente nos diferentes níveis de uma mensagem. Por exemplo, uma mensagem repetida, que no nível técnico e no nível semântico será pura redundância, pode ser altamente informativa a nível de eficácia, justamente porque inesperada. Eventualmente poderia julgar-se que os níveis técnico, semântico e de eficácia, são etapas sucessivas do processo comunicativo. Que a resolução do problema técnico é prévia à do problema semântico e esta à do problema da eficácia. Mas não é assim. O nível semântico pode constituir, e constitui normalmente, um factor redundante importante na decifração técnica da mensagem. Chega-se frequentes vezes à decifração acústica de uma palavra, pelo significado que lhe associamos. Por sua vez, há circunstâncias em que se visa mais a exactidão semântica do que a eficácia da mensagem. Se alguém se dirigir a uma outra pessoa para lhe dar uma ordem e esta não obedecer, o emissor da mensagem pode então limitar-se a estabelecer que a mensagem era clara, que o destinatário entendeu bem a mensagem ou ordem veiculada.

A redundância desejável é, assim, uma variável que depende de muitos factores. A novidade e exactidão que têm de compor qualquer mensagem conciliam-se em grau e modos diferentes.

<sup>11</sup> - Sobre linguagens especializadas, veja-se António Fidalgo, "A economia e a eficácia dos signos", em particular a secção intitulada "Os signos à medida. As linguagens especializadas." Texto disponível em [www.bocc.ubi.pt](http://www.bocc.ubi.pt)

## 5 Os quadros variáveis da redundância

A redundância que necessariamente envolve toda a informação e, em maior ou menor grau, enforma qualquer mensagem não é fixa, nem constante. Ou seja, o conjunto de possibilidades no seio do qual se dá ou obtém informação varia à medida de múltiplos tipos de circunstâncias e de estratégias. Tomemos novamente uma língua natural como exemplo. Em princípio, a língua é um quadro estabelecido de possibilidades linguísticas. Gramáticas e dicionários delimitam os elementos e as formas de expressões bem formadas. No entanto, há áreas ou actividades em que se permite, e até se incentiva, ultrapassar e infringir as regras estabelecidas, ou ir para além do quadro de possibilidades dado à partida ao falante. A poesia, o calão, são exemplos de áreas linguísticas onde a mutabilidade da língua é visível. Ou seja, o quadro de certeza que demarca as possíveis incertezas é ele mesmo passível de alterações, modificações, e de reenquadramento.

As noções que a linguística e filosofia contemporâneas têm introduzido no estudo das línguas e da linguagem, como contextos, actos de fala, jogos de linguagem, estratégias, representam contributos importantes para perceber a mutabilidade dos quadros de referências em que a comunicação é feita. O modelo rígido e único de Descartes cedeu o passo a modelos maleáveis, configuráveis, em que não conta o tal fundamento sólido e indubitável, mas em que predominam as concepções estratégicas.

A crítica de Charles Sanders Peirce à ideia cartesiana de certeza pode contribuir para enriquecer a leitura filosófica da teoria ma-

temática da comunicação. Peirce não considera que a clareza ou a certeza de uma ideia se relacione com a probabilidade de uma mensagem, mas antes com o modo como uma ideia se desenvolve e se revela nas consequências que acarreta no proceder e no agir de quem a tem. O que importa todavia aqui realçar é a noção de contexto que a máxima pragmatista impõe na concepção da ideia: “considera quais os efeitos, que podem ter certos comportamentos práticos, que concebemos que o objecto da nossa concepção tem. A nossa concepção dos seus efeitos constitui o conjunto da nossa concepção do objecto.”<sup>12</sup>

Mesmo compreendendo a informação como unidades de escolha entre múltiplas possibilidades, à maneira de Shannon, a interpretação pragmatista viabiliza uma compreensão mais rica, porque mais plural, de comunicação, na medida em que as mensagens poderão ser inseridas em diferentes conjuntos de redundância, isto é, de certeza. Os quadros redundantes da informação cruzam-se e entrecruzam-se e, assim, aumenta a incerteza, e *eo ipso* a informação.

<sup>12</sup> Charles Sanders Peirce, “Como tornar as nossas ideias claras”. Texto disponível em [www.bocc.ubi.pt](http://www.bocc.ubi.pt).