

variedade semântica - semantic manifold

ao [Cazuza](#) e ao Valupi



[an index] represents the objects independently of any resemblance to them, only by virtue of real connections with them.

C. S. Peirce

A estátua deste monstro era de prata em vulto de homem agigantado, de vinte e sete palmos em alto, tinha os cabelos de café, e as ventas dos narizes muito disformes, e os beiços grossos, e toda a fisionomia do rosto tristonha e mal assombrada. Tinha na mão uma bisarma a modo de segur de tanoeiro, mas c'ó cabo muito mais comprido, com o qual diziam os Sacerdotes ao povo que a noite passada matara a serpe tragadora da cônica funda da casa do fumo, por querer roubar a cinza dos sacrificados (...)

Peregrinação, Fernão Mendes Pinto

José Pinto Casquilho¹

¹ Centro de Ecologia Aplicada Baeta Neves (CEABN/UTL) ; Centro de Estudos de Comunicação e Linguagens (CECL/UNL). josecasquilho@gmail.com

Variedade semântica é uma expressão que existe em português, e reporta-se seja a um espaço [semântico](#) (Turunen, 2008) ou a um enquadramento [metafísico](#) (Zillig, 2007) em qualquer caso a propósito das multiplicidades de sentido e de significado.

Neste trabalho que agora escrevo trata-se antes de mostrar uma estrutura matemática na acepção da expressão inglesa *semantic manifold*, conceito que se aplica, por [exemplo](#), no reconhecimento de padrões e recuperação de imagem (Ohbuchi e Shimizu, 2008). O caso que se vai expôr baseou-se em resultados da [análise](#) teórica da composição em proporções de *habitats* em mosaicos de paisagem.

Em matemática, uma *variedade* é um *conjunto* de pontos num espaço, ligados por relações. Uma *variedade diferenciável real* pode ser [definida](#) formalmente, ou entendida através de um modelo, por exemplo uma superfície contínua suavemente imersa no espaço tridimensional euclideano, ou então, baixando uma dimensão: uma curva no plano cartesiano. Já uma *variedade complexa* é a generalização natural de uma superfície de [Riemann](#), onde as coordenadas locais de um ponto são números complexos (Kodaira, 1986)².

A *teoria da Informação* [ou](#), equivalentemente, a *teoria matemática da comunicação*, foi apresentada num [artigo](#) de Claude Shannon, editado em 1948. Na *teoria da Informação*, a [informação](#) associada a um acontecimento que ocorre com probabilidade p é uma quantidade, um número real, que podemos interpretar como a ordem de grandeza do tempo médio de reconstituição, ou repetição, desse acontecimento - quanto mais raro o acontecimento, maior será o seu tempo médio de reconstituição e portanto maior o seu valor de informação, e ainda maior será o efeito-surpresa associado à sua realização. Uma mensagem, ou um acontecimento, é tão mais informativo quanto mais reduzir a incerteza naquele contexto.

Existem outras acepções para o conceito de *informação*: um estado ordenado seria um estado cuja descrição, reprodução ou preparação, exige uma quantidade elevada de informação³. Mais concretamente, o valor ou medida da informação de uma mensagem é um tema original de Hartley, posteriormente retomado por Shannon, e relaciona-se com a ideia de que a informação resolve, ou reduz, a incerteza num espaço de acontecimentos⁴, e a *entropia*

² Kunihiro Kodaira. 1986. **Complex Manifolds and Deformation of Complex Structures**. Springer, Berlin, 2005

³ Ilya Prigogine e Isabelle Stengers. 1993. Ordem/desordem. Enciclopédia Einaudi, vol. 26 : **Sistema**, Imprensa Nacional Casa da Moeda, Lisboa.

⁴ George J. Klir. 2006. **Uncertainty and Information**. John Wiley and Sons Inc , Hoboken.

estatística é a medida da incerteza de uma variável aleatória⁵ que caracteriza esse espaço, sendo a *probabilidade* uma medida de possibilidade ou verosimilhança de um acontecimento. Mais pragmaticamente Weaver define a informação como a medida da livre selecção de uma mensagem⁶, de entre um conjunto.

A incerteza ou entropia da informação, é expressa na célebre fórmula de Shannon, interpretável com um *valor médio* de uma variável aleatória indicadora de um espaço de acontecimentos, ou de símbolos, X , cujos elementos x ocorrem com probabilidade $p(x)$, sendo $y = -\log p(x)$ o valor de informação do acontecimento x . Quanto maior o valor numérico de $H(X)$ maior a incerteza sobre a variável aleatória X .

$$H(X) = - \sum_{x \in X} p(x) \log p(x)$$

A teoria da informação baseou-se na eficiência da transmissão de um código [binário](#) (Touretzky, 2004) e a entropia é a medida da quantidade de informação num processo desse tipo, que pode ser [generalizado](#) para vários processos, simultâneos ou encadeados, e aí torna-se um caso especial de *informação mútua média* (Gray, 2008), onde se define uma *fonte de informação* como um modelo matemático de uma entidade física que produz uma sucessão de símbolos de um alfabeto de uma maneira aleatória. A entropia estatística é usada como medida de [complexidade](#) em sistemas biológicos (Carter, 2007).

A crítica mais comum que é feita a esta medida, $H(X)$, é que ela se reporta exclusivamente à expressão sintáctica do objecto e não ao seu conteúdo semântico: falta-lhe uma interpretação semântica que a possa acoplar ao mundo real⁷, porque os diferentes acontecimentos ou símbolos x não são distinguidos por nenhuma valorização intrínseca, mas apenas pela sua probabilidade de ocorrência enquanto acontecimentos de um espaço estandarizado.

Foi [Peirce](#) quem primeiro distinguiu três níveis na apreciação de um *signo*, uma relação [triádica](#), aspectos derivados de uma propriedade tricotómica genérica e fundamental da qual também se podem inferir as dimensões sintáctica, semântica e pragmática que lhe estão associadas. Determinar a significação de um enunciado equivale a definir o seu valor

⁵ Thomas M. Cover e Joy A. Thomas. 2006. **Elements of Information Theory**. John Wiley and Sons Inc, Hoboken.

⁶ Miguel R. Alsina. (1989) 1995. **Los Modelos de la Comunicación**. Editorial Tecnos SA, Madrid.

⁷ John F Sowa, 2007. Language games, a foundation for semantics and ontology. **Game Theory and Linguistic Meaning**, Aithi Vaiko Pietarinen ed, Elsevier Limited.

semântico⁸, ser capaz de apreender aquilo para que remete a sua expressão, em função da língua comum aos interlocutores.

A relação entre signos e valores foi [analisada](#) por Sharov (1998) no âmbito da biossemiótica, concluindo o autor que a noção de valor (de utilidade) não é tratável na Física, mas é muito importante para compreender o fenómeno da vida, e só seria adequadamente manejada sob o conceito de *fecho semântico*, tradução de *semantic closure* um tema de Patee [debatido](#) como princípio, na genética. O conjunto *fecho semântico* também é [definido](#) na lógica matemática como o conjunto de todas as fórmulas de um espaço que têm consequência semântica global, e pode ser entendido talvez como o termo da clausura de uma linguagem relativamente à atribuição dos valores de verdade de todas as expressões formuláveis num dado espaço de símbolos. Pessoalmente, não entendo o conceito de *fecho semântico* para além de um dispositivo operacional, que permite enquadrar e analisar um trecho de uma linguagem, porque antes subscrevo, junto com outros, que a *semiose*, o processo de significação, é uma gestação [infinda](#), e que a semiótica⁹, como teoria da significação, é uma ciência aberta.

Regressemos então ao exercício que vos apresento, que se pode designar como uma generalização, ou extensão, da fórmula de Shannon. Admitamos que temos k , seja 3 ou 4, para exemplificar, casos ou acontecimentos possíveis de um dado espaço de acontecimentos, descendo ao concreto: seja uma tábua de queijos e o acontecimento 'morder o queijo'. Queremos compôr a tábua de queijos com proporções desses queijos por forma a que a informação seja máxima, considerando que os queijos são diferentes mas têm igual valor? É fácil: tudo em partes iguais, volumes iguais dos queijos presentes na tábua, e a entropia, medida pela fórmula $H(X)$, toma o valor máximo, perante o acontecimento 'morder aleatoriamente os queijos'. Agora: e se admitirmos que os queijos têm valores diferentes, seja o preço, a apetência, ou outra coisa? Como devemos compôr a tábua de queijos por forma a que se possa dizer que verifica algum critério de equilíbrio baseado nesses números?

Antes do mais uma objecção metodológica: estou a admitir que posso reduzir o valor ao preço, um número real, coisa que é matéria muito controversa¹⁰. Noutro [texto](#) defendi, e mantenho, que a redução numérica máxima que podemos fazer para representar o *valor* de um objecto complexo é um *vector* de números reais ou um *número complexo*, que tem duas dimensões.

⁸ Adriano D. Rodrigues. 1996. **Dimensões Pragmáticas do Sentido**. Edições Cosmos, Lisboa.

⁹ José A. Mourão e Maria A. Babo. 2007. **Semiótica: Genealogias e Cartografias**. MinervaCoimbra.

¹⁰ Oscar Wilde: a cynic is a man who knows the *price* of everything but the *value* of nothing

A redução de um *valor* a um número real não pode significar mais do que remeter a um índice. Em rigor a expressão *valor real* identificada com *número real* só existe na categorização interna da linguagem matemática. Trata-se aliás de um problema longamente debatido, através duma tensão dialéctica entre dois conceitos. Assim foi com Marx que, na Economia Política, estabeleceu a dicotomia *valor de uso/valor de troca* sendo o primeiro intrinsecamente subjectivo mas relativo a um uso real, e o segundo interpretado como medida, valor médio, o número de horas de trabalho(s) necessário(s) para produzi-lo num dado contexto social. Também na Linguística, Saussure afirmava que¹¹, tal como na economia, se está em face da noção de *valor*, onde se trata de um sistema de equivalência entre coisas de ordens diferentes: numa entre um *trabalho* e um *salário*, noutra entre um *significado* e um *significante*, constituindo-se dois eixos distintos, duas dimensões. É com Fontanille e Zildeberg que se actualiza o conceito de valor como *estrutura tensiva*, com duas dimensões ou valências, o mais das vezes associadas¹² respectivamente a *intensidade* (ou energia) e *extensão*, ou, correlativamente, valores de absoluto e valores de universo, uma estrutura elementar para [interpretar](#) um objecto semiótico.

Prosseguindo na construção da variedade semântica que aqui apresento, entendida como conjunto matemático, vai-se então admitir que existe um conjunto de números reais associados a índices de valor de k *objectos*, num espaço de *objectos* do mesmo tipo ou natureza, ou um espaço de acontecimentos que gera esses *objectos*, com determinadas probabilidades, e onde acontecimento é interpretado como uma mudança óbvia de propriedades com efeitos semânticos.

Seja então $W=\{w_1, w_2, \dots, w_k\}$ o conjunto dos índices de valor dos k *objectos* num dado espaço e seja $X=\{x_1, x_2, \dots, x_k\}$ o conjunto das suas probabilidades de ocorrência, ou presença, recordando que $\sum_i x_i = 1$. Uma generalização imediata da fórmula de Shannon obtém-se como $H_w = -\sum_i w_i x_i \log x_i$ através de um produto interno de três vectores descrito como $(x, y)_w = \sum_i x_i y_i w_i$. Trata-se, no fim de contas, de avaliar a entropia indexada pelo conjunto W de índices de valor.

Esta fórmula continua a ser interpretável como o cálculo de um valor médio de uma variável aleatória, que noutro lugar designámos por *valor informativo* de um espaço de acontecimentos. Ora, enquanto que na ausência de W a entropia máxima acontecia quando os acontecimentos geradores dos *objectos* tinham igual probabilidade, agora, com a função H_w , a localização do ponto de equilíbrio já vai depender do conjunto de números reais W , assim variando no espaço através de uma parametrização, descrevendo uma variedade

¹¹ Ferdinand de Saussure. 1971 (1916). **Curso de Linguística Geral**. Publicações D. Quixote, Lisboa, 1999.

¹² Jacques Fontanille e Claude Zildeberg. 1998. **Tensão e Significação**. Discurso Editorial, São Paulo, 2001

diferenciável, uma variedade de equilíbrios que é descrita pelas equações em baixo e que pode ser resolvida numericamente¹³:

$$x_j^* : \sum_{i=1}^k e^{\frac{w_j - w_i}{w_i} x_j^{w_i}} = 1; x_i^* = e^{\frac{w_j - w_i}{w_i} x_j^{w_i}}; 0 < x_j^* < e^{-1}$$

Com isto não se faz mais que acrescentar ao tema *variedade semântica* um caso com descrição matemática, no sentido de *semantic manifold*, ou *variété des équilibres*¹⁴, um exemplo baseado numa generalização da função de Shannon.

Outros casos são igualmente possíveis, com base noutros germes funcionais. Cada caso é como se fora um poema matemático. A partir do momento em que aquelas fórmulas estabelecem dependências funcionais entre conjuntos de índices de valor e as proporções de equilíbrio, entrámos a aflorar o campo semântico, onde também se resolve o problema inverso a menos de uma standardização: dadas as proporções de equilíbrio obtêm-se os valores de W que lhe correspondem: W^* . Citando Saussure: *a elipse é exactamente o suplemento do valor*¹⁵, caminhando numa hipertopia [talvez](#) ...



¹³ José Casquilho. 1999. **Ecomosaico : índices para o diagnóstico de proporções de composição** . Tese de Doutoramento. Instituto Superior de Agronomia, Lisboa.

¹⁴ Michel Demazure. 1989. **Catastrophes et Bifurcations**. Édition Ellipses, Paris.

¹⁵ F. de Saussure cit in Anne Hénault. 1992 (1997). **História Concisa da Semiótica**. Parábola Editorial, São Paulo, 2006.