

# *Considerações sobre a Teoria da Informação Semântica em An Outline of a Semantic Theory of Information de Bar-Hillel e Carnap*

*Ralph Leal Heck*

Faculdade Católica de Fortaleza

## *I. Influências e desdobramentos*

A teoria da informação semântica apresentada por Bar-Hillel e Carnap parte da ideia de que há um valor informacional no conteúdo semântico de sentenças verdadeiras de uma linguagem L, dada a totalidade de possibilidades que podem ser exaustivamente avaliadas através de uma interpretação desta linguagem. (cf. BAR-HILLEL; CARNAP, 1952. pp.2-3). A concepção semântica pressuposta no artigo é extencional e se baseia em um cenário com falantes ideais, tanto no que compete ao ruído, quanto à relação entre os sujeitos e a valoração das sentenças, i.e. não é levantada a dimensão epistêmica ou pragmática da informação ou como eles dizem: “informação pragmática ‘ideal’” na qual destinatários “ideais” recebem apenas mensagens acerca de conhecimentos empíricos. (cf. BAR-HILLEL; CARNAP, p.3).

A ideia de uma teoria da informação semântica provém de um desdobramento do trabalho de Shannon (SHANNON, 1948.), tendo como pano de fundo o trabalho sobre lógica indutiva e probabilidade de Carnap (CARNAP, 1950.) somada a forte influência na construção dos conceitos componentes da teoria, a

partir da interlocução com os trabalhos de Popper<sup>1</sup> (cf. POPPER, 1972. pp.243-244.), em especial, a *Lógica da Pesquisa Científica* (POPPER, 1999.) em que Popper define conteúdo empírico, probabilidade, falseabilidade e, especialmente, informatividade.

Há ainda outras duas publicações acerca desta teoria da informação semântica feitas por Bar-Hillel e Carnap, como aponta Elias (cf. ELIAS, 1954. pp.230-231), mas estes textos se debruçam mais sobre o aspecto indutivo do que probabilístico *a priori* das sentenças que, de fato, é apresentado como fundamento para a aplicação indutiva da teoria.

O impacto da teoria da informação semântica foi amplo. Tomemos, por exemplo, a expansão feita por Hintikka (cf. HINTIKKA, 1970a.) ao ampliar esta teoria a uma linguagem poliádica e potencialmente infinita avaliando-se o valor informacional das sentenças a partir de um procedimento recursivo de análise de constituintes semânticos de uma dada sentença, criando a distinção entre informação de superfície (*surface information*) e informação de profundidade (*depth information*). Outro exemplo, é a discussão acerca do paradoxo da informatividade de contradições e sua respectiva solução adotada por Floridi (Cf. FLORIDI, 2011. pp.111-114.). Estes desdobramentos demonstram a importância da compreensão e a da aplicabilidade das teses contidas no *An Outline of a Semantic Theory of Information*.

## 2. Definições, funções e operações

Bar-Hillel e Carnap definem como ponto de partida de sua teoria: 1) um universo do discurso finito e 2) uma linguagem monádica e finita L. A partir de então, definem as principais relações sintáticas e semânticas da linguagem. Dentre as quais, vale ressaltar:

Os Q-predicadores, que são a aplicação dos predicados primitivos sobre todas as entidades designáveis no universo do discurso. Eles ocorrem como conjunções de predicados negados e não-negados. Uma conjunção exaustiva de Q-predicadores acerca de cada constante individual forma uma Q-sentença acerca de cada entidade designável.

---

<sup>1</sup> No artigo *Testability and Meaning*, Carnap substitui o conceito de verificabilidade por confirmabilidade baseado na influência da *Lógica da Pesquisa Científica* de Popper, conforme a nota em (CARNAP, 1936. pp.425-426).

Uma descrição-de-estado (*state-description*): para cada estado possível do universo  $W$  do discurso, a sentença que contém a totalidade das descrições verdadeiras acerca do estado de  $W$  é designado por “descrição-de-estado”. Em termos da linguagem  $L$ , uma descrição-de-estado é a conjunção de  $n$   $Q$ -sentenças, para  $n$  entidades no universo do discurso.

O Alcance (*Range*), definido a partir de uma proposição  $p$  qualquer por meio dos estados de descrição possíveis que satisfazem  $p$  em  $L$ . O alcance pode ser representado pela disjunção dos estados de descrição que  $L$ -implicam  $p$ .

Após estas preliminares, o primeiro conceito pré-sistemático de informação é a função  $In(\dots)$ . Ela se baseia na noção de implicação entre sentenças de  $L$  e representa um conceito conjunto-teorético que trata “a informação como uma classe de algo” (*ibidem*, p.7). A partir desta noção, eles definem a relação de implicação mínima de uma sentença  $L$ -verdadeira (tautológica) e de implicação máxima de uma sentença  $L$ -falsa (contraditória).

A noção pre-sistemática  $In(\dots)$  é explicada pela noção de negações de estados-de-descrição implicadas por sentenças de  $L$ . Tais negações são chamadas de elementos-conteúdo (*content-elements*), além disso, Carnap e Bar-Hillel introduzem a função  $Cont(\dots)$ . Esta função mede a totalidade dos estados-de-descrição excluídos pela sentença, i.e., a totalidade dos elementos-conteúdo de uma sentença.

Em outras palavras, dado os estados-de-descrição possíveis de  $W$ , podemos definir o conteúdo informacional de uma sentença  $p$  da seguinte forma:

$$Cont(p) = \{w \in W \mid w \models \neg p\},$$

Isto é: a classe de todos aqueles  $w$  que são excluídos por  $p$  (cf. BAR-HILLEL; CARNAP. p.11). Deste modo, o conteúdo de  $p$  é medido pelo conjunto dos estados de descrição de  $W$  que são incompatíveis com  $p$ . Nos casos de tautologia  $t$  e contradição  $c$ ,  $Cont(t) = \emptyset$  e  $Cont(c) = W$ .

Entretanto, a função  $Cont$  não é capaz de definir adequadamente a soma de conteúdo entre mensagens. Daí Carnap e Bar-Hillel introduzirem o conceito pré-sistemático de quantidade de informação *in* (com  $i$  minúsculo). Este conceito autoriza operações algébricas com os elementos-conteúdo de sentenças. Embora na situação em que duas mensagens compartilham parte de seus conteúdos informativos a soma destes conteúdos pode ser menor que o conteúdo da conjunção entre as sentenças (cf. BAR-HILLEL; CARNAP, 1952. pp.11-13.)

A resolução é a derivação da medida quantitativa *cont* (medida de conteúdo) por meio de *Cont* e *in. cont* se baseia na função de atribuição de probabilidade  $m(p)$ , esta última é definida *a priori* com base na quantidade de estados-de-descrição que tornam  $p$  verdadeira. Visto que, *a priori*,  $W$  tem estados equiponderáveis e que, para toda sentença contingente  $p$ , há um conjunto de estados de  $W$  que torna  $p$  verdadeira e outro conjunto de estados de  $W$  que torna  $p$  falsa ( $Cont(p) = \{w_1, \dots, w_n\}$ ). Assim, a toda sentença  $p$  de  $L$  atribui-se uma probabilidade  $m(p)$  calculada a partir da soma das probabilidades de todos os estados que tornam  $p$  verdadeiro. O limite superior do valor absoluto de  $m(p)$  é a soma das probabilidades de todos estados-de-descrição de  $W$  que é 1 (para tautologias) e o limite inferior é 0 (para contradições). A partir daí, temos a função:

$$cont(p) = 1 - m(p)^2$$

Esta função faz uso do que Barwise (cf. BARWISE, 1997.) chama de Princípio da Relação Inversa, em que se afirma que há uma relação inversa entre a probabilidade de enunciados (ou eventos ou situações possíveis (cf. DRETSKE, 1981)) e seu grau de informatividade. Segundo Floridi (cf. SOMMARUGA, 2009. p.42), Popper foi o primeiro filósofo a enunciar este princípio, mas apenas após a publicação do trabalho de Shannon foi possível construir uma teoria da informação semântica enquanto tal.

O que se tem em vista com esta função é a quantidade relativa de possibilidades excluídas pela sentença  $p$  como sua proporção de conteúdo informacional. Tal que  $1 \geq cont(i) \geq 0$  se  $p$  é contingente,  $cont(p) = 1$  quando  $p$  é uma contradição<sup>3</sup> e  $cont(p) = 0$  quando  $p$  é uma tautologia. Assim, seja  $m(i) = 2/5$  para a ocorrência da sentença  $i$ , então seu conteúdo nos fornece um valor de  $3/5$ , o que significa que  $3/5$  dos estados de  $W$  são excluídos em razão da verdade de  $p$ .

Deste modo, *cont* é definido pela probabilidade de todos os estados que são excluídos por não satisfazerem certo conteúdo:  $cont(p) = m(\neg p) = 1 - m(p)$ .

<sup>2</sup> Alternativamente,  $m(i) = 1 - cont(i)$  e  $cont(\neg i) = m(i)$ . (Cf. BAR-HILLEL; CARNAP, 1956. p.15)

<sup>3</sup> É justamente este princípio que leva ao paradoxo da contradição como aponta (FLORIDI, 2004b): “como, afinal, uma sentença contraditória pode ser tomada como mais informativa que uma sentença contingente?”

O que significa o conteúdo informacional repousa na abordagem de um espaço probabilístico definido *a priori* a partir da linguagem formal escolhida e de uma mensuração de conteúdo com base no princípio de relação inversa.

É claro que as funções acima são apenas o ponto de partida do que se pretende calcular nesta teoria da informação semântica. De fato, assim como sentenças elementares são avaliadas, devem as sentenças complexas encontrarem justificção de suas operações entre subfórmulas atômicas *i* e *j* de uma fórmula molecular:

$$(a) m(i \wedge j) = m(i) \times m(j) \text{ sse } i \text{ e } j \text{ não têm predicados em comum.}$$

$$(b) m(i \wedge j) \leq m(i) \leq m(i \vee j)$$

$$(c) m(i \vee j) = m(i) + m(j) - m(i \wedge j)$$

$$(d) m(i \vee j) = m(i) + m(j) \text{ sse } i \wedge j \text{ é uma contradição em } L$$

$$(e) m(i \wedge j) = m(i) + m(j) - m(i \vee j)$$

$$(f) m(i \wedge j) = m(i) + m(j) - 1 \text{ sse } i \vee j \text{ é uma tautologia em } L$$

Há, entretanto, duas questões sobre *cont* que motivam Carnap e Bar-Hillel a construir a função *inf* de medida de informação complementar a *cont*.

A primeira questão é sobre o conteúdo de sentenças *i* e *j*, tomadas como indutivamente independentes<sup>4</sup>, onde a quantidade de informação de *i*∧*j* deveria ser equivalente à soma da quantidade de informação entre *i* e *j*, isto é, a função da quantidade de informação semântica deveria ser aditiva sob a conjunção quando as sentenças são indutivamente independentes. O problema é que é possível derivar o seguinte teorema dos teoremas anteriores:  $cont(i \wedge j) \leq cont(i) + cont(j)$ . O que contradiz nossas intuições. Outra questão é que se *i* e *j* são sentenças básicas com diferentes predicados primitivos, então a expressão  $cont(i \rightarrow j)$  ou  $cont(j/i) = \frac{1}{2} \cdot cont(i)$ , ao invés de calcular a informatividade de *j* e de *i*, remove *j* do cálculo, uma vez que a probabilidade de *i* não afeta a probabilidade de *j* e vice-versa:  $cont(i/j) = cont(i) = 1 - \frac{1}{2} = \frac{1}{2}$ . Para resolver ambos os problemas,

<sup>4</sup> Definimos independência indutiva entre duas sentenças quando a probabilidade de uma sentença, dada a ocorrência da outra sentença, permanece a mesma. Em contraste, a independência de conteúdo ocorre quando, dadas as sentenças *i* e *j*, o conteúdo semântico de *i* é disjuntivo ao conteúdo semântico de *j*. Isto é, *i* e *j* são independentes com relação ao conteúdo se  $Cont(i) \cap Cont(j) = \emptyset$ .

eles propõem uma função que preserve o aspecto informativo de cada sentença, quando elas são expressas em conjunto ou quando uma se segue da outra, introduzindo a função  $inf$  (medida de informação). Tal que:

$$inf(i) = \log_2 \frac{1}{1 - cont(i)} = \log_2 \frac{1}{m(i)} = -\log_2 m(i)$$

Onde a equação significa o resultado negativo da operação de logaritmo na base 2 sobre a soma das probabilidades a priori dos estados de descrição no alcance de  $i$ . Sendo o limite de  $inf$  para sentenças contingentes variando entre 0 e infinito.

A base dois representa a possibilidade de valoração bipolar das sentenças. Em que o resultado gerado pelo cálculo significa um ganho parcial acerca de qual descrição-de-estado do mundo é o caso. A escolha da base do logaritmo representa tanto na semântica dos autores quanto na teoria matemática da comunicação de Shannon o aspecto binário das mensagens. Para a teoria de Shannon, o sistema binário (0 e 1) de transmissão de sinais. Para a teoria semântica, as valorações *verdadeiro* e *falso*. O que significa que em princípio é possível aplicar não só a função  $inf$ , mas toda a teoria da informação semântica em semânticas polivalentes, da mesma forma que seria possível aplicar a teoria de Shannon a um sistema enário de transmissão de sinais.

Segundo a interpretação de van Rooij (SOMMARUGA, 2009. p.170-172) em uma aproximação entre as teorias de Bar-Hillel e Carnap e de Shannon, a sentença  $p$  possui determinado valor em bits de informação. Onde cada bit de informação corresponde a uma escolha entre duas opções disponíveis (verdadeiro ou falso). Suponhamos um caso pré-determinado de 16 resultados equiprováveis (ou 16 mundos possíveis construídos a partir da combinação dos indivíduos com os predicados da linguagem), onde, para uma resposta precisa de qual das 16 opções é o caso, seria necessário um total de 4 bits, i.e., uma sentença  $q$  qualquer, que expressasse precisamente uma das 16 opções resultaria em  $inf(j) = 4$ . O aumento do valor em bits significa que a sentença aumenta a certeza acerca de qual estado do mundo (*descrição-de-estado*) é o caso. Esta preocupação com o aumento da previsibilidade e a diminuição das possibilidades (ou da incerteza) do que é o caso é análoga à equação elaborada por Shannon, exceto que o evento a ser previsto em sua teoria é a mensagem que está sendo enviada.

### 3. Interpretações e aplicações

Avaliando as funções em conjunto,  $m(p)$  retorna a probabilidade de  $p$  ser o caso, uma vez que,  $inf(p)$  retorna a quantidade de redução da incerteza acerca de qual  $w$ , em  $w \in W$ , é o caso, dada uma afirmação  $p$ . Ou seja, qual dos estados-de-descrição de  $L$  é o caso. Vale ressaltar que tautologias têm  $inf(T) = 0$  e  $cont(T) = 0$  já que  $m(T) = 1$  e contradições tem  $inf(\perp) = \infty$  e  $cont(\perp) = 1$ , já que  $m(\perp) = 0$ . Além disso, a teoria prevê o cálculo de sentenças nas formas normais disjuntivas e conjuntivas: Seja  $C_n$  uma conjunção de  $n$  sentenças básicas com  $n$  predicados primitivos distintos e  $D_n$  uma disjunção de  $n$  sentenças básicas com  $n$  predicados primitivos distintos,  $cont(C_n) = 1 - (1/2)^n$  e  $cont(D_n) = (1/2)^n$ . Para o caso de formas normais com as  $n$  sentenças atômicas ocorrendo em todas as disjunções  $[i_D]$  e conjunções  $[i_C]$ , temos:

$$cont(i_C) = \frac{m}{2^n}, \text{ com } m \text{ igual à } m\text{-ésima sentença da forma normal conjunta.}$$

$$cont(i_D) = 1 - \frac{m}{2^n}, \text{ com } m \text{ igual à } m\text{-ésima sentença da forma normal disjunta.}$$

Por fim, Bar-Hillel e Carnap ampliam o cálculo de informação semântica à lógica indutiva no contexto de corroboração ou infirmação de evidências com base em certas hipóteses. O ponto de partida é o conceito de grau de confirmação de uma hipótese, face uma dada evidência. Os teoremas são derivados a partir dos teoremas anteriores observando-se o quanto uma dada evidência acrescenta à determinada hipótese. Mas, a função padrão de probabilidade  $m$  (*a priori* e uniformemente distribuída) não é adequada para a indução científica. Precisando ser modificada por probabilidade indutiva e ser acrescida do critério de relevância instancial (*instantial relevance*) (cf. CARNAP, 1953.) que corresponde à mudança de probabilidade dos predicados em face do registro anterior do número de entidades que instanciam determinado predicado:

$$\frac{m(e \wedge i \wedge h)}{m(e \wedge i)} > \frac{m(e \wedge h)}{m(e)}$$

Tal que,  $h$  e  $i$  sejam sentenças contendo um predicado molecular factual  $P$  completamente expresso ( $h$  para hipótese e  $i$  para uma instância da propriedade  $P$ ), embora  $h$  e  $i$  não tenham constantes em comum, nem ocorrentes em  $e$ , e  $e$  seja tomada como uma sentença molecular não-L-falsa.

Com isto, se constrói o cálculo de estimativa estatística de informação, bastando para justificar as funções  $cont^*$  e  $inf^*$ , similarmente às funções  $cont$  e  $inf$  descritas acima. O conceito de probabilidade empregado é baseado em uma distribuição inicialmente equinúmerica, mas posteriormente balanceada a partir das ocorrências *a posteriori* das propriedades instanciadas por entidades, em que certas sentenças ganham maior peso probabilístico por descreverem instanciações de propriedades que já tenham algum registro de instanciações. Isto demonstra sua aplicabilidade no contexto da produção de teorias científicas em conformidade a filosofia da ciência desenvolvida por Carnap e pela tradição do Círculo de Viena da qual fazia parte, após as críticas de Popper à visão procedimental da ciência inicialmente advogada pelo Círculo (cf. CHALMERS, 1997.).

### *Conclusão*

Deste modo, a teoria da informação semântica desenvolvida por Bar-Hillel e Carnap se mostra uma ferramenta útil, tanto para a análise informacional de um espaço lógico concebido *a priori* (que pode ser explorado de modos diversos ao ser povoado com outras entidades além de estados-de-descrição como fez Dretske), quanto no contexto de um cenário indutivo, no qual as probabilidades são distribuídas em função do volume de corroborações de outras sentenças acerca dos mesmos predicados.

Além disso, o vínculo apresentado com a teoria matemática da comunicação de Shannon aponta para a possibilidade de se explorar nexos entre as duas teorias. E entre teorias ainda mais recentes como as teorias algorítmicas da informação (CHAITIN, 1997.) e as teorias pragmáticas da informação (cf. SOMMARUGA, 2009).

### *Referências*

- BAR-HILLEL, J. CARNAP, R. 1952. *An Outline of a Theory of Semantic Information*. Massachusetts: MIT Press.
- BARWISE, J., SELIGMAN, J. 1997. *Information Flow: The Logic of Distributed Systems*. Cambridge: Cambridge University Press.
- CARNAP, R. 1936. Testability and Meaning. *Philosophy of Science*, vol. 3, n. 4. pp. 419-471.



- CARNAP, R. 1945. The Two Concepts of Probability: The Problem of Probability. *Philosophy and Phenomenological Research*. Vol.5 No.4. pp.513-532.
- CARNAP, R. 1950. *Logical Foundations of Probability*. Chicago: University of Chicago Press.
- CHAITIN, Gregory J. 1997. *Algorithmic Information Theory*. 3ªed. Cambridge: Cambridge University Press.
- CHALMERS, A. F. 1997. *O que é a Ciência afinal?* 2ª Ed. São Paulo: Brasiliense.
- DRETSKE, F.I. 1999. *Knowledge and the Flow of Information*. Stanford: CSLI Publications.
- ELIAS, P. 1954. Review of: An Outline of a Theory of Semantic Information by Rudolf Carnap; Yehoshua Bar-Hillel. *The Journal of Symbolic Logic*, Vol. 19, No. 3. pp. 230-233.
- FLORIDI, L. 2004. Outline of a Theory of Strongly Semantic Information. *Minds and Machines*, vol.14 n.2. pp.197-222.
- FLORIDI, L. 2011. *The Philosophy of Information*. Oxford: Oxford University Press.
- HINTIKKA, J. 1970a/b. Surface Information and Depth Information. In: HINTIKKA, J.; SUPPES, P. (orgs). *Information and Inference*. Dordrecht: D. Reidel Publishing Company. pp.263-297.
- HINTIKKA, J. 1970a. On Semantic Information. In: HINTIKKA, J.; SUPPES, P. (orgs). *Information and Inference*. Dordrecht: D. Reidel Publishing Company. pp.3-27.
- POPPER, K. 1972. *Conjecturas e Refutações*. Brasília: Ed.UnB.
- POPPER, K. 1999. *A Lógica da Pesquisa Científica*. São Paulo: Cultrix.
- SHANNON, C. E. 1993. in: *Collected Papers*. Sloane, N. J. A.; Wyner, A. D. (eds.) New York: IEEE Press.
- SOMMARUGA, G. (ed.). 2009. *Formal Theories of Information: From Shannon to Semantic Information Theory and General Concepts of Information*. Berlim: Springer.
- WITTGENSTEIN, L. 2009. *Philosophical Investigations* [1953] 4ª ed. Oxford: Blackwell.

## Resumo

Esta apresentação identifica os principais conceitos envolvidos na teoria semântica da informação na obra *An Outline of a Semantic Theory of Information* de Bar-Hillel e Carnap. Início indicando as influências e desdobramentos sobre a elaboração desta teoria, seguido dos principais conceitos que os autores usarão para tratar o conteúdo semântico das sentenças. O ponto de partida são operações conjunto-teóricas, inicialmente através das funções *In* e *Cont*. Mas a dificuldade com a adição os leva a especificar o conceito de informatividade semântica por meio das funções *cont* e *m*, que avaliam o

espaço lógico-probabilístico sentencial. Sendo inversamente definíveis, onde a alta probabilidade de verdade de uma sentença implica em seu escasso conteúdo informacional. Com isto, os autores introduzem a função *inf* como a contribuição de uma sentença na identificação de qual descrição-de-estado possível é o caso. A partir disto, eles ajustam a teoria para tratar probabilidades experimental-indutivas entre hipóteses e evidências.

**Palavras-chave:** linguagem, espaço lógico, filosofia da informação, teoria da informação semântica, Bar-Hillel e Carnap.

### **Abstract**

*This presentation identifies the main concepts involved in the semantic theory of information in Bar-Hillel and Carnap's An Outline of the Semantic Theory of Information. It begins indicating the influences and developments on the elaboration of this theory, followed by the main concepts that the authors will use to treat the semantic content of the sentences. The starting point are set-theoretic operations, initially through In and Cont functions. But the difficulty with addition leads them to specify the concept of semantic information through the functions cont and m, which evaluate the logical-probabilistic sentential space. Being inversely defined where the high probability of truth of a sentence implies in its scarce informational content. With this, the authors introduce the inf function as the contribution of a sentence in identifying which possible state-of-description is the case. From this, they adjust the theory to treat experimental-inductive probabilities between hypotheses and evidence.*

**Keywords:** language, logical space, philosophy of information, theory of semantic Information, Bar-Hillel and Carnap.