

Pontifícia Universidade Católica de São Paulo (PUC/SP)

Programa de estudos pós-graduados em Tecnologias da Inteligência e Design Digital (TIDD)

Tarcísio de Sá Cardoso

Orientação e co-orientação: Lucia Santaella e Demi Getschko

CONSTRUÇÃO DE SIGNIFICADO NA WEB SEMÂNTICA

Esta pesquisa se propõe a realizar uma interpretação semiótica da construção de significados na Web Semântica, tendo em vista: esclarecer o que a web semântica se propõe a resolver; como ela se diferencia dos atuais serviços web no que diz respeito ao seu mecanismo de representação do conhecimento; que tipo de representação a web semântica pode fornecer para a chamada “internet das coisas” (SANTAELLA, 2008). Representação, entendida como relações ou operações entre signos (CP 3.540), opera sempre sistemicamente, podendo transitar desde um nível sociocultural até um nível meramente mecânico e regido estritamente por rígidas regras de inferência. Contudo, em um mundo cada vez mais conectado, as representações do nível mecânico podem cada vez mais acompanhar e moldar as formulações de linguagem humanas, tendo em vista um diálogo ainda mais refinado de interpretação e processamento artificial em forma de serviços de internet. A web semântica é uma das apostas do consorcio internacional W3C para melhorar a “interpretabilidade” (BERNERS-LEE, et al., 2001) e conseqüentemente a qualidade dos serviços oferecidos pela internet, através de um padrão internacional de representação do conhecimento (*knowledge representation*), uma vertente da inteligência artificial.

Este estudo busca comparar a Lógica de Descrição (*description logic*), modelo adotado pela web semântica, com o sistema semiótico proposto pelo lógico norte americano Charles Sanders Peirce (PEIRCE, 1931-1935). Partimos do

pressuposto de que as tentativas de sistematizar uma infraestrutura de internet capaz de transmitir informação com significado tanto para humanos quanto para máquinas tem íntima relação com as ideias de representação, inferência, mediação, sentido, semiose e significação. Nesta linha de raciocínio, o presente estudo vai relacionar tais conceitos na teoria semiótica peirceana com o modelo da web semântica, tendo em vista as contribuições que este diálogo entre semiótica e web semântica pode fornecer para a compreensão dos limites do significado proposto por esse modelo.

Os seguintes recortes temáticos são abordados em cada capítulo: caracterização do cenário atual e do estado da arte em estudos sobre web semântica (capítulo 1); explicação geral sobre a lógica da descrição e o modelo de representação da web semântica (capítulo 2); contextualização histórica dos problemas de lógica de descrição, especialmente a partir da incompletude da aritmética, teoremas da decidibilidade, computabilidade, inteligência artificial, informática aplicada à representação do conhecimento (capítulo 3); internet, cibercultura, folksonomy e comportamento orgânico em rede (capítulo 4); teoria da representação na semiótica peirceana (capítulo 5); crítica semiótica do modelo da web semântica e delimitação do alcance da representação nesse sistema formal (capítulo 6); algumas consequências tecnológicas, políticas e sociais da web semântica para a web de serviços e web das coisas (capítulo 7).

PROBLEMA

Tendo em vista esclarecer quais são os limites de significação que a web semântica promete sistematizar, esta pesquisa busca confrontar o sistema formado entre as ontologias computacionais, baseada na lógica da descrição, com a *folksonomy* fundada na espontaneidade do “tagueamento” orgânico. Tendo em vista refletir sobre a constituição dos modelos de internet atual, esta pesquisa propõe incluir um olhar semiótico para o fenômeno da internet 3.0 e seus processos de significação, buscando responder às perguntas: como ocorre a

representação do conhecimento no seio da web semântica? Qual a relação entre o sistema formal da lógica de descrição e o sistema de construção colaborativa da *folksonomy*? Que tipo de serviços os sistemas baseados na representação do conhecimento podem proporcionar? Como um olhar semiótico para os signos da web semântica pode ajudar a esclarecer o funcionamento da própria dialogia *Ontology-Folksonomy*, tão emblemática para a web atual? Até onde vão os limites de representação que um sistema artificial na rede mundial é capaz de produzir? Que caminhos para a construção de novos paradigmas culturais podem ser sugeridos por uma abordagem semiótica para a engenharia de softwares em termos de geração de “semântica” para sistemas web?

DESENVOLVIMENTO PRELIMINAR

Do ponto de vista formal, um estudo em lógica, representação e significado pode ter muito a ganhar com o edifício filosófico de Charles Sanders Peirce e nossa hipótese apenas faz sentido a partir das contribuições deste autor. Em um texto intitulado “A lógica dos relativos”¹ no livro *Reasoning and the logic of things: the Cambridge conference lectures of 1898* (PEIRCE, 1992), Peirce explica resumidamente o sistema de notação lógica dos grafos existenciais, elaborando um mecanismo proposicional no qual a ênfase está no conceito de identidade. Segundo Peirce (1992: 154), a lógica das relações só admite quatro tipos de relação de identidade: *medad* (que seria como um verbo impessoal, como por exemplo “chover”), *mônada* (verbo neutro, incompleto de significação, pois precisa de um sujeito para se referir completamente a um fato, como, por exemplo, o verbo “cair”), *díada* (verbo ativo simples, que precisa de dois sujeitos para completar sua proposição, como por exemplo “obedecer”) e *tríade* (que precisa de três sujeitos para significar, como o verbo “dar”).

¹ Tradução do autor, do original: *Relation logic*.

Apesar da ênfase no conceito de tríade, o mais completo modo de identidade, representação e significação, Peirce afirma que há uma gradação inclusiva entre os diversos níveis de relação. Desse modo, a díada inclui o conceito de mônada, assim como o de tríade inclui a díada. Mais do que isso, Peirce postula uma não ruptura entre os diferentes tipos de identidade, como se houvesse um contínuo universal, no qual as diferentes gradações fazem parte de um mesmo e único processo, chamado por ele de crescimento dos signos, ou simplesmente semiose.

A web semântica, em sua proposta de montar uma gramática universal de significados computacionais por meio do recurso novo chamado *ontologia*, pretende organizar a informação disponível na internet atribuindo significados de acordo com um modelo convencionado, portanto arbitrário. A ideia de conferir semântica aos softwares web por meio de ontologias computacionais é realmente um avanço em termos de compartilhamento de interpretação de dados. O problema geralmente atribuído pelos pesquisadores (CATTUTO et. al., 2006; PERIN, 2007 etc.) é que esse sistema arbitrário deve ser originado por quem gera o conteúdo, e não por quem usa esses dados. Nesse sentido, a web semântica se opõe, por exemplo, à lógica da interatividade tão facilmente encontrada na chamada *folksonomy*, que tira informação diretamente do uso de signos. Essa oposição parece insolúvel, haja vista os próprios propósitos anticognitivos (BREITMAN, et al., 2010 pp. 31-33) do sistema da web semântica.

O impasse do quadro atual, se considerado o sistema semiótico de Peirce, parece ter relação direta com a inter-relação do conceito de terceiridade, secundidade e primeiridade (CP 1.23-27). Vale lembrar que a terceira categoria nasce do conflito gerado pela segunda. Esse conflito reclama por uma terceiridade capaz de superá-lo via representação, pois mediação nasce do diálogo entre o mundo das representações e o mundo das ações. A experiência da objeção é sempre no aqui-agora, mas há eventos que permanecem, que são regulares, que estão sob

alguma forma de lei. Se não houver repetição, não há lei, e nada permanece. E se nada permanece, não se pode conhecer². Daí que a terceiridade está ligada a um princípio de lei, e a própria cognição fundamenta-se nas observações atentas das regularidades, mirando a previsibilidade. Neste sentido sinequista e pragmaticista, não cabe mais a cisão entre o universo da secundidade dos eventos, considerado em sua face puramente bruta, e o universo do contínuo das leis, considerada sua regularidade cega. Ao contrário de estabelecer um corte, a semiótica de C. S. Peirce parece reclamar por uma conciliação entre ambos. A passagem em que Peirce explica seu pragmatismo fazendo alusão à íntima relação do pensamento com os “portões” da percepção e da ação deixa claro que essas esferas não são tão estranhas quanto parece.

The elements of every concept enter into logical thought at the gate of perception and make their exit at the gate of purposive action; and whatever cannot show its passports at both those two gates is to be arrested as unauthorized by reason. (CP 5.212)

Por mais que percepção pareça um tema bastante distante do nosso objetivo nesta pesquisa, do ponto de vista exclusivo do pensamento e da representação do conhecimento, parece que Peirce está interessado no processo pelo qual essa representação tem origem. Em um texto de 1878, intitulado *Como tornar claras nossas ideias*, o raciocínio é explicado como um processamento que opera por associação de ideias, mas depende de um mecanismo de observação no qual essas ideias tenham origem.

A maquinaria do pensamento só pode proceder à transformação do conhecimento, mas nunca originá-lo, a menos que seja alimentado por fatos da observação (PEIRCE, 1993 p. 51)

² Cabe lembrar que conhecimento tem relação com previsibilidade e na própria ideia de prever está à ideia de futuro (IBRI, 1992).

Neste sentido de continuidade, pretende-se levantar a hipótese de que, do ponto de vista do sinequismo (PEIRCE, 1931-1935), isto é, da doutrina do *continuum* que emerge da abordagem filosófica de Chales S. Peirce, a lógica da ontologia – caracterizada na web semântica pela descrição de um conteúdo por meio de regras que o classificam em uma hierarquia bem formada – e a lógica da espontaneidade da *folksonomy* – caracterizada pelo modelo do imprevisto na interação entre sistemas da internet e uma coleção de usuários – não são conceitos tão antagônicos, como poderia parecer a um olhar ingênuo para o debate. À luz da semiótica peirceana e do preceito sinequista, não parece haver fissura entre uma lógica gramatical e uma lógica de uso espontâneo, sendo elas partes de um mesmo processo de amadurecimento do signo. Ao contrário do que já foi afirmado por pesquisadores da web semântica (PERIN, et al., 2007 p. 11), pode ser justamente esse ambiente intersticial o terreno mais fértil para a internet.

De qualquer modo, uma compreensão mais afinada sobre como ocorre a representação na web semântica parece ser fundamental até mesmo para entender o potencial dessa rede e sua relação com os crescentes serviços web. Acredita-se que, ao entender os limites do atual modelo, pode-se considerar a relação entre prós e contras de um padrão tão promissor – uma vez que pretende organizar o conteúdo disponível da web – e tão controverso – uma vez que, mesmo já estando disponível, quase ninguém usa a web semântica por menosprezar sua proposta. Traduzir a proposta da web semântica para o terreno semiótico pode permitir elucidar o que ainda é impreciso para a grande maioria dos profissionais de internet e para a comunidade em geral. Pode ainda, sugerir propostas heurísticas para os próprios membros dos consórcios que trabalham em aperfeiçoamentos dos padrões atuais da internet.

CONSIDERAÇÕES FINAIS

Neste ponto, cabe ressaltar as diferenças de Ontologia e Folksonomy. Ontologia em web semântica diz respeito a um tipo de arquivo que contém informações hierarquizadas e relacionadas entre recursos da internet (BERNERS-LEE, et al., 2001). A grande vantagem desse mecanismo, como já foi sugerido, refere-se à sua capacidade de permitir deduções com base em premissas. Dadas as regras (“Todo A é B” e “C é A”), chega-se a conclusões (“Logo, C é B”) e essas conclusões podem poupar muito trabalho mental humano, podem relacionar muitos conteúdos de forma automatizada e podem, do ponto de vista dos softwares web, criar uma unificação das linguagens específicas.

Já o termo Folksonomy se refere a uma geração orgânica de taxonomia, baseada nas interações diretas dos usuários, o que permite marcação (tag) em texto plano e a geração de inferências estatísticas, baseadas fundamentalmente na coleta de dados relativos ao processo de tagueamento pelos usuários humanos da web. Sua representação mais comum está associada à visualização de *tag clouds*, nuvens de termos nas quais os termos mais relevantes para o coletivo (mais buscados pelos usuários) aparecem com maior destaque, em contraste com os termos menos relevantes, que aparecem menores.

A dialogia *Ontologia versus Folksonomy* tem uma relação direta com o futuro dos sistemas computacionais “semânticos” na internet. Uma abordagem com foco em ontologias computacionais tende a ser altamente eficaz em termos de tratamento de dados e inteligência assistida, proporcionando uma dupla dependência criada entre serviços eficientes dos sistemas web e conforto para o usuário, mas com dificuldades no que diz respeito à socialização e ao funcionamento orgânico. Já uma abordagem centrada nos mecanismos de geração de conhecimento colaborativo, através da chamada Folksonomy, tende a ser altamente eficaz em termos de inteligência compartilhada e organicidade, mas extremamente limitada em termos de inferências lógicas e relacionamento

de informações de natureza diversa, com conseqüências diretas para a modelagem de agentes artificiais inteligentes.

O amadurecimento da web semântica parece requerer que essas duas abordagens – ontologia e *folksonomy* –, aparentemente antagônicas, se relacionem de modo menos alheio. A questão de como conciliar um mecanismo “alorreferencial” bem formado com uma lógica autorreferencial orgânica parece crucial para o futuro próximo da internet. Essa disputa ora tende a uma conciliação fundamentada em prévia hierarquização convencionalizada (centrado implicitamente numa gramaticalidade ontológica), ora tende a um ambiente colaborativo orgânico capaz de fazer emergir uma auto-organização a partir do caos (com tendência folksonômica). Um meio-termo ou uma conciliação consistente parece que ainda não foi alcançado pelos estudos realizados até o momento.

Tendo em mente alcançar tal ponto conciliatório, esta pesquisa estuda os meios formais para se empreender uma espécie de “ontologização semântica”, uma tentativa de conciliar a espontaneidade da web social (sin-signos permeados por intenções parcialmente livres por parte dos usuários) com a gramaticalidade da web semântica (legi-signos regidamente formais). Busca-se através de esboços formais e modelos aplicáveis, uma forma de tradução dos signos fornecidos pelos usuários para as redes de significado padronizadas pela web semântica.