

EM BUSCA DE UMA SEMÂNTICA DO DIGITAL, OU “AS THEY MAY THINK”

RESUMO - Discute as reais possibilidades de processamento “semântico” por parte dos computadores, conforme a proposta da Web Semântica (Berners-Lee, 2001). A partir desta proposta de agregar “semântica” aos conteúdos disponibilizados na Web de modo a potencializar seu processamento por programas “agentes inteligentes”, são analisadas questões como: o que é Semântica em diferentes acepções, a arquitetura e funcionamento interno dos computadores e sua real capacidade de processamento semântico de conteúdos. O objetivo é chegar ao conceito de uma possível “semântica” computacional, conseguir uma maior clareza quanto ao que se pode esperar, o que é possível e o que é ilusório ou metafórico, da proposta da Web Semântica e até onde esta poderá potencializar computadores para processamento “inteligente”.

Carlos Henrique Marcondes
Doutor em Ciência da
Informação, pesquisador do
CNPq, professor do
Departamento de Ciência da
Informação e Programa de
Pós-graduação em Ciência da
Informação Universidade
Federal Fluminense
(UFF)
marcon@vm.uff.br

Palavras-chave: Web Semântica; semântica computacional; Ontologia; inferência computacional; ciências cognitivas; ensino de Ciência da Informação.

LOOKING FOR A DIGITAL SEMANTIC, OR “AS THEY MAY THINK”

ABSTRACT- The real possibilities of semantic processing by computers are discussed. Having as a start point the Semantic Web proposal (Berners-Lee, 2001) of embody meaning to the Web content in order to enable its processing by software agents, questions are analyzed as: what does really means semantics, what are its different senses, how computer internal architecture and operation enables semantic processing, what are computers real possibilities to semantic processing? The aim is to define a feasible concept of computational semantic, thus making clear what to hope for, what is mystic or metaphoric, of the Semantic Web proposal and how it enables real intelligent processing by computers.

Keywords: Semantic Web. computational semantic. Ontology. computational reasoning. Information Science education.

Most of the Web's content today is
designed for humans to read, not for computer
programs to manipulate meaningfully
(Berners-Lee et al., 2001)

1 INTRODUÇÃO

Qual o real potencial de processamento inteligente por parte dos computadores? Os teóricos da Inteligência Artificial consideram uma questão de tempo e desenvolvimento tecnológico para que os computadores realizem de forma trivial tarefas inteligentes. A proposta da Web Semântica herda da Inteligência Artificial boa parte desta pretensão. O que ele tem de real e factível? Poderá esta proposta endereçar os problemas do excesso de informações advindos com a cultura e a economia digitais?

Segundo um sítio que faz avaliações diárias¹ “The Indexed Web contains at least 6.89 billion pages (Tuesday, 10 July, 2012)”. Mais do que este número gigantesco, o que mais impressiona é a enorme mudança cultural, econômica e social trazida com a Web. Desde ponto de encontro para os mais variados e exóticos tipos de relacionamento, passando por mecanismo de mobilização política (DELANY, 2009), plataforma para comércio eletrônico, mídia para arte e cultura, para a ciência (CYBERINFRASTRUCTURE, 2007), a Web se tornou hoje a plataforma universal para a produção, armazenamento, disponibilização, acesso e compartilhamento de registros da cultura e do conhecimento humanos.

A Web vem se tornando na prática a biblioteca universal com a qual a humanidade vem sonhando há milênios, o repositório universal da cultura e do conhecimento humanos de hoje e do passado. No entanto, é um repositório cujo dinamismo é caótico, onde não há um dono ou gerente que o alimente com informações, mas, ao contrário, milhões de alimentadores, que o fazem de forma descentralizada e sem nenhum controle. Este caráter de repositório universal de registros da cultura e do conhecimento humanos, juntamente com o crescimento recente destes pelo desenvolvimento da Ciência e Cultura contemporâneos, fazem com que hoje o avanço do conhecimento possa se fazer tanto pela descoberta de fenômenos ainda desconhecidos, quanto pela descoberta de *relacionamentos não previstos entre registros de conhecimento e de cultura* (hoje, cada vez mais, registrados na própria Web). Uma área inteira de conhecimento vem se

¹ <http://www.worldwidewebsize.com/>, consultado em 24 de julho de 2012.

estabelecendo justamente para identificar e desenvolver métodos para agenciar computadores para realizar mineração de dados e mineração em textos de artigos científicos digitais, na perspectiva de identificar relações novas e inusitadas entre o conhecimento já registrado na Web. Esta é uma área de pesquisa intensiva, chamada de mineração de dados na literatura (BATH, 2004) ou descobertas em literaturas relacionadas (KOSTOFF et al., 2008), importante para a CI -Ciência da Informação -, que sempre trabalhou com conhecimento registrado. Portanto registrar informação em formatos suscetíveis de processamento “semântico” por computadores, como é a proposta da Web Semântica, torna-se uma área de pesquisa relevante, facilitando o uso destas técnicas e metodologias.

Este crescimento vertiginoso, descentralizado, segundo alguns, caótico, e a dependência crescente da cultura humana em seu sentido mais amplo, da plataforma Web para as mais variadas atividades, tornam o problema do acesso aos diferentes recursos aí disponibilizados uma questão extremamente crítica, inédita na evolução da cultura humana. Mas, como diria Marx (1975, p. 183), “a humanidade se propõe somente os objetivos que pode alcançar”², a questão da explosão informacional vêm sendo reconhecida e seu enfrentamento endereçado de forma coordenada pelo W3C - World Wide Web Consortiun. Por ser uma rede que tem sua história e desenvolvimento imbricados com o desenvolvimento do computador, a máquina que é o símbolo da modernidade e parece ser a solução para todos os problemas, sempre que se pensa em soluções para os problemas descritos, estas passam prioritariamente pelo próprio computador.

Embora a cobertura dos atuais mecanismos de busca seja ampla e tenha altos graus de completeza³, estes ainda assim são toscos e ineficazes no que diz respeito à precisão e a “semântica”, ou seja, compreenderem clara e inequivocamente as pretensões de um usuário ao fazer uma consulta. Indexam a Web extraíndo palavras do texto das páginas visitadas e armazenando em suas bases de dados registros com estas palavras

² Tradução nossa.

³ <http://www.searchenginewatch.com>.

associadas à URL da página onde a palavra foi encontrada. As palavras são extraídas uma a uma, isoladas, sem qualquer informação contextual, sem qualquer controle terminológico. Os mecanismos de busca são eficazes em revocação, mas falham em precisão. Além dessas deficiências inerentes à indexação automática praticada pelos mecanismos de busca, esses não são capazes, por diversos motivos (FORD; MANSOURIAN, 2006), de indexarem conteúdos no segmento chamado “Deep Web” (BERGMAN, 2001). Este é o segmento da Web que mais cresce e que agrupa fontes de informação de maior autoridade, como agências governamentais e internacionais, universidades, bibliotecas, centros de pesquisa, etc.

A iniciativa Web Semântica (Berners-Lee et al., 2001) objetiva como solução para esses problemas delinear uma nova Web em que documentos eletrônicos não sejam, como hoje, meras cópias eletrônicas de documentos textuais impressos, voltados para serem lidos somente por pessoas; Documentos conteriam *conhecimento* registrado, sob a forma de metadados legíveis por programas – os “agentes inteligentes” -, de forma tal que possamos agenciar estes programas para a realização de tarefas ditas “inteligentes” e mais sofisticadas.

Nas palavras do seu criador, Tim Berners-Lee, a proposta da Web Semântica é assim descrita:

The Semantic Web is not a separate Web but an extension of the current one, in which information is given well-defined meaning, better enabling computers and people to work in cooperation.” (BERNERS-LEE ; HENDLER; LASSILA, 2001, p. 3).

Berners-Lee, Hendler e Lassila, (2001) delinea uma visão ampla e ambiciosa para a Web Semântica, relacionando-a diretamente com o futuro da cultura humana, crescentemente uma cultura digital:

The Semantic Web is not “merely” the tool for conducting individual tasks that we have discussed so far. In addition, if properly designed, The Semantic Web can assist the evolution of human knowledge as a whole” (BERNERS-LEE; HENDLER; LASSILA, 2001, p. 3).

Para que este cenário seja viável, registros de conhecimento em diferentes domínios devem estar formalizados dentro de uma compreensão comum, padronizada,

representados e disponibilizados publicamente em formato legível por programas – nos chamados metadados e ontologias -, para que seu conteúdo possa ser utilizado por “agentes inteligentes” para realizarem “raciocínio automático” (BERNERS-LEE ; HENDLER; LASSILA, 2001, p. 5).

Quaisquer que sejam as soluções, tudo parece indicar que elas vão passar por um incremento nas capacidades de processamento “semântico” dos computadores. Portanto, aumentar o poder “semântico” dos computadores se torna um objetivo significativo para endereçar as questões levantadas anteriormente. Até que ponto isto é possível? Podem os computadores compreender significados ou, pelo menos, simular sua compreensão? Qual o alcance de propostas como as da Web Semântica?

Há quase 70 anos atrás, aspectos parciais desta problemática já vinham se delineando e haviam sido percebidos por Vannevar Bush (1945), coordenador do esforço científico americano no pós Segunda Guerra Mundial. Bush observou, de sua posição privilegiada, a chamada “explosão informacional”, o crescimento do número de publicações científicas para além das capacidades de leitura e absorção individual pelos pesquisadores das novidades publicadas em sua área. Compreendeu como esta situação poderia comprometer o desenvolvimento científico, vital para a economia e a segurança dos EUA. Em seu artigo “As we may think”, Bush propõe uma solução tecnológica para o problema, na forma de uma máquina, chamada por ele de Memex, que usava as tecnologias mais avançadas da época como o microfilme e a televisão, para associar conteúdos de forma semelhante ao cérebro humano. Bush é considerado um dos fundadores da Ciência da Informação e sua concepção do Memex como uma antevisão dos computadores digitais, hipertexto e da Web atuais. Com a proposta da Web Semântica o problema fundacional da CI, a explosão informacional, é retomado na atualidade, agora com uma dimensão muito mais crítica.

O computador começou como uma máquina destinada a fazer cálculos numéricos com grande velocidade e precisão⁴ (BRETON, 1991). Com o desenvolvimento tecnológico

⁴ A propósito, ENIAC, a sigla do primeiro computador, significa “Electronic Numeric Integrator and Calculator”.

estes vieram a se transformar, a partir da década de 60, em *máquinas de processamento de conteúdos*. Esta época marca também o surgimento da CI e seus esforços de instrumentalizar o nascente computador para o processamento “semântico” da crescente quantidade de conteúdos nele armazenados em virtude do aumento da capacidade de armazenamento das memórias. Shera (1957) e Luhn (1958) são pioneiros na proposta do uso do computador para processamento “semântico”, para indexar e fazer sumários automáticos de textos, baseados na frequência ou ocorrência de determinadas palavras ou expressões ou no uso de códigos que indicassem o sentido contido nos textos (“semantic codes”).

Computadores entenderem significados como propõe o projeto da Web Semântica, não é uma proposta nova. Veltman (2004) vê uma continuidade entre esta proposta e a da Inteligência Artificial das décadas de 60 e 70 do século XX, de construir um computador capaz de ter um comportamento “inteligente”. Esta, por sua vez, seria um desdobramento de uma proposta anterior, a da Cibernética, que segundo Dupuy (1996) se propunha a desvendar a última fronteira do conhecimento humano, o funcionamento do cérebro. Na raiz desse projeto está a visão antiga, mas que começa a tomar corpo na segunda metade do século XIX, de que *pensar é calcular*, que é possível automatizar o pensamento e que este “é redutível a métodos lógicos” (TENÓRIO, 2001, p. 30). O que seria um comportamento “inteligente”? A lógica e a matemática, método supremo de raciocínio humano, e que tinham provado sua validade a partir das descobertas científicas e da visão mecanicista de universo instaurada a partir da Modernidade, eram os modelos óbvios.

O matemático Hilbert no século XIX, propõe a questão de descobrir se seria possível provar a verdade ou falsidade de uma proposição em lógica *através de um método automático, cálculo lógico ou dedução*. Esta proposta levada às últimas conseqüências, tem como corolário a crença de que o aumento da capacidade dos computadores poderá inclusive torná-los mais inteligentes que nós. A pretensão de Hilbert foi derrubada pelas conclusões do matemático Gödel, com seu Teorema da Incompletude, que provou que dentro de um sistema fechado de preposições lógicas não

se pode provar a consistência de todas elas; é sempre possível construir algumas, a partir dos axiomas do próprio sistema, que não podem ser provadas, resultando portanto num sistema cuja verdade ou falsidade é “indecidível” (DUPUY, 1996, p. 29).

Apesar do Teorema da Incompletude, a proposta de Hilbert resultou num modelo abstrato de máquina capaz de deduzir por um cálculo automático a verdade ou falsidade de um conjunto de proposições em lógica, as chamadas “máquinas de Turing”, matemático inglês da primeira metade do século XX. A máquina de Turing, um procedimento automático para resolver problemas lógicos, torna-se um modelo abstrato para os computadores eletrônicos que começaram a ser construídos na década de 40, introduzindo na sua concepção a idéia de algoritmo ou programa⁵.

Um testemunho inequívoco, Herbert Simon, um dos pais da IA - Inteligência Artificial – dizia em 1965: “No espaço de vinte anos, as máquinas serão capazes de realizar todas as tarefas, sejam elas quais forem, que os homens podem realizar” (DUPUY, 1996, p. 37). Apesar dos sonhos ufanistas da IA, muito pouco se avançou nesta área. Autores como Dupuy deixam transparecer que o projeto da IA fracassou. “Para quem escreve, com uma atitude a priori de simpatia e até de gratidão ... a história da cibernética, o sentimento que finalmente prevalece é o de uma grande decepção” (DUPUY, 1996, p. 195).

A literatura em Ciência da Computação, desde suas raízes Cibernéticas, esta cheia de afirmações, mais ou menos naturalizadas, acerca da “inteligência computacional”, “raciocínio automático”, “ontologia”, da capacidade de “inferência”, da capacidade “semântica” dos computadores, etc. Visões da “inteligência” computacional veiculadas por filmes como “2001 – uma odisséia no espaço”, “Inteligência artificial”, entre outros, alimentam visões ficcionais, fantasiosas e não científicas sobre a real capacidade dos computadores realizarem tarefas inteligentes. Junto a visões ufanistas de muitos pesquisadores da área de Inteligência Artificial, ajudam a criar um mito com relação às reais capacidades inteligentes dos computadores. O quanto estas expressões representam da real capacidade dos computadores, o quanto ela têm de uma visão antropomórfica ou de sentido metafórico? Qual o real alcance dessa proposta?

⁵ Máquinas de Turing, http://en.wikipedia.org/wiki/Turing_machine.

O objetivo deste artigo é tentar desvendar estas questões, separando o que elas têm de mistificação ou metáfora. Espera-se permitir que profissionais da área de CI possam ter clareza das reais possibilidades de agenciar computadores nos projetos que endereçam os problemas atuais trazidos pela “explosão da informação” no ambiente Web e tornar a CI parte integral da construção de uma Sociedade da Informação (SARACEVIC, 1996, p. 42).

O artigo foi desenvolvido a partir de pesquisa bibliográfica nas áreas de ciência da computação, história da computação e das ciências cognitivas, lingüística e filosofia. Está estruturado da seguinte maneira: a seção 2 seguinte discute a proposta e Web Semântica e seus pressupostos, além de diferentes definições de semântica, para chegar a formulação de “semântica proposta pelo projeto da Web Semântica; a seção 3, a partir dessa formulação, procura discutir as reais capacidades “semânticas” de um computador a partir de sua arquitetura e de seu funcionamento interno; finalmente a seção 4 discute os resultados a que se chegou, as limitações da proposta de “semântica” do projeto da Web Semântica e apresenta conclusões.

2 A WEB SEMÂNTICA

Nesta seção será exposta e discutida a proposta da Web Semântica a partir da formulação de seus criadores. Serão delineadas também algumas aplicações baseadas na sua visão e nas tecnologias hoje chamadas de tecnologias semânticas daí derivadas. Estas aplicações permitirão montar o cenário para discutir as possibilidades e chegar-se a uma definição do que seria uma “semântica computacional”. São delineados e discutidos além disso, os pressupostos e as condições em que se busca uma tal “semântica computacional”.

A proposta da Web Semântica coloca uma visão que serve de norte para as pesquisas atuais em tecnologias semânticas computacionais. Este esforço mundial é coordenado pela W3C, através da seção, W3C Semantic Web Activity (2010). A proposta delineada por Berners-Lee acentua a questão de que conteúdos na Web Semântica teriam “significado” para programas e que web atual passaria de uma web de “links” para uma

web de significados (HEFLIN, et al., 2003, p. 29). Isso seria concretizado através de uma Web que contivesse não somente dados, mas sim dados estruturados, na forma de asserções formatadas em RDF (1999) – Resource Description Framework -, no que Berners-Lee (1998a) chama de “Assertion Model”, que permitiria processar estas asserções usando a Lógica formal.

Para ilustrar as potencialidades de sua proposta Tim Berners-Lee (2001) descreve uma aplicação, um programa “agente inteligente” capaz de ser programado para atingir o objetivo de conciliar as agendas de dois irmãos que têm que levar a mãe idosa para fazer exames médicos e exercícios físicos prescritos por seu médico com as agendas de outros médicos, de clínicas especiais e de outros tipos de provedores de serviços, todas mantidas por outros “agentes inteligentes”. “Agentes inteligentes” (FRANKLIN; GRAESSER, 1996) são programas surgidos a partir da proposta da Web Semântica, encarregados de cumprir determinadas tarefas não previamente programadas, para as quais devem ser capazes de interagir com serviços disponíveis na Web, e com outros agentes, “compreendendo” sua semântica de funcionamento específica, avaliando sua possível utilidade para a tarefa para a qual foi encarregado e agenciando-os quando pertinente.

Um exemplo de aplicação dos conceitos e tecnologias da Web Semântica no domínio científico específico da biomedicina é a Semantic Network, esquema classificatório de supercategorias do UMLS – Unified Medical Language System – Metathesaurus. A Semantic Network é formada por 153 supercategorias de tipos de entidades biomédicas e por 54 tipos de relações existentes entre estas entidades. Determinados tipos de relação ligam determinados tipos semânticos uns aos outros, todos com semântica definidas em termos do conhecimento biomédico, Segundo a documentação da UMLS Semantic Network: “The purpose of NLM's Unified Medical Language System (UMLS ®) is to facilitate the development of computer systems that behave as if they "understand "the meaning of the language of biomedicine and health”. UMLS Fact Sheet (2006).

A UMLS Semantic Network serve de suporte para que aplicações identifiquem em determinado artigo científico biomédico afirmações como “*Encurtamento dos telômeros*

(Antecedente, um fenômeno) está associado a (tipo de relação) senescência celular (Consequente, outro fenômeno)". A partir daí a aplicação pode identificar outros artigos que confirmem ou que questionem esta afirmação, ou identifiquem se esta é uma afirmação cientificamente inédita, indicando uma possível novidade ou descoberta científica. (MARCONDES et al, 2009).

Vai-se precisar a seguir de que pressupostos está se partindo e qual o sentido de semântica que se busca discutir: Segundo Berners-Lee, Hendler e Lassila (2001, p. 3), *"The Semantic Web will bring structure to the meaningful content of Web pages, creating an environment where software agents roaming from page to page can readily carry out sophisticated tasks for users"*. Assim, a semântica que se busca aqui deve ser entendida a partir do formulado nesta citação, como sendo:

Definição 1: "semântica computacional" seria a capacidade de programas "agentes inteligentes" de interagir com diferentes tipos de serviços disponíveis na Web (e não só com um tipo específico), compreender as mensagens que descrevem seu funcionamento e de modo a poder agir de acordo, em princípio, utilizando o serviço ou recurso

A compreensão de significados restrita a mensagens é somente parte das discussões mais gerais acerca de teorias que procuram dar conta da questão do significado e da semântica. Segundo Speaks (2010, p. 1), teorias do significado podem ser divididas em teorias semânticas, as que discutem somente o significado de expressões de uma linguagem, e teorias fundacionais, as que discutem a partir de que fatos expressões têm determinado significado. Essas últimas se baseiam fortemente em aspectos pragmáticos dos agentes ilocutivos, como intenções, crenças, convenções, etc. É claro que estas teorias, por pressuporem contextos mais amplos, complexos e diálogos em situações específicas destes agentes, não são as mais adequadas para a discussão do significado de mensagens disponíveis em páginas Web, sem a presença interativa desses seus agentes, que poderiam corrigir qualquer mal-entendido.

A linguagem humana como mecanismo de comunicação, cooperação e articulação social não tem paralelo em relação ao de outras espécies. Ela é responsável por uma

verdadeira mudança de qualidade na evolução da espécie humana, viabilizando um alto grau de intersubjetividade, que ao se autonomizar passamos a chamar de Cultura:

As mutações que viabilizaram a linguagem humana foram as últimas alterações biológicas significativas na espécie humana. Biologicamente hoje somos praticamente idênticos aos nossos antepassados de 100.000 anos atrás. A partir deste período o grande vetor evolutivo da humanidade tem sido a *cultura*” (MARCONDES, 2010).

A relação pensamento simbólico-linguagem humana é direta. A linguagem viabiliza a intersubjetividade. Ao discutir as bases biológicas do conhecimento, Manturana e Varela (2001) falam das amebas como possuindo mecanismos simples de sensibilidade ao meio ambiente, que lhes permitem tão somente identificarem a direção da luz no meio líquido em que vivem e moverem-se nesta direção, onde as chances de encontrarem alimento, sob a forma de algas microscópicas, é maior. Ao contrário dos organismos primitivos, que a um “input” do meio ambiente reagem de uma única maneira, só chegando ao que Pierce chamava de Secundidade, a espécie humana reage aos “inputs” do seu meio ambiente de uma forma altamente complexa: este “input” é cotejado e avaliado com outros “inputs” que estão na memória de curto prazo configurando uma determinada situação, com o conhecimento e experiência anteriores do indivíduo, até gerar um interpretante, num processo que Pierce chama de semiótico ou de Terceridade (ECO, 1995). Esse é o processo de *interpretação*, característica da comunicação humana, que ganha nova complexidade na medida que artefatos da cultura humana como documentos, se autonomizam dos seus criadores e de seus contextos de criação (MARCONDES, 2010).

Segundo Morris (1976, p. 17), *“Podem-se estudar as relações dos signos com os objetos aos quais eles são aplicáveis. Chamar-se-á esta relação a dimensão semântica da semiose”*. Na mesma linha, Fodor (2007, p. 1), um dos criadores da Teoria Computacional da Mente (HORST, 2009), afirma que Semântica : *“é a parte da gramática que se preocupa com as relações entre os símbolos da língua e as coisas no mundo a que eles referem, ou sobre as quais mantêm condições de verdade”*.

Segundo Chierchia (2003, p. 40) existem três abordagens para a Semântica: a abordagem representacional ou mentalística, segundo a qual o significado é o modo como

representamos mentalmente – a imagem mental ou conceito - aquilo que dizemos; esta abordagem esta ligada à Teoria Computacional da Mente, segundo a qual a mente é uma máquina de processar símbolos, tal como um computador. A segunda abordagem, a pragmático-social entende a linguagem como uma prática que instaura relações sociais; a semântica de uma expressão estaria nas condições que permitem que ela se realize efetivamente na prática. Por fim, a abordagem denotacional enfatiza o estado de coisas no mundo que uma expressão da linguagem denota; segundo este enfoque, a semântica de uma expressão seria tão somente o que ela denota, as condições de verdade do que a expressão denota. O autor, ao apresentar as diferentes abordagens, enfatiza que a Lógica apresenta as potencialidades de unificar os pontos fortes das três abordagens (CHIERCHIA, 2003, p. 48).

Frege, matemático e lógico alemão da 2ª. metade do século XIX, inaugura a prática de utilizar a lógica para compreender a linguagem, um caminho trilhado por muitos lógicos, filósofos e lingüistas a partir de então. Ao tratar da Semântica, Frege (1948) distingue entre *sentido* (ou semântica) de uma expressão lingüística e *referência* (ou denotação) da expressão, as suas condições de verdade no contexto em que a expressão é formulada. Haveria, segundo Frege, um único sentido para uma dada expressão, mas poderia haver várias referências, dependendo dos diferentes contextos em que a expressão é formulada.

Tarski (1944, p. 345), seguindo a tradição inaugurada com Frege, de empregar a Lógica para o tratamento da linguagem, afirma que: “*Semantics is a discipline which, speaking loosely, deal with certain relations between expressions of a language and the objects (or “states of affairs”) “referred to” by those expressions*”.

Preocupado em formalizar esta relação, Tarki afirmava que as condições de verdade, ou *semântica*, de uma expressão em lógica (formada por signos como letras expressando preposições, conectadas pelos conectivos lógicos *E*, *OU*, *NÃO*, *SE...ENTÃO* e *SE E SOMENTE SE* (ver (CERQUEIRA; OLIVA, 1982)), chamada de interpretação desta expressão, é uma estrutura lógica $\langle D, I \rangle$, onde *D* é um conjunto de objetos e *I* é uma função de interpretação, que relaciona as constantes existentes nas proposições com

indivíduos do conjunto D. As condições de verdade, ou *semântica*, para Tarski, eram condições lógicas que *se acrescentavam* as preposições de uma linguagem. Se nas preposições “x é marido de y”, em lógica $\text{Marido}(x,y)$, substituímos as variáveis x e y pelas constantes “João” e “Maria”, a proposição $\text{Marido}(\text{João},\text{Maria})$ será verdade se e somente se houver duas instâncias no conjunto D, João e Maria, tais que João seja efetivamente casado com Maria. Uma interpretação sob a qual um conjunto de preposições é verdadeiro é chamada de *modelo*.

Chomsky (1975) faz uma contribuição significativa e original à Semântica, na mesma linha das tentativas anteriores de empregar a Lógica para formalizar a linguagem. Ele propõe um modelo de gramática de uma linguagem como tendo uma estrutura profunda, semântica, universal a todas as línguas humanas, uma característica não social, mas biológica da espécie humana; a estrutura profunda é complementada por uma estrutura superficial, as línguas que falamos.

Todas as teorias semânticas pressupõem uma linguagem e discutem em que medida “As expressões da nossa linguagem “significam alguma coisa, e isto é o que nos permite comunicar” (CHIERCHIA, 2003, p. vii). Entretanto o que se procura elucidar neste artigo não é a capacidade expressiva de linguagens em geral, nem mesmo das linguagens da Web Semântica em si, como RDFS e OWL, nem a capacidade destas de expressar o mundo (ontologia) e sim, *em que medida pode-se falar numa “semântica computacional”, em que medida mensagens nestas e em outras linguagens podem (e como podem) ser interpretadas e inteligíveis por programas, podem ter seu significado apreendido por estes*. No mais, está se falando das mesmas mensagens cujos significados são, no fundo, necessários para fazer a economia, a cultura e a ciência na Web funcionarem. São semelhantes às mensagens amplamente estudadas na CI, mensagens sociais autonomizadas, separadas no tempo e no espaço de seus agentes ilocutivos no âmbito geral da cultura, ou do que Popper chama de 3º. Mundo.

Segundo Speaks (2009) as teorias que dão conta da semântica de mensagens em determinada linguagem, por sua vez, podem ser subdivididas nas que discutem a semântica de proposições (ou mensagens), muitas delas baseadas nas condições de

verdade ou *referência* das mensagens e nas que baseiam sua análise mais além das proposições, nas chamadas estruturas profundas da linguagem, subjacentes às mensagens, como proposto por Chomsky. As condições de verdade de mensagens, ou sua *referência*, também vão ser tomadas aqui como *a priori verdadeiras*; não haveria sentido em serviços ou recursos disponíveis na Web veicularem mensagens falsas sobre eles próprios. A própria proposta de arquitetura da Web Semântica prevê níveis encarregados da prova – “proof” - e da garantia de confiança – “thrust” (IANELLA, 2010) das afirmações feitas nos níveis lógico e ontológico, ou seja, de afirmações encontradas em páginas Web como nos exemplos que estão servindo de base para esta discussão. No entanto deve ser feita a ressalva que mensagens oriundas de serviços ou recursos disponíveis na Web possam ser contraditórias entre si; não poderia ser de outra maneira, não haveria coordenação capaz de garantir esta questão, num ambiente descentralizado como a Web. Esta se baseia portanto, ao contrário de um sistema computacional isolado - a hipótese do mundo fechado -, onde se pode afirmar que “if it cannot prove a statement to be true, the statement is false” (LEGG, 2009, p. 413), na chamada “hipótese do mundo aberto” (DRUMMOND; SHEARER, 2006) na qual afirmações contraditórias não são necessariamente uma verdadeira e outra falsa.

3 AS BASES DE UMA “SEMÂNTICA” COMPUTACIONAL

Nesta seção vai-se procurar compreender, a partir da análise do funcionamento básico e da arquitetura dos computadores, quais suas reais potencialidade de processamento semântico. Como, a partir dos moldes estabelecidos pela **Definição 1**, poderemos chegar a compreender como se daria uma “semântica computacional”?

Um computador, embora carregue na sociedade contemporânea um poder simbólico, talvez como nenhum outro artefato já construído pelo homem, é construído, ainda hoje, segundo uma concepção que remota aos primórdios da sua invenção na década de 40 do século XX. O ENIAC, considerado primeiro computador, criado em cerca de 1943, nada mais era que uma calculadora na qual haviam sido substituídos dispositivos mecânicos por outros, funcionalmente equivalentes a estes, mas baseados na recém

descoberta tecnologia das válvulas eletrônicas, muito mais eficaz e rápida. A evolução do ENIAC foi o EDVAC⁶, uma máquina já construída segundo uma concepção totalmente nova. Enquanto o ENIAC não deixava de ser uma calculadora eletrônica, em que as operações tinham que ser comandadas uma a uma, o EDVAC contava com um componente inovador, uma *memória*, onde eram gravados não só dados a serem processados ou resultado de processamento, como nas calculadoras, mas também um *programa*, uma seqüência de passos a serem executados para a consecução de uma dada tarefa. Assim, quando queremos calcular a média aritmética de três números usando uma calculadora, devemos alimentar os três números um a um, acumulando-os e dividir o total acumulado por três. No EDVAC, como em todos os computadores até hoje, os passos para se calcular a média aritmética são inscritos na memória, formando um programa, capaz de calcular a média aritmética não de três, mas de quaisquer números.

O conceito de programa armazenado na memória, iniciado com o EDVAC, é devido a um matemático húngaro chamado Von Neumann, que trabalhou na sua construção; este conceito, a chamada arquitetura de Von Neumann, revolucionou a concepção do computador, é base e paradigma para a construção de computadores até hoje. O computador se transformou a partir daí numa máquina de uso geral extremamente versátil, um dispositivo elétrico-eletrônico-mecânico – o chamado “hardware” - que executa um certo número (finito) de funções implementadas neste “hardware”, como somar, dividir, mover dados entre a memória e um dispositivo de entrada-saída e vice-versa, desviar para outro trecho do programa, etc., só que de diferentes maneiras, em diferentes seqüências, diferentes número de vezes; tudo isto comandado por um texto em que estas as funções do “hardware” são evocadas, cada uma através de seu código, um *programa* – o “software”. A memória, por ser *apagável*, permite que um programa possa ser substituído por outro. O computador se comporta de acordo com o programa que esta em execução na sua memória em determinado instante.

Um computador é portanto comandado por um programa armazenado em sua memória, escrito numa linguagem que evoca as funções inscritas no seu “hardware”. No

⁶ A sigla significa Electronic Discrete Variable Automatic Computer

início da evolução dos computadores, esta linguagem era formada de códigos que evocavam diretamente as funções do “hardware”, numa correspondência uma a uma, a chamada linguagem de máquina. As linguagens de máquina são herméticas, têm os códigos das funções que evocam dependentes dos fabricantes – são linguagens ditas proprietárias -, difíceis portanto de serem programadas, tornando os computadores difíceis de usar. Ainda nos primórdios da história da computação foram substituídas, a princípio pela linguagem “assembler”, que mantém ainda uma correspondência uma a uma com as funções inscritas no “hardware” que evocam. Por não serem diretamente executáveis, ou seja, por não corresponderem diretamente às funções inscritas no “hardware”, um programa em linguagem “assembler”, antes de poder ser executado num determinado computador, deve ser traduzido para a linguagem de máquina. Quem faz isto é um programa especial, o programa tradutor, para o qual o programa escrito em linguagem assembler serve de entrada e que gera como saída, a tradução deste mesmo programa para linguagem de máquina.

Baseado neste princípio de tradução, logo foram desenvolvidas diferentes linguagens, chamadas de “alto nível”, muito mais próximas à linguagem humana, e portanto, muito mais fáceis de serem programadas, como Fortran, Cobol, Pascal, etc. Os programas tradutores das linguagens de alto nível – chamados de compiladores – inauguram outra concepção que preside o funcionamento dos computadores até hoje e que foi estendida a um outro tipo de programa que exerce função análoga aos programas compiladores, os sistemas operacionais: o “software” em geral, incluindo hoje principalmente sistemas operacionais, mas também compiladores de diferentes linguagens, funciona como um *intermediário* entre os usuários humanos e as idiosincrasias, dificuldades e tecnicismo do “hardware”, traduzindo a linguagem humana para os códigos/funções do “hardware”.

Mas não é da semântica desta linguagem, que comanda o funcionamento do computador, que estamos falando. Essa seria uma semântica em que as expressões da linguagem de programação correspondem diretamente com às instruções do “instruction set” do seu processador porque se limita à semântica da comunicação humano-

computador com aqueles (poucos) profissionais cujo objetivo é programar o computador. A semântica que buscamos é mais geral, é a semântica da comunicação humano-computador ou computador-computador no contexto dos seres humanos e computadores em ação, se comunicando mutuamente.

Como fazer o computador “compreender” sentidos registrados “inputs”, em mensagens de entrada, em dispositivos computacionais como ontologias, ou seja, como fazer um computador reagir “semanticamente”, como se compreendesse mensagens e conteúdos não previamente programadas, encontráveis na Web? Que semântica um computador extrai ou “compreende”, uma vez programado, das interações com seus usuários? A linguagem natural, usada por seres humanos, está ligada a outras dimensões, a dimensão sintática, que normaliza as construções de expressões válidas na linguagem, e a dimensão pragmática, que discute o agir social pretendido ou induzido a partir da elocução ou compreensão de uma expressão da linguagem natural. Por exemplo, quando um motorista humano vê um sinal de trânsito vermelho, ele pára o seu veículo. A semântica está ligada à teleologia, ao agir intencional, não aleatório, a partir de uma compreensão, à pragmática. Na verdade, uma linguagem, através de suas três dimensões, é um mecanismo de agir social no mundo.

Um computador é incapaz de designar objetos do mundo como um ser humano e tem somente uma capacidade muito limitada de interagir com eles. Um computador convencional, que não esteja ligado a dispositivos mecânicos como um braço robótico, só pode interagir como um recorte do mundo que é o seu próprio ambiente, sua memória e, em especial, seus periféricos, os dispositivos com os quais se comunica com seres humanos ou a outros computadores na rede ao qual está ligado. Talvez uma semântica computacional possa ser baseada no interagir de um computador com o recorte do mundo que é o seu próprio ambiente. Ou seja, o que ele faz a partir da “compreensão” das instruções da linguagem de programação que o comandam e da interação homem-máquina. Para isto teremos que compreender melhor como trabalha conceitualmente um computador:

Definição 2 - Um modelo conceitual do funcionamento de qualquer computador é composto de duas entidades: *dados* (de entrada e de saída) e *processos* ou programas. Inclui como funções: a- uma função de entrada de dados, b- um mecanismo ou função de processamento, que transforma ou processa os dados de entrada e os transforma em dados de saída, e c- uma função de saída de dados, que exhibe o resultado do processamento dos dados de entrada pela função de processamento

A noção matemática de função é bastante ilustrativa do funcionamento do computador: é como se tivéssemos a função matemática $y=2x$; esta função tem como entrada um determinado número, processa esta entrada multiplicando cada número entrado por 2 e exhibe como saída o dobro do número entrado. No exemplo citado, os dados de entrada a- seriam os três números para os quais se deseja calcular a média aritmética, a função de processamento seria b- o programa ou algoritmo que calcula a média aritmética e os dados de saída seriam c- o resultado da média aritmética dos três números. A função de entrada é executada por dispositivos de entrada como teclado, “mouse”, escaner, disco rígido, etc; a função de processamento é executada pela conjunção do programa armazenado na memória com as funções inscritas no “hardware” que abriga o programa; as funções de saída são executadas por dispositivos de saída como monitor, impressora, disco rígido, etc. É deste artefato que queremos discutir a capacidade semântica, que queremos que compreenda a semântica de mensagens conforme, a **Definição 1** anterior.

Como então interagir com diferentes serviços e recursos que “agentes inteligente” encontram ao executar as tarefas para as quais foram encarregados? Como “compreender” a semântica das mensagens que estes serviços e recursos emitem? Já que no nosso modelo computacional *programas processam dados*, a **Definição 1** de semântica esta ligada à dimensão pragmática, não a pragmática clássica, humana, mas a uma pragmática computacional, ou seja, *como o programa deve se comportar, como processar uma determinada mensagem?* Segundo o modelo delineado anteriormente, dada uma entrada, *qual deve ser o processamento executado e qual deve ser a saída.* A

manifestação da “compreensão” de uma mensagem por um programa “agente inteligente” é pois, a execução de determinado processamento (e não de outro) a partir da sua entrada.

Na verdade o potencial de processamento semântico por computadores, comparado com o que conhecemos por processamento semântico humano, é extremamente reduzido. Computadores, desde a sua criação, trabalham dentro de um paradigma de processamento seqüencial de instruções armazenadas na memória. O processador, o “cérebro” de qualquer computador, sai de fábrica com um elenco predefinido de instruções, seu “instruction set”. Por mais versátil e “inteligente” que um computador possa parecer, todos os programas, sejam eles um “video game”, um programa de desenho, um “sistema especialista” ou uma planilha eletrônica, se constituem de um texto seqüencial formado de códigos de instruções do “instruction set” (INTEL, 2003) ou código de máquina do computador; ele não é capaz de fazer nada além disso. O computador realiza essas instruções uma a uma, recuperando-a da memória onde o programa esta armazenado, reconhecendo-a e acionando o circuito eletrônico capaz de realizá-la. Os manuais de arquitetura de computadores ensinam que as instruções do “instruction set” de qualquer processador se dividem em instruções lógicas, aritméticas e de movimentação de dados. Para as finalidades deste artigo vamos fazer uma classificação diferente.

A capacidade de processamento “semântico” de conteúdos armazenados na memória de um computador se restringe ou esta baseada tão somente em três operações: a- *identificar conteúdos* definidos por padrões de bits acessos e apagados, b- *desviar para outras seqüências de instruções* do programa de acordo com estes conteúdos e c- *relacionar conteúdos* armazenados em endereços de memória distintos uns com os outros, como explicado a seguir:

a- Identificar conteúdos significa, em nível de linguagem de máquina, comparar o conteúdo de dois endereços de memória onde estes conteúdos estão armazenados, verificar se suas seqüências de bits acesos e apagados são idênticas ou não. Assim, se determinada endereço de memória contém o conteúdo

“01000001” – a letra A em código ASCII⁷, e outro endereço de memória contém o conteúdo “01000010 – a letra B em código ASCII, estes conteúdos não são idênticos.

b- Desviar para outras seqüências de instruções de instruções do programa de acordo com estes conteúdos significa, em nível de linguagem de máquina, executar instruções especiais de desvio da seqüência de instruções do programa, de acordo com o conteúdo de determinado endereço de memória; estas instruções especiais são por exemplo, o “IF ... THEN ... ELSE...”. Assim, na seqüência de instruções de um programa, quando é encontrada uma instrução

```
IF A= “BRASIL” THEN Seq1 ELSE Seq2,
```

se o conteúdo do endereço de memória A contiver a seqüência de bits acesos e apagados que formam a palavra “BRASIL” a próxima instrução do programa a ser executada será a primeira a seguir ao parágrafo do programa indicado pelo rótulo “Seq1”; em caso contrário, a próxima instrução do programa a ser executada será a primeira a seguir ao parágrafo do programa indicado pelo rótulo “Seq2”.

Outras formas de instruções de desvio da seqüência de instruções do programa são as instruções “CASE”, que avalia o conteúdo da memória designado por A e desvia para diferentes seqüências do programa, por exemplo:

```
CASE
```

```
    A=“BOLÍVIA” Seq2  
    A=“PARAGUAI” Seq3
```

```
    ...
```

```
END CASE.
```

c- Relacionar conteúdos armazenados em endereços de memória distintos uns com os outros pode ser feito de maneira implícita ou explícita. Ao relacionarem conteúdos armazenados na memória, computadores podem fazê-lo na forma de relações implícitas, feitas através de instruções no texto de programas, ou na forma de relações explícitas, também armazenadas como conteúdos na memória. Quando um programa acessa um endereço de memória contendo, por exemplo, o conteúdo “Brasil”, e acessa também outro endereço de memória contendo o conteúdo “193.000.000”, e mostra uma tela que informa que a população do Brasil atualmente é 193.000.000 de habitantes, esta relacionando dois conteúdos de memória *implicitamente*. Por outro lado, quando, através de um programa navegador, “clicamos” um hiperlink ao navegarmos numa página da Web e

⁷ ASCII – American Standard Code for Information Interchange, código formado por 8 bits resultando em 256 combinações, através do qual todos os caracteres, sinais de pontuação e outros sinais são representados; ver em <http://pt.wikipedia.org/wiki/ASCII>

acessamos outra página, este hiperlink é uma relação *explícita* entre as duas páginas.

Essa última operação - relacionar conteúdos -, esta na base da capacidade de “inferência” computacional. Desde que os filósofos gregos desenvolveram a forma de raciocínio conhecido como *silogismo lógico* (SMITH, 2009), que o papel das relações para viabilizar a inferência é reconhecido. De fato, no conhecido exemplo de silogismo lógico, nas proposições “Todo o homem é mortal; Sócrates é homem; portanto, Sócrates é mortal”, o poder do raciocínio inferencial se sustenta essencialmente na relação estabelecida entre os três termos “homem”, “Sócrates” e “mortal” e *independentemente* do conteúdo das proposições. Aliás, este é irrelevante; alguém poderia dizer que: “Todo A é X; a é A; portanto, a é X” é verdadeiro porque se baseia na *forma* como se dá relação entre os termos, *independentes de conteúdo* das proposições:

[...] o que esta em questão, quando um lógico ou matemático constrói suas expressões não é aquilo que eventualmente poderiam referir, mas tão somente a maneira como se articulam as unidades que compõe tais expressões. Então, é a sintaxe ou forma, que constitui a preocupação essencial de tal comportamento teórico (CERQUEIRA; OLIVA, 1982, p. 10).

Portanto na medida em que elementos de um artefato computacional estejam interrelacionados, como num base de dados ou numa ontologia, um “agente de software” é capaz de realizar “inferências” automáticas, com base nessas relações, seguindo os “links” entre os nós de uma rede semântica. Numa ontologia de veículos a motor por exemplo, que tenha informações sobre classes de veículos como carros, ônibus, motocicletas, etc., e suas partes, como chassis, motor e rodas, e dividida em subclasses como veículos de passageiros e veículos de transporte, um “agente de software” poderia “inferir” que uma instância dessa ontologia, um “scooter”, tem também chassis, motor e rodas. Isso é possível porque instâncias, classes, subclasses e partes são interligadas por relações.

Na medida que artefatos computacionais se tornem altamente estruturados e tenham seus componentes interrelacionados de diversos modos semanticamente distintos, mais se amplia o potencial de “inferências” automáticas de “agentes de

software”. Aparte a metáfora sobre a capacidade de raciocínio dos computadores, essa é a essência da proposta da Web Semântica. Relacionar conteúdos permite uma ampliação da capacidade de “semântica computacional” além de meramente reconhecer um conteúdo, para também relacionar um conteúdo com outro, derivando ou “inferindo” um novo conteúdo não explícito de outro previamente identificado. Desta forma, ampliá-se a **Definição 1** como se segue:

Definição 3 - compreender as mensagens seria então ou identificar conteúdos, como exposto em a-, ou “inferi-los” (como exposto no parágrafo anterior) a partir de conteúdos já reconhecidos.

Qualquer veleidade de processamento dito “semântico” por computadores tem, ainda hoje, necessariamente que se basear nas três operações descritas. As propostas avançadas da Web Semântica, apesar de seu apelo metafórico por agentes de software processando de forma “inteligente” o conteúdo de documentos, no fundo, baseiam-se nessas três operações básicas.

Vimos que nosso modelo conceitual do computador é formado de duas entidades, dados e programa. A semântica computacional ou “compreensão” de um dado de entrada, segundo nossa **Definição 1**, estará portanto inscrita, *em maior ou menor grau*, conforme expõe Uschold (2003), nos programas, pelos programadores que os escreveram e que sabem, em última instância, o que deveria ser feito a partir da recepção de determinado dado de entrada. É inevitável portanto, como expõe o autor, que, pela própria natureza dos computadores, pelo menos um mínimo de capacidade “semântica” esteja inscrita no código dos programas, sendo portanto inseparável destes.

Se no entanto, nosso dado de entrada, ao invés de ser um dado simples, como “João”, é um dado estruturado como aparece na tabela de um banco de dados conforme ilustrado na figura seguinte.

Figura 1 – Tabela de banco de dados representando funcionários e seus cargos

	Matricula	Nome	Cargo
	03115543	João da Silva	Gerente

Ou então um outro tipo de dado estruturado, como o seguinte predicado binário, representado em Lógica de 1ª. ordem:

Figura 2 – Predicado binário especificando que “João” tem o cargo de “Gerente”.

<p>- cargo(João, Gerente)</p>

Assim estaria especificado que o cargo do Funcionário João é Gerente.

Em que grau nossa semântica computacional esta inscrita no próprio código dos programas ou já inscrita nos próprios dados de entrada ou pode ser obtida a partir destes. Esta questão tem implicações significativas para a questão que estamos discutindo.

Imaginemos que uma agência de viagem deseja construir um programa “agente inteligente” que percorra páginas Web para procurar as ofertas de preços de passagens aéreas para a agência oferecer aos seus clientes. Na forma mais tosca de ser construído, este programa inicializaria uma memória usada para comparação como o conteúdo “oferta de passagem aérea” e toda vez que encontrasse (isto é, obtivesse como entrada) uma página que contivesse em seu texto estas quatro palavras, guardaria o valor seguinte

à palavra “aérea”, juntamente com o endereço Web da página. No final do processamento exibiria o endereço da página Web cujo valor fosse o menor entre os encontrados. Neste caso a semântica computacional deste programa estaria totalmente inscrita no seu código: só teriam significado para o programa páginas Web com o conteúdo “oferta de passagem aérea” e destas, ele só se interessaria pelo menor valor.

Se, no entanto, ao invés desse programa inicializar uma memória para comparação com o conteúdo “oferta de passagem aérea”, ele permitisse que seu usuário entrasse, ele próprio digitando o conteúdo desta memória, por exemplo, “ofertas de hospedagem”, no início de funcionamento do programa. Um programa que funcionasse dessa forma teria um grau de flexibilidade superior à versão anterior, pois poderia obter pelo menos parte de sua “semântica” de “compreensão” dos dados de entrada diretamente de outros dados de entrada, fornecidos como parâmetro para o seu funcionamento; a outra parte naturalmente, a exibição do menor preço, estaria ainda inscrita no código do programa.

Segundo a **Definição 2** de modelo computacional, as linguagens da Web Semântica são *dados*, objetivam estruturar conteúdos e não *processos*, que especificam passos de processamento de dados, como as linguagens computacionais em geral; mesmo linguagens orientadas a objeto com as quais são desenvolvidas aplicações na Web como C++ e Java, são linguagens processáveis, dependem, em maior ou menor grau, de processadores – “software” e “hardware” – que processem suas instruções.

Já as linguagens da Web Semântica visam não processamento, mas estruturação de conteúdos. Segundo a Teoria dos Sistemas, a estrutura e complexidade de um sistema são dadas pelo *número de relações* entre seus componentes (SANTAELLA; VIEIRA, 2008, p. 36). Por isso, a arquitetura da Web Semântica visa estruturar conteúdos (documentos) de forma crescentemente mais complexa, através de linguagens que expressam mais complexas e diferenciadas relações, como explicitado na proposta de Berners-Lee Hendler e Lassila (2001). Desta forma um nível sintático, com conteúdos estruturados em linguagem XML, serve de base para níveis crescentemente estruturados, aumentando assim o potencial de “inferência computacional” ou processamento semântico: XML

Schema, RDF, RDF Schema e ontologias, estruturadas em linguagem OWL. A estruturação de conteúdos segundo os padrões da proposta da Web Semântica é assim descrita:

XML provides a surface syntax for structured documents, but imposes no semantic constraints on the meaning of this documents.

XML Schema is a language for restricting the structure of XML documents and also extends XML with datatype.

RDF is a datamodel for objects (“resources”) and relations between them, provides a simple semantic for this datamodel, and this datamodels can be represented in an XML syntax.”

RDF Schema is a vocabulary for describing properties and classes with a semantic for generalization-hierarchies of such properties and classes.

OWL adds more vocabulary for describing properties and classes: among others, relations between classes (e.g. disjointness) cardinality (e.g. “exactly one”), equality, richer typing of properties, characteristics of properties (e.g. symmetry), and enumerated classes. (OWL Ontology Web Language Overview, 2004. p.3).

Segundo de Berners-Lee, Hendler e Lassila (2001), o padrão RDF seria fundamental e marcaria um ponto de inflexão na seqüência de linguagens apresentada, pois ali estaria o primeiro “lampejo” do que seria uma “semântica” inteligível por programas: “In short, XML allows users to add arbitrary structure to their documents but **says nothing about what the structures mean. Meaning is expressed by RDF**⁸, which encodes it in sets of triples, each triple being rather like the subject, verb and object of an elementary sentence”.

Outra questão importante ligada à Semântica diz respeito à capacidade expressiva de uma linguagem, a que entidades ou fenômenos do mundo ela é capaz de expressar. Muitas vezes dizemos que um sentimento é inexpressível com palavras. As linguagens têm uma capacidade limitada de expressar fenômenos, especialmente linguagens artificiais como as linguagens computacionais em geral, ou linguagens gráficas como os modelos conceituais construídos segundo o (meta)modelo Entidades-Relacionamentos (CHEN, 1976). A “semântica” inteligível por programas se daria porque o RDF seria, no ordenamento das linguagens da WEB mencionado, a primeira linguagem que *pressupõe* uma semântica; esta é composta de sujeito, predicado e objeto de uma

⁸ Grifo nosso.

assertiva, os elementos que estão pressupostos no seu metamodelo. Este é definido nas palavras de Guizzardi (2005, p. 36) como:

What is referred by structure of a language can be accessed via the description of the specification of conceptual model underlying the language, i.e., a description of the worldview embedded in the language's modeling primitives. In (Milton & Kamierczak, 2004), this is called the ontological metamodel of the language, or simply, the ontology of the language. From a philosophical standpoint, this view is strongly associated with Quine (Quine, 1969), who proposes that an ontology can be found in the ontological commitment of a given language, that is, the entities the primitives of a language commit to the existence of. For example, Peter Chen's Entity Relationship model (Chen, 1976) commits to a worldview that accounts for the existence of three types of things: entity, relationship and attribute.

Assim, um metamodelo expressa os pressupostos semânticos – o significado -, de cada elemento do vocabulário da linguagem. RDF marcaria uma mudança de qualidade, já que, ao contrário do XML, possui um metamodelo. Segundo o metamodelo ontológico da linguagem RDF (ou seja, que semântica pode ser expressa com os elementos da linguagem RDF), sempre se poderá identificar numa assertiva RDF sobre que entidade é feita a assertiva (o sujeito), que tipo de assertiva esta sendo dito sobre esta entidade (o predicado) e o quê especificamente esta sendo dito sobre a entidade (o objeto) (RDF PRIMER, 2004).

No amplo espectro de formalismo para representar conteúdos que vêm sendo usados desde que o computador se tornou um instrumento privilegiado de processar representações do mundo e tendo em vista a proposta da Web Semântica, uma primeira sistematização e ampliação da discussão anterior de uma possível “semântica computacional” classificaria a capacidade dos computadores “compreenderem” mensagens expressas por linguagens cuja expressividade varie entre os seguintes pólos:

- permitam *semânticas externalizadas, públicas, concensadas e padronizadas*, em oposição a *semânticas locais, particulares, internas, inscritas e prisioneira do código de programas*;

- privilegiaria também *semânticas formalizadas*, na qual os elementos da linguagem têm um papel inequívoco, como o (meta)modelo conceitual E-R e os modelos conceituais construídos com ele como o FRBR ou o CIDOC-CRM, ou as

representações e ontologias construídas com RDF, RFDS e OWL em oposição a semânticas informais, inscritas em textos inteligíveis somente por pessoas, como conjunto de metadados Dublin Core, cujos significados são dados em última instância pelas diversas comunidades de usuários que o utilizam;

- por fim, privilegiaria linguagens *processáveis por computador*, como as já mencionadas representações e ontologias construídas com RDF, RFDS e OWL em oposição a modelos conceituais construídos com o (meta) modelo conceitual E-R mas não processáveis por computadores.

Fica claro então que, para otimizar a proposta da Web Semântica,

Definição 4 - uma “semântica computacional” teria que ser: a- minimamente inscrita no código dos programas; b- inteligível por programas (formal) e c- o máximo disponível publicamente (na própria Web), consensada, compartilhada, padronizada, independente de programas.

Percebe-se que na proposta da Web Semântica vêm ganhando importância bases de conhecimento independentes de programas como destaca Guarino (1995, p. 625), em que os relacionamentos entre conteúdos são explícitos e têm semântica própria, as chamadas ontologias, que acentuam o potencial de “semântica computacional”, conforme descrito na **Definição 4**.

Numa ontologia computacional em que a classe de Veículos a motor tem como subclasses, como na figura:

Figura 4 – ontologia de veículos (RDF Primer, 2004)

```
<rdf:Description rdf:about="#Veículos_a_Motor">
  <rdf:type rdf:resource="http://www.w3.org/2000/01/rdf-schema#Class"/>
</rdf:Description>
<rdf:Description rdf:about="#Veículos_de_Carga">
  <rdfs:subClassOf rdf:resource="#Veículos_a_Motor"/>
  <rdf:type rdf:resource="http://www.w3.org/2000/01/rdf-schema#Class"/>
</rdf:Description>
<rdf:Description rdf:about="#Veículos_de_Passageiros">
  <rdfs:subClassOf rdf:resource="#Veículos_a_Motor"/>
  <rdf:type rdf:resource="http://www.w3.org/2000/01/rdf-schema#Class"/>
</rdf:Description>
<rdf:Description rdf:about="#Onibus">
  <rdf:type rdf:resource="http://www.w3.org/2000/01/rdf-schema#Class"/>
  <rdfs:subClassOf rdf:resource="#Veículos_de_Passageiros"/>
</rdf:Description>
<rdf:Description rdf:about="#Van">
  <rdfs:subClassOf rdf:resource="#Veículos_de_Passageiros"/>
  <rdf:type rdf:resource="http://www.w3.org/2000/01/rdf-schema#Class"/>
</rdf:Description>
<rdf:Description rdf:about="#MiniVan">
  <rdf:type rdf:resource="http://www.w3.org/2000/01/rdf-schema#Class"/>
  <rdfs:subClassOf rdf:resource="#Van"/>
</rdf:Description>
<rdf:Description rdf:about="#MiniVan_LCD_2089">
  <rdf:type rdf:resource="#MiniVan"/>
</rdf:Description>
```

Se a classe Veículos a motor tem as propriedades descritas na figura seguinte, um programa “agente inteligente” poderia “inferir” que a MiniVan de placa LCD_2089, por ser uma Van, que por sua vez é um Veículo a motor, tem como componentes motor, chassis e carroceria.

Figura 5 – propriedades das classes da ontologia de veículos

```
<rdf:Description rdf:about="#Tem_Componente">
  <rdfs:domain rdf:resource="#Veiculos_a_Motor"/>
  <owl:inverseOf rdf:resource="#E_Componente"/>
  <rdf:type rdf:resource="http://www.w3.org/2002/07/owl#ObjectProperty"/>
  <rdfs:range rdf:resource="#Componentes"/>
</rdf:Description>
<rdf:Description rdf:about="#E_Componente">
  <rdfs:domain rdf:resource="#Componentes"/>
  <owl:inverseOf rdf:resource="#Tem_Componente"/>
  <rdf:type rdf:resource="http://www.w3.org/2002/07/owl#ObjectProperty"/>
  <rdfs:range rdf:resource="#Veiculos_a_Motor"/>
</rdf:Description>
<rdf:Description rdf:about="#Componentes">
  <rdf:type rdf:resource="http://www.w3.org/2000/01/rdf-schema#Class"/>
</rdf:Description>
<rdf:Description rdf:about="#Motor">
  <rdfs:subClassOf rdf:resource="#Componentes"/>
  <rdf:type rdf:resource="http://www.w3.org/2000/01/rdf-schema#Class"/>
</rdf:Description>
<rdf:Description rdf:about="#Chassis">
  <rdf:type rdf:resource="http://www.w3.org/2000/01/rdf-schema#Class"/>
  <rdfs:subClassOf rdf:resource="#Componentes"/>
</rdf:Description>
<rdf:Description rdf:about="#Suspensão">
  <rdf:type rdf:resource="http://www.w3.org/2000/01/rdf-schema#Class"/>
  <rdfs:subClassOf rdf:resource="#Componentes"/>
</rdf:Description>
<rdf:Description rdf:about="#Carroceria">
  <rdf:type rdf:resource="http://www.w3.org/2000/01/rdf-schema#Class"/>
  <rdfs:subClassOf rdf:resource="#Componentes"/>
</rdf:Description>
```

Não parece muito, mas uma ontologia computacional pode ser bem mais complexa. É um modelo formal de um determinado domínio, consensado dentro de uma comunidade, contendo as entidades, as relações entre estas e regras para caracterizar cada categoria de entidade e para que estes relacionamentos se dêem. Diversas definições de ontologias podem ser encontradas na literature, como por exemplo:

[...] an ontology is a partial conceptualization of a given knowledge domain, shared by a community of users, that has been defined in a formal, machine-processable language for the explicit purpose of sharing semantic information across automated systems. (JACOB, 2003, p. 20).

Por ser um artefato altamente estruturado, um arquivo codificado em linguagem OWL, onde as relações existentes no domínio modelado são explícitas, uma ontologia, quando processada por programas, permite a estes realizar “inferências computacionais”.

Segundo nossa **Definição 2** de modelo computacional, uma ontologia, como nossa *veiculos.owl*, é um *dado*, altamente *estruturado*, contendo o conhecimento explícito do domínio de veículos a motor, com seus elementos estruturados e altamente interrelacionados, permitindo assim um alto grau de “inferências computacionais” por *programa*. Estes, os “agentes inteligentes”, conteriam, inscritos em seus códigos, os procedimentos para seguir as relações explícitas, padronizadas na linguagem OWL ou seja, realizar “inferências computacionais”.

4 CONCLUSÕES: O MUNDO DIGITAL, VERDADE, MITOS, SONHOS E VIAGENS

Identificar conteúdos ou “inferi-los” a partir de conteúdos já reconhecidos, independentemente (e apesar) desses conteúdos. É nisso que se baseia em última instância o potencial semântico e de inferência dos computadores e, em consequência, a proposta de semântica da Web Semântica. Esta questão é deixada clara por Tim Berners-Lee (1998b) nas formulações iniciais da proposta:

The basic assertion model provides the concepts of assertion (property) and quotation. (This is provided by the RDF Model and Syntax Specification). This allows an entity-relationship-like model to be made for the data, giving it the *semantics of assertions propositional logic*⁹.

A lógica proposicional ou lógica formal a que Berners-Lee se refere é, como já foi visto (CERQUEIRA; OLIVA, 1982) é a ciência das formas de inferência, *independentes de conteúdo* das asserções. Sua semântica são as tabelas de verdade das asserções ou proposições unidas pelos conectivos lógicos, nas quais proposições são representadas por letras minúsculas como p, q, r, etc., *independentes de seu conteúdo*. Há bastante tempo que os teóricos da linguagem identificaram a importância da pragmática dos falantes, de suas intenções e do contexto para a compreensão de significados, muito além e de forma muito mais complexa que a manipulação de relações da lógica formal. Seguindo

⁹ Destaque nosso

Wittgenstein e Austin, Searle (1971) propõe uma teoria dos atos da fala; Habermas propõe uma pragmática universal (COSTA, 2003, p. 43).

A limitação semântica da Web Semântica é inerente ao paradigma computacional atual, o do processamento de símbolos, o chamado cognitivismo ou representacionismo (HARNAD, 1990). Este paradigma, muitas vezes usado para explicar os processos cognitivos em seres humanos (HORST, 2009), está na raiz da visão de que a mente humana funciona como um computador já discutida na Introdução. Segundo o cognitivismo um processo cognitivo (humano ou simulado em máquinas) consistiria na manipulação ordenada de um conjunto de símbolos, como são feitas as inferências da lógica formal já vistas (CERQUEIRA; OLIVA, 1982) ou nas máquinas de Turing. A manipulação ordenada de símbolos faz com que, no limite, um símbolo se relacione a ouro. No entanto os símbolos devem *referir-se* à realidade. A cognição e a semântica humanas vão muito além da simples manipulação de símbolos. É Harnad (1990, p. 335) que pergunta:

How can the semantic interpretation of a formal symbol system be made *intrinsic* to the system, rather than just parasitic on the meanings in our heads? How can the meanings of the meaningless symbol tokens, manipulated solely on the basis of their (arbitrary) shapes, be grounded in anything but other meaningless symbols? The problem is analogous to trying to learn Chinese from a Chinese/Chinese dictionary alone

Críticos como Gärdenfors (2004) questionam diretamente o caráter “semântico” da Web Semântica: “The Semantic Web is not semantic. It is good for syllogistic reasoning, but there is much more to semantics than syllogisms”. Outro crítico, Shirky (2003), declara enfaticamente caráter utópico da proposta, o esforço necessário e as dificuldades de sua implementação:

The Semantic Web, with its neat ontologies and its syllogistic logic, is a nice vision. However, like many visions that project future benefits but ignore present costs, it requires too much coordination and too much energy to effect in the real world, where deductive logic is less effective and shared world view is harder to create than we often want to admit.

Em contextos limitados, agentes de busca de preços citados por Uschold (2003) já implementam a **Definição 1** de semântica. A proposta da Web Semântica só faz ampliar este potencial.

Uma formulação recente do W3C apresenta uma visão bem mais realista da proposta da Web Semântica em comparação com as apresentada originalmente por Tim Berners-Lee (1998a, 1998b, 2001) nos textos já comentados:

The Semantic Web is a web of data. There is lots of data we all use every day, and it is not part of the web. I can see my bank statements on the web, and my photographs, and I can see my appointments in a calendar. But can I see my photos in a calendar to see what I was doing when I took them? Can I see bank statement lines in a calendar?

Why not? Because we don't have a web of data. Because data is controlled by applications, and each application keeps it to itself.

The Semantic Web is about two things. It is about common formats for integration and combination of data drawn from diverse sources, where on the original Web mainly concentrated on the interchange of documents. It is also about language for recording how the data relates to real world objects. That allows a person, or a machine, to start off in one database, and then move through an unending set of databases which are connected not by wires but by being about the same thing. (W3C Semantic Web Activity, 2010).

Atenção, a citação fala de uma web de *dados*, e não uma web de significados. É uma mudança significativa da proposta original.

Dos três tipos de inferência lógica, a dedução, a indução e a abdução (REICHERTZ, 2004), a proposta da Web Semântica só implementa a dedução (VELTMAN, 2004), sob a forma de silogismo. Embora sistemas computadores fechados em projetos de IA possam implementar raciocínios indutivos e mesmo abduativos (COLTON; STEEL, 1999), a proposta da Web Semântica é a de um sistema aberto. Como enfatiza Shirky (2003):

The Semantic Web specifies ways of exposing these kinds of assertions on the Web, so that third parties can combine them to discover things that are true but not specified directly. This is the promise of the Semantic Web -- it will improve all the areas of your life where you currently use syllogisms.

As aplicações do silogismo, ou *dedução* a partir de premissas, são limitadas, conforme enfatiza este autor. As áreas de aplicação onde existe conhecimento

universalmente válido, critérios de verdade absolutos, independentes de contexto, são as ciências e os negócios. Demo (2009, p. 2) ressalta que: “Nos contextos hermenêuticos, tomam-se ambigüidades como naturais e próprias da dinâmica semântica, já que nenhum conceito pode ser definido sem supor outros conceitos ainda não definidos”. Toda a área de cultura, que valoriza o valor das interpretações, as singularidades, contextos específicos, históricos e culturais, a evolução do conhecimento ao longo do tempo, são pouco suscetíveis ao uso da lógica formal (VELTMAN, 2004), que busca um critério único e absoluto de verdade. Por fim, além de tudo que foi dito, esse tipo de inferência, o silogismo, pode conduzir às chamadas falácias, como por exemplo: "Se há carros então há poluição. Não há carros. Logo, não há poluição", ou “Se choveu então a terra fica molhada. A terra esta molhada, então choveu”.

Críticos como Floridi (2009), consideram também as capacidades “semânticas” da proposta da Web Semântica limitadas. Este autor aposta mais nas capacidades interativas e comunicacionais da Web 2.0 como capazes de potencializar os verdadeiros elementos semânticos ligados nas redes computacionais, as pessoas.

As capacidades “semânticas” da Web Semântica são limitadas e devem ser compreendidas dentro dessas limitações, sem maiores ilusões e fantasias. Isto não significa que a proposta seja inválida; de fato, ela encerra muitas potencialidades, incapazes de serem atingidas na Web atual.

Assim como nas discussões sobre a origem do conhecimento entre Empiristas e Racionalistas, ambas as posições, a de que o conhecimento provem somente dos sentidos do sujeito ou a de que o conhecimento já preexiste na mente do sujeito, foram superadas por Kant (1991) ao afirmar que o conhecimento depende da capacidade de conhecer “a priori” do sujeito, a capacidade de processamento “semântico” por parte dos computadores não pode depender somente dos programas ou somente dos dados, ela deve encontrar uma capacidade de processamento “a priori” de dados por parte dos programas. Isso é inevitável, dada a arquitetura e ao paradigma de funcionamento atual dos computadores. “[...] toda significação depende de significações prévias elaboradas na

linguagem [...]” (DEMO, 2009, p. 2). Ou seja, alguma “semântica” é inevitavelmente interna aos programas, não pode ser publicizada em artefatos como ontologias.

Sintetizando visões anteriores de vários autores, Obrst (2010) apresenta a evolução dos diferentes sistemas de organização do conhecimento, ordenado-os de acordo com seus graus de expressividade semântica. No maior grau de expressividade semântica, que o autor chama de “strong semantic”, estão as ontologias formais expressas como teorias lógicas que usam lógicas modais para exprimir suas propriedades. O autor também fala em incrementos semânticos devidos à integração de ontologias. Pode-se pensar em ontologias de aplicação integradas a ontologias de alto nível, formando uma complexa rede semântica, assim como a integração de ontologias com mecanismos de garantia de confiança nas informações (a camada “trust” da Web Semântica) disponibilizadas por diferentes sistemas de informação, que equivaleriam ao que na semântica de Tarski seriam as funções de interpretação da teoria lógica expressa pela ontologia de aplicação. Pode-se então antever um complexo sistema semântico, embora baseado somente na concepção de semântica da lógica dedutiva.

Apesar das limitações discutidas acima, no paradigma da Web Semântica, *relações* são a essência, o “coração” mesmo (SHET; ARPINAR; KASHYAP, 2003), da capacidade de computadores realizarem “inferências computacionais”, ou seja, deduções.

When the relationships among data are fully accessible to our machines, our machines will be able to help us browse those relationships and interpret the data as well as assess the appropriateness of the data for our intended purposes. (MILLER; SWICK, 2003, p. 8).

Em recente publicação a importância das relações semânticas, não só entre documentos, como na proposta original da Web Semântica, mas também entre todos os tipos de dados - “linked data” -, é enfatizada (BIZER; HEATH; BERNERS-LEE, 2009).

Na medida em que elementos de um artefato computacional estejam interrelacionados, como num base de conhecimento ou numa ontologia, um “agente de software” é capaz de realizar “inferências computacionais”, com base nessas relações. Portanto torna-se um objetivo de pesquisa e uma contribuição importante da área de CI para a proposta da Web Semântica (KHOO; NA, 2007) a identificação e sistematização de

um conjunto básico de relações semânticas, dentre as diferentes propostas existentes (WÜSTER, 1981), (DAHLBERG, 1978, 1995), (PERREAULT, 1994), (SMITH; GRENON, 2004), (SMITH et al., 2005). Se estas relações forem de aplicação geral e independentes de domínios específicos, sua “semântica” poderá ser especificada usando lógica e representada em formato “inteligível” por programas, de modo a poderem ser usadas de forma generalizada, e assim, apesar das ressalvas, ampliar as potencialidades da Web “Semântica”.

Agradecimentos: este artigo apresenta os resultados iniciais de pesquisa desenvolvida durante estágio sênior apoiada pela CAPES, realizado na Universidade Carlos III de Madrid sob a orientação do Prof. Jose Antonio Moreiro González.,

REFERÊNCIAS

BATH, Peter A. Data Mining in Health and Medical Information. *Annual Review of Information Science and Technology*, v. 3, n. 1, 2004. p. 331-369.

BERNERS-LEE, Tim. Semantic Web Road Map. W3C, 1998. Disponível em: <<http://www.w3.org/DesignIssues/Semantic.html>>. Acesso em 18 maio 2005.

BERNERS-LEE, Tim. Web Architecture from 50,000 feet. W3C, 1998. Disponível em: <<http://www.w3.org/DesignIssues/Architecture.html>>. Acesso em 18 maio 2005.

BERNERS-LEE, Tim; HENDLER, James; LASSILA, Ora. The semantic web. *Scientific American*, May, 2001. Disponível em: <<http://www.scian.com/2001/0501issue/0501berners-lee.html>>. Acesso em: 24 maio 2001.

BIZER, Christian; HEATH, Tom; BERNERS-LEE, Tim. Linked Data - The Story So Far. *International Journal on Semantic Web and Information Systems (IJSWIS)*, v. 5, n. 3, p. 1-22, 2009. Disponível em: < <http://tomheath.com/papers/bizer-heath-berners-lee-ijswis-linked-data.pdf> >. Acesso em 26 jan. 2011.

BRETON, Philippe. *História da informática*. São Paulo : Ed. UNESP, 1991. 260 p.

BUSH, Vannevar. As we may think. *The Atlantic Monthly*, July, 1945. Disponível em: <<http://www.theatlantic.com/unbound/flashbacks/computer/bushf.html>>. Acesso em: 1 jun. 2004.

- CHIERCHIA, Gennaro. *Semântica*. Campinas, SP: Editora da UNICAMP, 2003.
- DUPUY, Jean-Pierre. *Nas origens das ciências cognitivas*. São Paulo: Ed. UNESP, 1996. 228 p.
- CHEN, Peter Pin-Shan. The Entity-Relationship Model-Toward a Unified View of Data. *ACM Transactions on Database Systems*, v. 1, n.1. p.9-36, March 1976. Disponível em: <<http://csc.lsu.edu/news/erd.pdf>>. Acesso em 15 fev. 2008.
- CHOMSKY, Noan. Aspectos da teoria da sintaxe. In: *Textos selecionados*. São Paulo: Abril Cultural, 1975 p. 233-290. (Os Pensadores, 49).
- COLTON, Simon; STEEL, Grahah. Artificial Intelligence and Scientific Creativity. In: AISB 99 Symposium on Artificial Intelligence and Scientific Creativity, Edinburgh, Scotland, 1999, *Proceedings... Society for the Study of Artificial Intelligence and Simulation of Behaviour*, 1999. Disponível em: <<http://www.dai.ed.ac.uk/homes/simonco/papers/AISBQ99.html>>. Acesso em 26 jan. 2008.
- COSTA, Claudio. *Filosofia da linguagem*. 2. ed. Rio de Janeiro: Jorge Zahar ed., 2003. (Filosofia passo a passo).
- DAHLBERG, Ingetraut. Conceptual structures and systematization. *International Forum on Information and Documentation*, v. 20, n. 3, July, 1995.
- DAHLBERG, Ingetraut. *Ontical structures and universal classification*. Bangalore: Sarada Ranganathan Endowment for Library Science, 1978.
- DEMO, Pedro. *Ontologia e Web Semântica*. 2009. (Texto para discussão apresentado quando da visita do autor ao PPGCI/UFF, em agosto de 2009).
- DRUMMOND, Nick; SHEARER, Rob. The Open World Assumption or Sometimes its nice to know what we don't know. University of Manchester, 2006. (apresentação .PPT). Disponível em: <<http://www.cs.man.ac.uk/~drummond/presentations/OWA.pdf>>. Acesso em 19 jan. 2011.
- ECO, Umberto. *Os limites da interpretação*. São Paulo: Perspectivas, 1995. 315 p. (Coleção Estudos).
- FLORIDI, Luciano. Web 2.0 vs. the Semantic Web: A Philosophical Assessment. *Episteme*, v. 6, p. 25-37, 2009. Disponível em <<http://www.philosophyofinformation.net/publications/pdf/w2vsw.pdf>>. Acesso em: 1 ago. 2009.

FODOR, Jerry. Semântica: uma entrevista com Jerry Fodor. *ReVEL*, v. 5, n. 8, 2007. Disponível em: <http://www.revel.inf.br/site2007/_pdf/8/entrevistas/revel_8_entrevista_jerry_fodor.pdf>.

Acesso em: 28 jan. 2011.

FRANKLIN, Stan; GRAESSER, Art. Is it an Agent, or just a Program?: A Taxonomy for Autonomous Agents. In: International Workshop on Agent Theories, Architectures, and Languages, 3, Proceedings... [S.e]: Springer-Verlag, 1996. Disponível em: <<http://www.msci.memphis.edu/~franklin/AgentProg.html>>. Acesso em 18 jan. 2011.

FREGE, Gottlob. Sense and Reference. *The Philosophical Review*, v. 57, n. 3, p. 209-230, May, 1948. Disponível em: <<http://www.jstor.org/stable/2181485>>. Acesso em: 1 dez. 2009.

GUIZZARDI, Giancarlo. Ontological foundations for structural conceptual models. Enschede, The Netherlands: CTIT, Telematica Institut, 2005.

HARNAD, Stevan. The symbol grounding problem. *Physica*, v. 42, p. 335-346, 1990. Disponível em: <<http://www.lizardphunk.org/skrivut/gvammen/Harnad%20symbol%20grounding.pdf>>. Acesso em 22 jan. 2011.

HEFLIN, J.; HENDLER, J.; LUKE, S. SHOE: A blueprint for the Semantic Web. In: FENSEL, D; HENDLER, J.; LIEBERMAN, H.; WAHLSTER, (Eds.), *Spinning the Semantic Web: Bringing the World Wide **Web** to its full potential*. Cambridge, MA: MIT Press, 2003. p. 29-63.

HORST, Steven. The Computational Theory of Mind. In: ZALTA, Edward (Ed.). *The Stanford Encyclopedia of Philosophy* (Winter 2009 Edition). Disponível em <<http://plato.stanford.edu/archives/win2009/entries/computational-mind/>>. Acesso em 25 nov. 2010.

HUBERT DREYFUS, DEMISTIFICADOR DA INTELIGÊNCIA ARTIFICIAL, ANTE EDWARD FEIGUENBAUM, ESPECIALISTA EM SISTEMAS ESPECIALISTAS. In: PESSIS-PARTERNAK, Guitta. *Do caos à inteligência artificial*. São Paulo: Ed. UNESP, 1993 p. 207-232.

IANELLA, Renato. Semantic Web Architecture. *Semantic Identity*, 2010. Disponível em : <[http://semanticidentity.com/Resources/Entries/2010/9/8_Semantic_Web_Architectures_\(Whitepaper\)_files/semweb-arch-wp-2010.pdf](http://semanticidentity.com/Resources/Entries/2010/9/8_Semantic_Web_Architectures_(Whitepaper)_files/semweb-arch-wp-2010.pdf)>. Acesso em : 19 jan. 2011.

JACOB, Elin K. Ontologies and the Semantic Web. *Bulletin of the American Society for Information Science and Technology*, April/May, 2003. p. 19-22.

KANT, Immanuel. *Crítica da Razão Pura*. São Paulo : Nova Cultural, 1991. (Os pensadores, 7).

KOSTOFF, Ronald N.; BRIGGS, Michael B. ; SOLKA, Jeffrey L. ; RUSHENBERG, Robert L. Literature-related discovery (LRD): Methodology. *Technological Forecasting & Social Change*, v. 75, p. 186–202, 2008.

KHOO, C.; NA, J. C. Semantic Relations in Information Science. *Annual Review of Information Science and Technology*, p. 157-228, 2007.

LUHN, H. P. The automatic creation of literature abstracts. *IBM Journal of Research and Development*, v. 2, n. 2, p. 159–165, 1958.

MARCONDES, Carlos Henrique; MENDONÇA, Marília Alvarenga Rocha; MALHEIROS, Luciana Reis; COSTA, Leonardo Cruz da; SANTOS, Tatiana. Cristina Paredes. Bases ontológicas e conceituais para um modelo do conhecimento científico em artigos biomédicos. *RECIIS*, v. 3, n. 1, p.19-30, 2009. Disponível em: <<http://www.reciis.cict.fiocruz.br/index.php/reciis/article/view/240/251>>. Acesso em: 8 abr. 2009.

MARCONDES, Carlos Henrique. Linguagem e Documento: fundamentos evolutivos e culturais da Ciência da Informação. *Perspectivas em Ciência da Informação*, v. 15, n. 2, 2010. p. 2-21. Disponível em <<http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/viewFile/1019/729>>. Acesso em 14 set. 2010.

MARX, Carlos. Prólogo de La contribución a la crítica de la economía política. In: MARX, Carlos; ENGELS, Frederico. *Obras Escogidas*. Moscou: Editorial Progreso, 1975. p. 181-185.

MATURANA, Humberto R.; VARELA, Francisco J. *A árvore do conhecimento: as bases biológicas da compreensão humana*. São Paulo: Palas Athena, 2001.

MILLER, Eric; SWICK, Ralph. An overview of W3C Semantic Web Activity. *Bulletin of the American Society for Information Science and Technology*, p. 8-11, April/May 2003.

MORRIS, Charles W. *Fundamentos da teoria dos signos*. São Paulo: Ed. Da Universidade de São Paulo, 1976.

OBRST, Leo. *Ontologies & the Semantic Web for Semantic Interoperability*. 2010. Apresentação .ppt. Disponível em <www.web-services.gov/OntologiesSemanticWebSemInteropSICOP909-Obrst.ppt>. Acesso em 27 jul. 2010.

PERREAULT, Jean. Categories and relators: a new schema. *Knowledge Organization*, v. 21, n. 4, p. 189-198, 1994.

RESOURCE DESCRIPTION FRAMEWORK (RDF) Model and Syntax Specification. W3C, 1999. Disponível em <<http://www.w3.org/TR/PR-rdf-syntax/>>. Acesso em 18 maio 2005.

RDF PRIMER. MANOLA, Frank; MILLER, Eric (eds.). W3C, 2004. Disponível em: <<http://www.w3.org/TR/2004/REC-rdf-primer-20040210/>>. Acesso em: 15 dez. 2011.

REICHERTZ, J. Abduction, deduction and induction in qualitative research. In: Dans U. Flick, E. von Kardorff, & I. Steinke (Eds). *A companion to qualitative research*. London: Sage, 2004. p. 159-164.

SARACEVIC, Tefko. Ciência da informação: origem, evolução e relações. *Perspectivas da Ciência da Informação*, v. 1, n. 1, p. 41-62, jan./jun. 1996.

SEARLE, J. R. What is a speech act? In: SEARLE, J. R (Ed.) *The Philosophy of Language*. Oxford: Oxford University Press 1971. p. 44–46. Disponível em: <<http://openlearn.open.ac.uk/file.php/3183!/via/oucontent/course/60/reading2.pdf>>. Acesso em 21 jan. 2010.

SHIRKY, C. The Semantic Web, Syllogism, and Worldview. In: “*Networks, Economics, and Culture*” mailing list, November 7, 2003. Disponível em: <http://www.shirky.com/writings/semantic_syllogism.html> . Acesso em 18 jan. 2011.

SHET, A.; ARPINAR, I. B.; KASHYAP, V. Relationships at the heart of semantic web: modeling, discovering and exploiting complex semantic relationships. In: M. Nikraves, B. Azvine, R. Yager and L.A. Zadeh (eds.), *Enhancing the Power of the Internet - studies in fuzziness and soft computing* (Springer-Verlag, 2003). Disponível em cgsb2.nlm.nih.gov/~kashyap/publications/relations.pdf. Acesso em 18 set 2006.

SMITH, Barry; CEUSTERS, Werner; KLAGGES, Bert; KÖHLER, Jacob; KUMAR, Anand; LOMAX, Jane; MUNGALL, Chris; NEUHAUS, Fabian; RECTOR, Alan L.; ROSSE, Cornelius. Relations in biomedical ontologies. *Genome Biology* 2005, v. 6, n. R46, April 2005. Disponível em <<http://genomebiology.com/2005/6/5/R46>>. Acesso em 4 dez. 2005.

SHERA, J. H., Kent, A; PERRY, J. W. (ed.). *Information systems in documentation*. NY: Interscience Publishers, 1957. (Advances in Documentation and Library Science, v. 1).

SMITH, Barry; GRENON, Pierre. The cornucopia of formal-ontological relations. *Dialectica*, v. 58, n. 3, p. 279-296, 2004. Disponível em: <<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.71.9983&rep=rep1&type=pdf>>. Acesso em: 16 abril 2004.

SMITH, Robin. Aristotle's Logic. In: Edward N. Zalta (Ed.). *The Stanford Encyclopedia of Philosophy* (Spring 2009 Edition). Disponível em:

PontodeAcesso, Salvador, V.6, n.2, p. 35-73, dez 2012

www.pontodeacesso.ici.ufba.br

<<http://plato.stanford.edu/archives/spr2009/entries/aristotle-logic/>>. Acesso em 22 jan. 2011.

SPEAKS, Jeff. Theories of Meaning. In: Edward N. Zalta (Ed.). *The Stanford Encyclopedia of Philosophy* (Winter 2010 Edition). Disponível em: <<http://plato.stanford.edu/archives/win2010/entries/meaning/>>. Acesso em: 24 dez. 2010.

TENÓRIO, Robinson. *Computadores de papel: máquinas abstratas para um ensino concreto*. 2 ed. São Paulo: Cortez Editora, 2001. (Questões de nossa época, 80).

TARSKI, Alfred. The semantic concept of truth: and the foundations of semantic. *Philosophy and Phenomenology research*, v.4, n. 3, p. 341-376, March 1944.

UNIFIED MEDICAL LANGUAGE SYSTEM FACT SHEET. 2006. Disponível em: <<http://www.nlm.nih.gov/pubs/factsheets/umls.html>>. Acesso em: 23 março de 2007.

USCHOLD, Michael. Where are the Semantics in the Semantic Web? *AI Magazine*, v. 24. n. 3, 3 2003. Disponível em <<http://www.aaai.org/ojs/index.php/aimagazine/article/view/1716/1614>>. Acesso em 22 out. 2010.

VELTMAN, Kim H. Towards a Semantic Web for Culture. *Journal of Digital Information*, v.4, n. 4, p. 3-15, 2004. Disponível em: <<http://jodi.ecs.soton.ac.uk/Articles/v04/i04/Veltman/>>. Acesso em: 24 nov. 2004.

W3C SEMANTIC WEB ACTIVITY. 2010. Disponível em : <<http://www.w3.org/2001/sw/>>. Acesso em: 18 jan. 2011.

WUSTER, E. L' étude scientifique générale de la terminologie, zone frontalière entre la linguistique, la logique, l' ontologie, L' informatique et les sciences des chose. In: RONDEAU, G.; FELBER, E. (Org.). *Textes choisis de terminologie*. Québec: GIRSTERM, 1981. (Fondements théoriques de la terminologie, v. I). p. 57-114.