

Semantic Web とオントロジーの課題

武田英明

国立情報学研究所

takeda@nii.ac.jp

<http://www.kasm.nii.ac.jp/~takeda/>

Some Issues on Semantic Web and Ontologies

Hideaki Takeda

National Institute of Informatics

Abstract: In this article, we discuss issues on Semantic Web (SW) from a perspective of artificial intelligence. It looks that Semantic Web is technically not new at all. But there are two challenges in SW technically. There is a difference on what kind of information is handled, i.e., the traditional knowledge representation (KR) assumes that information is consistent while SW is not. As a result, SW necessarily needs proof and trust that are not explicit in KR because of consistency of information. The other feature of SW is on the process of knowledge creation. Since Web is naturally interactive with respect to information creation, so does with respect to knowledge creation. It means that human is the loop of knowledge creation. For example, if some people share a URI, it means that they can share the concept indicated by the URI. Human activities can assure knowledge and knowledge sharing.

1. はじめに

本稿では Semantic Web の可能性と課題を人工知能研究の側面から考察する。Semantic Web は人工知能の側面から見ると総じて技術的には常識的（より厳しく言えばありきたり）であるが社会的には大胆な計画であるといえよう。以下ではこの2点について考察してどのような可能性がありまた課題があるかについて述べる。

2. 技術的側面

Semantic Web ではあっと驚く技術が提示されているわけではない。そこにあるのはこれまでの人工知能の研究で行われてきたことが基本となっている。その意味でありきたりといえよう。すなわち Semantic Web が求めているものは、これまで研究されてきた知識表現そのものである。新たにつくらないという意味では手堅い方法なのかもしれないが、人工知能のいつかた道という批判はある。たとえば Semantic Web の源流は Tim Berners-Lee のいう最初の Web の提案¹であるとされているが、そこでのネットワークは意味ネットワークそのものであり

意味ネットワークの抱える問題はそのまま引き継ぐことになる。

すなわち人工知能と同じ道をたどるだけであろうか。筆者は2つの点で新しい方法が提示されていると考えている。

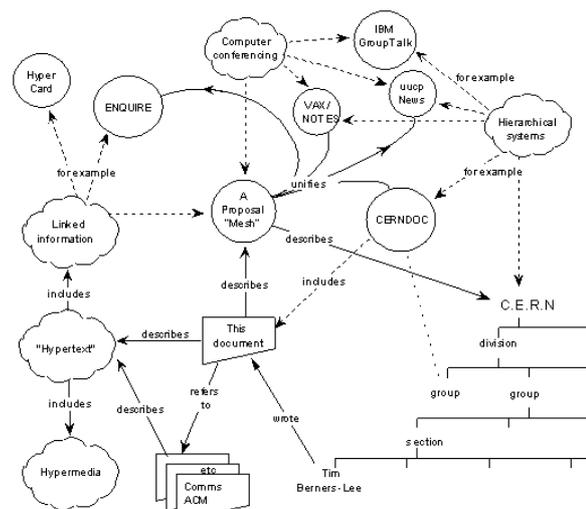


図1 Semantic Web の源流

¹<http://www.w3.org/History/1989/proposal.html>

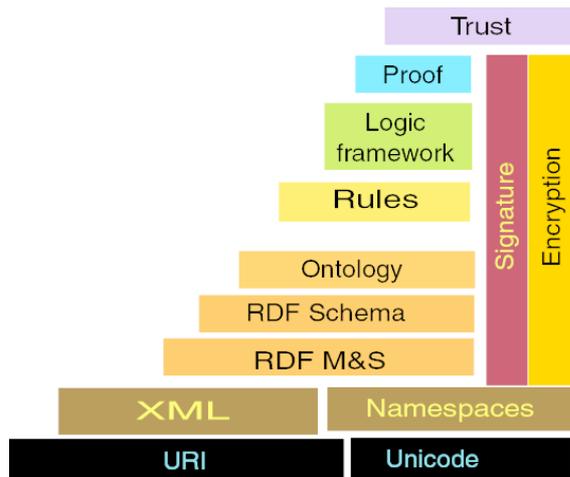


図2 Semantic Web の”layer cake”

2.1 開かれた世界と閉じた世界

人工知能の知識表現は基本的には閉じた世界を前提としていた。すなわち“完全な世界”というものをどっかに想定する(別に本当に書けなくてもよい)ことで、その上での表現能力と推論能力を研究するという立場であった。そこでは完全性とか一貫性が重要なポイントとなる。

ところが Web の世界はそもそも“完全な世界”など想定しうる世界ではない。範囲を定めることもできないし、矛盾する記述もしばしばである。完全性とか一貫性を定義することはできない。このような世界での知識表現を考えるというのが Semantic Web であり、それはまさに新しい挑戦である。

知識表現は Classic 系列の言語から Ontolingua, OKBC そして DAML+OIL と表現の完全さや一貫性重視から共有の可能性といった開かれた問題の対応へと移行している。逆に Web は HTML から XML, RDF, RDFS, DAML+OIL と記述の深さを追及している。

挑戦というだけでは解決策はないのであろうか。

私は Semantic Web は 2 つの点で解決の方向性を示していると考え、その一つは無秩序なデータをどう扱うかという点である。Layer Cake と呼ばれる Semantic Web の階層性である(図 2²参照)。この階層の上位には証明(proof)と信頼(trust)層がある。これは無秩序な世界でのデータの取り扱いには必須であろう。

知識表現ではどうであったか。もちろん、考えていなかったわけではなく信念の論理といった研究はなされてきた。しかし実際的な問題解決にはつながらないように見える。それはなぜか。知識表現とい

う問題を考えるときに意識的あるいは無意識的に整合的な世界を想定している。論理で記述するときには暗黙にそのような前提をおいていることが多い³。論理的に整合的な世界であれば実は証明も計算手段でしかない。また信念の論理でもたとえエージェント間で矛盾があったとしても個々のエージェントは整合的である。すなわち、信頼・証明・知識は整合的な世界では一つに縮約されている(図 3 参照)。したがって信頼の問題も本質的な議論にならなかったと思われる。

現実的なデータを扱うためには Semantic Web の階層で示される証明・信頼が必要であり、知識表現もこのような機能をもつことが要請されているということである。

2.2 Web を通じた Human-in-the-Loop の実現

前項では全般的な問題提起であったが、その解決の道はあるのであろうか。それは WWW のもつインタラクティブ性、すなわちユーザをプロセスに取り込んだ知識表現記述にあると思われる。WWW のもつインタラクティブ性とはユーザが自らオーサとなって記述を追加することで参加できるという仕組みである。データ提供者と利用者の区別は本質的にはなく誰でも参加できる。このことは一見記述の不安定さ、信頼のなさを作り出しているように見えるが、巨視的にみると安定を作り出すことが期待される。Semantic Web でもこの特性を利用している。RDF(S)あるいは DAML+OIL である概念を記述して共有するかについてはどうであろうか。果たして概念の意味を書いているのだろうか。

もちろんそんなことはない。実は概念の実体は URI に過ぎない。もちろん他の概念との関係で定義可能な概念はこれらの言語の記法でその概念と関連付けられるが、それ以外の概念は単に URI が示されるだけである。定義もない概念が共有可能なのだろうか。URI はそれ自身何も意味しない(URI に示されることにはなにもなくてもよい)。しかし、この URI を多くの人があるラベルをもって参照すれば別である。それらの人々は結果的にその概念を共有したことになる。意味の共有そのものはその利用する人たちに任せきりであるが、URI の共有ということで客観的な示すことができる(他のユーザやコンピュータが理解可能)。

これは人間が介在することで可能になったわけである。同様のことが証明や信頼といった層に可能なように思える。たとえば、Web ページの依存関係や量

² Tim Berners-Lee
<http://www.w3.org/2002/Talks/09-lcs-sweb-tbl/>

³ 例えば Hintikka の認識論理などがよい例であろう。

が結果的に証明や信頼を作り出すということも考えられよう。実際 google はそのような手段で “よい” ページを探している。これもあくまで人間がページを作っている人間をループにいれた仕組みだからうまくいくと考えられる⁴。

これはこれまでの知識表現では考えられなかった別の意味での開かれた問題である。

3. 社会的な側面

もうひとつの重要な側面は Semantic Web がどう社会的に受け入れられるかということである。Semantic Web 普及の鍵は以下の 4 つであろう。

1. 仕様の確定と展開(推論など)
[Semantic Web 関係者]
2. 利用者に普及すること
[Semantic Web 関係者, Application 関係者, 一般ユーザ]
3. よいオントロジーの提供
[Semantic Web 関係者, Application 関係者]
4. よいドメインを見つけること
[Application 関係者]

このような “卵と鶏” 的問題は新しいプロトコルを普及させるときには常に起こる問題であるが, Semantic Web の場合は本質的である。というのは, Semantic Web の場合は先に述べたように共有するプロセスが含まれてこそ Semantic Web は機能するからである。“よい” オントロジーとは “正しい” オントロジーではなく, “みんなが共有できる” オントロジーである (もちろん, その上で “正しい” ければなおよいのであるが)。

難しい問題であるが, 新しい有意義な挑戦であると思われる。いわば社会的実験であるといえよう。一箇所で集中的に構築された知識ベース (いわば “実験室” 知識ともいえよう) の成果が必ずしも芳しくないことを考え合わすと, 逆にこのようなボトムアップ的アプローチに期待するというのは楽観的すぎようか。

4. おわりに

本稿では人工知能研究の立場から Semantic Web の課題と可能性について考察した。Web が社会的インフラとなっている現在, 多少なりとも人工知能での知見が生かされて Web が発展すればそれは人工知能の分野としても有意義なことであろう。Semantic

⁴ もちろんすでに google でも問題になっているように意図的な攪乱はありえるだろう (というか不可避か)。このような楽観主義, 性善説的方法には当然是非がある。

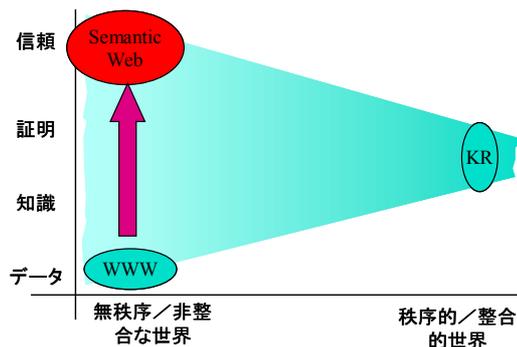


図3 KR と Semantic Web

Web はそのような可能性をもったものだと考えられる。