# Where is the Semantics on the Semantic Web?

## Ontologies and Agents Workshop
## Autonomous Agents

**Montreal, 29 May 2001**

## Mike Uschold

**Mathematics and Computing Technology**

**Boeing   Phantom   Works**

# Acknowledgements

**Material from this lecture was drawn from many fruitful discussions with:**

- Peter Clark
- John Thompson
- Rob Jasper
- Anita Tyler
- Dieter Fensel
- Frank vanHarmlen
- Michael Gruninger

*BOEING*

# The Evolving Web

- **Locating Resources**
  - *free text & keyword* search → *semantic* search

- **Web Users**
  - primarily *humans* → both humans and *machines*

- **Web Tasks & Services**
  - a place to *find* things → a place to *do* things

**Semantics is the Core Requirement**

- web content with *no semantics* → *with semantics*

# Agents and the Semantic Web

- **Semantic Web:** killer 'app' for agents?

- **Agents need to communicate and understand meaning.**
  - Advertise and require capabilities
  - Locate meaningful information resources on web & combine them in meaningful ways to perform tasks
  - How to interpret communication acts?

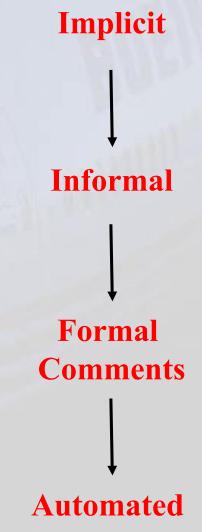- **But what do we mean by the Semantic Web?**

# TBL's Vision

- **Extension of current web;**

- **Layered, extendible, composable;**

- **Meta-data, Ontologies, KBs, Agents, WWKB**
  - Inference, proofs, queries

- **'Semantics' – in machine processible form.**

# What do we mean by 'Semantics'?

- **Semantics of What?**
  - language?, term?, expression?
  - communication protocol?
  - **domain ontology & markup!**

- **Plicity:** Are the semantics **im**plicit or **ex**plicit?

- **Formality:** How are semantics expressed?

- **Semantics Processing:** Who are they for?
  - human only – fully manual
  - human and computer – partially automated
  - computer only – fully automated

**Implicit**

↓

**Informal**

↓

**Formal Comments**

↓

**Automated**

*BOEING*

5

# Examples

- **Implicit:** *based on human consensus, shared understanding*

    - Typical XML tags
        - `<price>       200     </price>`
        - `<address>      …      </address>`
        - `<delivery-date> … </delivery-date>`

    - Used by screen-scrapers, wrappers

    - Rife with ambiguity.

- **Informal:** *only humans can use*     (until NLP solved)

    - Text specification document for HTML e.g. `<h2>`

    - UML semantics document

    - Java language definition, for compiler writers

    - Still ambiguous

*BOEING*

# Examples

- **'Formal Comments'**
    - Semantics of FIPA ACL 'inform' in modal logic
    - Formal definitions in any requirements spec (e.g. Z)
    - Many axioms in Ontolingua ontologies
    - Much less ambiguous
    - Still error-prone, human in the loop.

- **Automated**
    - RDF(S), DAML+OIL term definitions
      e.g. `mammal, date`
    - How does the machine process the semantics?

# Machine Processible Semantics

- **How can an agent learn the meaning of a term?**

- **Procedural Semantics**
  - How does an agent system know what to do when it sees the term '`inform`'
  - The (possibly informal) semantics of '`inform`' is embedded in a procedure by a human.
  - The system places a call to the procedure when it encounters '`inform`'.
  - The 'meaning' of '`inform`' is what happens when this procedure is called.

- **Machine processible semantics? – perhaps.**

# Machine Processible Semantics

- **Learning the meaning of a term from a formal declarative specification of the semantics…**

- **General case: no assumptions, nothing shared**
  - all symbols might as well be in 'Greek' script
  - no knowledge of language syntax, or semantics
  - Cryptography, impossible to automate
  - So, we have to cheat…

- **We must make some assumptions…**

_BOEING_

# Assumptions: language

- **Shared language syntax and semantics,**
  - e.g. KIF, RDF(S), DAML+OIL

- **But: may have incompatible assumptions in conceptualization.**

  - Time point, vs. time interval
  - Agent can **never** incorporate meaning of new term in its axioms.

# More Assumptions: compatibility

- **Logical compatibility as well as language.**

- **But: Different people build different ontologies for the same domain.**

  - Two terms, same meaning, or vica versa;

  - Same concept modeled at different level of detail;

  - Different language primitives used for same concept;
    - e.g. red an attribute, or RedThings a class.

- **Computationally intractable to determine if two terms actually mean the same thing.**

  - I.e. have same set of models

# More Assumptions: sharing

- **Term explicitly mapped to a shared concept**
  - Encounter new term, leprechaun, a subclass of mammal.
  - 'mammal' defined in shared animal ontology in OIL.

- **Machine can learn something about meaning.**
  - I.e. there are now more things that it cannot be.
  - Still plenty of scope for ambiguity;
  - Definition of mammal in OIL can never be complete.

- **Can do some inference**
  - e.g. for search application looking for content about mammals.

*BOEING*

# Processing Semantics

- **Relies on a formal semantics of OIL to infer semantics of terms and expressions in OIL.**

- **OIL semantics is for humans**
  - it helps build inference engines;
  - not machine processible.

- **Humans may still embed some meaning in code**
  - May be dangerous to do so – or –
  - May be necessary to do so…

- **The shared concept referred to may not be formally defined (e.g. Dublin Core terms)**

*BOEING*

# Enter:
Opinion and Speculation Mode

# When is Semantic Web Needed?

- **Good Question!** <span style="color:red">**Where are the use cases?**</span>

- **No case made for search, at least not for humans. Google works brilliantly!**

- **Build it and they will come! Or will they?**

- <span style="color:red">**Analogy:**</span> **So what if my toaster can talk to my washing machine!**
  - What would they say?
  - Does this improve my life?

# Law of the Semantic Web?

*The more agreement there is, the less it is necessary to have "machine sensible semantics".*

- **E.g. <h2> in HTML specification;**

- **No need to do inference;**

- **Just embed the semantics in the browsers.**

# Two Show Stoppers

- **Mapping**
  - There will never be global standards
  - Mapping will always be necessary
  - Hard to automate
  - Time-consuming to do manually

- **Markup**
  - Noone will do this unless it is painless.
  - Can't get anywhere without it.

*BOEING*

# How to Cope?

- **Mapping**
  - Get agreement where possible, standards in limited communities and scope;
  - Create mappings as necessary;
  - Do lots of research!

- **Markup**
  - Many good statistical techniques from IR
    - Limited to putting things in buckets, not fine grained semantic markup
  - Markup for 'free' – ala Hendler's recent paper "Agents on the Semantic Web" (or similar)

# Summary:
# Where IS the semantics?

- **Often just in the human.**

- **Informally in specification documents.**

- **Embedded in implemented code.**

- **Formal Comments to help humans understand and/or write code.**

- **Formally encoded for machine processing**

- **In the representation language specification**

# Summary:
# Characterizing the Semantic Web

- **Purpose, Benefits, Mechanisms of semantics**
  - **Needs a lot more work!**

- **What has the semantics?**
  - Language? Terms? Communication protocols?

- **Representing and Processing semantics**
  - Implicit or Explicit?
  - Formal or Informal?
  - For human or for computer?

- **Agreement and Sharing of semantics**
  - Does agreement reduce need for explicit semantics?

*BOEING*