

# SemML: Creating a Standardized Markup Language for Semantic Networks

Brian Harrington and Pia-Ramona Wojtinek  
Department of Computer Science  
University of Oxford

20 September 2011



## HOW STANDARDS PROLIFERATE: (SEE: A/C CHARGERS, CHARACTER ENCODINGS, INSTANT MESSAGING, ETC.)



image courtesy xkcd comics (<http://www.xkcd.com/927/>)

# Why SemML?

- A standard for reading/writing/sharing semantic networks
- A standard geared towards the Semantic Computing community
- A standard that does what we need it to do

# This Presentation

- SemML is a work in progress
- v0.1 - Does what WE want
- v0.2 and beyond - Make it do what YOU want

# This Presentation

- SemML is a work in progress
- v0.1 - Does what WE want
- v0.2 and beyond - Make it do what YOU want
  - But we need to know what you want it to do

# Semantic Relatedness: A Grand Challenge for Semantic Computing?

- Semantic Relatedness - A standard task in NLP/AI
- How *related* are word pairs
  
- Comparison with human judgements
- Traditionally: Computers very good at similarity, humans very good at relatedness

# Semantic Relatedness: A Grand Challenge for Semantic Computing?

- Semantic Relatedness - A standard task in NLP/AI
- How *related* are word pairs
  - Baker - Lawyer: More *similar*
  - Baker - Bread: More *related*
- Comparison with human judgements
- Traditionally: Computers very good at similarity, humans very good at relatedness

- Extending Semantic Relatedness beyond words



Article Discussion

## North Korea – Russia relations

From Wikipedia, the free encyclopedia

**North Korea and Russia** established diplomatic relations on October 12, 1949 proclaimed. Close allies during the Cold War, the relations between them cool some importance again during the 2000s.

Russia has sunk considerable amounts of capital into numerous large-scale, I peninsula, such as oil and gas pipelines and Trans-Korean and Trans-Siberian economic revitalization of the Russian Far East, and in the case of a new Kai be severely damaged. The two states share a border along the lower Tuman I 1880 when the Tsar bought lands from China.

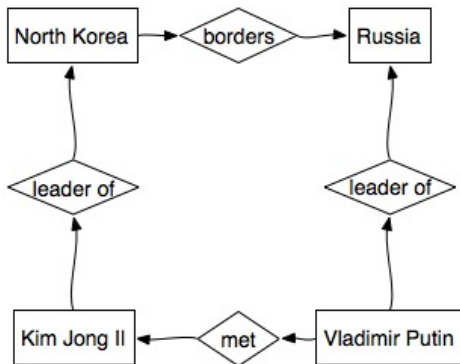
**Contents** [hide]

- 1 History
- 1.1 Soviet Union
- 1.2 1948-1953





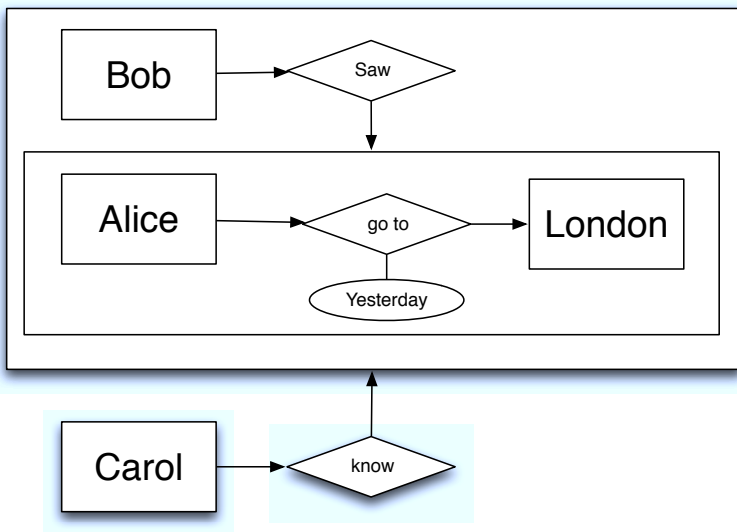
- Extending Semantic Relatedness beyond words
  - Requires a Semantic Interlingua



# SemML

- XML based standard
- Encodes features of semantic network
- Easy to translate between network types

# SemML: Nested Recursive Structure



# SemML: Nested Recursive Structure

```

<schema xmlns="http://www.w3.org/2001/XMLSchema"
  targetNamespace="http://www.isc-home.org/SemML"
  xmlns:netml="http://www.isc-home.org/SemML">

  <element name="Example">
    <primitive label="Alice" type="Person" />
    <primitive label="Bob" type="Person" />
    <primitive label="Carol" type="Person" />
    <primitive label="London" type="Location" uri="http://en.wikipedia.org/wiki/London" />
    <concept label="Alice-goto-London">
      <relation label="goto1" type="action" temporal="past" strength="1.0">
        <text>go to</text>
        <object>Alice</object>
        <subject>London</subject>
        <source>example1</source>
        <attribute label="yesterday1" type="temporal">Yesterday</attribute>
      </relation>
    </concept>
    <concept label="Bob-see-X">
      <relation label="see1" type="action" temporal="past" strength="1.0" continuity="temporary">
        <text>see</text>
        <object>Bob</object>
        <subject>Alice-goto-London</subject>
        <source>example1</source>
      </relation>
    </concept>
    <concept label="Carol-know-X">
      <relation label="know1" type="action" temporal="current" strength="1.0" continuity="indefinite">
        <text>know</text>
        <object>Carol</object>
        <subject>Bob-see-X</subject>
        <source>example1</source>
      </relation>
    </concept>
  </element>

```

# SemML: Relational Strengths

- Can represent *saliency* or *confidence*
- Allow filtering of extraneous or tangential information from the network

# SemML: Temporality

- Each relation has:
  - Temporality: past, present or future
  - Continuity: incidental, temporary or permanent

# SemML: Resource Integration

- Nodes (atomic or complex) can have associated `uri`, which can be linked to Wikipedia, DBPedia, WordNet or other structured resource

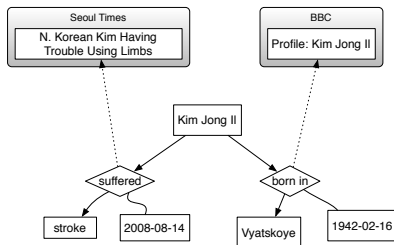
# SemML: Source Management

- Each relation has an associated *source* (original document from which fact was derived).
  - sources have corresponding *trust* score
  - can adjust the strength of relations, or even turn off relations from a given source

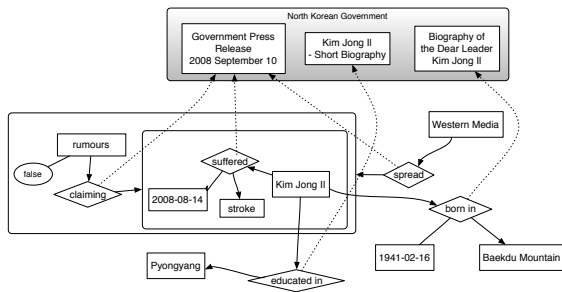




# SemML: Managing World Views



# SemML: Managing World Views



# SemML: What's Next?

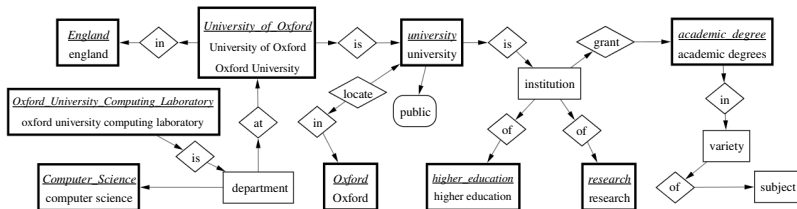
- Standardize feature set for 0.1
- Create translation systems for existing Semantic Software
- Create filters to import resources





The **Oxford University Computing Laboratory** is the Computer Science department at Oxford University in England.

The **University of Oxford** is a public university located in Oxford.



# SemML: What's Else?

- We need your help
  - What features do you need/want
  - What resources would you like to see converted
  - Do you have anything to share in SemML format?
- SemML belongs to the Semantic Computing Community
- What does the community need?