Technology transfer from bioinformatics?

Speechacts in ontology-based IS

Fabian Neuhaus[1] Bill Andersen[2]

[1]National Institute of Standards and Technology, Gaithersburg, MD

[2]Ontology Works Inc., Balitmore, MD

Dec 3, 2008

OIC call for paper

'New approaches are required to enable greater flexibility, precision, timeliness and automation of analysis in response to rapidly evolving threats. Ontology-based technology as applied in the areas such as bioinformatics has demonstrated the possibility of gains along all of these dimensions. The time is ripe to extend these gains to other spheres.'

Why is it so hard to repeat the success of ontology based technologies in biology in other domains?

Kinds of knowledge

Biology:
Scientific knowledge: more or less general laws
No interest in individuals, except as evidence for or against laws

Intelligence community:
Intelligence community keeps track of individual persons of interest
Much of the knowledge can not be understood as laws

Reasoning with instances

Biology:
Only interested in type-level reasoning (TBox reasoning)

Intelligence community:
Need to reason with instances (ABox reasoning)

- OBO Format: instance reasoning impossible
- OWL DL: instance reasoning still too slow

Time

Biology:
Laws are timeless

Intelligence community:
Knowledge is time-relative
(i: "Bill is in Virginia on 12/03/08")
(ii: "All members of OWI are in Baltimore on 12/06/08")

- OBO Format: impossible
- OWL DL: possible, but rather convoluted

Unreliable Sources

Biology:
Ontologies represent the scientific consensus
Represented knowledge is consistent.
Ontologies are taken by the user as whole as true.

Intelligence community:
Information come from unreliable sources.
Available information might be inconsistent.

- Inconsistencies render classical reasoning useless (that includes OWL DL reasoners)

Independent assertions of the same fact

Assume that source A and source B are confirming independently fact X.

Biology:
The knowledge base represents X. (References to A and B might be included but are not accessible to reasoners.)

Intelligence community:
Need to represent that X has been verified independently by A and B. Need for reasoning with metadata.

- Neither OBO Format nor OWL DL can handle statements of the form "Source A asserts that X".

Security

Biology:
The ontologies represent the scientific consensus, ontologies are open to everybody.

Intelligence community:
Need for multi-level security access control.

Conclusion I

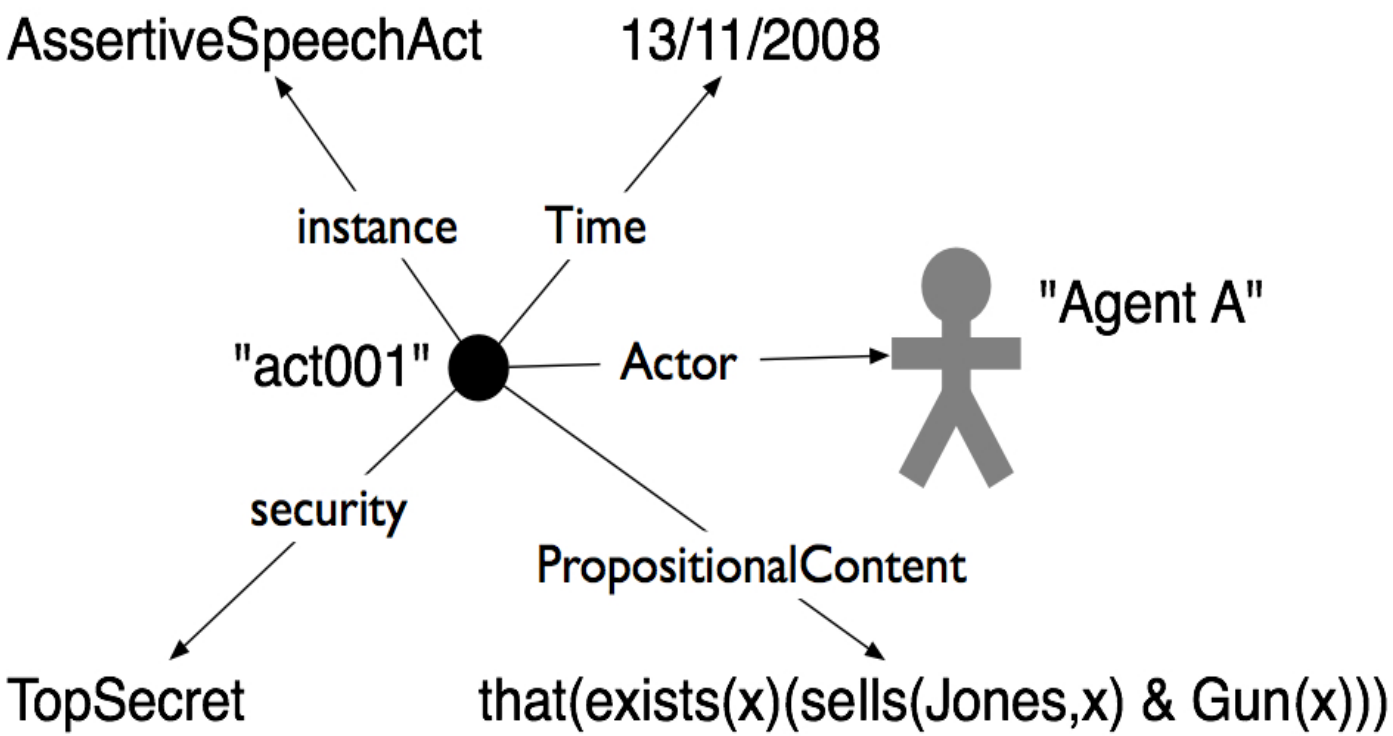
What can the intelligence community adapt from biologists?

Not much.

Goals

- Distinguishing assertions with the same propositional content
- Reasoning with knowledge from potentially unreliable sources
- Enabling security access control

Speech act analysis: example



Advantages

Different speech acts can have the same propositional content

Inconsistencies blocked since they are embedded in the propositional content

The support relationship (simplified)

A assertive speech act supports the assertion of it's propositional content.

If $S_1, \ldots, S_n$ are speech acts, and S_1 supports A_1, S_2 supports A_2, etc., then the sequence $(S_1, \ldots, S_n)$ supports the the proposition that $(A_1 \wedge \ldots \wedge A_n)$.

If a sequence of speech acts $(S_1, \ldots, S_n)$ supports the proposition that A, and A is logically consistent, and A logically entails B, then $(S_1, \ldots, S_n)$ supports the proposition that B.

Typical query

A user with no security clearance asks: 'Does Jones sell weapons?'

The system understands: 'Find all (minimal) sets of unclassified speech acts that support that Jones sells weapons.'

Reasoning in this framework

Because of how the "supports" relationship is defined, inconsistent sequences of speech acts do not support anything. (Support is a 'relevant' relationship between speech acts and propositional content.)

Security restrictions are handled by the reasoner in the same framework as the other kind of reasoning, they translate into conditions that speech acts have to meet.

contact: Fabian Neuhaus, fneuhaus@nist.gov