

The SILK 2 System and RIF Dialect: Semantic Rules Grow Up

Benjamin Grosf*

June 24, 2010

30-min. presentation at SemTech 2010**

* Vulcan Inc., benjaming@vulcan.com

<http://www.mit.edu/~bgrosof/>

For More Info: <http://silk.semwebcentral.org>

<http://projecthalo.com>

** 2010 Semantic Technology Conference, San Francisco, CA, USA

<http://www.semanticconference.com>



© This presentation is copyright Benjamin Grosf and Vulcan Inc.
All rights reserved. (Limited copyright granted to Semantic Universe.)



Outline and Overview

- **SILK research program within Vulcan's Project Halo**
 - Language and system for advanced knowledge representation (KR)
 - Digital Aristotle vision: question-answering for science
 - Scalability for social structured knowledge: entry, reuse, querying
- **SILK Language, Hyper Logic Programs KR, and RIF-SILK**
 - Expressive features: defaults, actions, higher-order, frames, webized, interchange
 - Advanced defaults: prioritized conflict handling, argumentation, hyper rules, tractability
- **SILK System**
 - Reasoning: layered architecture, transformations, tabling
 - Knowledge Acquisition (KA) and UI: interchange; editing, explanation
- **Potential application areas in business and government**
 - Horizontal: policies, workflows; ontology mapping, knowledge integration
 - Vertical: e-commerce, defense intelligence, trust, biomed, financial, mobile
- **Demo of default hyper rules in SILK GUI: edit, query, explain**
- **Conclusions**
 - Higher-abstraction KR closer to human cognition and social pragmatics
 - Radically extends expressive power of SQL, RDF(S), SPARQL, OWL-RL, RIF-BLD
 - Remedies major limitations of semantic web's current KR foundation

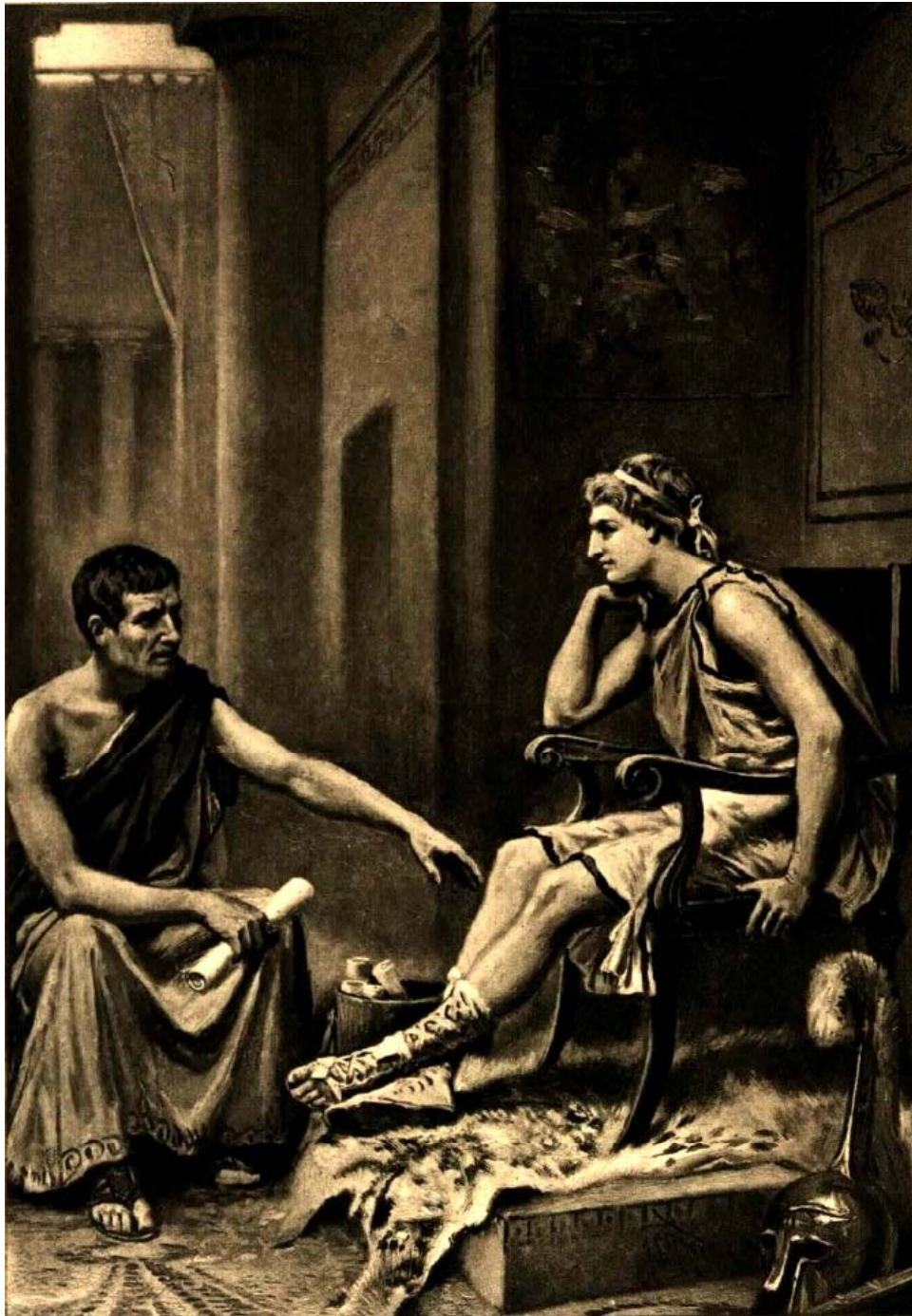
Vulcan's Project Halo

- Vision of **Digital Aristotle**: question-answering for science
 - Put the bulk of the world's scientific and similar knowledge on-line
 - Answer questions, act as personal tutor, with deep reasoning
- College-level **science** selected as initial domain focus
 - Good metrics available: textbook-type exam Q's. Initial domain task focus is:
 - **Advanced Placement Exam (AP)** in Biology, Physics, and Chemistry
 - Taken by USA high-school students to get credit for 1st-year college courses
- **AURA AI expert system developed (2004-)**
 - Controlled Natural Language, GUI, Frame-based KR, Problem-Solving
 - Students as users – formulate questions, formulate knowledge
- **Semantic MediaWiki+ developed (2007-)**
 - Leading semantic wiki. Open-source. Simple rules, light ontologies.
- **SILK developed (2008-)**
 - Largest* rule research program in USA. Multi-institutional: primarily via contractors.
 - A knowledge representation (KR) system: reasoner, language, UI, interchange.



* (that we're aware of)





Aristotle Tutoring Alexander

Image in public domain (copyright has expired), downloaded from
http://commons.wikimedia.org/wiki/Image:Aristotle_tutoring_Alexander_by_J_L_G_Ferris_1895.jpg

SILK Contributors current/past (partial list)

- Vulcan (Benjamin Grosf, Mark Greaves, Dave Gunning)
- Stony Brook University (Michael Kifer; students H. Wan, S. Liang, P. Fodor)
- Raytheon BBN Technologies (Mike Dean, Mark Burstein, C. Andersen, B. Ferguson)
- SRI International (Vinay Chaudhri, David Martin, Ken Murray)
- Cycorp (Keith Goolsbey, Doug Lenat, Jon Curtis, Ben Rode)
- Automata (Paul Haley)
- Boeing (Peter Clark)
- Ontoprise GmbH (Daniel Hansch, Jurgen Angele)
- Terrance Swift, consultant
- Richard Fikes, consultant (Stanford University)
- University of Toronto (Sheila McIlraith, S. Sohrabi, H. Ghaderi)
- Texas Tech University (Michael Gelfond, D. Inclezan)
- University of Texas (Bruce Porter, Ken Barker)
- University of Amsterdam (Bert Bredeweg)
- University of Freiburg (Georg Lausen)
- University of Michigan (Michael Wellman)
- Raphael Volz, consultant



SILK's Goals

- Address fundamental requirements for scaling Semantic Web to widely-authored Very Large KBs in business and science that answer questions, proactively supply info, and reason powerfully
- Expressiveness + Semantics + Scalability
 - Push the frontier. Language and system.
- **Better Knowledge Representation (KR)**
 - Expressive power: defeasibility, higher-order. E.g., causal processes in AP Biology.
 - Performance scalability of reasoning, including knowledge updates
- **More effective Knowledge Acquisition (KA)**
 - + By Subject Matter Experts (SMEs), not programmers or knowledge engineers
 - + Collaboratively – incorporate large #s of SMEs in KB construction & maintenance
 - + Leveraging the Web
- **Better KR also for sake of better KA**
 - Web knowledge interchange (with merging) for scalability of collaborative KA
 - The underlying KR is the target for KA: **“The KR is the deep UI”**
 - Understandability via semantics and expressiveness



Raise abstraction level closer to the user's natural language and cognition



Expressiveness “Brittleness” Areas Targeted

- **Defaults/Exceptions/Defeasible** (*incl. nonmonotonic reasoning, theory revision, argumentation, truth maintenance*)
 - A kinematics problem situation has standard earth gravity, and no air resistance. [physics AP]
 - A given organism has the anatomy/behavior that is typical/normal for its species, e.g., a bat has 2 wings and flies. [bio AP]
 - Price info for an airplane ticket on Alaska Air’s website is accurate and up to date. [e-shopping]
 - ❖ **Practical reasoning almost always involves a potential for exceptions**
- **Hypotheticals**
 - If Apollo astronaut Joe golfed a ball on the moon, then standard earth gravity would not apply. [negative hypothetical] [*conflict* between defaults, resolved by *priority* among them]
 - If I had swerved my car 5 seconds later than I did, I would have hit the debris in the left lane with my tire. [*counterfactual*]
- **Actions and Causality**
 - If a doorkey is incompletely inserted into the keyhole, turning the key will fail. [*precondition*]
 - During the mitotic stage of prometaphase, a cell’s nuclear envelope fragments [biology AP]
 - After a customer submits an order on the website, Amazon will email a confirmation and ship the item. [Event-Condition-Action (*ECA*) rule] [policy]
- **Processes (i.e., representing and reasoning about processes)**
 - Mitosis has five stages; its successful completion results in two cells. [compose] [partial description]
 - If Amazon learns that it will take an unexpectedly long time to stock an ordered item, then it emails the customer and offers to cancel the order without penalty. [exception handling]
 - A Stillco sensor-based negative feedback thermal regulator is adequate to ensure the overnight vat fermentation of the apple mash will proceed within desired bounds of the alcohol concentration parameter. [science-based business process]

Ubiquitous in science, commonsense, business, etc. All are interrelated.

Declarative Logic Programs (LP) is the Core KR in today's world ... including the Semantic Web

- **LP is the core KR of structured knowledge management today**
 - Databases
 - Relational, semi-structured, RDF, XML, object-oriented
 - SQL, SPARQL, XQuery
 - Each fact, query, and view is essentially a rule
 - Semantic Rules
 - Rule Interchange Format (RIF): -BLD, -Core
 - RuleML standards design, including SWRL
 - Semantic Ontologies
 - RDF(S)
 - OWL-RL (= the Rules subset). E.g., Oracle's implementation of OWL.
- **The Semantic Web today is mainly based on LP KR**
 - ... and thus essentially equivalent to semantic rules
 - **You might not have realized that!**



SILK's new KR: *Hyper* Logic Programs

- Extended LP that is the first to combine key advanced features
- Defaults + Higher-Order + External Actions/Events/Queries
- + Webized, Frames, (clean) Negation and NAF, Equality, Functions, Skolems, Aggregates, Integrity Constraints, Lloyd-Topor, ...
- Tightly interoperates with very broad case of first order logic (FOL), too
 - Any clause, not just Horn. Sound from FOL viewpoint.
- Transforms knowledge from higher to lower abstraction levels
 - Higher is good for KA; lower is good for reasoning (code reuse, optimization)
- Tractable computationally – complexity is same as Horn LP
 - Polynomial time -- similar to relational DBMS -- if there's no recursion thru functions
 - Retains pragmatic quality of LP: "intuitionistic" – lack general "reasoning by cases"
- New approach to defaults
 - *Argumentation theory*: ~20 "meta-" rules specify debate principles for defeat
 - Much more expressive: higher-order, equality, ...
 - Much easier to implement: ~20 rules instead of 1000's of lines of code
 - Much more efficient: eliminates expensive knowledge recompilation step
 - *Hyper rules*: omni-directional clauses, prioritized handling of multi-way conflicts

SILK's KR Approach, continued

- **KR Language**
 - Syntax: ASCII presentation syntax, abstract syntax, RIF dialect (RIF-SILK)
 - Semantics: model theory, proof theory. Closely related to the transformations (above).
- **Knowledge Interchange**
 - Via load, or query, or event. E.g., embed a SPARQL query in the body of a rule.
 - KR languages: SPARQL, RDF(S), SQL, ODBC; SILK, RIF, OWL(-RL), Cyc, AURA
- **Reasoning system**
 - Backward inferencing primarily -- i.e., query answering
 - Tabling saves and reuses computation from previous subqueries
 - Supports fast updating and forward inferencing
 - Good efficiency/scalability of performance
- **Synergizes 20 years of LP research progress**
 - Courteous defaults and external actions/queries cf. IBM Common Rules, SweetRules
 - Higher-order cf. HiLog, Common Logic
 - Negation-As-Failure cf. well founded
 - Performance optimizations from DBMS, Prolog, BRMS, AI
- **Extensive requirements analysis, use cases, benchmarking**

Representational Uses for Defaults and Higher-Order

Defaults (cf. Courteous, with Prioritization)

- Negation
- Pragmatic knowledge/reasoning has potential for exceptions and revision
 - Learning and science: may falsify previous hypotheses after observation or communication
- Debate and trust: priorities from authority, reliability, recency
- Updating, merging, change: increase modularity/reuse in KA/KB lifecycle
- Process causality: persistence, indirect ramified effects, interference
- Hypotheticals, e.g., counterfactuals
- Inheritance: more-specific case overrides more-general case
- Policies, regulations, laws – the backbone of society and institutions
- Natural language understanding (NLU) aspects: e.g., co-reference

Higher-Order (cf. Hilog and reification)

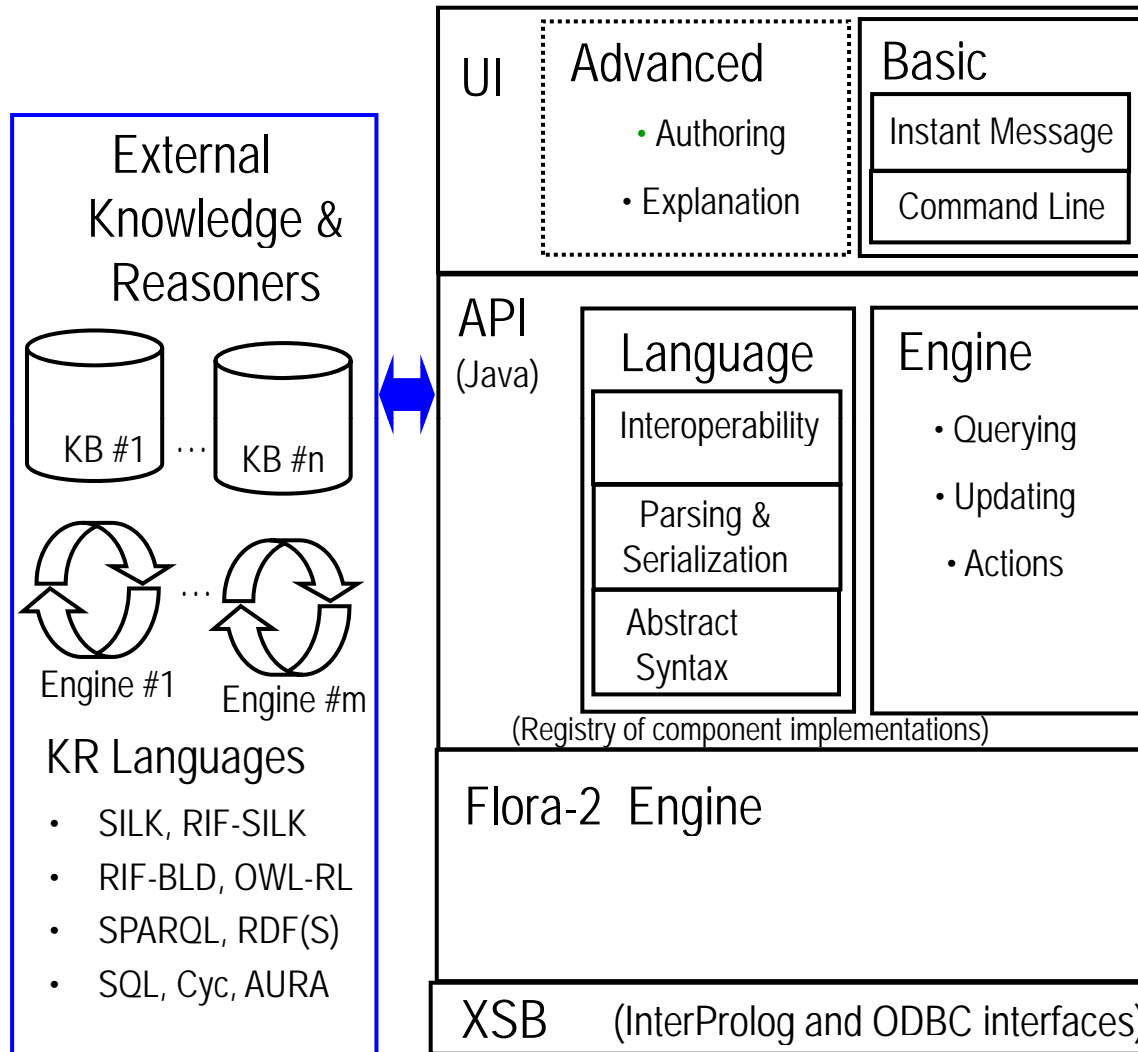
- Meta- knowledge and meta- reasoning, generally
- Ontology mapping, KB translation, KR macros, reflection, NLU aspects
- Provenance, multi-agent belief, modals, many aspects of context

RIF-SILK Dialect

- **It's expressively powerful RIF** (RIF = W3C Rule Interchange Format standard)
 - New dialect defined using RIF's Framework for Logic Dialects (FLD)
 - Extends (supersedes) RIF-BLD (Basic Logic Dialect) and RIF-Core
 - These are based essentially on Horn LP
 - Notably: adds defaults and external actions (side-effectful)
 - Needed for most of today's business applications of (non-semantic) rules
 - Retains "Grade AAA" semantics – model-theoretic
 - Retains computational scalability of Horn LP
- **Status**
 - Draft specification – public (initial version 12/2009, current 2/2010)
 - Semantics section is in progress (summarizes previous theory papers)
 - Implemented translator (bidirectional) is in current SILK system
 - Under discussion with W3C: role in next steps of RIF overall

-
- **SHOW briefly: RIF-SILK Dialect specification document**

SILK Architecture today (V2.1+)



- API Functionality
 - Higher-order defaults reasoning, combines many other advanced KR features
 - SILK and external KR language support integrated tightly with reasoning engine
- UI Functionality
 - Graphical, tabular
 - For Knowledge Engineers
- *Future Items*
 - *UI: SME-friendlier, English (NL)*
 - *KR: probabilistic, parallelization, more interchange KRs*
- Test Sets Focus
 - Defaults, Process
 - AP esp. Biology

Ecology Ex. of Causal Process Reasoning in SILK

```
/* Toxic discharge into a river causes fish die-off. */
/* Initial facts, and a constraint that fish count is unique */
@[strict] occupies(trout,Squamish);
@[strict] fishCount(0,Squamish,trout,400); // first argument is an integer time
@[strict] neg (fishCount(?s,?r,?f,?C1) and fishCount(?s,?r,?f,?C2)) <== ?C1 != ?C2;
/* Action/event description that specifies causal change, i.e., effect on next state */
@[tag->tdf1] fishCount(?s+1,?r,?f,0)
                <== occurs(?s,toxicDischarge,?r) and occupies(?f,?r);
/* Persistence ("frame") axiom */
@[tag->pef1] fishCount(?s+1,?r,?f,?p) <== fishCount(?s,?r,?f,?p);
/* Action effect axiom has higher priority than persistence axiom */
@[strict] silk:overrides(tdf1,pef1).
/* An action instance occurs */
@[id->UhOh, strict] occurs(s1,toxicDischarge,Squamish);
```

As desired: |= fishCount(1,Squamish,trout,400)

|= fishCount(2,Squamish,trout,0)



In Frame syntax: subject[property -> object] stands for property(subject,object).



E-Commerce Ex. of Causal Process Reasoning

```
/* E-commerce delivery logistics. */
```

```
/* Initial fact, and constraint that location is unique */
```

```
@[strict] loc(0,PlasmaTV46,WH_LasVegasNV); // first argument is an integer time
```

```
@[strict] neg(loc(?s,?item,?posn1) and loc(?s,?item,?posn2)) <== ?posn1 != ?posn2;
```

```
/* Action/event description that specifies causal change, i.e., effect on next state */
```

```
@[tag->mov] loc(?s+1,?item,?addr) and neg loc(?s+1,?item,?warehouse)
```

```
<== shipment(?s,?item,?warehouse,?addr) and loc(?s,?item,?warehouse);
```

```
/* Persistence ("frame") axioms about location */
```

```
@[tag->peLoc] loc(?s+1,?item,?posn) <== loc(?s,?item,?posn);
```

```
@[tag->peLoc] neg loc(?s+1,?item,?posn) <== neg loc(?s,?item,?posn);
```

```
/* Action effect axiom has higher priority than the persistence axioms */
```

```
silk:overrides(mov,peLoc)
```

```
/* An action instance occurs */
```

```
@[id->deliv57, strict] shipment(1, PlasmaTV46, WH_LasVegasNV, Nine_Fog_St_SeattleWA);
```

As desired: |= loc(2, PlasmaTV46, Nine_Fog_St_SeattleWA);

|= neg loc(2, PlasmaTV46, WH_LasVegasNV);



In Frame syntax: subject[property -> object] stands for property(subject,object).



Complex AP Biology Examples

- Causal process reasoning is a large portion of AP Biology, often requiring multi-step causal chains and/or multiple grain sizes of description to answer a question.
- Several such complex examples drawn from exams or textbooks have been successfully represented in SILK. E.g.:
 - "A researcher treats cells with a chemical that prevents DNA synthesis from starting. This treatment traps the cells in which part of the cell cycle?"
The correct answer is: G1 [which is a sub-phase of interphase]
 - "In some organisms, mitosis occurs without cytokinesis occurring. This will result in:
 - a. cells with more than one nucleus
 - b. cells that are unusually small.
 - c. cells lacking nuclei.
 - d. destruction of chromosomes.
 - e. cell cycles lacking an S phase."The correct answer is: a. [two nuclei form in a cell, but no new cell wall splits the cell]
 - "Suppose the typical number of chromosomes in a human liver cell was 12. [Notice this is counterfactual; there are actually 46]. What would the typical number of chromosomes in a human sperm cell be?"
The correct answer is: 6 [half of the number in the liver and most other organs]

Potential Applications in Business and Government

- **Horizontal**

- Policies and policy-based workflows
 - Monitor, report, react, handle exceptions, execute, enforce, customize
 - Trust: confidentiality, authorization, compliance, governance
- Ontology mapping/mediation and knowledge integration
 - Perspective: the mappings themselves constitute ontological knowledge. E.g., a dictionary.

- **Vertical**

- E-commerce: shopping & advertising, contracts, customer care, catalogs
- Defense: intelligence, operations
- Financial: reporting, regulatory compliance
- Biomed: pharma, e-science, clinical records and guidance, insurance
- Mobile: personalize communication

- **Many use cases in RIF, RuleML, SWSL documents & prototypes**

- E.g., employ defaults or other features not yet well supported commercially

DEMOS

- SHOW DEMO of hyper rules in SILK GUI: edit, query, explain
- Also: Demo'd at ISWC-2009 and RuleML-2009 conferences
 - Scenario of environmental watchdog group's monitoring workflow
 - Recognize toxic discharge into Ohio River watershed from sharp decline in fish count
 - Alert news media, government agencies, citizens social network
 - Reactive: standing queries trigger external actions upon update events
 - Load imported RDF(S) and RIF-BLD
 - Externally query SPARQL, and Excel via ODBC
 - *This demo won an award at RuleML-2009, essentially for best system*
- Aim to make videos of both these demos and post on SILK website

Remedying FOL Semantics' Lack of Scalability

- Hyper LP handles conflict robustly – get consistent conclusions

- Whereas FOL is a “Bubble” – it’s perfectly brittle semantically in face of contradictions from quality problems or merging conflicts.

- Any contradiction is totally contagious – the conclusions all become garbage

E.g., OWL beyond the RL subset suffers this problem. So does Common Logic. (Technically, RIF-BLD and RDF(S) are defined via FOL semantics too, although their typical implementations are essentially LP.)

A KB with a million or billion axioms formed by merging from multiple Web sources, is unlikely to have zero KB/KA conflicts from:

- Human knowledge entry/editing
- Implicit context, cross-source ontology interpretation
- Updating cross-source
- Source trustworthiness

- *Hyper LP’s approach provides a critical advantage for KB scalability*

- semantically, as well as computationally

FOL: A Bubble

Extreme sensitivity to conflict limits its scalability in # of axioms and # of merges



Left:

<http://www.dailymail.co.uk/sciencetech/article-1199149/Super-slow-motion-pictures-soap-bubble-bursting-stunning-detail.html>

Above:

http://img.dailymail.co.uk/i/pix/2007/11_03/BubblePA_468x585.jpg

Conclusions

- **Radically extends the KR power of W3C OWL, SPARQL, RIF-BLD – and of SQL**
 - Defaults and robust conflict handling – *cope with knowledge quality and context*
 - Higher-order and flexible meta-reasoning – *elevate meta-data to meta-knowledge*
 - Actions and events, cf. production rules and process models – *activate knowledge*
- **Raises the KR abstraction level for business users (SMEs) and NL KA/UI**
- **Use cases in biomed, business policies, ontology mapping, e-commerce, ...**
- **Redefining the KR playing field for Semantic Web, business rules, and rule-based process management**
 - Defaults and Higher-Order – yet retain computational web scalability
 - Escape from FOL Bubble– yet retain grade-AAA model-theoretic semantics
- **Motto: “Transforming Knowledge”**
 - Composes a set of KR transformations for ...
 - Expressive extensions – language and semantics
 - Translations between KR/syntaxes, for interchange
 - Reuse of previous algorithms and implementations



<http://silk.semwebcentral.org>



Acknowledgements

- **SILK contributors**
 - (previously listed)
- **Other contributors to several key previous KR efforts**
 - RuleML and SWSL (Semantic Web Services Language) standards designs
 - Notably: Harold Boley and Said Tabet
 - SweetRules and Flora-2 systems
 - Notably: Guizhen Yang

SILK – Transforming Knowledge

Thank You



Disclaimer: The preceding slides represent the views of the author only.
All brands, logos and products are trademarks or registered trademarks of their respective companies.



© This presentation is copyright Benjamin Grosf and Vulcan Inc.
All rights reserved.

