Rulelog as Theoretical Foundation for Universal Health Exchange Language†

Benjamin Grosof*

June 25, 2013

Universal Health Exchange Language Workshop‡

Webconference session


† Work partly supported by Vulcan, Inc., http://www.vulcan.com

‡ http://www.new-health-project.net/2013/06/12/universal-health-exchange-language-workshop-encinitas-june-25-26-2013/
Goal: Model the relevant info

- **Preamble**
  - *Focus here on theoretical foundations*
  - *For health info, clinical decision support, research*
  - *Bootstrap existing/legacy tech*

- **Data (e.g., triples or cells in a database)**
- **Schemas & other ontologies: vocabulary and definitional info**
- **Mappings between ontologies/schemas, e.g.**
  - Info integration, interoperability
- **Policies (in broad sense), e.g.**
  - Info access authorization, incl. privacy/confidentiality/security
  - Treatment protocols
  - Regulations/law, contracts, reporting: e.g., around insurance
- **Causality, e.g., process descriptions**
- **HCI, esp. knowledge acquisition (KA) from English/natural-language**
  - Science
  - Policies
Logical Representations Available

• There’s one main family of logical syntax available: higher-order logic (HOL) formulas (we’ll call this “rich” knowledge)

• Two main semantics available, each with a set of reasoning techniques
  • Note: Both use hilog FOL, not “real” higher-order, for practicality. It transforms into FOL.
    1. LP-based (declarative logic programs, with well founded semantics) – “the 99%”
      • Core of world’s structured info management, incl. databases, business rules, SPARQL, OWL RL
    2. FOL-based – “the 1%”. Designed for math. Used in verification.

• Rulelog – the current acme of (1.)
  • RIF-Rulelog is in draft as industry standard submission by RuleML to W3C and Oasis
  • Supersumes Datalog, SQL, SPARQL, Xquery, semantic production/ECA rules, OWL RL

• Common Logic (CL) – the current acme of (2.)
  • ISO standard. Partial implementations avail.
  • Supersumes RDF(S), OWL 2, FOL.
Textual Logic

- Logic-based text interpretation and text generation (i.e., NLP)
- On top of HOL formulas; text generation targets HOL
- Key idea: **textual terminology**
  - NL phrase corresponds 1-1 to logical term
  - Ontology thus emerges automatically, essentially
- Key idea: rapid interactive disambiguation
  - Of: parse, quantifiers, coreference, word-sense
  - An English sentence is interpreted into a single HOL formula, e.g., in Rulelog
- Enables cost-effective deep-capture KA from English into rich logic
  - Cf. recent Vulcan pilot study. E.g., ~~$1000. per page of text document.
More about Key Expressive Features

- **The ff. are all about meta knowledge**

- **Hidlog: rule id’s + hilog** (relies on logical functions; includes reification)
  - For mappings, e.g., ontological, knowledge interchange, NLP
  - For provenance, incl. KB mangement, versioning, collaboration support
  - For modalities, e.g., “should” (deontic) in policies, “can” (alethic) in causality
  - As foundation for defeasibility

- **Defeasibility (a.k.a., exceptions; the core of logical nonmonotonicity)**
  - For policies, causality, NLP
  - Represents change in the world and change in knowledge (K). K is empirical, evolves.
    - Key for social scalability

- **Restraint – semantically represent pragmatic incompleteness**
  - For computational scalability, well-behavior
  - Voluntary, i.e., bounded rationality. Involuntary, e.g., remote query failure.
  - Selectiveness of reasoning-by-cases (srbc)

- **Rulelog has all of the above. CL has hilog only.**
  - Rulelog is the first logical KRR that is defeasible + tractable + rich
  - CL/FOL is perfectly brittle in face of conflict from errors, confusions, tacit context
    - Entails garbage
Probabilistic Uncertainty

• Probabilistic uncertainty is representable in Rulelog, using meta features
• Reason about hilog atoms of the form: prob(formula, pvalue)
• Example of reasoning about cancer risk of patients
  • @\{genericRisk\}
    prob(cancer(will(be(diagnosed(in(?patient(within(5(years)))))))), 0.11) :-
    age(of(?patient(is(in(interval(55(to(60))))))).
  • @\{smokerRisk\}
    prob(cancer(will(be(diagnosed(in(?patient(within(5(years)))))))), 0.28) :-
    age(of(?patient(is(in(interval(55(to(60)))))))) and
    history(of(?patient))(includes(heavy(smoking))).
    _overrides(smokerRisk, genericRisk). /* smokerRisk rule has higher priority than
genericRisk */
• Research is needed into how to yet more tightly couple with rich logic
  yet achieve scalability
  • Challenge for Markov Logic Networks and other FOL-based probabilistic approaches
  • Interesting direction: probabilistic extension of Rulelog, generalizing Probabilistic LP
Rulelog Implementation Stack

• **UI, esp. KA**
  - Textual logic
  - Logic-oriented, esp. explanation and advanced debugging

• **Engine for reasoning**
  - Extensions for knowledge interchange
  - Extensions for omniformity (→ full Rulelog expressiveness)
  - Flora-2 (most Rulelog expressiveness) – open source
  - XSB Prolog (LP expressiveness) – open source

• **Available implementations for highly expressive Rulelog:**
  - Flora-2 reasoner: open source implementation for most expressiveness (gd. research code)
    - On top of XSB Prolog reasoner: open source (gd. research code, used commercially)
  - Additional extensions for all of the above: (omniformity, UI, and knowledge interchange)
    - Were implemented in Vulcan’s SILK but that’s not avail. (tho’ may be, esp. for research)
    - Commercial start-up: [Coherent Knowledge Systems](#) (formed by members of former SILK team)
For More Info

- See/listen-to the recent Ontolog Forum talk (6/20/13) by Benjamin Grosof

- It and other detailed material about Rulelog are available via the Coherent Knowledge Systems website’s Publications page
  - [http://coherentknowledge.com/publications](http://coherentknowledge.com/publications)
Thank You

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