High-Power Logical Representation via Rulelog for Neuro-Symbolic *(position paper)*

Benjamin Grosof*
Coherent Knowledge**

Presentation (30-min.)
July 17, 2016
at for NeSy '16 (Neuro-Symbolic Workshop)
Held New York, NY, USA

** Web: http://coherentknowledge.com
* Email: *firstname dot lastname* at coherentknowledge.com

Outline

- Why to combine KRR (symbolic) + ML (incl. neuro)
  - Esp. from viewpoint of ML
- KRR requirements for combining w/ ML
  - Expressiveness, scalability
- Rulelog KRR as a candidate for such combination
  - Features, Ergo implementation, applications. Pros/cons.
- Directions for future research

The Core of AI

Artificial Intelligence (AI)

Machine Learning (ML)

Knowledge Representation and Reasoning (KRR)

KRR is required for Cognitive Computing too.

Logic-based KRR's Roles in AI

- Complements ML … in sense of induction from data
  … to enable ML in broader sense
- The power of cultural transmission
  - "Evolution's lesson" (Wolfgang Bibel)
- Accumulate knowledge coherently
- Communicate with humans: expertise, questions
- "Inject" ML results into predictable software

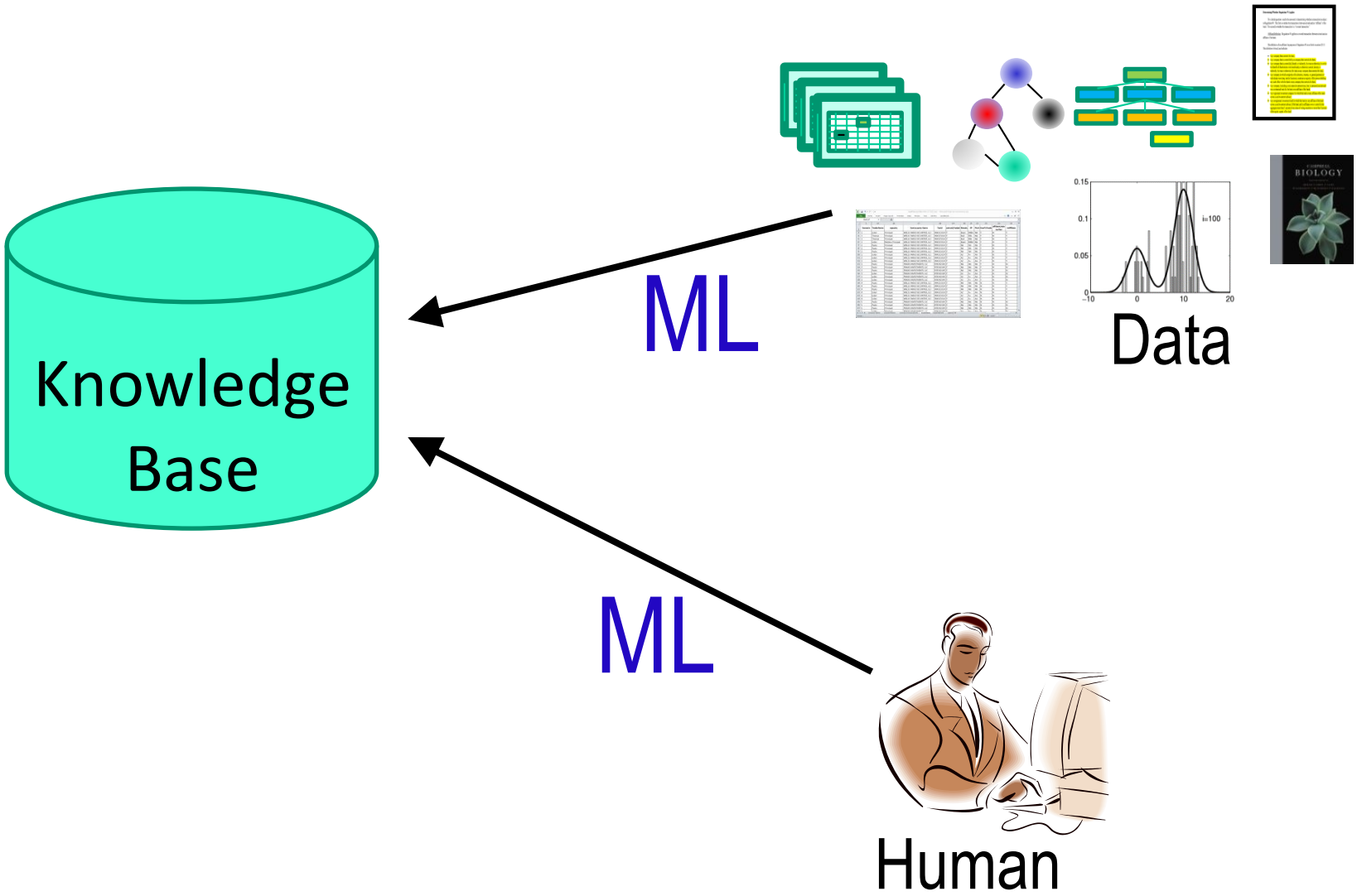
Why Combine ML with KRR

2 ways it's useful or even required,
from the viewpoint of KRR, i.e., "for KRR's sake":

1. KB construction: ML is useful to *supply* knowledge

2. *Improve the process of knowledge acquisition*
   - (Can view this as supplying a kind of meta-knowledge)
   - From manual entry of knowledge, e.g., encoding NL into rules
   - From knowledge interchange

Why Combine ML with KRR – diagram



Why Combine KRR with ML (I)

10 ways it's useful or even required,
from the viewpoint of ML, i.e., "for ML's sake":

1. The *prediction* step of ML requires reasoning
   - This could be pulled by an ML system via backchaining
   - Why not hook up various external programs such as reasoners, to NNs to evaluate some nodes/functions?

2. The *target* of ML is a representation

3. Getting business *value* from ML requires reasoning for analysis and decisions

Why Combine KRR with ML (II)

4. KRR is required to *combine* results of ML from
   a. Multiple episodes
   b. Multiple sources
   c. Multiple methods

5. KRR is required to *accumulate* knowledge coherently
   - Weakness of ML today
   - Think cultural transmission

Why Combine KRR with ML (III)

6. KRR is required to *explain* knowledge understandably to humans
   - Weakness of ML today
   - Needed for humans to trust an automated system
   - Often part of required/desired analysis functionality for own sake

Why Combine KRR with ML (IV)

7. Reasoning to *supply derived facts* for ML to chew on as training examples or background info
   - This could be pulled by an ML system via backchaining

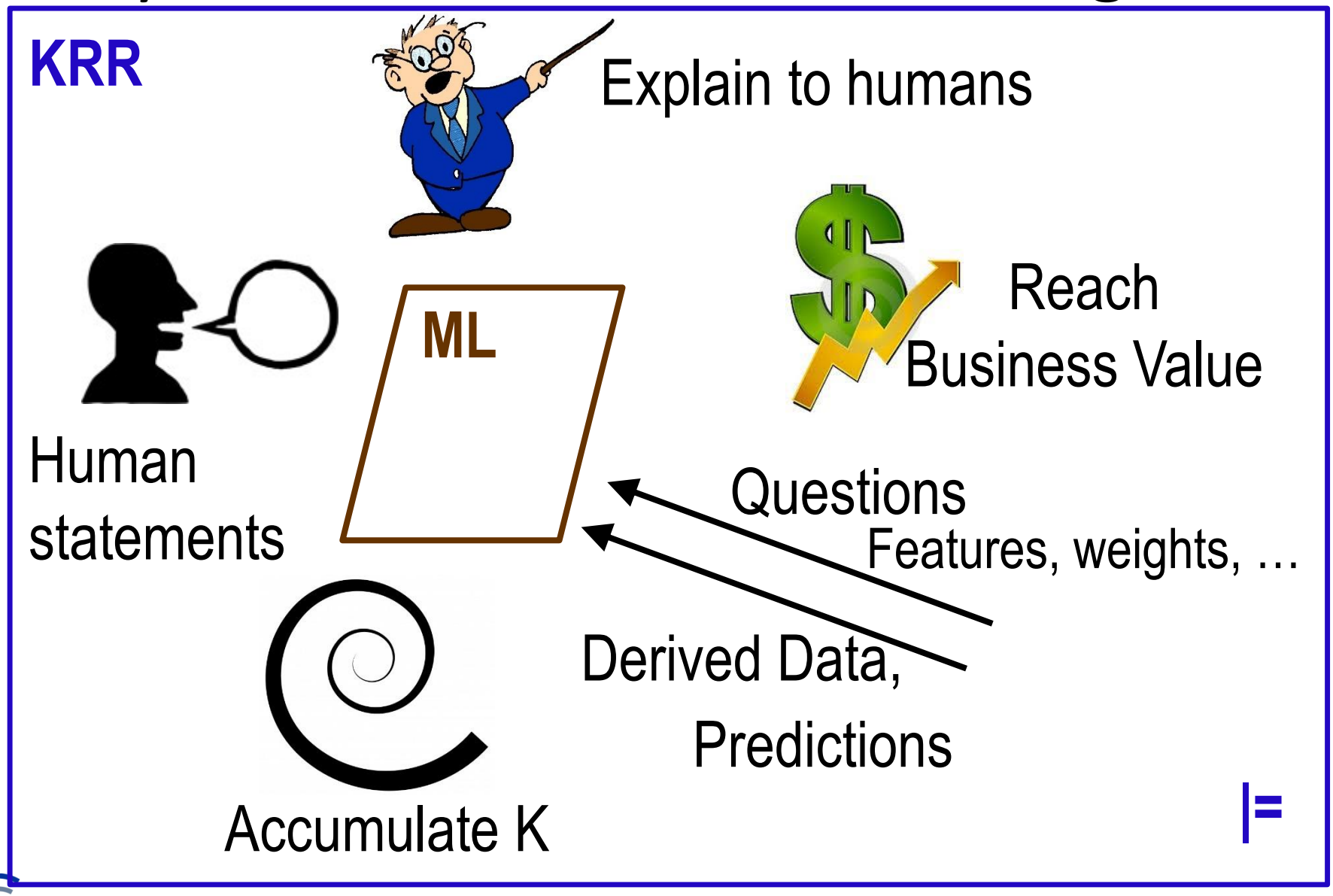
8. *Humans know stuff* beyond what's available via ML training data, and such knowledge is often complex to state / enter
   - KRR methods for entry/often more cost-effective than programming

Why Combine KRR with ML (V)

9. Reasoning is desirable to *pose questions* (tasks) to ML
   - as reasoning (sub)goals from KRR

10. Reasoning is desirable to *provide sets of relevant features, parameters, and/or weights* to ML

Why Combine KRR with ML – Diagram



Outline

- Why to combine KRR (symbolic) + ML (incl. neuro)
  - Esp. from viewpoint of ML
- KRR requirements for combining w/ ML
  - Expressiveness, scalability
- Rulelog KRR as a candidate for such combination
  - Features, Ergo implementation, applications. Pros/cons.
- Directions for future research

KRR Expressive Requirements for combination with neuro/ML (I)

Highly flexible expressiveness:

- Higher-order syntax
- Quantified formulas
- Strong meta (statements about statements)
- Numeric uncertainty (including weighting)
  - (thus differentiability)
  - Including probabilistic and fuzzy
- Defeasibility (exceptions)
  - Treat the evolving character of knowledge and of the world

KRR Scalability Requirements for combination with neuro/ML (II)

- Scalable computationally
  - To large amounts of asserted and concluded knowledge, i.e., “volume” and “velocity”

- Scalable “socially” to multiplicity of diverse knowledge, i.e., “variety”
  - To multiplicity of diverse ML/etc. sources, e.g., org.’s
  - To multiplicity of diverse ML/etc. algorithmic methods
  - To multiplicity of diverse underlying ML data samples

Outline

- Why to combine KRR (symbolic) + ML (incl. neuro)
  - Esp. from viewpoint of ML
- KRR requirements for combining w/ ML
  - Expressiveness, scalability
- Rulelog KRR as a candidate for such combination
  - Features, Ergo implementation, applications. Pros/cons.
- Directions for future research

Coherent Knowledge: Company Overview

- Fundamentally new kind of logic/rule based AI software platform for advanced analytics: flexible deep reasoning + natural language processing
- Radical business benefits: accuracy/competence, cost, agility, transparency
- Company offers: software product Ergo + professional services for custom solution dev
  - Capabilities: engine + development environment, for executable knowledge bases (logic/rules) embedded in apps
- World-class founder team: created many industry-leading logic systems & standards
  - Extensive experience applying logic systems to financial, regulation/policy, and other domains
  - Former/current professors at Stony Brook University and MIT

Michael Kifer, PhD
Principal Engineer
Prof., Stonybrook Univ.
Winner, 3 ACM & ALP test-of-time research awards.

Benjamin Grosof, PhD
CTO & CEO
Prof., MIT Sloan. DARPA PI.
Advanced AI Prog. Mger. for Paul Allen.
Creator, IBM Common Rules.

Theresa Swift, PhD
Principal Engineer
Co-lead dev, mission-critical rules system, US Customs.
Co-Architect, XSB Prolog.

Paul Fodor, PhD
Senior Engineer
Prof., Stonybrook Univ.
IBM Watson team.

Janine Bloomfield, PhD
Director of Operations
Sr. Scientist, Climate Change, Environmental Defense Fund.
Data Science at Yale, US Forest Service.

Coherent Knowledge

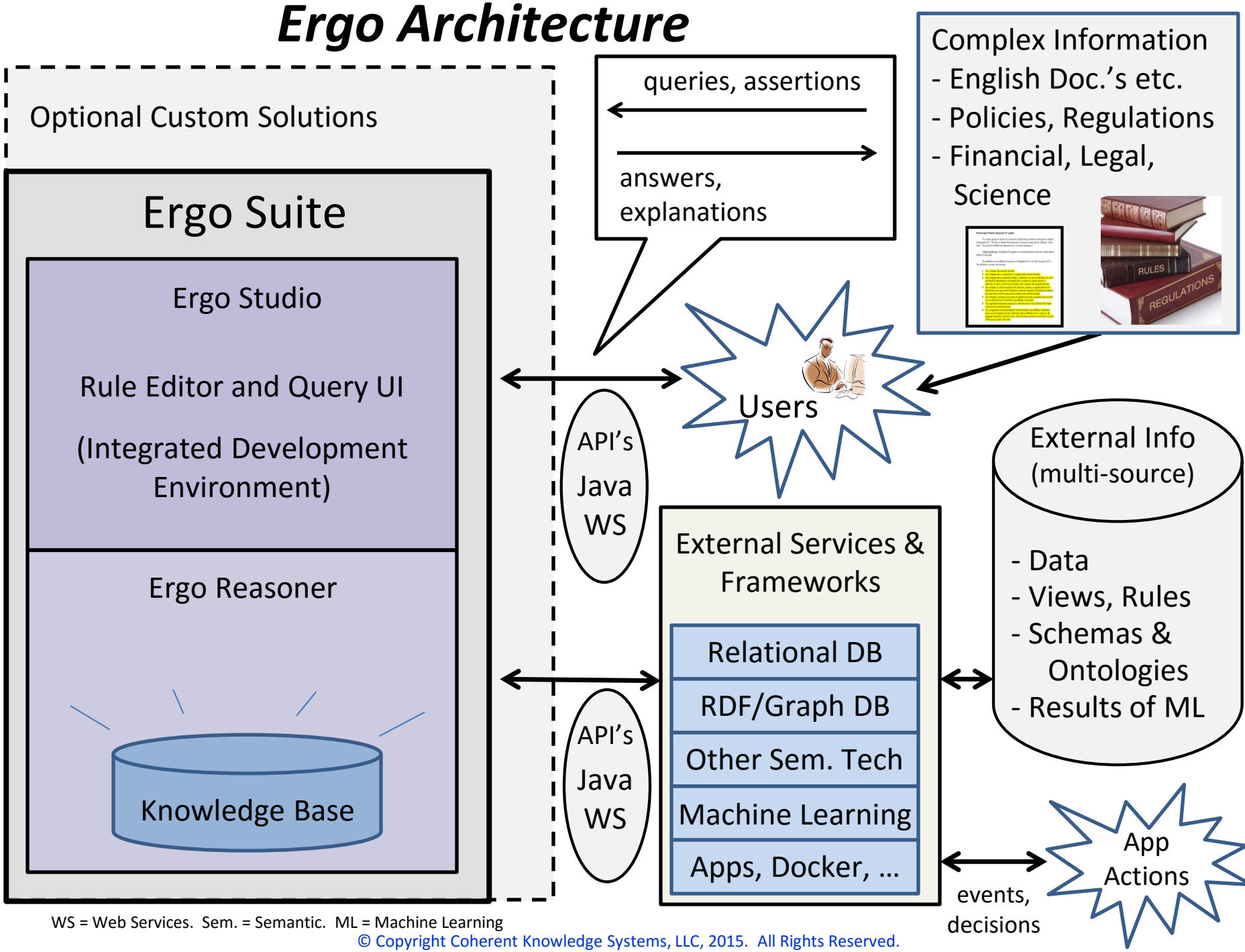
Rulelog and Textual Rulelog

- *Rulelog* is a kind of logical knowledge representation and reasoning (KRR)
  - A major research advance in KRR theory & algorithms, which culminated in 2012
- **Ergo is the most complete & highly optimized implementation available**
- Rulelog features very high/flexible expressiveness:
  - **Higher-order, general quantified formulas** (with logical chaining);
  - **Defeasibility (i.e., exceptions and argumentation)**;
  - Provenance, probabilistic, restraint bounded rationality, and more
- Yet Rulelog reasoning scales well: polynomial-time, as in databases
  - Millions of sentences concluded/asserted on a single processor
  - Up to trillions by orchestrating database etc. systems in distributed settings
- *Textual* Rulelog extends Rulelog with natural language processing (NLP)
  - Logic itself is utilized to map between English syntax and logic syntax
  - ErgoText templates aid knowledge entry and explanation generation

Series of Advances → Rulelog's Core Expressive Features

- Well-founded semantics; basic tabling algorithms
  - *Undefined* for paradox; smart cacheing; intuitionistic disjunction
- Higher-order syntax (Hilog); frame syntax
  - Associated optimizations of LP tabling etc. algorithms
- Statement id's for meta; argumentation meta-rules for defeasibility; provenance
- General formulas with all usual classical connectives and quantifiers (omniformity)
- Restraint bounded rationality
  - Use 3rd truth value *undefined* for "don't-care"
  - Radial, skipping; naf unsafety; external-query unsafety, unreturn

Ergo Architecture



WS = Web Services. Sem. = Semantic. ML = Machine Learning

Ergo Makes Sentences Executable

- If *something* is true then *something else* must be true. Written as:

then if

something_else :- *something*

- Example of <u>executable</u> Ergo sentence:

\(The individual affiliate threshold for transaction under Regulation W by ?Bank with ?Counterparty is ?Amount\) :-

\(?Counterparty is deemed an affiliate of ?Bank under Regulation W\) \and
\(?Bank has capital stock and surplus ?Capital\) \and
\(the threshold percentage for an individual affiliate is ?Percentage\) \and
?Amount = ?Capital * ?Percentage/100.

ErgoText

- ErgoText:

\(The proposed transaction ?Id by ?Bank with ?Affiliate of $?Amount is a RegW covered transaction\)

- ErgoText Template:

```
template(headbody,
        \(The proposed transaction ?Id by ?Bank with ?Affiliate of $?Amount
          is a RegW covered transaction\),

        covered(proposed(transaction))(by(?Bank))(with(?Affiliate))
          (of(amount(?Amount)))(having(id(?Id)))
        ).
```

- The templates are self-documenting

Textual Rulelog (III)

- Almost any NL sentence can be represented as a logical sentence
  - Leverages the <u>logical quantifiers</u> feature of Rulelog
  - Ex.: "each large company has some talented CEO"
    - forall(?x)^((?x \isa \(large company\)) ==>
      exists(?y)^((\?x has ?y\) \and
      (?y \isa \(talented CEO\)))).

Why is the proposed transaction prohibited by Regulation W?

3. *(continued) Why is the aggregate-affiliates limit $10 million?*

Why 'What proposed transactions are prohibited by RegW? Show '('Pacific Bank','Maui Sunset',23.0) ?

Edit Operations

- RegW prohibits the proposed transaction by Pacific Bank with Maui Sunset of $23.0 million
  - The proposed transaction by Pacific Bank with Maui Sunset of $23.0 million is a RegW covered transaction
  - There is a limit of $10.0 million for any proposed RegW covered transaction by Pacific Bank with Maui Sunset
    - There is an aggregated-affiliates limit of $10.0 million for any proposed RegW covered transaction by Pacific Bank with any affiliate
      - The aggregated total of previous RegW covered transactions by Pacific Bank with all affiliates is $490.0 million
      - The maximum threshold for aggregate RegW covered transactions by Pacific Bank with all affiliates is $500.0 million
        - The capital stock and surplus of Pacific Bank is $2500.0 million
        - The RegW threshold percentage for aggregate affiliates is 20.0 percent
        - $500.0 million is $2500.0 million multiplied by 20.0 percent
      - The limit of $10.0 million is the result of subtracting the previous RegW covered transactions total of $490.0 million from the RegW threshold $500.0 million
    - There is an individual-affiliate limit of $250.0 million for any proposed RegW covered transaction by Pacific Bank with Maui Sunset
    - The overall RegW limit of $10.0 million is the lesser of $10.0 million and $250.0 million
  - The proposed transaction of $23.0 million is greater than the RegW limit of $10.0 million

Business Benefits of Textual Rulelog for Analytics

- Deep in reasoning & knowledge
- Accurate
- Transparent, with explanations
- More Cost-Effective & Agile
- More Automated
- Easy to modify,
  end users empowered
- Greater Integration
- Greater Reusability
- What-if analyses

Analyze

Explain

Make Decisions

Monitor & Alert

Answer Questions

Application Areas for Rulelog (I)

- *Commercially, to date:*
- Financial regulatory/policy compliance
- Defense intelligence analysis
- Info integration: defense, financial, supply chain
- E-commerce pricing/promotion policies

Application Areas for Rulelog (II)

- *Explored in research, to date*
  … & promising commercially as further sub-areas

- Confidentiality policies: security, social media, HIPAA

- Financial/business reporting: XBRL

- Contracts: e-commerce, financial instruments, license agreements, large construction

- Health: treatment guidance, insurance

Coherent Knowledge

Application Areas for Rulelog (III)

- *Explored in research, to date*
  … & promising commercially further areas

- Education/e-learning: personalized tutoring

- NL understanding and conversational interfaces

- Workflow / business process management: helpdesk, personal communications

KRs for Neuro-Symbolic: Candidates and Pros/Cons (I)

- Rulelog limitations:
  - Lacks "reasoning-by-cases", a.k.a. it is "intuitionistic"
    - Only concludes a disjunction if it concludes one of the disjuncts
  - Not yet optimized for numeric uncertainty reasoning
- Classical logic (first-order / higher-order):
  - Has reasoning-by-cases. But …
  - Lacks defeasibility
    - Brittle in face of conflicting/evolving K
    - Lacks social scalability
  - Lacks computational scalability, restraint
  - Lacks rule id's – hook for important kinds of meta

KRs for Neuro-Symbolic: Candidates and Pros/Cons (II)

- Markov Logic Networks
  - Attractively flexible and principled in numeric uncertainty
  - Has reasoning-by-cases
  - Much less computationally scalable than Rulelog
    - Though there are some tractable cases, they are restrictions
  - Lacks some key strong meta features, e.g., higher-order, id's
- Answer Set Programs (ASP)
  - Has reasoning-by-cases
  - Much less brittle than classical logic
  - Lacks computational scalability
  - Lacks some key strong meta features, e.g., rule id's

Outline

- Why to combine KRR (symbolic) + ML (incl. neuro)
  - Esp. from viewpoint of ML
- KRR requirements for combining w/ ML
  - Expressiveness, scalability
- Rulelog KRR as a candidate for such combination
  - Features, Ergo implementation, applications. Pros/cons.
- Directions for future research

Some Rulelog Lessons for Neuro-Symbolic

- Flexible expressiveness of rules
  - Doable practically. Very beneficial.
  - Specific constructs: HiLog, rule id's, defeasible, quantifiers, meta
  - Extend to NL via templates and parsing+quantification
    - Human-machine logic
  - Need to optimize probabilistic –ish weighted
- Explanation
  - Doable practically. Very beneficial.
  - Rule notion is very accessible to people.
  - Present justification digraph (with cycles) as tree with drill-down

Directions for Future Research

- Hook up Rulelog implementations to NN/ML systems
  - Potential applications include:
    - Compliance and fraud
    - NL understanding
      - Intelligence analysis
      - Search
- Analyze & compare more KRs tha are probabilistic/weighted + logic
  - e.g., in statistical relational
  - e.g., probabilistic soft logic

Future Research Directions: Core Technology and Experiments

- Feed derived data from Rulelog to NN (or ML)
  - Could be push or pull
- Combine NN results with other knowledge
  - E.g., with human-authored complex knowledge
    - Terminology mappings
    - Source trustworthiness
- Combine NN word-vector distributed representations with Textual Rulelog

Future Research Directions: on Rulelog KRR itself

- Optimize, and study further: Rulelog reasoning with uncertainty that is numerically weighted, including probabilistic and fuzzy

- Extend Rulelog's expressiveness to selective reasoning-by-bases

Thank you.

Deep Reasoning for Advanced Analytics

http://coherentknowledge.com