Learning to Reason: from Question Answering to Problem Solving

Michael Witbrock
Broad AI Lab, University of Auckland School of Computer Science
m.witbrock@auckland.ac.nz

Presented Remotely
History of AI
Two Kinds of Processing

System 1: skill.
- Fast,
- Frequent,
- Automatic,
- Unconscious,
- Stereotyped
- Invariant,
- Emotional

- Mammals, Birds, Fish, Insects, Reptiles

System 2: reason.
- Slow
- Infrequent,
- Deliberate,
- Conscious,
- Compositional,
- Flexible
- Rational

- Mostly Humans, apes, cetaceans, octopuses, parrots
Two Kinds of Learning

Skills.
• Training
• Require many examples
• Transfer is limited
• Performed

Knowledge
• Instruction
• Can be acquired by reading, listening, watching, looking
• Can be reused widely
• Reasoned with
SKILLS

Classification

- Generative
  - Reinforcement Learning
  - Deep and Convolutional
But what happened to knowledge and reasoning?

&

How can we combine them with skills and learning?

“This falls significantly short of human-problem solving, including question-answering: it does not recursively decompose problems for solution, it does not follow that decomposition to assemble answers, and it does not store and apply salient background knowledge for decomposition, partial solution, or answer composition.”
What does that mean for progress towards broader and general AI?
Symbolic AI

Applying knowledge that is stored in a use-agnostic form to solve new problems

One way to think about knowledge, from a machine-learning point-of-view, is stored inductive bias: it reduces the amount of new data required to learn an applicable solution method.
Types of Symbolic Knowledge (Text Like)

Mathematical Logic

Careful Legal Language

Formal Narrative

Fiction

Computer Programs

Scientific Writing

Informal Narrative

Poetry & Song

High precision and reliable reuse, Low expressiveness

Low precision and usability

High expressiveness
Types of Symbolic Knowledge (picture-like)

High precision and reliable reuse, Low expressiveness

Low precision and usability, High expressiveness
Types of Symbolic Knowledge (data-like)

Serialised Data Structures
- WSML Annotated Feeds
  - High precision and reliable reuse, Low expressiveness

Relational DBs
- NoSQL
  - Low precision and usability, High expressiveness

Knowledge Graphs
- Observations
  - Low precision and usability, High expressiveness
Fundamental Operation of Reasoning (recursive and exploratory)

• Interpret problem to be solved
• Transform it into a set of simpler problems, using knowledge
• Solve the simpler problems (by knowing the answer, applying a skill, or breaking them down)
• Use the solutions of the simple problems to assemble the solution to the larger problem
What does reasoning look like to a computer

https://www.cyc.com/cyc-technology-overview/

What days of the week can’t Hugo Jederman use his Segway to commute on.

(Read Lime Scooter)
Is Honolulu a City?

- It’s a city if it’s the capital of a country
  - Is Honolulu the capital of a country?
    - Nope

- It’s a city if it has more than 70000 inhabitants
  - X is number of inhabitants of Honolulu
    - X > 70000

Between 1845 and 1893 yes.

351800 is the number of inhabitants of Honolulu (2016)

351800 > 70000

It’s a city because >70000
Kind of trivial in pseudo logic, BUT

- Hundreds of thousands of millions of facts or problem transformation methods might apply at each step [intractable]
- Much as computers might like it, humans have not converted most facts and problem solving methods into mathematical logic (or computer programs) [impractical]
What might we do?

• Pretend reasoning is a game, and apply reinforcement learning to it
• Pretend logic is a natural language, and translate into it
• Learn to reason directly with text, as we did (kind of) in the example
• Learn to modify text so it’s more like logic

• Work on problems that are already in logic (Mizar & HOLStep mathematical formalisation, WikiData, parts of systems biology, e.g. OBO)
Question Answering as a Proxy for NLU and Reasoning
Spider: A Large-Scale Human-Labeled Dataset for Complex and Cross-Domain Semantic Parsing and Text-to-SQL Task


Table 2: Accuracy of Exact Matching on SQL queries with different hardness levels.
Improving Natural Language Inference Using External Knowledge in the Science Questions Domain

Xiaoyan Wang, Pavan Kapanipathi, Ryan Musa, Mo Yu, Kartik Talamadupula, Ibrahim Abdelaziz, Maria Chang, Achille Fokoue, Bassem Makni, Nicholas Mattei, Michael Witbrock

SciTail dataset

<table>
<thead>
<tr>
<th>Model</th>
<th>Dev</th>
<th>Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decomp-Attn (Parikh et al. 2016)</td>
<td>75.4</td>
<td>72.3</td>
</tr>
<tr>
<td>DGEM* (Khot, Sabharwal, and Clark 2018)</td>
<td>79.6</td>
<td>77.3</td>
</tr>
<tr>
<td>DeIstTe (Yin, Roth, and Schütze 2018)</td>
<td>82.4</td>
<td>82.1</td>
</tr>
<tr>
<td>BiLSTM-Maxout (Mihaylov et al. 2018)</td>
<td>-</td>
<td>84.0</td>
</tr>
<tr>
<td>match-LSTM (Wang and Jiang 2015)</td>
<td>88.2</td>
<td>84.1</td>
</tr>
</tbody>
</table>

Our implementation

match-LSTM (GRU) 88.5 84.2
match-LSTM+WordNet* (Chen et al. 2018) 88.8 84.3
match-LSTM+Gmatch-LSTM* (ConSeqNet) 89.6 85.2
AI2 Reasoning Challenge (ARC) multi-choice selected to be unanswerable by IR. Knowledge added to query representation, inputs rewritten to improve performance, entailment of answer-augmented query by retrieved evidence.
BERT encoding; pointers into sentences to decompose; fixed decomposition patterns; exactly two sub-questions; base QA system is BERT reading comprehension model (Devlin et al., 2019) algorithmic answer composition; chooses best one to answer.
The reasoning these systems do is very limited
End-to-End Differentiable Proving

Tim Rocktäschel
University of Oxford
tim.rocktaschel@cs.ox.ac.uk

Sebastian Riedel
University College London & Bloomsbury AI
s.riedel@cs.ucl.ac.uk

Abstract

We introduce neural networks for end-to-end differentiable proving of queries to knowledge bases by operating on dense vector representations of symbols. These neural networks are constructed recursively by taking inspiration from the backward chaining algorithm as used in Prolog. Specifically, we replace symbolic unification with a differentiable computation on vector representations of symbols using a radial basis function kernel, thereby combining symbolic reasoning with learning subsymbolic vector representations. By using gradient descent, the resulting neural network can be trained to infer facts from a given incomplete knowledge base. It learns to (i) place representations of similar symbols in close proximity in a vector space, (ii) make use of such similarities to prove queries, (iii) induce logical rules, and (iv) use provided and induced logical rules for multi-hop reasoning. We demonstrate that this architecture outperforms ComplEx, a state-of-the-art neural link prediction model, on three out of four benchmark knowledge bases while at the same time inducing interpretable function-free first-order logic rules.
Questions in CLEVR test various aspects of visual reasoning including attribute identification, counting, comparison, spatial relationships, and logical operations.

Q: Are there an equal number of large things and metal spheres?
Q: What size is the cylinder that is left of the brown metal thing that is left of the big sphere?
Q: There is a sphere with the same size as the metal cube; is it made of the same material as the small red sphere?
Q: How many objects are either small cylinders or red things?

Each question in CLEVR is represented both in natural language and as a functional program. The functional program representation allows for precise determination of the reasoning skills required to answer each question.

Sample chain-structured question:
Filter color
Filter shape
Unique
Relate
Filter size
Unique
Query color

What color is the cube to the right of the yellow sphere?

Sample tree-structured question:
Filter color
Unique
Relate
Filter size
Unique
Relate
Filter shape
Count

How many cylinders are in front of the tiny thing and on the left side of the green object?


https://cs.stanford.edu/people/jcjohns/clevr/
Deep Reasoning for QA

- Recursively transform queries into subproblems to be solved, over as many levels as necessary
- Learn the transformation operations
- Learn the answer assembly operations
Enabling Inference at Scale

• Treat deduction, abduction and inductive steps, and other reliable skills, as steps in a game
• Learn to be very good at that game
Freeing inference from logic

- Recalling the slide earlier about the precision / expressiveness trade-off
- Learn to produce inferences that are as reliable as possible using language-like representations
- Where possible, bound the unreliability and report it with explanations
Interoperable Symbolic Representations

- Using modern Deep Learning techniques for language (sequence) modelling (e.g. GPT2), form representations for logic, natural language, and programming and query languages that aid downstream tasks
  - E.G. Optimise BERT or GPT2 for entailment tasks
- Apply to tasks in understanding legal documents, question answering, and understanding biomedical texts and biological systems
Automatic Production of Precise Representations

• Exact, logical or programmatic knowledge representations have strong efficiency and accuracy benefits when available.

• How can computers build these representations themselves, with little or no human intervention, from existing knowledge resources, text, and other training data
Knowledge Transformation and Capture for Problem Solving

• Learn to transform/rewrite text sources so that they can be more reliably used for problem solving steps
• Learn to elicit material in reasoning-ready form from human beings in task contexts
What’s it for?

- Broad AI: solving these challenges should move us significantly closer to AI systems that can be applied to a broad and heterogeneous set of problems. Move beyond the narrow (but still often superhuman AI skills of today).
Graph2seq: Graph to sequence learning with attention-based neural networks
K Xu, L Wu, Z Wang, Y Feng, M Witbrock, V Sheinin
Broad AI Lab

- Learning-based general artificial intelligence
- Complex problem solving with Natural Language (NL)
- Combine Deep Learning revolution in AI with symbolic AI to give computers the powers of understanding and integration
- Challenging area of AI research, with high potential commercial impact.
  - Near-term advances in understanding text, diagrams and tables so they can be automatically repurposed and combined to answer questions
- Secondary Focus: AI for the benefit of human civilisation
- Michael Witbrock
  - Ph.D, Computer Science, Carnegie Mellon University
  - Previous:
    - Distinguished Research Staff Member & Manager of Reasoning Group at IBM Research, Yorktown New York
    - Vice President for Research, Cycorp Inc
    - Principal Scientist, Lycos Inc.
    - Research Scientist, JustSystem Pittsburgh Research Center
  - Research: Deep Learning, Automated Reasoning, Natural Language Understanding
  - Current affiliations: University of Auckland, Robust.AI, Epistemic.AI, AI for Good Foundation AI4Good.org (founder)
  - Current collaborations: IBM Research AI (USA), Northwestern University (USA)