Toward an Architecture for Never-Ending Language Learning

Andrew Carlson, Justin Betteridge, Bryan Kisiel, Burr Settles, Estevam R. Hruschka Jr., and Tom M. Mitchell

School of Computer Science
Carnegie Mellon University
Humans learn many things, for many years, and become better learners over time.

Why not machines?
Never-Ending Learning

Task: acquire a growing competence without asymptote
- over years
- learning multiple functions
- where learning one thing improves ability to learn the next
- acquiring data from humans, environment

Many candidate domains
- Robots
- Softbots
- Game players
NELL: Never-Ending Language Learner

Inputs:
- Initial ontology
- Handful of examples of each predicate in the ontology
- The web
- Occasional interaction with human trainers

Task:
- Run 24x7, forever
- Each day:
  - Extract more facts from the web to populate initial ontology
  - Learn to read better than yesterday
Ontology

123 Categories

City
Country
Athlete
Company
Sports Team
Economic Sector
Emotion

55 Relations

LocatedIn
HeadquarteredIn
PlaysFor
TeamInLeague
PlaysSport
OperatesInEconomicSector
Why do this?

- Case study in never-ending learning
- Potential for new breakthroughs in natural language understanding
- Producing the world’s largest structured KB
Bootstrapped Pattern Learning
(Brin 98, Riloff and Jones 99)

- Canada
- Egypt
- France
- Pakistan
- Sri Lanka
- Argentina
- Planet Earth
- North Africa
- Student Council

- X is the only country
  - home country of X
- invasion of X
  - elected president of X
Without proper constraints, a never-ending bootstrap learner will “run off the rails.”

How can we avoid this?
Solution Part 1: Coupled Learning of Many Functions

Diagram:
- **LocatedIn** connects to City
- **HeadquarteredIn** connects to Company
- City connects to Country, Company, and Sports Team
- Country connects to City and Athlete
- Company connects to City, Sports Team, and Athlete
- Athlete connects to City, Company, and PlaysFor
- Sports Team connects to City, Company, and PlaysFor
Exploiting Mutual Exclusion

Positives:
Canada
Egypt
France

invasion of X
elected president of X

Planet Earth
North Africa
Student Council

Negatives:
Europe
London
Florida
Baghdad

nations like X
countries other than X

Pakistan
Sri Lanka
Argentina
Coupled Pattern Learner: Type Checking

X, which is based in Y

Pillar, San Jose  OK

Type Checking Arguments:
... companies such as Pillar ...
... cities like San Jose ...

Inclined pillar, foundation plate  Not OK
Solution Part 2: Multiple Extraction Methods

Textual Extraction Patterns
- Mayor of X

List Extraction

Morphology Classifier
- “-son” suffix likely to be a last name

Rule Learner
- An athlete who plays for a team that plays in the NBA plays in the NBA
NELL architecture

Data Resources (e.g., corpora)

Knowledge Base

beliefs

candidate facts

Knowledge Integrator

Subsystem Components

CPL  CSEAL  CMC  RL
## Learned Extraction Patterns

<table>
<thead>
<tr>
<th>Pattern</th>
<th>Predicate</th>
</tr>
</thead>
<tbody>
<tr>
<td>blockbuster trade for X</td>
<td>athlete</td>
</tr>
<tr>
<td>airlines, including X</td>
<td>company</td>
</tr>
<tr>
<td>personal feelings of X</td>
<td>emotion</td>
</tr>
<tr>
<td>X announced plans to buy Y</td>
<td>companyAcquiredCompany</td>
</tr>
<tr>
<td>X learned to play Y</td>
<td>athletePlaysSport</td>
</tr>
<tr>
<td>X dominance in Y</td>
<td>teamPlaysInLeague</td>
</tr>
</tbody>
</table>
## Example Morphological Features

<table>
<thead>
<tr>
<th>Predicate</th>
<th>Feature</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>mountain</td>
<td>LAST=peak</td>
<td>1.791</td>
</tr>
<tr>
<td>mountain</td>
<td>LAST=mountain</td>
<td>1.093</td>
</tr>
<tr>
<td>mountain</td>
<td>FIRST=mountain</td>
<td>-0.875</td>
</tr>
<tr>
<td>musicArtist</td>
<td>LAST=band</td>
<td>1.853</td>
</tr>
<tr>
<td>musicArtist</td>
<td>POS=DT_NNS</td>
<td>1.412</td>
</tr>
<tr>
<td>musicArtist</td>
<td>POS=DT_JJ_NN</td>
<td>-0.807</td>
</tr>
</tbody>
</table>
Example Learned Rules

• Athletes who play in the NBA play basketball.
• Teams that won the Stanley Cup play in the NHL.
• If an athlete plays for a team that plays in a league, then the athlete plays in that league.

(Solution Part 3: Discovery of New Constraints)
### 6 facts learned in the last week

<table>
<thead>
<tr>
<th>Predicate</th>
<th>Instance</th>
</tr>
</thead>
<tbody>
<tr>
<td>architect</td>
<td>Charles Moore ✓</td>
</tr>
<tr>
<td>park</td>
<td>Parque Nacional Conguillio ✓</td>
</tr>
<tr>
<td>kitchen item</td>
<td>oven safe skillet ✓</td>
</tr>
<tr>
<td>county</td>
<td>Woodbury County ✓</td>
</tr>
<tr>
<td>card game</td>
<td>cash bonus ✗</td>
</tr>
<tr>
<td>perception event</td>
<td>energy engineering ✗</td>
</tr>
</tbody>
</table>
NELL right now

- 314K beliefs
- 30K textual extraction patterns
- 486 accepted learned rules leading to 4K new beliefs
- 65-75% of predicates currently populating well, others are receiving significant correction
Lessons so far

• Key architectural ingredients:
  • Coupled target functions
  • Multiple extraction methods
  • Discovery of new constraints among relations

• We’ve changed the accuracy vs. experience curve from ___ to ____, but not to ___
The future

• Distinguish entities from textual strings
• More human involvement
• Ontology extension
• Planning
Thank you

Thanks to Yahoo! for M45 computing
Thanks to Jamie Callan for ClueWeb09 corpus
Thanks to Google, NSF, and DARPA for partial funding

Learn more at http://rtw.ml.cmu.edu