

The Future of AI: The View from AI2

Oren Etzioni

www.allenai.org

Allen Institute for Artificial Intelligence (AI2)



AI Present: Deep Learning Tidal Wave



The large ball crashed right through the table because **it** was made from foam.

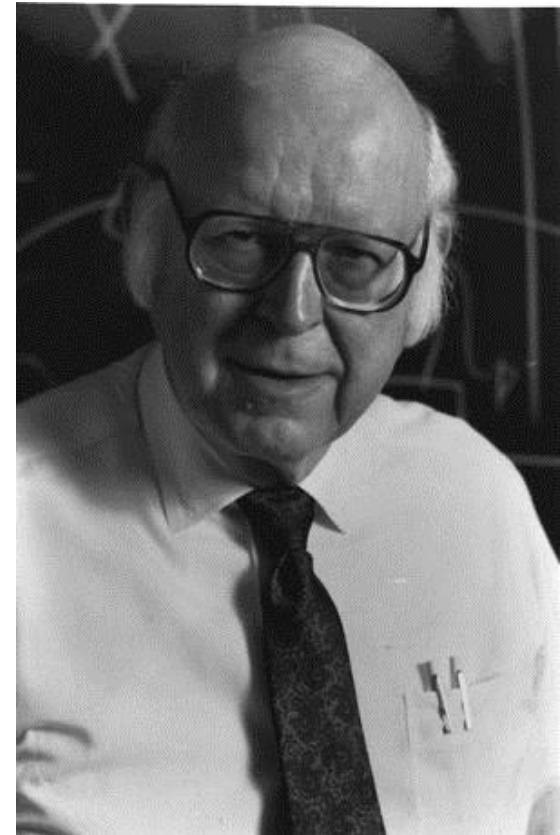
It = ?

**Common-sense
knowledge & tractable
reasoning are necessary
for NLP!**

**AI cannot be reduced to classification
(using realistic feature sets)**

**“You can’t play 20 questions
with nature and Win.”**

Allen Newell, 1973



What's Next?

Outline

- I. AI2 Methodology
- II. Euclid
- III. Aristo
- IV. Semantic Scholar Overview & Demo
- V. Conclusions

AI is Increasingly Fragmented

SAT solver MLN SVM
Drones
CNN **We need integrated AI Systems!**
NLP

Etc.



r, molers.dk

Build AI Programs that take written tests

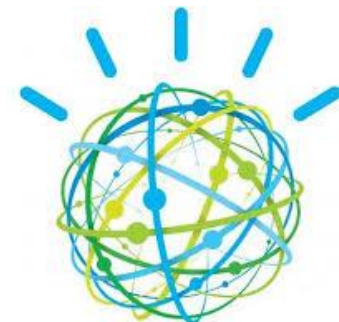
1. Externally-defined challenge tasks
2. Training data + **unseen test data** (“as is”)
3. Measurable progress, clear focus, ambitious goals

Key differences with Watson:

1. Deeper semantics & inference
2. Open model: publish, collaborate, open source

achieve
more

SAT[®]
AP[®]

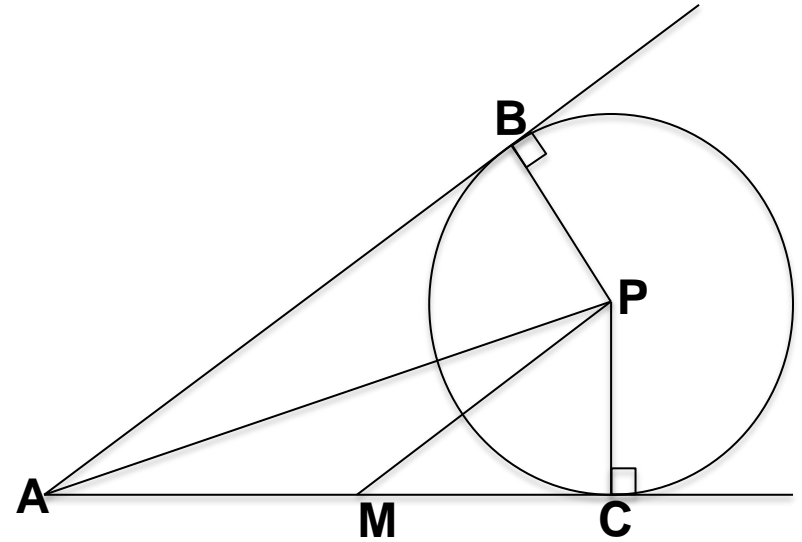


IBM WATSON

II. Euclid: Solving Geometry Questions

In the figure at the right, the circle has center P and radius r . Lines AB and AC are tangent to the circle at points B and C , respectively. If PM is the bisector of the tangent, and the measure of angle PMC equals the measure of angle MPC , what is the length of segment PA ?

- (a) (b) (c)
(d) (e)



$r\sqrt{3}$
 $r+1$

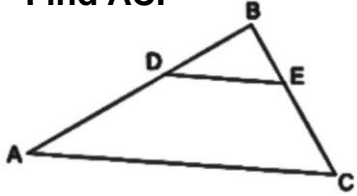
$r\sqrt{5}$
 $2r$

$r\sqrt{2}$

Multi-Modal Parsing Approach

Geometry Question

In triangle ABC, line DE is parallel with line AC, DB equals 4, AD is 8, and DE is 5. Find AC.

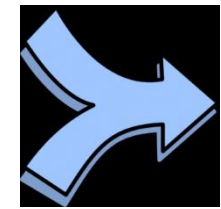


Semantic Forest
→

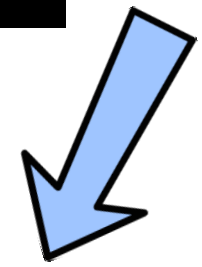
(Noisy) Relations
Parallel(line(DE), line(AC))
Parallel(line(AC), line(DB))
Equal(length(line(AD)), 4)
Equal(length(line(AD)), 8)
...

G-Aligner
→

X^D, R^D, d



Grounded Relations
Parallel (I1,I2)
Parallel (I2,I3)
Equal(length(I3),4)
Equal(length(I4),8)
...



Solver
←

Answer

Selected Relations
Parallel (I1,I2)
Equal(length(I1),4)
Equal(length(I3),4)
Equal(length(I2),8)
Length(I2) = ?

Details in [EMNLP15]

Numerical Solver

- Translate a logical form to a non-linear equation

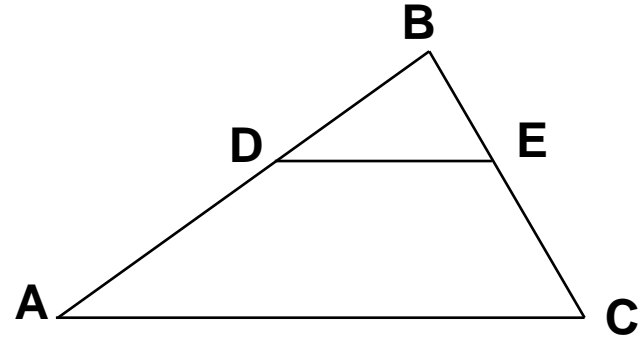
Formal Language	Equations
Equals(LengthOf(AB),d)	$(A_x - B_x)^2 + (A_y - B_y)^2 - d^2 = 0$
Parallel(AB, CD)	$(A_x - B_x)(C_y - D_y) - (A_y - B_y)(C_x - D_x) = 0$
LiesOn(B, AC)	$(A_x - B_x)(B_y - C_y) - (A_y - B_y)(B_x - C_x) = 0$
Perpendicular(AB,CD)	$(A_x - B_x)(C_x - D_x) + (A_y - B_y)(C_y - D_y) = 0$

Goal: Find an assignment to the variables that satisfies all the equations *simultaneously*

Optimization Result

“In triangle ABC, line DE is parallel with line AC, DB equals 4, AD is 8, and DE is 5. Find AC.”

a) 2 b) 4 c) 6 d) 8 e) 10



IsTriangle(ABC)	0.96
Parallel(AC, DE)	0.91
Parallel(AC, DB)	0.74
Equals(LengthOf(DB), 4)	0.97
Equals(LengthOf(AD), 8)	0.94
Equals(LengthOf(DE), 5)	0.94
Equals(4, LengthOf(AD))	0.31
Find(LengthOf(AC))	0.90

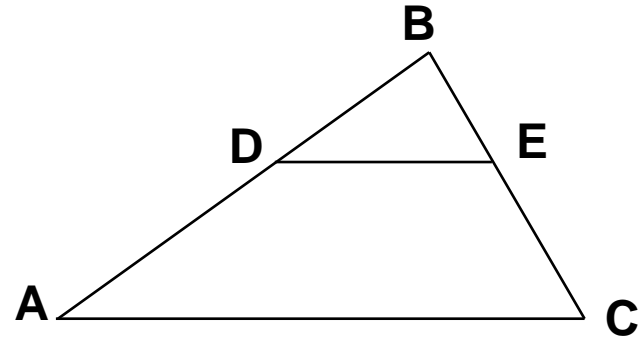
LiesOn(D, AB)	1.0
LiesOn(E, BC)	1.0
Parallel(AC, DE)	0.99
Parallel(AC, DB)	0.02

IsTriangle(ABC)	Parallel(AC, DE)
Equals(LengthOf(DB), 4)	Equals(LengthOf(AD), 8)
Equals(LengthOf(DE), 5)	Find(LengthOf(AC))

Optimization Result

“In triangle ABC, line DE is parallel with line AC, DB equals 4, AD is 8, and DE is 5. Find AC.”

- a) 2 b) 4 c) 6 d) 8 **e) 10**

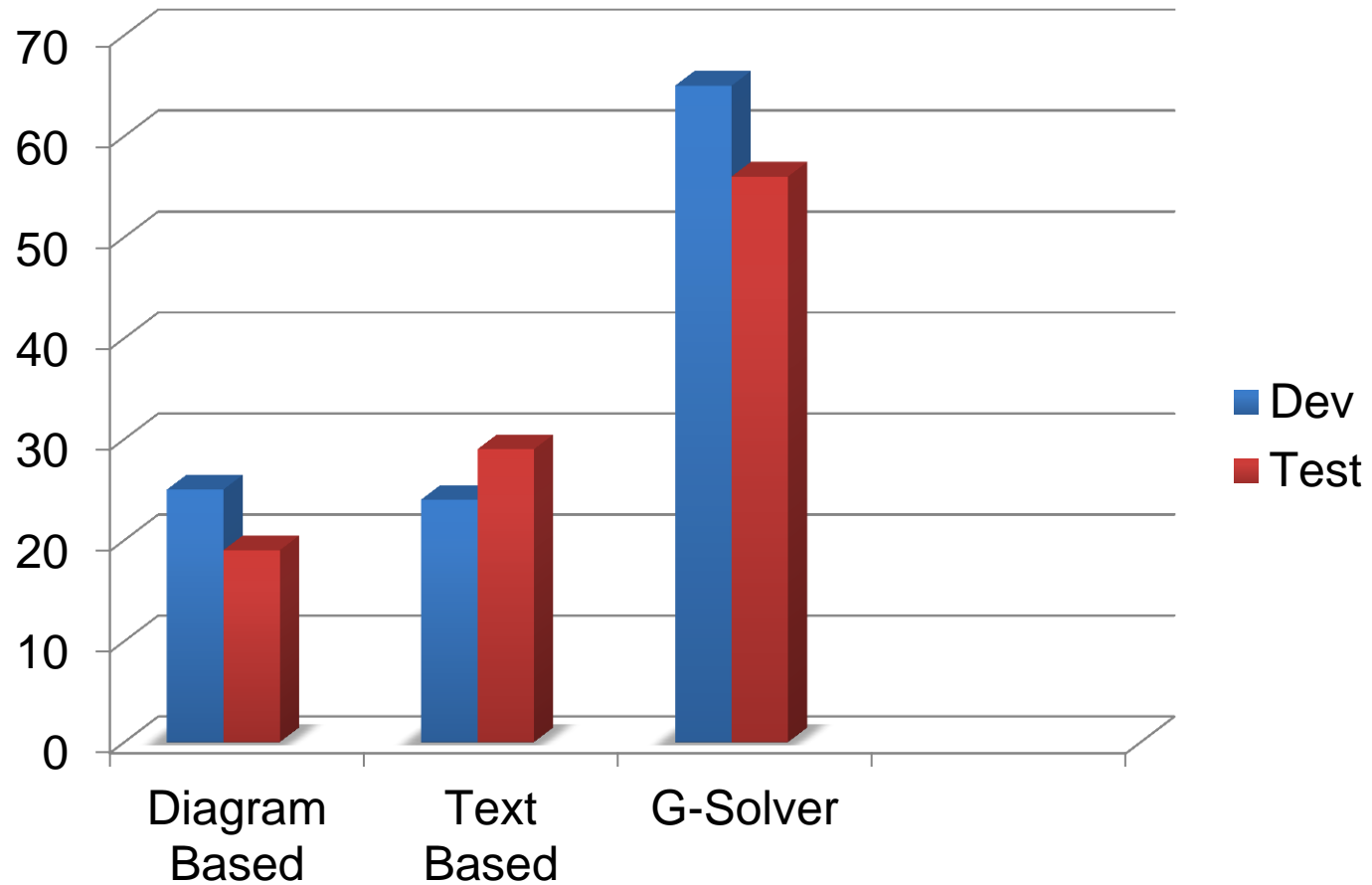


IsTriangle(ABC)	0.96
Parallel(AC, DE)	0.91
Parallel(AC, DB)	0.74
Equals(LengthOf(DB), 4)	0.97
Equals(LengthOf(AD), 8)	0.94
Equals(LengthOf(DE), 5)	0.94
Equals(4, LengthOf(AD))	0.31
Find(LengthOf(AC))	0.90

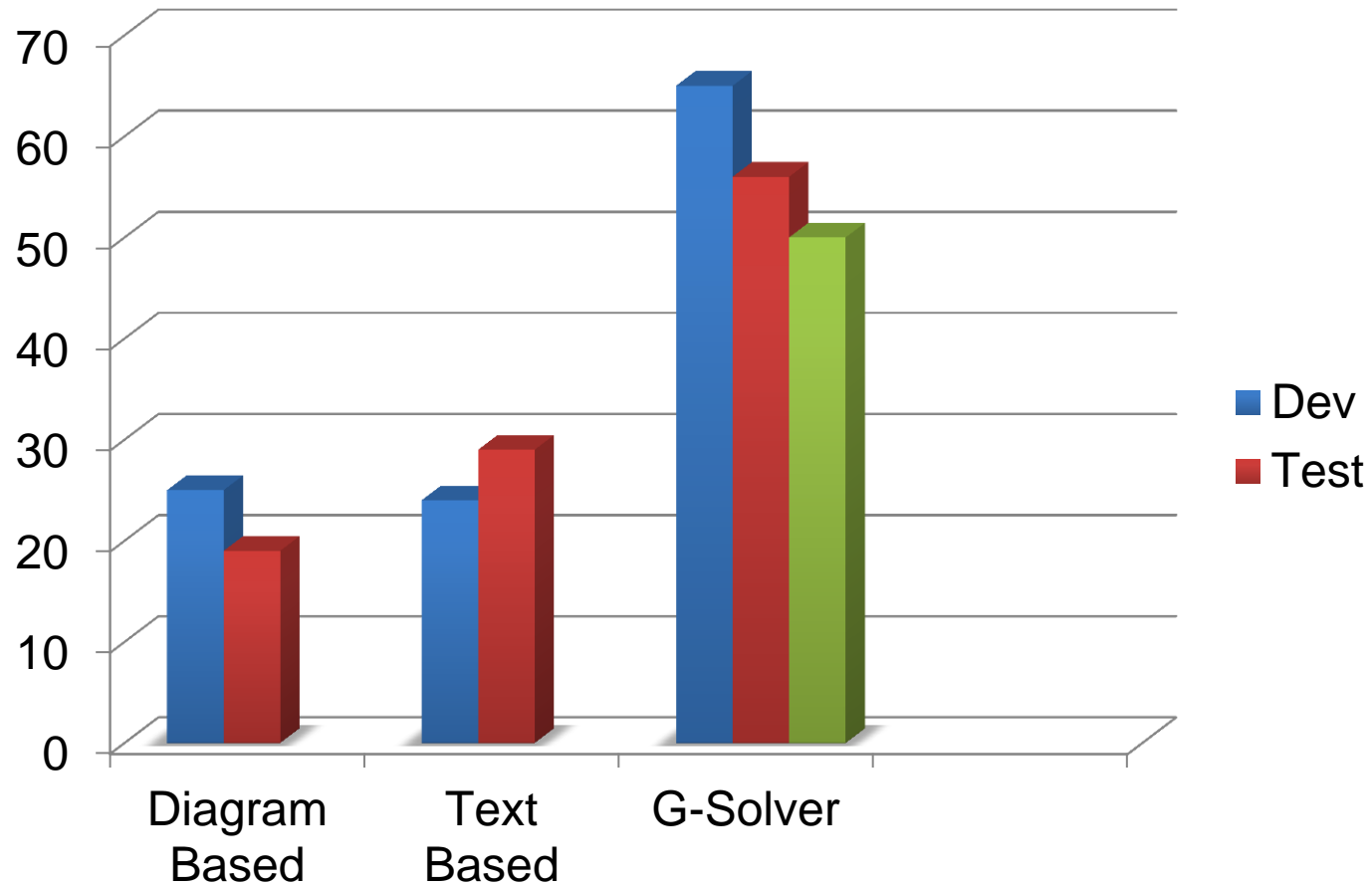
LiesOn(D, AB)	1.0
LiesOn(E, BC)	1.0
Parallel(AC, DE)	0.99
Parallel(AC, DB)	0.02

IsTriangle(ABC)	Parallel(AC, DE)
Equals(LengthOf(DB), 4)	Equals(LengthOf(AD), 8)
Equals(LengthOf(DE), 5)	Equal(LengthOf(AC), 10)

Evaluation



Evaluation



Just presented at [EMNLP2015]

III. Aristo: Elementary School Science Tests

THE UNIVERSITY OF THE STATE OF NEW YORK

GRADE 4

ELEMENTARY-LEVEL
SCIENCE TEST

WRITTEN TEST

JUNE 3, 2013

- 12 Which force causes a bicycle to slow down when the brakes are used?
- A friction
 - B electricity
 - C gravity
 - D magnetism

Why Elementary Science is Challenging

- **Natural Language Understanding** is often needed
 - “Shallow” NLP maxes out <60% (guessing = 25%)

- **Scientific and World Knowledge** is needed

*A ball is tossed in the air and **it** comes back down due to (A) gravity (B) ...*

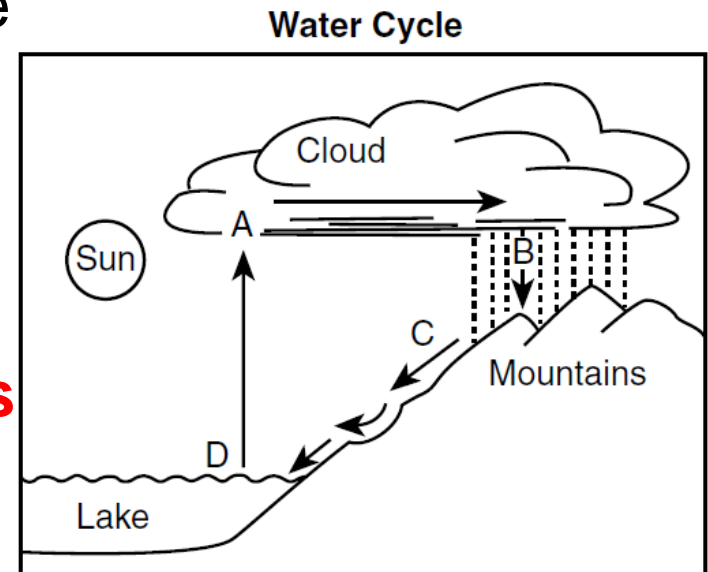
- **Linguistic Variability** is everywhere

allow heat transfer ~ is a conductor

get a better look at ~ view in more detail

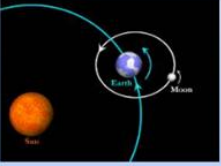








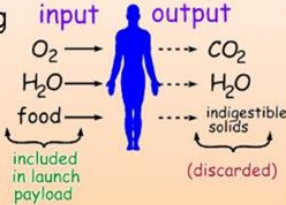
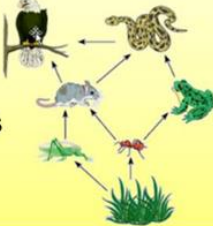

healthy lifestyle ~ maintain good health

- 45% of questions involve **diagrams**



Which letter shows runoff?

What does Aristo need to know to pass the exam?

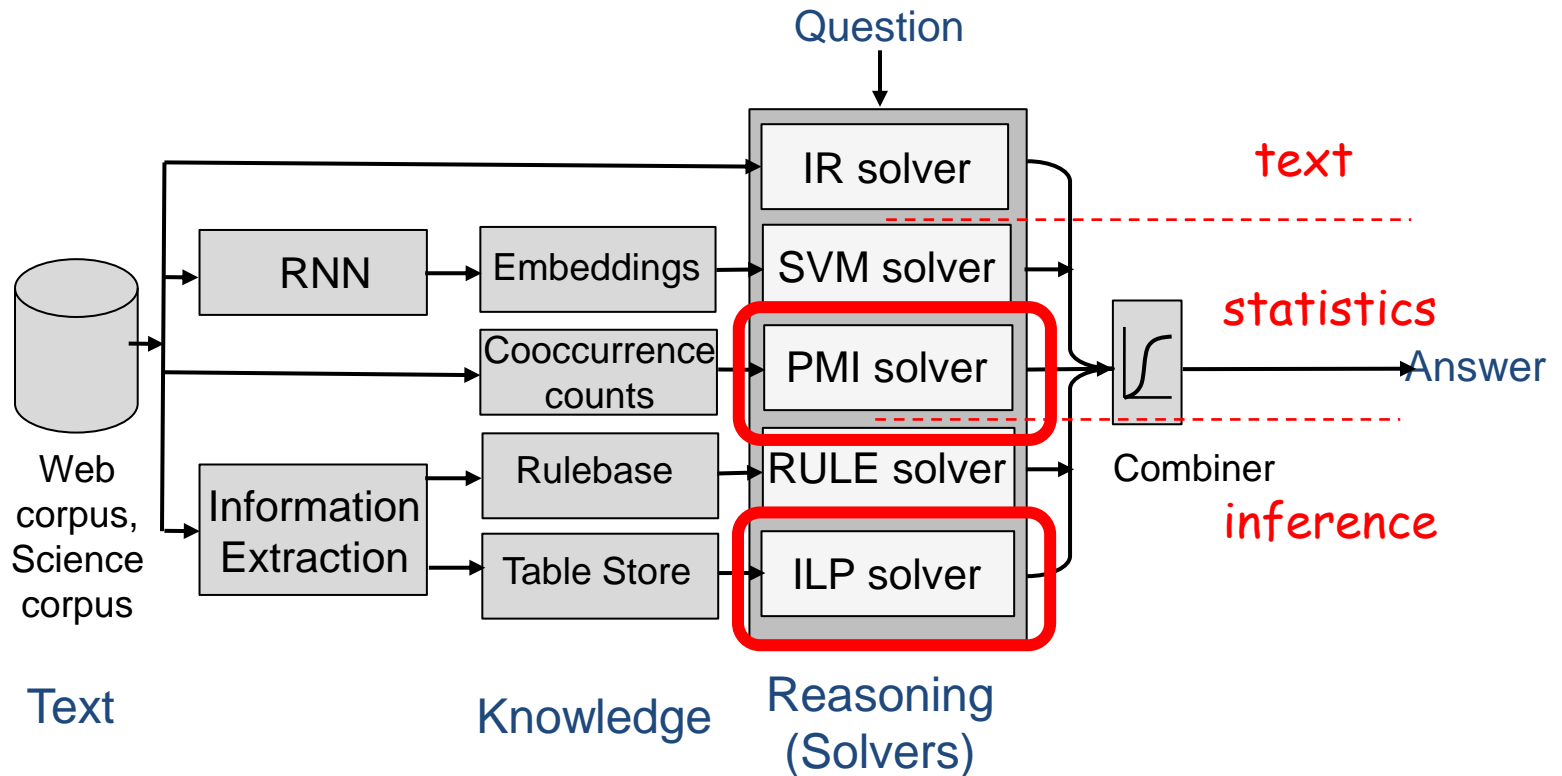
<p>Celestial Phenomena</p> <p>sun moon stars day/night, rotation revolution</p> 	<p>The Earth</p> <p>air water land weather precipitation erosion</p> 	<p>Matter</p> <p>solid/liquid/gas properties conductivity texture temperature measuring tools</p> 	<p>Energy</p> <p>forms energy transfer heat electricity chemical energy energy conversion</p> 
<p>Forces</p> <p>gravity magnetism force friction pull/pushing attraction</p> 	<p>Living things</p> <p>living nonliving characteristics animals plants fish</p> 	<p>Inheritance</p> <p>inherited traits resemblance acquired traits learned traits body features skills</p> 	<p>The Environment and Adaptation</p> <p>senses habitats behavior camouflage survival</p> 
<p>Continuity of Life</p> <p>life cycle life span offspring reproduction coloration mating</p> 	<p>Life Functions</p> <p>breathing growing eating food air water</p> <p>input</p> <p>output</p> <p>O_2 → CO_2</p> <p>H_2O → H_2O</p> <p>food → indigestible solids</p> <p>included in launch payload (discarded)</p> 	<p>Interdependence</p> <p>food web producers consumers decomposers predators prey</p> 	<p>Human Impact</p> <p>human activities environment ecosystem pollution conservation deforestation</p> 

Evolving Design of Aristo

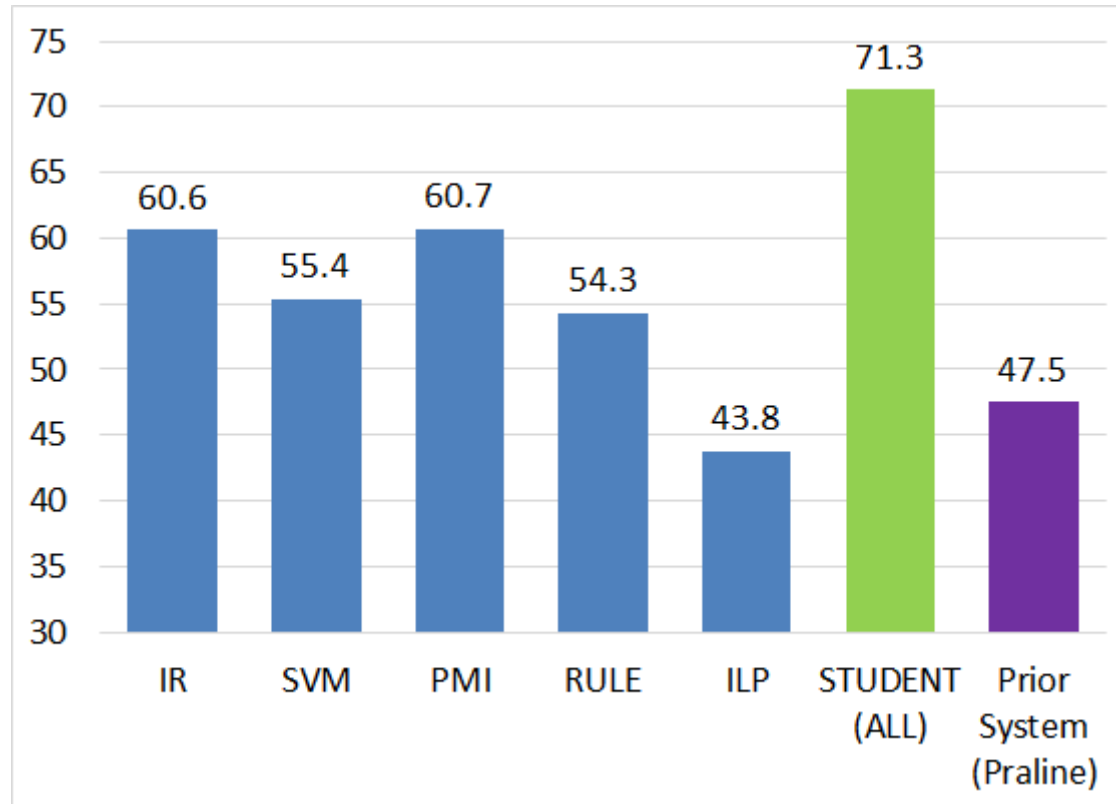
- **EMNLP 2015:** MLN “solver” operating on machine-extracted rules



- **2016:** Several levels of structure



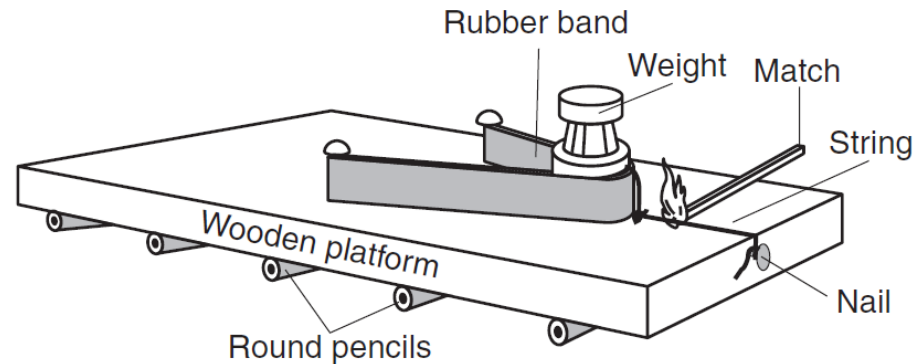
Current Results (test score on ND Questions)



Aristo

8th Grade – A Significantly Increased Challenge

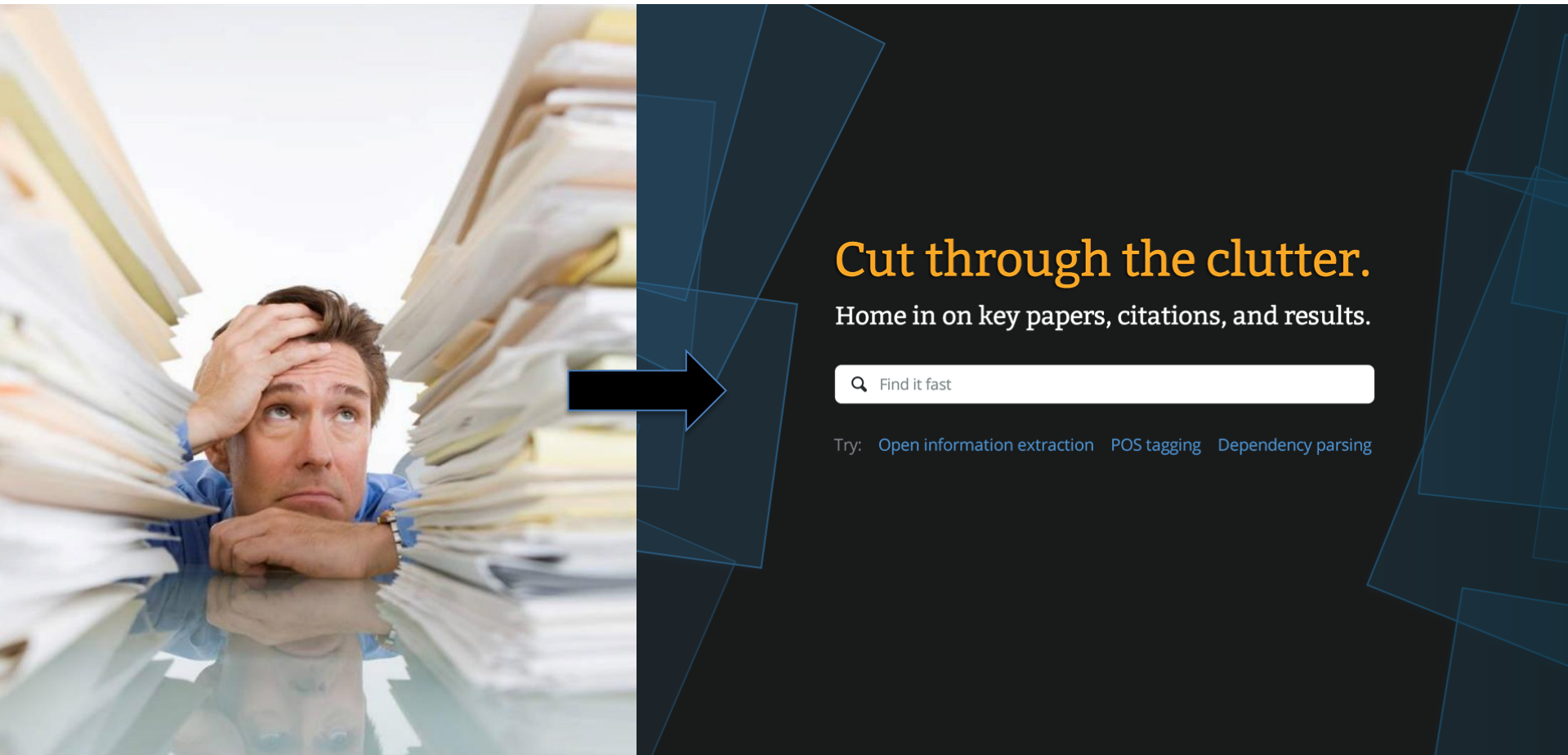
- 4th to 8th grade: major jump in challenges: 4 years of education and life experience for a student
- Majority of questions use diagrams
- Some are “AI hard”, e.g. (in 2006 8th grade exam)



Source: Adapted from, Constantine Constant, *Earth Science Workbook*, AMSCO, 1972

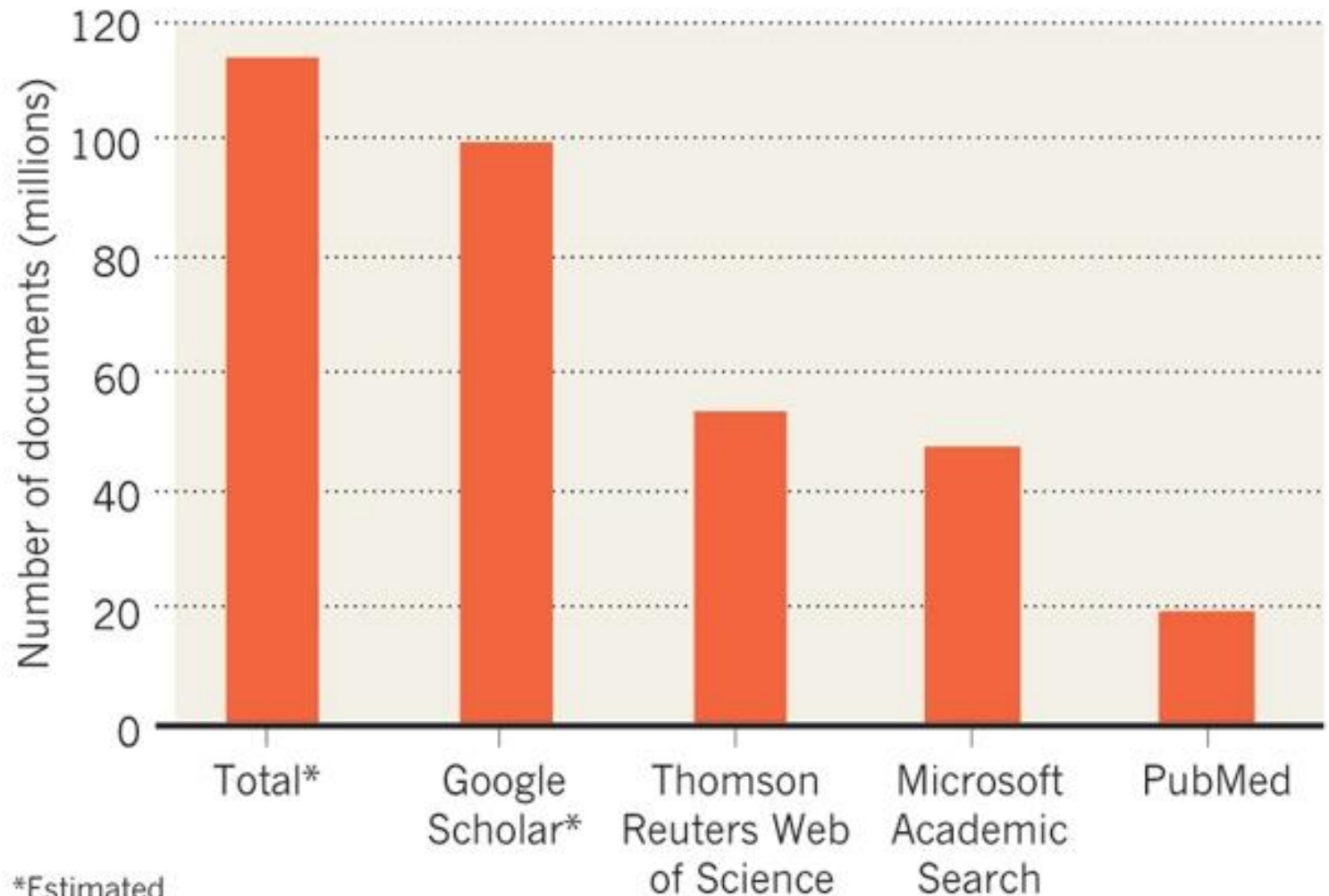
57 On the diagram above, draw an arrow to represent the direction the **wooden platform** will move when the lit match burns through the string and the weight is propelled from the platform. [1]

III. Semantic Scholar

A composite image illustrating the transition from information clutter to a search solution. On the left, a man in a blue shirt is buried up to his chest in a towering, chaotic stack of papers, looking stressed with his hand on his forehead. A large black arrow points from this scene towards the right, where a dark-themed search interface is displayed. The interface features the headline "Cut through the clutter." in orange, followed by the subtext "Home in on key papers, citations, and results." Below this is a white search bar with a magnifying glass icon and the placeholder text "Find it fast". At the bottom of the interface, there is a "Try:" section with links for "Open information extraction", "POS tagging", and "Dependency parsing". The background of the interface is dark with faint blue geometric shapes.

THE WEB OF SCHOLARSHIP

Around 114 million English-language scholarly documents (including papers, books and technical reports) can be found on the web.





"information extraction"



Scholar

About 139,000 results (0.12 sec)

Articles

Case law

My library

[PDF] [Maximum Entropy Markov Models for Information Extraction and Segmentation.](#)

[A McCallum](#), [D Freitag](#), [FCN Pereira](#) - ICML, 2000 - [courses.ischool.berkeley.edu](#)

Page 1. 1 Maximum Entropy Markov Models for **Information Extraction** and Segmentation Andrew McCallum, Dayne Freitag, and Fernando Pereira ... Named entity recognition: <ORG>Mips</ORG> Vice President <PRS>John Hime</PRS> – **Information extraction**: ...

Cited by 1126 Related articles All 50 versions Cite Save More

Any time

Since 2014

Since 2013

Since 2010

Custom range...

[Incorporating non-local information into information extraction systems by gibbs sampling](#)

[JR Finkel](#), [T Grenager](#), [C Manning](#) - ... of the 43rd Annual Meeting on ..., 2005 - [dl.acm.org](#)

Abstract Most current statistical natural language processing models use only local features so as to permit dynamic programming in inference, but this makes them unable to fully account for the long distance structure that is prevalent in language use. We show how to ...

Cited by 1129 Related articles All 23 versions Cite Save

Sort by relevance

Sort by date

- include patents
- include citations

[PDF] [Learning dictionaries for information extraction by multi-level bootstrapping](#)

[E Riloff](#), [R Jones](#) - AAAI/IAAI, 1999 - [aaai.org](#)

Abstract **Information extraction** systems usually require two dictionaries: a semantic lexicon and a dictionary of extraction patterns for the domain. We present a multilevel bootstrapping algorithm that generates both the semantic lexicon and extraction patterns simultaneously. ...

Cited by 753 Related articles All 18 versions Cite Save More

[PDF] [Open information extraction for the web](#)

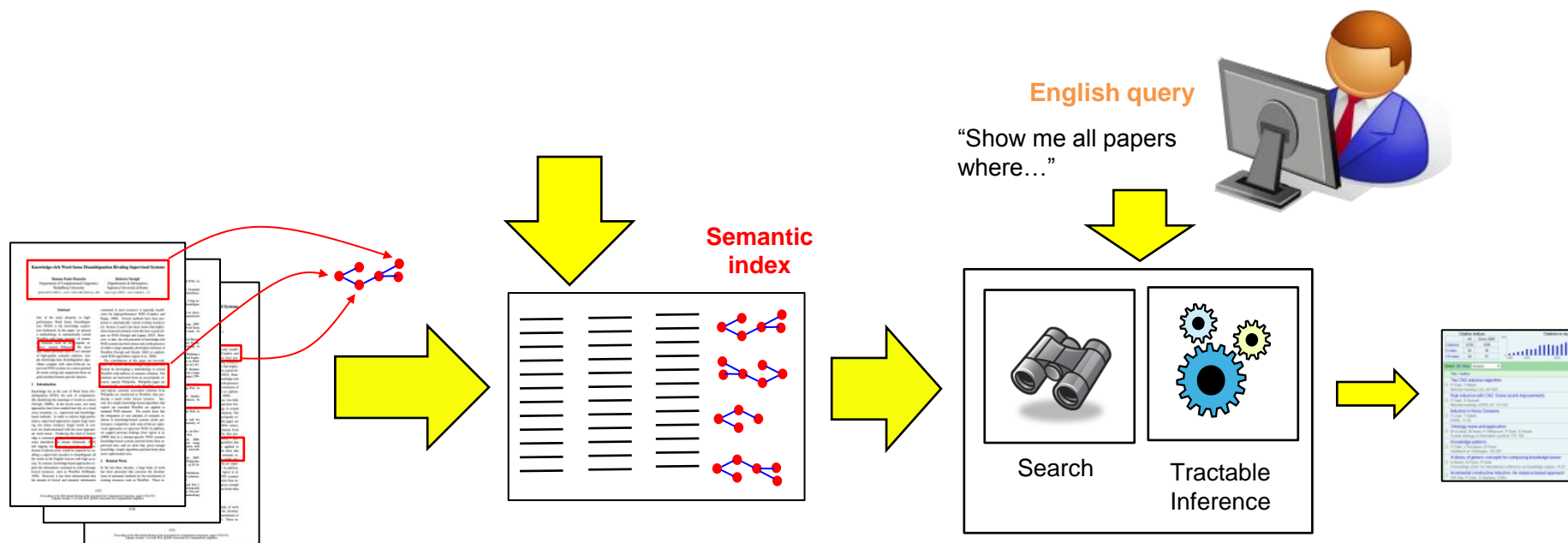
[M Banko](#), [MJ Cafarella](#), [S Soderland](#), [M Broadhead](#)... - IJCAI, 2007 - [aaai.org](#)

Abstract Traditionally, **Information Extraction** (IE) has focused on satisfying precise, narrow, pre-specified requests from small homogeneous corpora (eg, extract the location and time of seminars from a set of announcements). Shifting to a new domain requires the user to ...

Cited by 806 Related articles All 31 versions Cite Save More

Create alert

Leverage AI to Combat Information Overload



DEMO

Our Approach to Figure Understanding

Relation Extraction with Matrix Factorization and Universal Schemas

Sebastian Riedel
Department of Computer Science
University College London
s.riedel@ucl.ac.uk

Limin Yao, Andrew McCallum, Benjamin M. Marlin
Department of Computer Science
University of Massachusetts at Amherst
{lmayao, mcallum, marlin}@cs.umass.edu

Abstract

Traditional relation extraction predicts relations within some fixed and finite target schema. Machine learning approaches to this task require either manual annotation or, in the case of distant supervision, existing structured sources of the same schema. The need for existing datasets can be avoided by using a *universal schema*: the union of all involved schemas (surface forms predicted as in OpenIE, and relations in the schemas of pre-existing databases). This schema has an almost unlimited set of relations (due to surface forms), and supports integration with existing structured data (through the relation types of existing databases). To populate a database of such schema we present matrix factorization models that learn latent feature vectors for entity types and relations. We show that such latent models achieve substantially higher accuracy than a traditional classification approach. More importantly, by operating simultaneously on relations observed in text and in pre-existing structured DBs such as Freebase, we are able to reason about unstructured and structured data in mutually-supporting ways. By doing so our approach outperforms state-of-the-art supervision.

1 Introduction

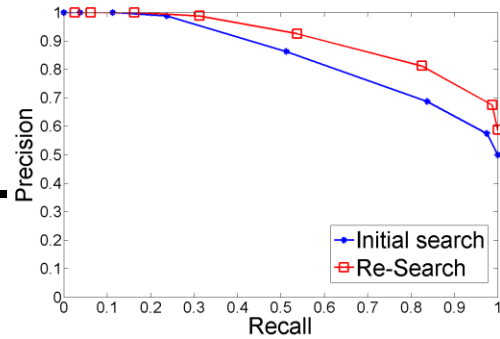
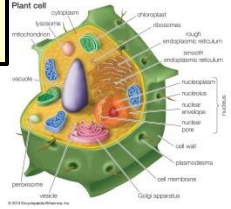
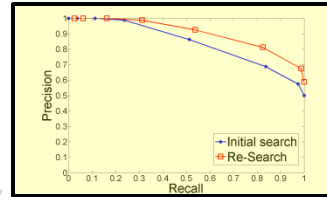
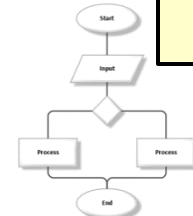
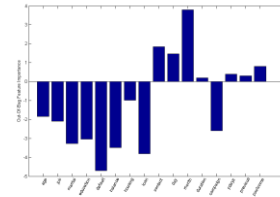
Most previous work in relation extraction uses a predefined, finite and fixed schema of relation types (such as *born-in* or *employed-by*). Usually some textual data is labeled according to this schema, and this labeling is then used in supervised training of an automated relation extractor, e.g. Colotta and Soares (2004). However, labeling textual rela-

tions is time-consuming and difficult, leading to significant recent interest in distant-supervised learning. Here one aligns existing database records with the sentences in which these records have been "rendered"—effectively labeling the text—and from this labeling we can train a machine learning system as before (Cohen and Kamlic, 1999; Mintz et al., 2009; Bunesco and Mooney, 2007; Riedel et al., 2010). However, this method relies on the availability of a large database that has the desired schema. The need for pre-existing datasets can be avoided by using language itself as the source of the schema. This is the approach taken by OpenIE (Etzioni et al., 2008). Here surface patterns between mentions of concepts serve as relations. This approach requires no supervision and has tremendous flexibility, but lacks the ability to generalize. For example, OpenIE may find *FERGUSON-historian-at-HARVARD* but does not know *FERGUSON-is-a-professor-at-HARVARD*. OpenIE has traditionally relied on a large diversity of textual expressions to provide good coverage. But this diversity is not always available, and, in any case, the lack of generalization greatly inhibits the ability to support reasoning.

One way to gain generalization is to cluster textual surface forms that have similar meaning (Lin and Pantel, 2001; Pantel et al., 2007; Yates and Etzioni, 2009; Yao et al., 2011). While the clusters discovered by all these methods usually contain semantically related items, closer inspection invariably shows that they do not provide reliable implications. For example, a typical representative cluster may include *historian-at-professor-at-scientist-worked-at*. Although these relation types are indeed semantically related, note that *scientist* does not necessarily imply *professor*, and *worked-at*

74

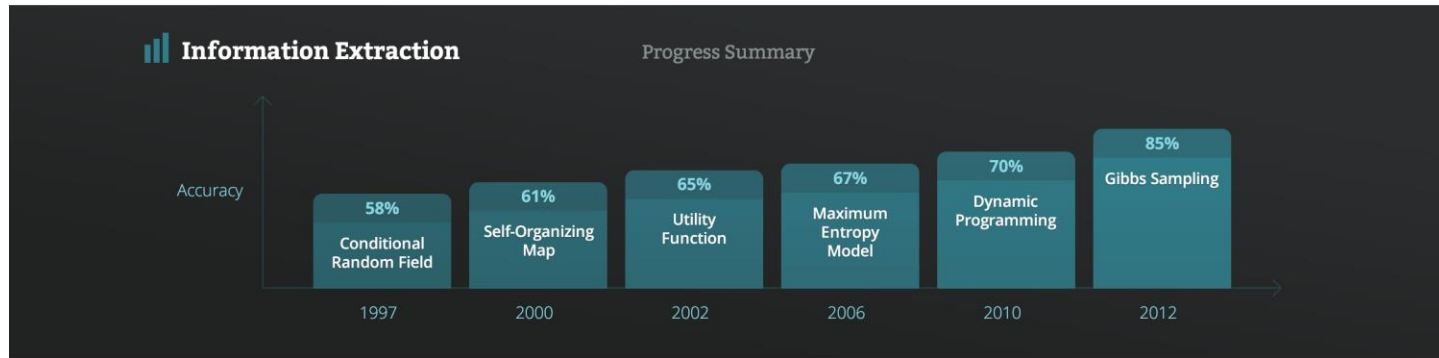
Proceedings of SIGACL/BET 2013, pages 74-84.
Atlanta, Georgia, 9-14 June 2013. ©2013 Association for Computational Linguistics



```
{
  "type": "Figure",
  "number": 4,
  "page": 6,
  "dpi": 100,
  "width": 827,
  "height": 1159,
  "caption": [464, 400, 691, 416],
  "caption": "Figure 4: predicted label probabilites.",
  "image": [427, 85, 729, 386],
  "image": [427, 85, 729, 386],
  "rotations": [
    {"rotation": 0, "text": "[622.016, 88.298, 645.698, 97.0056]", "text": "Label"},
    {"rotation": 0, "text": "[648.12, 88.298, 700.152, 97.0056]", "text": "probabilites"},
    {"rotation": 0, "text": "[702.573, 88.298, 718.71, 97.0056]", "text": "(%)"},
    {"rotation": 0, "text": "[445.871, 99.3676, 465.601, 108.075]", "text": "True"},
    {"rotation": 0, "text": "[468.822, 99.3676, 491.704, 108.075]", "text": "Label"},
    {"rotation": 0, "text": "[583.826, 99.3676, 623.46, 108.075]", "text": "Predicted"},
    {"rotation": 0, "text": "[642.26, 99.4838, 649.524, 108.143]", "text": "..."},
    {"rotation": 0, "text": "[651.946, 99.3676, 661.632, 108.075]", "text": "10"},
    {"rotation": 0, "text": "[664.853, 99.3676, 673.739, 108.075]", "text": "11"},
    {"rotation": 0, "text": "[676.16, 99.3676, 685.846, 108.075]", "text": "12"},
    {"rotation": 0, "text": "[690.689, 99.4838, 697.954, 108.143]", "text": "..."},
    {"rotation": 0, "text": "[702.797, 99.3676, 712.483, 108.075]", "text": "13"},
    {"rotation": 0, "text": "[714.904, 99.3676, 724.59, 108.075]", "text": "18"},
    {"rotation": 0, "text": "[445.871, 111.107, 456.041, 119.766]", "text": "..."},
    {"rotation": 0, "text": "[437.987, 122.176, 443.449, 130.835]", "text": "="},
    {"rotation": 0, "text": "[445.871, 122.176, 490.929, 130.835]", "text": "PastHistory"},
    {"rotation": 0, "text": "[493.351, 122.176, 509.487, 130.835]", "text": "[11]"},
    {"rotation": 0, "text": "[520.442, 122.176, 533.886, 130.835]", "text": "age"},
    {"rotation": 0, "text": "[583.826, 122.176, 628.885, 130.835]", "text": "PastHistory"},
    {"rotation": 0, "text": "[656.79, 122.176, 661.632, 130.835]", "text": "0"},
    {"rotation": 0, "text": "[664.854, 122.176, 673.74, 130.835]", "text": "95"},
    {"rotation": 0, "text": "[681.004, 122.176, 685.847, 130.835]", "text": "00"},
    {"rotation": 0, "text": "[707.64, 122.176, 712.483, 130.835]", "text": "00"},
    {"rotation": 0, "text": "[719.747, 122.176, 724.59, 130.835]", "text": "00"},
    {"rotation": 0, "text": "[437.987, 133.246, 443.449, 141.905]", "text": "="},
    {"rotation": 0, "text": "[445.871, 133.246, 490.929, 141.905]", "text": "PastHistory"},
    {"rotation": 0, "text": "[493.351, 133.246, 509.487, 141.905]", "text": "[11]"},
    {"rotation": 0, "text": "[520.442, 133.246, 530.128, 141.905]", "text": "63"},
    {"rotation": 0, "text": "[583.826, 133.246, 628.885, 141.905]", "text": "PastHistory"}
  ]
}
```

.json

Progress Summary in an Topic



FILTER RESULTS

CLASSIFICATION

- Survey
- Experimental
- Theoretical
- Software

YEAR

yyyy to yyyy

VENUES (21)

- HLT
- NAACL
- Workshop On Automatic Summarization
- ACL
- Workshop On Operational Factors In Practical, Robust Anaphora Resolution For Unrestricted Tests

6,962 results

Sort by: Relevance

Incorporating Non-Local Information Into Information Extraction Systems By Gibbs Sampling

Jenny Rose Finkel, Trond Grenager, Christopher D. Manning / ACL / 2005

Cited by 166 / Abstract / View PDF / Add to reading list

structure that is prevalent in language use. We show how to solve this dilemma with Gibbs sampling **information extraction** task. We show 10 runs of Gibbs sampling in the same CRF...

On-Demand Information Extraction

Satoshi Sekine / ACL / 2006

Cited by 22 / Abstract / View PDF / Add to reading list

At present, adapting an **Information Extraction** system to new topics is an expensive and slow for each new topic. We propose a new paradigm of **Information Extraction** which operates 'on demand

Multidocument Summarization Via Information Extraction

Michael White, Tanya Korelsky, and 4 others / HLT / 2006

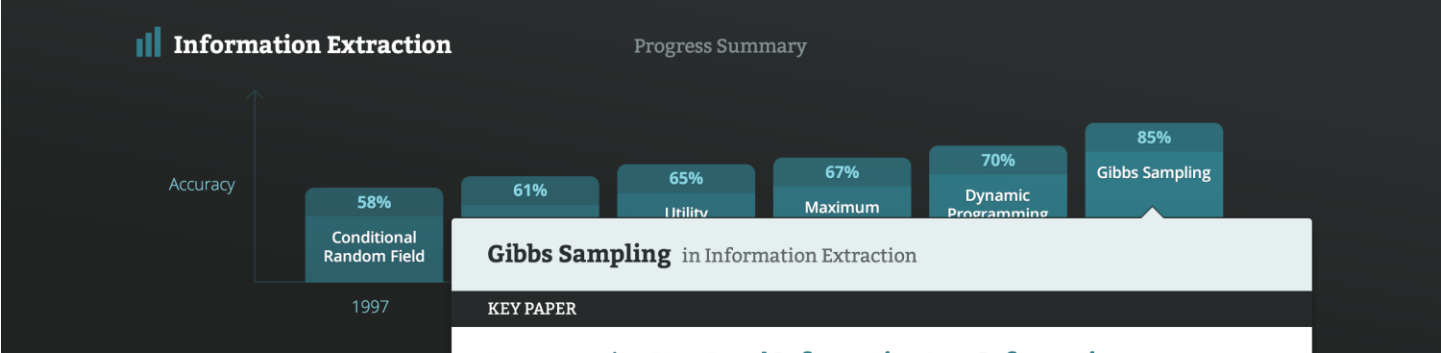
Cited by 11 / Abstract / View PDF / Add to reading list

We present and evaluate the initial version of RIPTIDES, a system that combines **information information extraction, extraction**-based summarization, and natural language generation to...

Confidence Estimation For Information Extraction

Aron Culotta, Andrew McCallum / NAACL / 2002

Provenance with Text Understanding



FILTER RESULTS

CLASSIFICATION

- Survey
- Experimental
- Theoretical
- Software

YEAR

yyyy to yyyy

VENUES (21)

- HLT
- NAACL
- Workshop On Automatic Summarization
- ACL
- Workshop On Operational Factors In Practical, Robust Anaphora Resolution For Unrestricted Tests

Gibbs Sampling in Information Extraction

KEY PAPER

Incorporating Non-Local Information Into Information Extraction Systems By Gibbs Sampling

Jenny Rose Finkel, Trond Grenager, Christopher D. Manning / ACL / 2012

An illustration of the effectiveness of **Gibbs sampling**, compared to Viterbi inference, for the two tasks addressed in this paper: the CoNLL named entity recognition task **which returned an accuracy rate of 85.54%**, and the CMU Seminar Announcements **information extraction** task. We show 10 runs of **Gibbs sampling** in the same CRF model that was used for Viterbi. For each run the sampler was initialized to a random sequence, and used a linear annealing schedule that sampled the complete sequence 1000 times. CoNLL performance is measured as per-entity, and CMU Seminar. Announcements performance is measured as per-token.

On-Demand Information Extraction

Satoshi Sekine / ACL / 2006

Cited by 22 / Abstract / View PDF / Add to reading list

At present, adapting an **Information Extraction** system to new topics is an expensive and slow for each new topic. We propose a new paradigm of **Information Extraction** which operates 'on demand

Multidocument Summarization Via Information Extraction

Michael White, Tanya Korelsky, and 4 others / HLT / 2006

Cited by 11 / Abstract / View PDF / Add to reading list


We present and evaluate the initial version of RIPTIDES, a system that combines **information information extraction, extraction-based summarization**, and natural language generation to...

Confidence Estimation For Information Extraction






Aron Culotta, Andrew McCallum / NAACL / 2002

Cut through the clutter.

Home in on key papers, citations, and results.

 gibbs sampling

gibbs sampling overview of applications

-  **gibbs sampling** in *information extraction*
162 papers, 41.7 average rank
-  **gibbs sampling** in *dependency parsing*
105 papers, 18.71 average rank
-  **gibbs sampling** in *parsing*
159 papers, 32.4 average rank
-  **gibbs sampling** in *machine translation*
163 papers, 30.24 average rank
-  **gibbs sampling** in *POS tagging*
87 papers, 23.31 average rank

Applications of a Technique

Gibbs Sampling

Applications for this technique



FILTER RESULTS

CLASSIFICATION

- Survey
- Experimental
- Theoretical
- Software

YEAR

yyyy to yyyy

VENUES (15)

ACL

Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data

EMNLP

CoNLL

NAACL

429 results

Sort by: Relevance

Incorporating Non-Local Information Into Information Extraction Systems By Gibbs Sampling

Jenny Rose Finkel, Trond Grenager, Christopher D. Manning / ACL / 2012

Cited by 166 / Abstract / View PDF / Add to reading list

structure that is prevalent in language use. We show how to solve this dilemma with Gibbs sampling information extraction task. We show 10 runs of Gibbs sampling in the same CRF...

Not-So-Latent Dirichlet Allocation: Collapsed Gibbs Sampling Using Human Judgments

Jonathan Chang / Proceedings of the NAACL HLT 2010 Workshop on Creating Speech ... / 2010

Cited by 1 / Abstract / View PDF / Add to reading list

Probabilistic topic models are a popular tool for the unsupervised analysis of text, providing both ... and cluster that annotation. This task simulates the **sampling** step of the collapsed **Gibbs** sampler

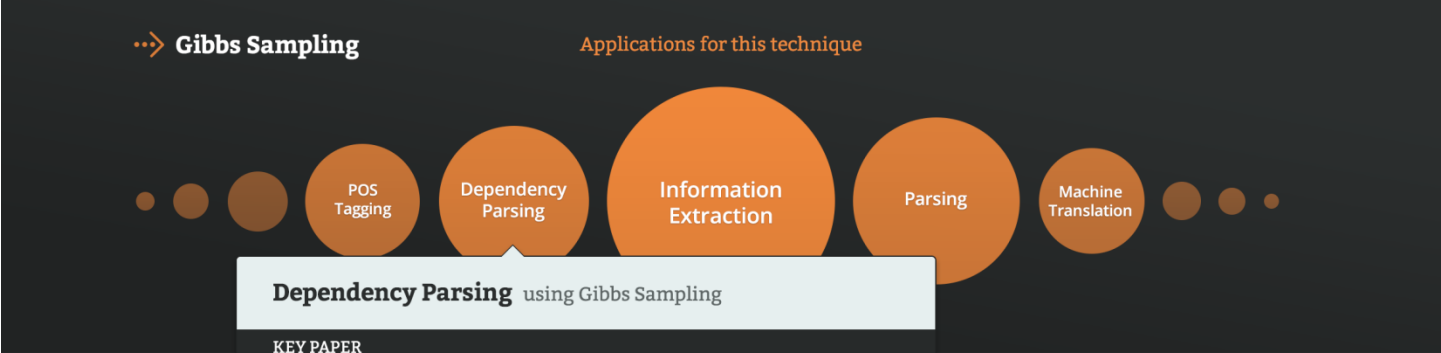
Sampling Alignment Structure under a Bayesian Translation Model

John DeNero, Alexandre Bouchard-Côté, Dan Klein / EMNLP / 2008

Cited by 31 / Abstract / View PDF / Add to reading list

We describe the first tractable **Gibbs sampling** procedure for estimating phrase pair frequencies
Abstract We describe the first tractable **Gibbs sampling** procedure for estimating phrase pair

Provenance with Diagram Understanding



Dependency Parsing using Gibbs Sampling

KEY PAPER

Unsupervised Dependency Parsing using Reducibility and Fertility features
David Marecek, Zdeněk Zabokrtsky / NAACL / 2012

Inference	CoNLL	Seminars
Viterbi	85.51	91.85
Gibbs	85.54	91.85
Sampling	85.51	91.85
	85.49	91.85
	85.51	91.85
	85.51	91.85
	85.51	91.85
	85.51	91.85
	85.51	91.85
	85.51	91.86
Mean	85.51	91.85
Std. Dev.	0.01	0.004

- FILTER RE
- CLASSIFIC
- Survey
- Exper
- Theor
- Softw
- YEAR
- yyyy

- Sort by:
- Into Information Extraction**
g / ACL / 2012
- to solve this dilemma with Gibbs sampling
sampling in the same CRF...
- Collapsed Gibbs Sampling Using**
orkshop on Creating Speech ... / 2010

- VENUES (15)
- ACL
- Proceedings of the NAACL
HLT 2010 Workshop on
Creating Speech and
Language Data
- EMNLP
- CoNLL
- NAACL

Cited by 1 / [Abstract](#) / [View PDF](#) / [Add to reading list](#)

Probabilistic topic models are a popular tool for the unsupervised analysis of text, providing both ... and cluster that annotation. This task simulates the **sampling** step of the collapsed **Gibbs** sampler

Sampling Alignment Structure under a Bayesian Translation Model
John DeNero, Alexandre Bouchard-Côté, Dan Klein / EMNMP / 2008

Cited by 31 / [Abstract](#) / [View PDF](#) / [Add to reading list](#)

We describe the first tractable **Gibbs sampling** procedure for estimating phrase pair frequencies
Abstract We describe the first tractable **Gibbs sampling** procedure for estimating phrase pair

Semantic Scholar to Launch in 2015

Sign up for notification here:

allenai.org/semantic-scholar.html



“It's the **absence** of AI technologies that is **already** killing people.”

The Semantic Scholar Vision

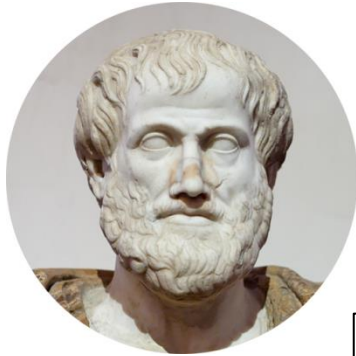
“What if a cure for an intractable cancer is hidden within the tedious reports on thousands of clinical studies? In 20 years' time, AI will be able to read — and more importantly, understand — scientific text. These AI readers will be able to connect the dots between disparate studies to identify novel hypotheses and suggest experiments which would otherwise be missed. AI-based discovery engines will help to find the answers to science's thorniest problems and ultimately revolutionize science.”

Allen Institute for Artificial Intelligence

[Wired Magazine, September, 2015](#)

AI2 Core Projects:

Science QA



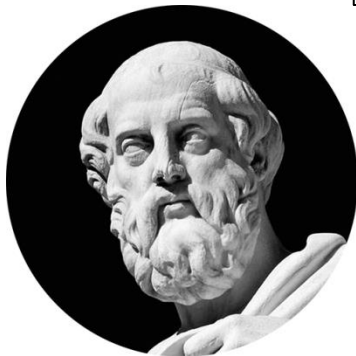
Aristo

Semantic
Search over
papers



**Semantic
Scholar**

Knowledge
from images
& diagrams



Plato

SAT Math QA



Euclid



1. Grand Challenge Problems
2. Data-driven rather than single mechanism
3. Ambitious goals, but measurable progress
4. **Idea:** augment Turing Test with battery of standardized tests to measure AI progress
5. Semantic Scholar = AI to help Scientists

The Future of AI: The View from AI2

Oren Etzioni

www.allenai.org

The View from AI2 (allenai.org)

