

Frame-Semantic Parsing

Nathan Schneider, University of Edinburgh

May 31, 2015 ■ FrameNet Tutorial at NAACL-HLT

FrameNet + NLP = <3

- We want to develop systems that understand text
- Frame semantics and FrameNet offer a linguistically & computationally satisfying theory/representation for semantic relations

Frame-semantic Parsing

SemEval Task 19 [Baker, Ellsworth, & Erk 2007]

- Given a text sentence, analyze its frame semantics. Mark:
 - ▶ words/phrases that are **lexical units**
 - ▶ **frame** evoked by each LU
 - ▶ **frame elements** (role–argument pairings)
- Analysis is in terms of groups of tokens.
No assumption that we know the syntax.

FrameNet SRL, Parsing: Early Work

- The original SRL paper actually used FrameNet (Gildea & Jurafsky 2002).
Also Thompson et al. 2003 (w/ frame ID),
Fleischman et al. 2003, Padó & Lapata 2005,
Erk & Padó 2006, Matsubayashi et al. 2009,
Fürstenauf & Lapata 2009.
- SemEval 2007 shared task (Baker et al. 2007): full-text annotations.
Best system by Johansson & Nugues.

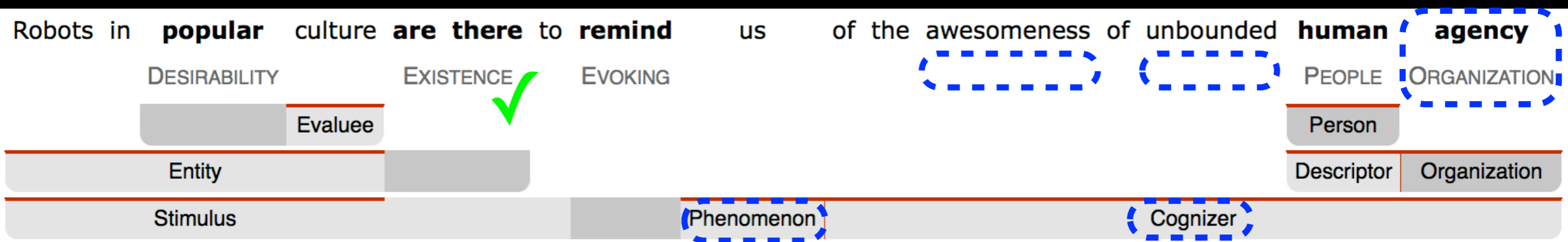
SEMAFOR

[Das, Schneider, Chen, & Smith 2010]

Robots in **popular** culture **are there** to **remind** us of the awesomeness of unbounded **human agency**

SEMAFOR

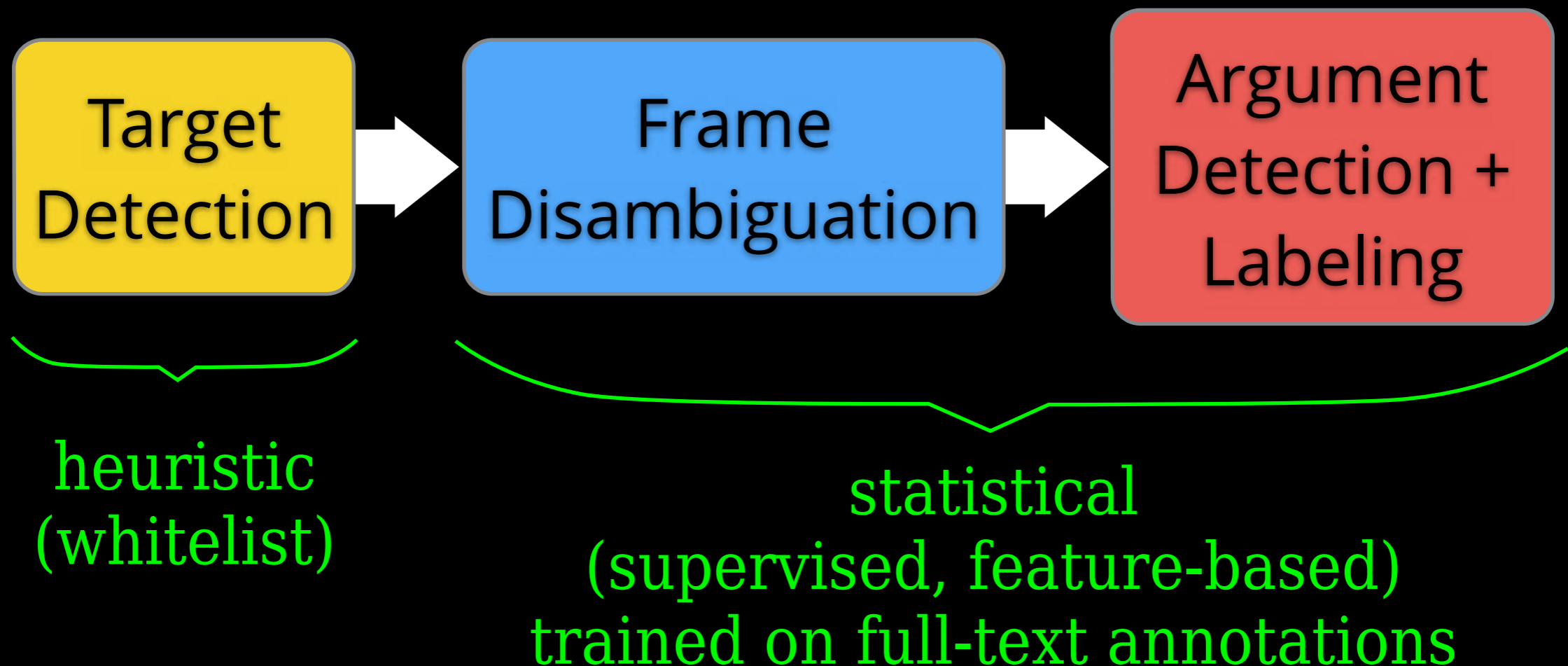
[Das, Schneider, Chen, & Smith 2010]



The SEMAFOR Pipeline

[Das, Schneider, Chen, & Smith 2010]

Preprocessing: syntactic dependency parser



Full-text Annotations

+ American National Corpus Texts

1. [Berlitz History of Greece](#)
2. [Berlitz History of Jerusalem](#)
3. [Berlitz History of Las Vegas](#)
4. [Berlitz Intro of Dublin](#)
5. [Berlitz Intro of Hong Kong](#)
6. [Berlitz Intro of Jamaica](#)
7. [Berlitz What to Do in Hong Kong](#)
8. [Berlitz Where to Go in Hong Kong](#)
9. [Children's home fund-raising letter](#)
10. [Children's home fund-raising letter](#)
11. [Goodwill fund-raising letter](#)
12. [Goodwill fund-raising letter](#)
13. [Goodwill fund-raising letter](#)
14. [Goodwill fund-raising letter](#)
15. [Goodwill fund-raising letter](#)
16. [Goodwill fund-raising letter](#)
17. [journal christine](#)
18. [journal patrick](#)
19. [journal ryan](#)
20. [journal.pbio.0020001](#)
21. [Slate magazine article: Entrepreneur as Madonna](#)
22. [Slate magazine article: Stephanopoulos Crimes](#)

+ AQUAINT Knowledge-Based Evaluation Texts

+ LUCorpus-v0.3

+ Miscellaneous

+ Texts from Nuclear Threat Initiative website, created by Center for Non-Proliferation Studies

+ Wall Street Journal Texts from the PropBank Project

<https://framenet.icsi.berkeley.edu/fndrupal/index.php?q=fulltextIndex>

Full-text Annotations

1. **Stephanopoulos** Analyzes His Own **CRIME**_{Committing_crime}
2. **THERE**_{Locative_relation} was **FORMER**_{Time_vector} **Clinton** aide **George Stephanopoulos** on **ABC 's This Week this morning** ,
furrow-browed and `` **HEARTBROKEN**_{Emotion_directed} with all the **EVIDENCE**_{Evidence} coming out '' against the
PRESIDENT_{Leadership} . Last **WEEK**_{Calendric_unit} , **WHEN**_{Temporal_collocation} the **Lewinsky** story was only **A**_{Quantified_mass}
FEW_{Quantified_mass} **HOURS**_{Measure_duration} **OLD**_{Age} , **Stephanopoulos** popped up on **Good Morning America** to demonstrate
his **CONCERN**_{Emotion_directed} . `` These are **PROBABLY**_{Likelihood} the most **SERIOUS**_{Importance} **ALLEGATIONS**_{Statement} yet
leveled against the **PRESIDENT**_{Leadership} . **THERE**_{Existence} 'S_{Existence} no **QUESTION**_{Point_of_dispute} that , if they 're true ,
they ... could **LEAD**_{Causation} to impeachment proceedings . ''

SEMAFOR

[Das, Schneider, Chen, & Smith 2010]

- SEMAFOR's models consist of features over **observable** parts of the sentence (words, lemmas, POS tags, dependency edges & paths) that may be predictive of **frame/role labels**
- Full-text annotations as training data for (semi)supervised learning
- Extensive body of work on semantic role labeling [starting with Gildea & Jurafsky 2002 for FrameNet; also much work for PropBank]

SEMAFOR

[Das, Schneider, Chen, & Smith 2010]

- State-of-the-art performance on SemEval'07 evaluation (outperforms the best system from the task, Johansson & Nugues 2007)
- On SE07: [F] 74% [A] 68% [F→A] 46%
On FN1.5: [F] 91% [A] 80% [F→A] 69%
[Das et al. 2014]
- BUT: This task is really hard. Room for improvement at all stages.

SEMAFOR Demo

<http://demo.ark.cs.cmu.edu/parse>

So Amelia Bedelia sat right down
and she drew those drapes.



How to improve?

- Better modeling with current resources
- Ways to use non-FrameNet resources
- Create new resources?



Karl Moritz
Hermann



Oscar
Täckström



Dipanjan Das

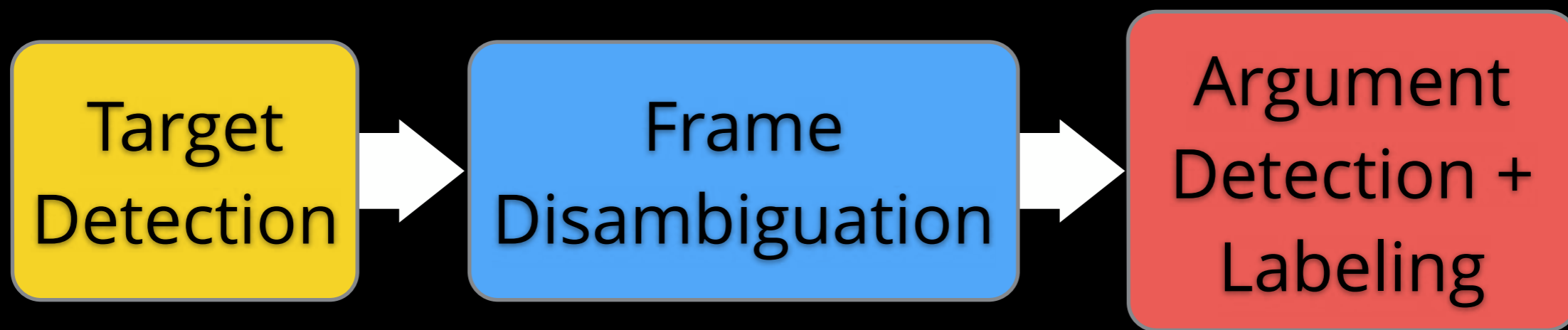


Sam
Thomson

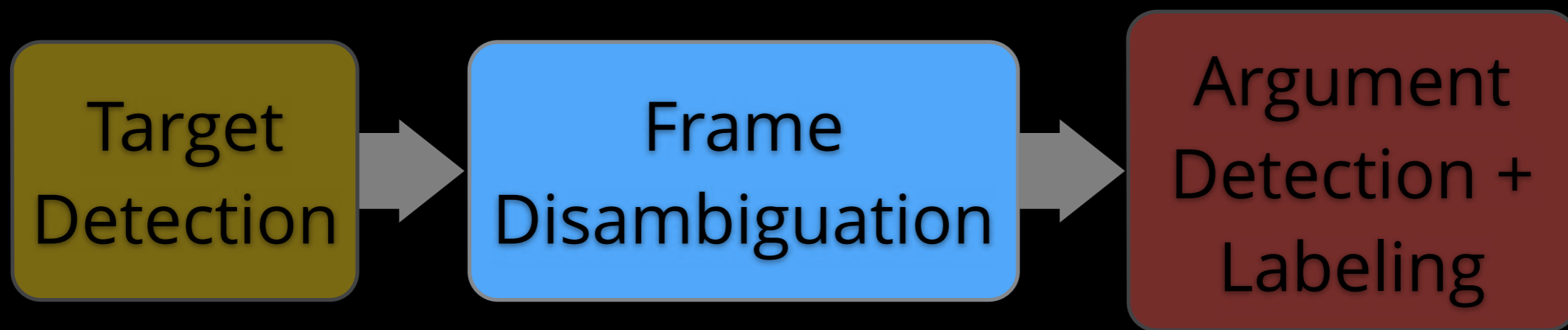


Meghana
Kshirsagar

Advances in Modeling



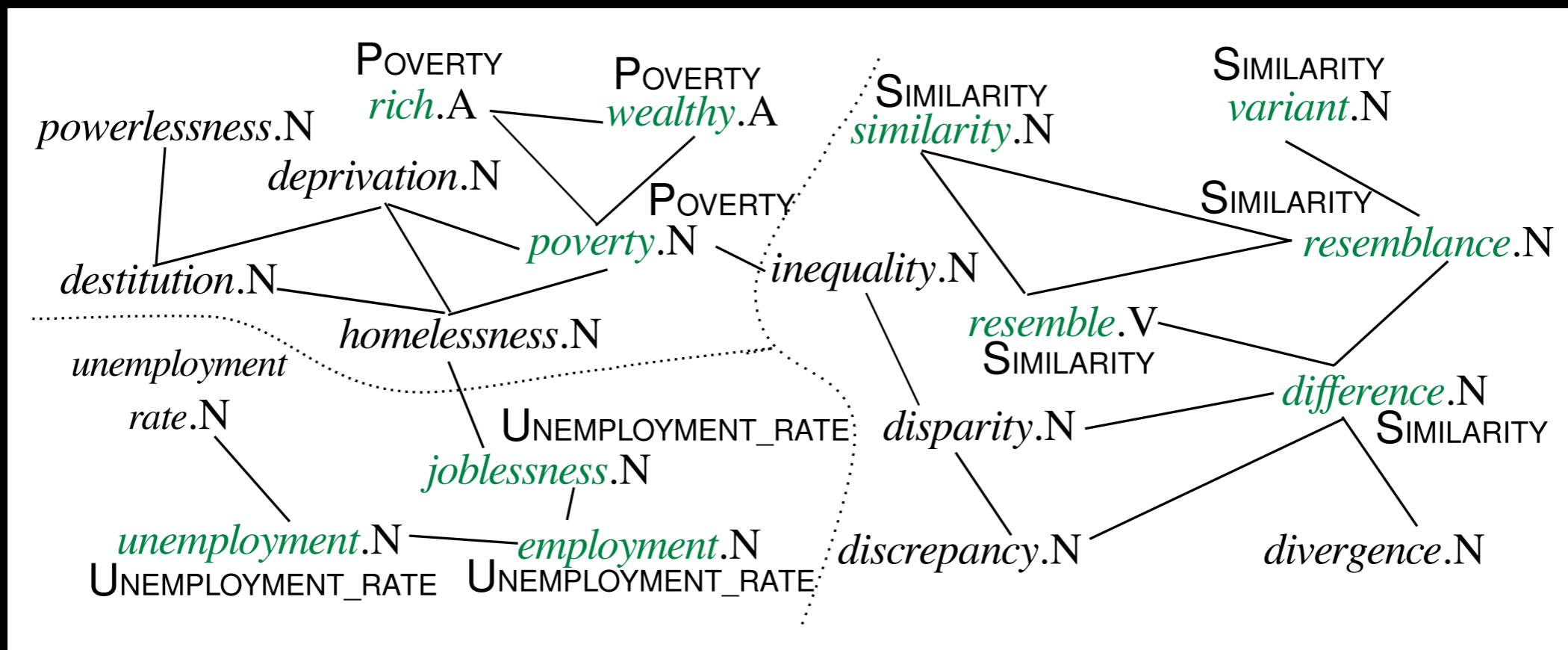
Advances in Modeling



Unknown Predicates

- **Problem:** Many frame-evoking predicates are seen neither in lexicon nor training data. How, then, to assign the correct frame?
- **Solution:** Propagate frame labels from known predicates to unknown predicates in a similarity graph. [Das & Smith 2011, 2012 / 2014]

Unknown Predicates



Word Representations

- **Problem:** With little training data, many features are too infrequent to be useful—particularly for rare/unseen words.
- **Solution:** Learn word embeddings that are predictive of frame labels (neural network).
[Hermann et al. 2014]

Advances in Modeling



Constraints on Argument Combinations

- **Problem:** A frame's arguments should not overlap, but this means classification decisions are not independent.
 - Also, some frames define hard Requires/Excludes constraints over role pairs.
- **Solution 1:** Beam search (approximate).
[Das et al. 2010 / 2014]
- **Solution 2:** Dual decomposition (exact).
[Das et al. 2012 / 2014]
- **Solution 3** (Google's variant of SEMAFOR): Label arguments with dynamic programming. [Täckström et al. 2015]

Constraints on Argument Combinations

Agent
Austria , once expected to **waltz** smoothly into the European Union , is elbowing its **partners** ,
Self_mover Manner Manner Goal **CONDUCT** Partner_1 Partner_2
treading on toes and pogo-dancing in a most un-Viennese **manner** .

Conclusion

- SEMAFOR system from CMU has been applied to tasks as diverse as stock prediction and spoken dialogue segmentation

<http://www.ark.cs.cmu.edu/SEMAFOR/>

<http://demo.ark.cs.cmu.edu/parse>

- Ongoing research at CMU, Google, & elsewhere!

Baker, C., Ellsworth, M., & Erk, K. (2007). SemEval-2007 Task 19: Frame Semantic Structure Extraction. SemEval.

Das, D., Chen, D., Martins, A. F. T., Schneider, N., & Smith, N. A. (2014). Frame-semantic parsing. *Computational Linguistics* 40(1), 9–56.

Das, D., Martins, A. F. T., & Smith, N. A. (2012). An exact dual decomposition algorithm for shallow semantic parsing with constraints. *SEM.

Das, D., Schneider, N., Chen, D., & Smith, N. A. (2010). Probabilistic frame-semantic parsing. NAACL-HLT.

Das, D., & Smith, N. A. (2011). Semi-supervised frame-semantic parsing for unknown predicates. ACL.

Das, D., & Smith, N. A. (2012). Graph-based lexicon expansion with sparsity-inducing penalties. NAACL-HLT.

Erk, K., & Padó, S. (2006). Shalmaneser - a toolchain for shallow semantic parsing. LREC.

Fleischman, M., Kwon, N., & Hovy, E. (2003). Maximum entropy models for FrameNet classification. EMNLP.

Fürstenauf, H., & Lapata, M. (2009). Semi-supervised semantic role labeling. EACL.

- Gildea, D., & Jurafsky, D. (2002). Automatic labeling of semantic roles. *Computational Linguistics*, 28(3), 245–288.
- Hermann, K. M., Das, D., Weston, J., & Ganchev, K. (2014). Semantic frame identification with distributed word representations. ACL.
- Johansson, R., & Nugues, P. (2007). LTH: semantic structure extraction using nonprojective dependency trees. SemEval.
- Kwon, N., Fleischman, M., & Hovy, E. (2004). FrameNet-based semantic parsing using maximum entropy models. Coling.
- Matsubayashi, Y., Okazaki, N., & Tsujii, J. (2009). A comparative study on generalization of semantic roles in FrameNet. ACL-IJCNLP.
- Padó, S., & Lapata, M. (2005). Cross-linguistic projection of role-semantic information. HLT/EMNLP.
- Täckström, O., Ganchev, K., & Das, D. (2015). Efficient inference and structured learning for semantic role labeling. *TACL*, 3, 29–41.
- Thompson, C. A., Levy, R., & Manning, C. D. (2003). A generative model for semantic role labeling. ECML.