Real-World Causal Relationship Discovery from Text

Constantine Lignos, Chester Palen-Michel, Oskar Singer, Pedro Szekely, Elizabeth Boschee
University of Southern California – Information Sciences Institute

USC Viterbi School of Engineering
Information Sciences Institute

Overview

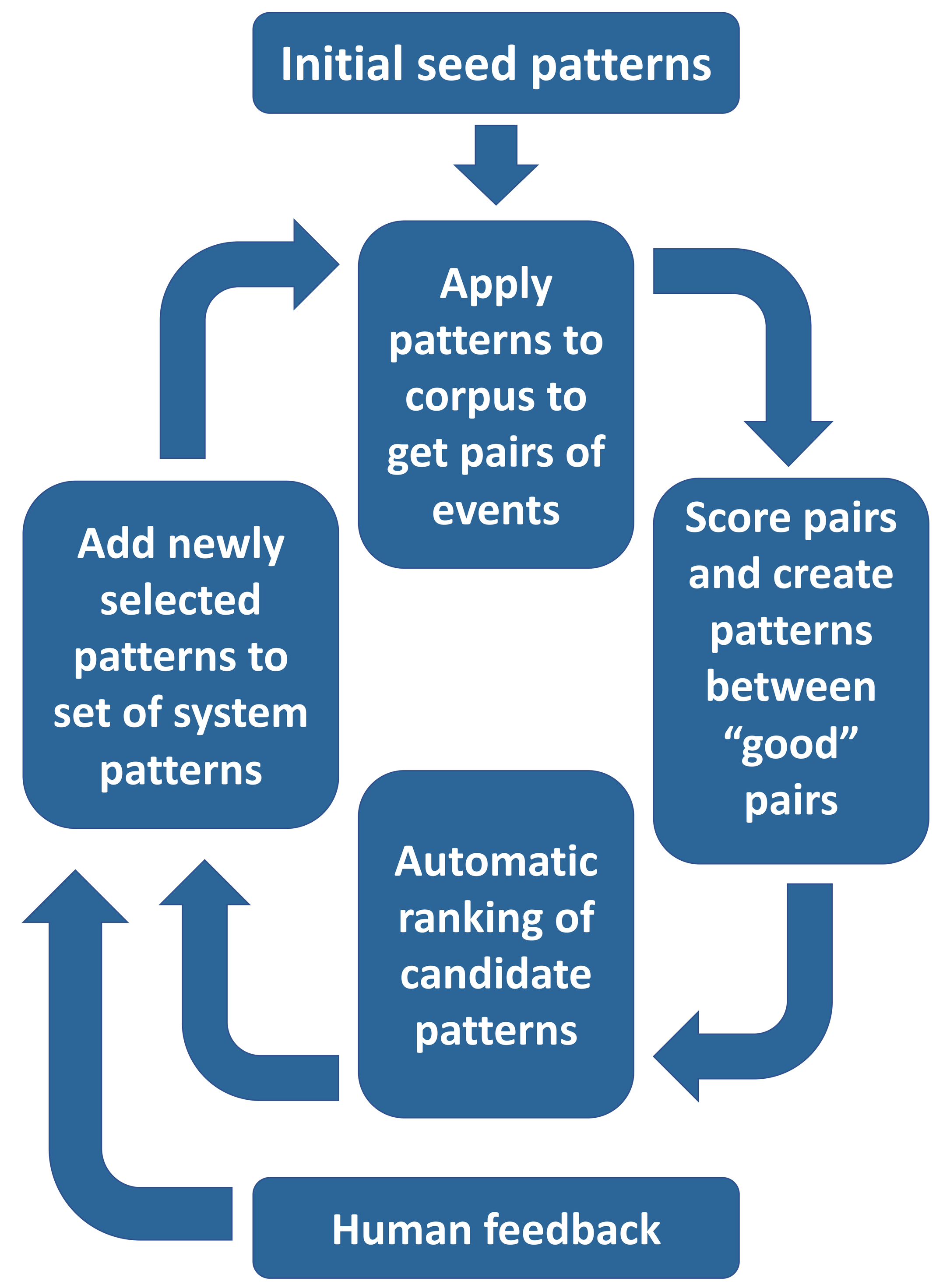
Hasn't this been done before?

- Causality extraction from text has been previously explored in both unsupervised and supervised settings **but…**
- rarely in the context of a real-world application
- We trained a simple model trained on natural language processing (NLP) community 2010 shared task [6]
  - Achieved a precision of 0.88 and recall of 0.91 on the shared task data
  - Precision plummeted to < 0.10, when run directly on a broad sample of newswire documents

Approach:

- A bootstrapping approach to causal pattern discovery
- A flexible end-to-end document processing pipeline that integrates a variety of customized and off-the-shelf NLP components

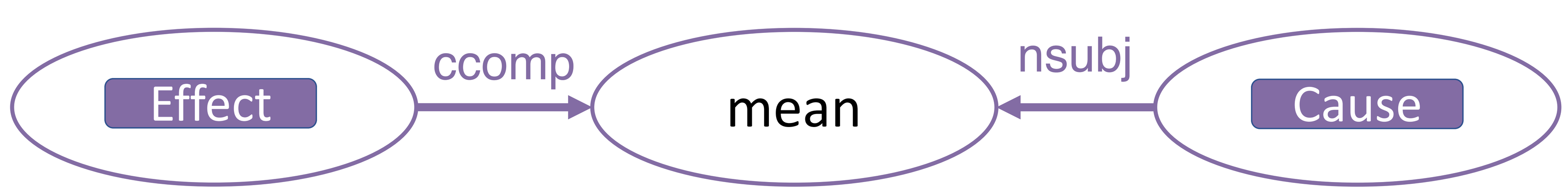
Causal Relationship Discovery



- Initial seed patterns from simple causal sentences
  - *The protests caused violence* → X cause Y
- Apply to corpus for pairs of events
  - (cause=*bombing;* effect=*fatalities*)
- Use ontological types
  - Ontology developed for DARPA's Causal Exploration program, which contains ~500 event types
  - E.g. (cause=*Attack;* effect=*Death)*
- Score pairs by combination of measures
  - Combining a measure of the "causality" of each individual word
  - The word pair together
  - Same for their ontological classes
- Pairs with scores above an experimentally-tuned threshold are deemed causally-likely
- Patterns are ranked by how causally likely the pairs they extract are
- Human feedback to filter best patterns

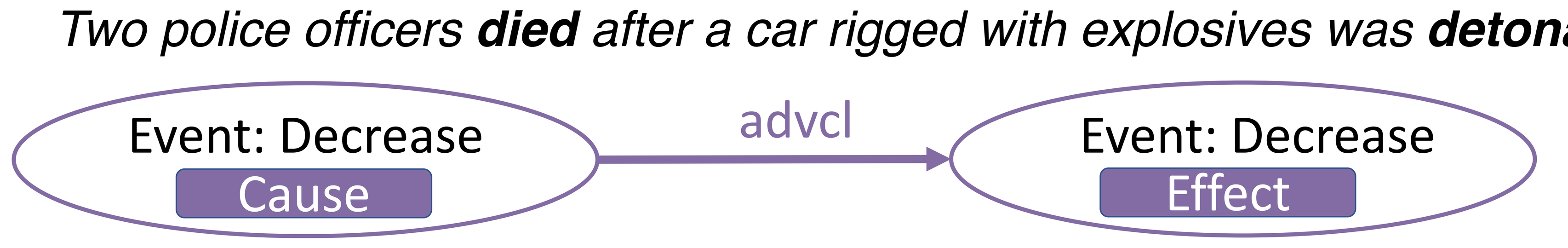
Evaluation

Below is a selection of patterns suggested by our system at an early iteration, along with examples of causal relationships they extract



*Safina's **ranking** means the Williams sisters will **be** in the same half of the draw.*

*Two police officers **died** after a car rigged with explosives was **detonated**.*



*Automakers are expected to **reduce** vehicle production by 25 percent from last year, when auto sales **fell** 18 percent from 2007 levels.*

Examples:

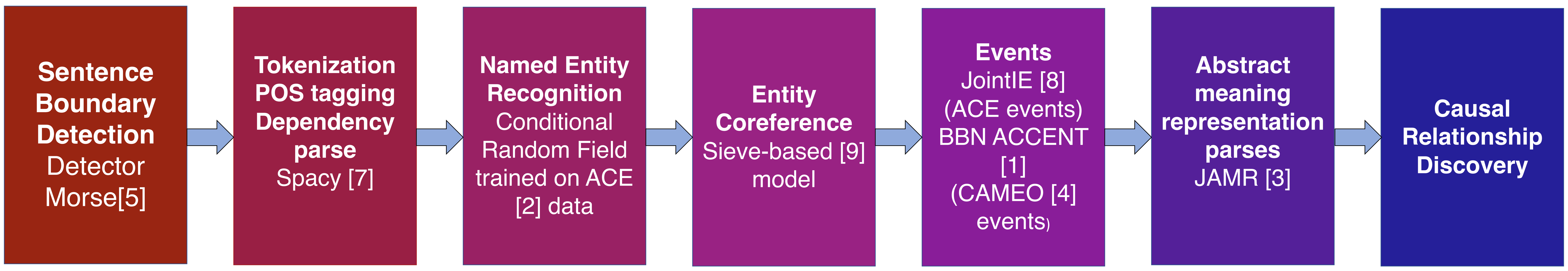
- The first and second are good, high-precision additions to our collection
  - The first applying to all possible pairs
  - The second restricted by ontological classes (Attack/Death)
- The third, however, is too general
  - It frequently produces false positives where the two events are merely coincident
  - For example: *Domestic traffic fell 4.6 percent while international trade fell 11.2 percent*

System performance:

- Evaluated our system on 1,000 newswire documents
- Baseline of 660 seed patterns developed by pattern-writing experts over several months
- Even on top of this strong baseline, our iteratively-discovered patterns improve recall by 6.6% while maintaining the precision of the seed patterns (70%)

End to End System

NLP pipeline to provide rich structure for applying causal patterns



New, flexible Python NLP framework (ISI VistaNLP) to integrate all components

- Implements named entity extraction, coreference, and causal pattern discovery components; integrates third-party components for additional tasks
- Enables simultaneous representation of multiple analyses of a document (e.g. events extracted by multiple systems using different ontologies)
- Allows merging of information extracted by all components to a single representation using a single ontology

References

1. Boschee, E., Lautenschlager, J., O'Brien, S., Shellman, S., Starz, J., Ward, M.: BBN ACCENT event coding evaluation. In: ICEWS Coded Event Data. Harvard Dataverse (2015), https://doi.org/10.7910/DVN/28075/GBAGXI
2. Doddington, G.R., Mitchell, A., Przybocki, M.A., Ramshaw, L.A., Strassel, S.M., Weischedel, R.M.: The automatic content extraction (ACE) program-tasks, data, and evaluation. In: Proc. LREC. vol. 2, p. 1 (2004).
3. Flanigan, J., Thomson, S., Carbonell, J., Dyer, C., Smith, N.A.: A discriminative graph-based parser for the abstract meaning representation. In: Proc. ACL. vol. 1, pp. 1426{1436 (2014)
4. Gerner, D.J., Schrodt, P.A., Yilmaz, O., Abu-Jabr, R.: Conict and mediation event observations (CAMEO): A new event data framework for the analysis of foreign policy interactions (2002), presented at the Annual Meeting of the ISA
5. Gorman, K.: DetectorMorse (2014), https://pypi.org/project/DetectorMorse/
6. Hendrickx, I., Kim, S.N., Kozareva, Z., Nakov, P., Seaghdha, D.O., Pado, S., Pennacchiotti, M., Romano, L., Szpakowicz, S.: SemEval-2010 task 8: Multi-way classification of semantic relations between pairs of nominals. In: Proc. International Workshop on Semantic Evaluation. pp. 33{38 (2010)
7. Honnibal, M., Montani, I.: spaCy 2: Natural language understanding with Bloom embeddings, convolutional neural networks and incremental parsing (2017)
8. Li, Q., Ji, H., Huang, L.: Joint event extraction via structured prediction with global features. In: Proc. ACL. vol. 1, pp. 73{82 (2013)
9. Raghunathan, K., Lee, H., Rangarajan, S., Chambers, N., Surdeanu, M., Jurafsky, D., Manning, C.: A multi-pass sieve for coreference resolution. In: Proc. EMNLP. pp. 492{501 (2010)
10. Riloff, E., Jones, R.: Learning dictionaries for information extraction by multi-level bootstrapping. In: Proc. AAAI. pp. 474{479 (1999)
11. Riloff, E., Jones, R.: A retrospective on mutual bootstrapping. AI Magazine 39(1), 51-61 (2018)

 This material is based upon work supported by United States Air Force under Contract No. FA8650-17-C-7715.