Natural Language Access to Big Data: Papers from the 2014 AAAI Fall Symposium

Discovering and Characterizing Emerging Events in Big Data (DISCERN)

Bonnie J. Dorr, Milenko Petrovic, James F. Allen, Choh Man Teng, Adam Dalton

Florida Institute for Human and Machine Cognition
{bdorr,mpetrovic,jallen,cmteng,adalton}@ihmc.us

Abstract

We describe a novel system for discovering and characterizing emerging events. We define *event emergence* to be a developing situation comprised of a series of sub-events. To detect sub-events from a very large, continuous textual input stream, we use two techniques: (1) frequency-based detection of sub-events that are potentially entailed by an emerging event; and (2) anomaly-based detection of other sub-events that are potentially indicative of an emerging event. Identifying emerging events from detected sub-events involves connecting sub-events to each other and to the relevant emerging events within the event models and estimating the likelihood of possible emerging events. Each sub-event can be part of a number of emerging events and supports various event models to varying degrees. We adopt a coherent and compact model that probabilistically identifies emerging events. The innovative aspect of our work is a well-defined framework where statistical Big Data techniques are informed by event semantics and inference techniques (and vice versa). Our work is strongly grounded in semantics and knowledge representation, which enables us to produce more reliable results than would otherwise be possible with a purely statistical approach.

1 Introduction and Background

Broad-scale detection and characterization of emerging events in streams of textual data (e.g., a natural disaster, a new scientific breakthrough, or a terrorist event) is a vital technology for the advancement of "Big Data." This is an important capability for a wide range of applications such as analysis of news stories, understanding of scientific papers, analysis of social interactions, and building medical records. In fact, any application that requires the processing of large textual data could very well benefit from the detection and characterization of emerging events.

Our team is investigating the development of a novel system for DIScovering and Characterizing EmeRging eveNts (DISCERN). Our approach transcends shallow-analysis techniques for event tracking and detection, such as those employed in the Rich Textual Entailment initiative (Dagan, Glickman, and Magnini, 2006), in that it applies deep semantic processing and inferencing to an unstructured textual input stream and is able to detect emerging events that may not be explicitly stated.

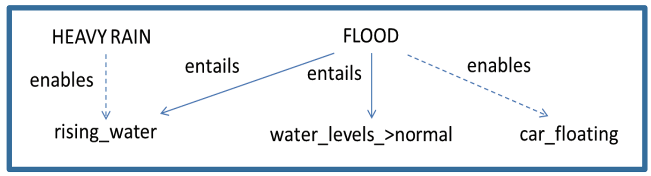

Figure 1: Sub-Events Associated with Heavy Rain and Flood

Consider the following two examples:

(1) *The heavy rains and rising water led to flooding conditions.*

(2) *I just saw a car floating down the street.*

We define *event emergence* to be a developing situation (e.g., flooding) that is potentially comprised of a series of *sub-events* (e.g., rising water). In (1), the *flooding conditions* are a superordinate event—referred to by some (e.g., Pustejovsky, 2013) as a "container"—for the sub-event *rising water* (which is enabled by the sub-event *heavy rains*). This necessitates a hierarchical organization of predicate argument representations. In (2), the *car floating down the street* is another sub-event that is an indicator of flooding conditions, although flooding is not explicitly mentioned.

Figure 1 depicts these two types of sub-events for the examples above, with solid lines used to link sub-events that are definitionally entailed by an emerging event, and dotted lines used to link sub-events that are indicators enabled by an emerging event.[1]

The notion of emergence may rely on a distributed setting, constructed from sub-events that occur in a stream of

[1] For illustration, only the links relevant to this discussion are shown in Figure 1. The figure is not complete (e.g. the link for "Heavy Rain enables Flood" is not shown) and it contains implicit links that could be inferred via constraints on the enables/entails relationships. in the graph.

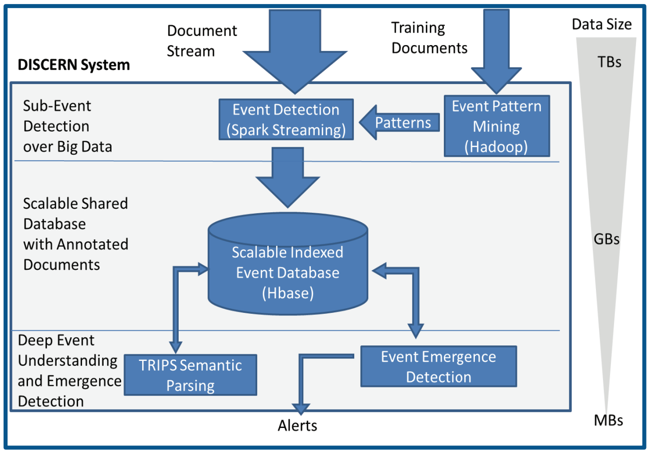

Figure 2: DISCERN System Architecture

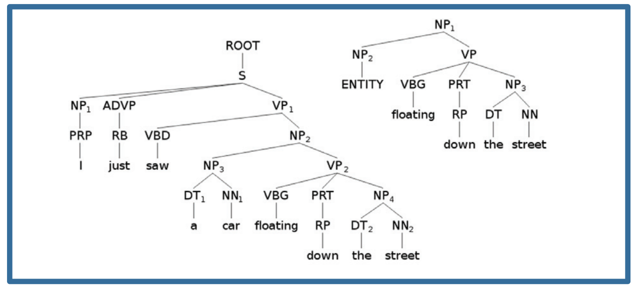

Figure 3: Example of a parse tree for a sub-event candidate, with an event-argument pattern indicated above.

seemingly uncoordinated and independent events. For example, one might imagine a social-media interaction that contains evidence that a *flooding event* may be emerging:

A: I am watching TV... ... It's raining really hard...
B: Anyone going to the party?... Water is rising much faster than I expected.
C: The pizza is burnt... There is a car floating out on the street!...

We address the following challenges:

- Automatic detection of (explicitly mentioned) sub-events from an input stream and filtering the content for downstream emerging event discovery.
- Discovery and characterization of (potentially implicit) emerging events through deep semantic linking of the sub-events detected above.

We focus on the extraction of knowledge from both verbs and other linguistic constructs such as resultatives, adjectivals, nominalizations, and pronominals (for event coreference)—as well as compositional structures that defy standard phrase-level approaches. Our approach enables a deeper level of understanding based on event semantics and hierarchical knowledge that enables detection of implicitly conveyed emerging events.

The overall architecture for the DISCERN system is shown in Figure 2.

2 Broad-Scale Detection of Sub-Events for Event Emergence Detection

We expect *event emergence* to arise when there is a sharp increase in the degree to which specific sub-events occur. Detecting sub-events and their compositional relationships in large volumes of streaming text requires: (1) characterization of a large variety of syntactically- and semantically-grounded sub-event patterns; (2) tracking the large number of such patterns over time and computing aggregated statistics to identify outliers.

We adopt a *hierarchical pattern-based approach* to extracting potential sub-events of emerging events, producing predicate-argument structures, and assigning relationships between them. We automatically construct sub-event patterns that are later used for filtering input so that downstream inference processes will be applied to a smaller set of sub-events.

2.1 Pattern Construction

Three components of hierarchical patterns are: (1) Syntactic structures (the full syntactic tree shown in Figure 3) produced by parsing text fragments that potentially contain sub-event mentions; (2) "Trimmed" sub-structures (smaller syntactic tree shown above the full parse in Figure 3) to represent indicative sub-event patterns for anomalous events; and (3) paraphrastic variants associated with lexical and phrasal items (available, but not shown in Figure 3). Each of these is described in turn below.

Sub-event Pattern Construction: One challenge in constructing such patterns is that sentences often contain content not directly related to a potential sub-event. Such content obstructs discovery of the structure intrinsic to a particular event. For example, in the sentence *City officials agreed to fund flood relief efforts, the mayor said Tuesday*, the phrase *the mayor said Tuesday* does not contain any additional content related to the main sentence. In our previous work on text summarization (Zajic et al., 2008; Zajic et al., 2007; Qazvinian et al., 2013), we identified a set of rules that succinctly capture the structure of such "extraneous" content as patterns on dependency trees using a system called Trimmer. Removing these structures will simplify the sentence, while preserving the core syntax structures related to the event. In addition, event syntax structures need to be generalized into patterns Figure 3 shows a syntactic structure from our running example, *I just saw a car floating down the street.* The pattern resulting from Trimmer's removal of the extraneous content (*I just saw*) is indicated in the upper right corner.

Another challenge is that the Trimmer approach was originally designed to eliminate portions of the tree based on what the system considers to be syntactic periphery

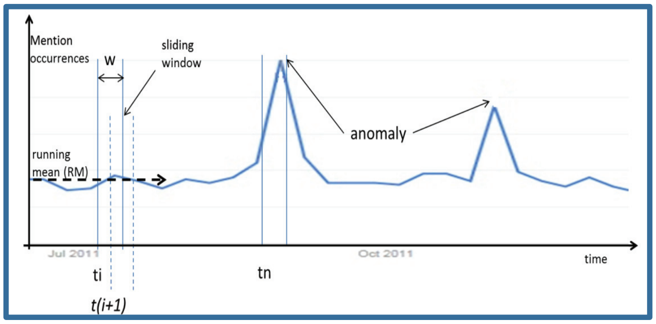

Figure 4: Anomaly detection in pattern/mention occurrences

(e.g., subordinate clauses are generally deleted), without taking into account the semantic contribution of these potentially useful sub-trees. Thus, a clause like *floating down the street* would automatically be deleted in our ongoing example in Figure 3. We extend Trimmer rule application to take semantic context into account, e.g., retaining the *car floating* sub-event, by leveraging anomaly detection—as described below.

A third challenge is the restrictive nature of lexicalized patterns, e.g., the pattern above for *car floating* contains lexical items *car* and *floating* – these need to be generalized, so that other similar patterns will be constructed, e.g., for *pig flying*. To address this, we use the patterns generated in the first step as a seed to find additional patterns in large text corpora. We expand the set of patterns from the previous step by mining for statistically similar patterns from large text corpora.

Anomaly-based Detection of Indicative Sub-event Patterns: Sub-events that are not entailed by an emerging event but are enabled by them are indicative in nature; these typically provide information on emerging events based on an anomalous number of mentions (as in the repeated mention of *rain* and *water* in our running example) or on an anomalous co-occurrence of terms (as in the *car floating* example above). An anomaly is detected when there is a spike in the number of such occurrences (Figure 4) (Vlachos et al., 2004). Such spikes are indicative of an anomaly, and Trimmer uses this as a feature for retention of that portion of the syntactic structure during the construction of sub-event patterns.

Incorporation of Parphrasatic Variants: We are investigating the use of the two paraphrasing techniques for capturing event-oriented relations including, among others: (1) coreference; (2) sub-event:

- Categorial Variation Detection to relate derivational variants that are paraphrastic, but are not necessarily the same part of speech (e.g., the verb *raining* and the noun *rain*—which may be leveraged to identify the raining activity as a potential indicator for the emerging *flood* event).

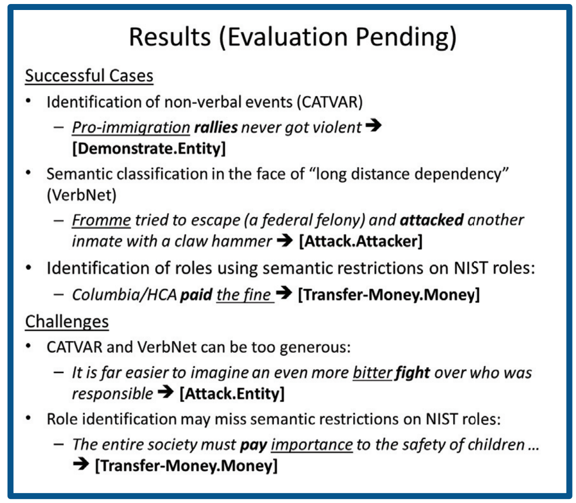

Figure 5: Preliminary results over NIST Event Evaluations Pilot dataset.

- Targeted paraphrasing to relate phrasal units that may be used in the same context (e.g., *heavy rains* and *rising water*—which may be leveraged to identify sub-event relationships between these two, ultimately leading to discovery of the sub-event relationship between *rising water* and *flood*).
  - We have extracted patterns associated with such relationships using the predicate-level annotations in ACE2005 and SemLink, as well as two additional sources described below from our prior work. Our approach transcends strictly verb-focused techniques that are inherent in most current event analysis frameworks (e.g., Ruppenhover et al., 2009; UzZaman, 2013). We extract knowledge not just from verbs, but from other syntactic renderings of events, e.g., nominalized events, adjectivals, and resultatives.

Even more robust paraphrase-based detections are possible using a combination of paraphrasing and categorial variations. For example, it is possible to tease apart semantic equivalence (event coreference) from other event-event relations. Consider the paraphrases below:

The flood killed 7 and injured 20
As a result of the flood, 7 died
Rising waters drowned 7 and wounded 20
The drowning of 7 people and injury of 20
Heavy rains pelted the city, killing 7 and injuring 20

In standard paraphrasing systems (Bannard and Callison-Burch, 2005), due to the similarity in argument structure and contextual usage, it easy to see how terms like *drowned* and *died* (and *injured* and *wounded*) would be equated paraphrastically. However, when these same features are used, there may be cases identified as paraphrastic events that are not an exact match, as one event may be a sub-portion of the other. Our approach involves extraction of approximate paraphrase patterns (akin to the high-recall/low-precision version of the paraphrase database at paraphrase.org) to provide fodder for the discovery of predicate-argument structures that express sub-event relations (*flood, rising waters; flood, heavy rains*) or causal relations (*flood, die; flood, drown; kill, die*). Categorial Variations are leveraged for additional features to relate terms, e.g., *injured* and *injury*.

As a preliminary step toward testing some of the components described, most notably, Categorial Variations, we applied the DISCERN system to the Event Detection problem in the recent NIST KBP Evaluation. Figure 5 summarizes our findings on the pilot data. The results of the final evaluations are currently pending NIST assessment.

2.2 Scalable Processing

Pattern Extraction: To support scalable pattern extraction, we combine structured-pattern detection with anomaly-based detection by introducing an explicit representation of event patterns which can be (1) mined from data and (2) efficiently extracted at scale using efficient structured pattern matching.

Pattern representations will affect scalability when there are a large number of patterns. As the number of patterns increases, the processing time correspondingly increases: each pattern needs to be evaluated on every tree. When the number of patterns is large and a tree matches only a few patterns, it is beneficial to index patterns so that we quickly access only those that apply to a given tree. This requires computing commonalities between multiple patterns.

Event-Relationship Graph: The main challenge in scaling up event emergence detection is achieving efficient access to a large number of sub-events and indicators. We make use of HBase (DB in Figure 2), a state-of-the-art semi-structured database for the Hadoop platform, which is based on BigTable data model (Chang et. al., 2008).

2.3 Related Work

Prior work on broad-scale sub-event detection has concentrated primarily on two areas: complex event processing (Wu et al., 2006; Cugola et al., 2012; Eckert et al., 2011) and statistical event extraction from natural text (Galitsky, 2013; Pradhan, 2005; Das et al.,2010).

A novelty of our approach is the use of explicit, linguistically motivated sub-event patterns—in contrast to the shallower, string-based techniques employed by purely statistical event-extraction systems. We are optimistic that this approach will be very efficient on large-scale data processing, having obtained promising results from prior experimentation with pattern extraction on graphs (Petrovic et al., 2005a, 2005b; Liu et al., 2006, 2011) and semi-

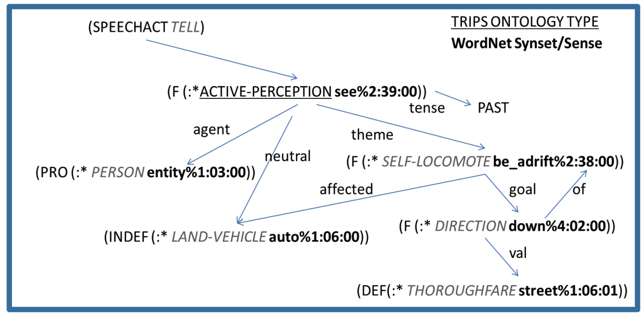

Figure 6: The Semantic Representation for *I saw a car floating down the street*.

structured data (Petrovic et al., 2003; Burcea et al., 2003; Wun et al., 2007).

3 Deep Event Understanding and Emergence Detection

We take advantage of a robust, broad coverage semantic parser to represent event semantics: TRIPS (Allen et al., 2008). TRIPS produces a representation that enables event understanding and event emergence detection.

3.1 Representing Semantics of Events

The TRIPS semantic lexicon provides information on thematic roles and selectional restrictions for about 5000 verbs. The parser constructs a semantic representation in a state-of-the-art logic (Manshadi et al., 2008; Manshadi and Allen, 2012).

Figure 6 shows the semantic analysis of "*I saw a car floating down the street*." Each node captures a predicate or entity in the sentence, as well as the ontology type of the node in the hand-built TRIPS ontology and the WordNet sense (Fellbaum, 1998) for the predicate/entity. Predicate-argument structure is indicated by links labeled with semantic roles. In addition, the parser extracts temporal information, which is critical for capturing ongoing situations that signal emerging events.

To attain broad lexical coverage beyond its hand-defined lexicon, the TRIPS parser uses a subsystem called WordFinder that accesses WordNet when an unknown word is encountered. This uses a hand-built mapping between the WordNet hypernym hierarchy and the TRIPS ontology. WordFinder uses the combined information from WordNet and the TRIPS lexicon and ontology to dynamically build lexical entries with approximate semantic and syntactic structures for words not in the core lexicon.

TRIPS has demonstrated capability in parsing both dialogues (e.g., Allen et al., 2007) and arbitrary text (e.g., Allen et al., 2008), which allows us to tackle informal communications such as the “emergent flood” example given earlier. The output of the parser is processed to identify temporally located events of interest. In this example, the event of interest is the car floating, and the evidence for this event is the fact that the speaker saw it. Furthermore, we know the floating event occurred prior to the time of the utterance. This stream of temporally located events serves as input to the inference component described next.

3.2 Acquiring Commonsense Models of Emerging Events

We acquire commonsense models of emerging events automatically by reading dictionary definitions (Allen et al, 2011; Allen & Teng, 2013). The Gloss system (Allen et al., 2011, 2013) is a system that uses the TRIPS parser to process the glosses in WordNet. These definitions provide a rich set of entailments, capturing many key relationships between events. For instance, WordNet contains the following definitions:

Drown: kill by submerging in water
Kill: cause to die

The Gloss system parses such definitions and then generates entailment axioms from the logical forms.

Figure 7 shows the parse of the definition for *drown*. Chaining from this definition and that of *kill* allows DISCERN to infer from *Rising waters drowned 7 people* that 7 people died. In addition, new ontology types are created for each new event (and other entities), producing a rich hierarchy to support inference.

3.3 Mini-Theories

From our own experience, simply chaining through lexical definitions has limited impact for identifying emerging events. Definitions are too vague and idiosyncratic by themselves. We address this in two ways. We first introduce hand-constructed axioms for a selective set of key concepts that link into our reasoning system (Allen & Teng, 2013). These hand-built axioms “activate” the axioms produced automatically.

In addition, we then organize the derived knowledge around *mini-theories*, each capturing specific knowledge about particular aspects of our world, e.g., life/death, working, eating, and commonsense temporal cycles, e.g., day/night, sleeping. Mini theories are constructed automatically by merging information from large clusters of definitions, all related in some way to the theory being constructed. By focusing on a single transition (e.g., life/death, asleep/awake) we are able to use heuristic techniques encoded in a probabilistic optimization process, using Markov Logic Networks (Richardson and Domingos, 2006) to construct the best concise consistent models.

A sample mini-theory regarding the process of drowning is shown in Figure 8. Note that we can have links between mini-theories. For instance, the mini-theory of flooding would involve things being submerged, which plays a key

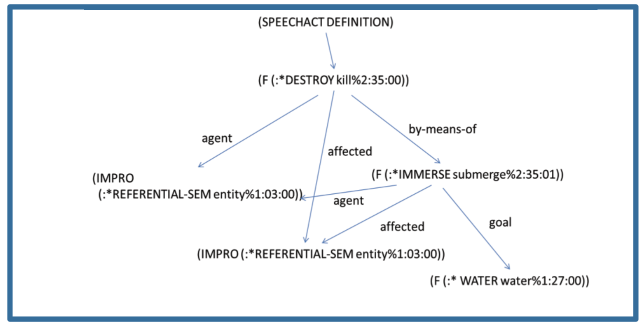

Figure 7: TRIPS parser output for drown defined as kill by submerging in water.

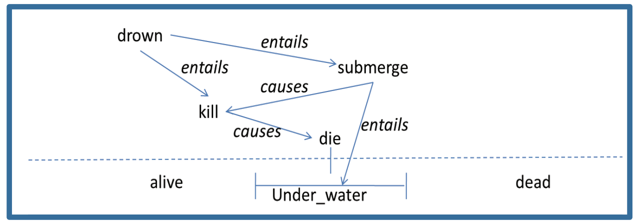

Figure 8: Mini-Theory of Drowning

role in the mini-theory of drowning. In the mini-theory for flooding, we would have key information such as the following:

- Raining entails that water is falling
- Rain can cause rising water levels
- A flood requires abnormally high levels of water

We employed heuristic approaches within a probabilistic framework yielding promising results for deriving mini-theories from clusters of related definitions. We encoded reasoning and heuristics in Markov logic and employed a probabilistic optimization technique to identify models that satisfy simplicity preferences.

We adopt this probabilistic optimization framework to perform an associative search through WordNet, as well as using CATVAR (Habash and Dorr, 2003) and other paraphrastic units (Madnani and Dorr, 2013), to identify the concepts related variously to the target central concept of the mini-theory. All these concepts are linked by their definitions, thus allowing us to construct a graph of their dependencies, within which we find key relations using an algorithm similar to PageRank.

3.4 Detecting Event Emergence

Consider the following sample input stream relating to an emerging flood event. Unlike those given earlier, this example is more complex in that it consists of a series of messages (with timestamps) posted by several people located within the same general area.

18:39 (Joan) I am watching TV.
19:00 (Mike) It's been raining really hard.
19:02 (Joan) Cats and dogs all day!
19:13 (Michelle) I had lamb curry for dinner.
19:15 (Mark) There are six inches of water in the yard.

19:21 (Michelle) It's pouring like mad.
19:32 (Jessica) I've been developing pictures in the darkroom all day.
19:34 (Billy) I have a burst pipe.
19:40 (Jessica) I haven't seen any rain.
20:04 (News) Water level at Wahoo River is five feet above normal.
20:13 (Billy) The whole kitchen got flooded!
...
<rain, more rain, a lot of rain>
...
23:17 (Alice) Water is seeping in around the door!
23:32 (Bob) There is a car floating in the middle of the street!
23:34 (News) A flock of sheep drowned in the heavy rains and rising water.

With respect to the emerging *flood* event in the example, the messages *I am watching TV* and *I had lamb for dinner* are not relevant, while most of the rest can be fit together into a mini-theory of "flood," much more coherently than into a mini-theory of, for example, "wildfire." Thus, "flood" is much more probable than "wildfire" in this example. Note that the messages that are irrelevant to "flood" could be supporting other events that are developing in parallel to the flood (e.g., lamb is becoming trendy). At any given point, there can be multiple main and secondary mini-theories actively engaged according to how well the candidate sub-events can be accounted for by each mini-theory.

We will now examine a number of challenges in emerging event detection.

Paraphrasing: As mentioned in Section 2.1, paraphrase handling is pivotal in identifying implicit event-event relationships that can be expressed in a diverse number of ways, as is often the case in free text. The methods discussed above for detecting categorial variations (e.g. raining—rain) and targeted paraphrasing (e.g. heavy rain—rising water—drown—flood) are used to link sub-events into event mini-theories.

Identifying Spurious Evidence: The example contains information seemingly contrary to an emerging flood:

19:40 (Jessica) I haven't seen any rain.

as well as information seemingly supportive of an emerging flood:

20:13 (Billy) The whole kitchen got flooded!

However, taking into account the larger context, we see that both cases are spurious, neither supporting nor undermining the bigger picture of an emerging flood:

19:32 (Jessica) I've been developing pictures in the darkroom all day.
19:40 (Jessica) I haven't seen any rain.
19:34 (Billy) I have a burst pipe.
20:13 (Billy) The whole kitchen got flooded!

The event structures and mini-theories provide the backbone for inferring that, although on the surface these pieces of evidence appear to be relevant to *flood* (as an emerging widespread event), there are alternative, more plausible explanations and thus they likely do not contribute to *flood* either positively or negatively. (It is of course possible, but less likely, that Billy's unlucky kitchen sustained flooding independently from *both* a burst pipe and the rain.) To properly interpret the messages in their larger context, we need discourse management across messages/sub-events and the continuity and relational structure provided by the causal mini-theories.

Precursor Events and Filling in Gaps: Consider these two messages:

19:15 (Mark) There are six inches of water in the yard.
19:34 (News) Water level at Wahoo River is five feet above normal.

In these cases, there is no indication of flooding *yet*, but by approximately 23:00 when Alice and Bob are posting we can infer that, since it has not stopped raining (hard) in the last several hours, the water levels in Mark's yard and at Wahoo River are likely to have increased substantially.

Often the incoming information is incomplete. We can only obtain bits and pieces of what is going on, with gaps in between reported sub-events. This is especially prevalent for informal sources. For example, rising water is not explicitly mentioned in the above data stream, even though the water must have been rising. With the commonsense knowledge captured in mini-theories we are able to infer the missing pieces and fill in the gaps. This involves reasoning about properties of the world, including space, time, measurement and matter (e.g. rain/water).

Indicator Events: By 23:00 we are starting to see initial evidence of a flood:

23:17 (Alice) Water is seeping in around the door!
23:32 (Bob) There is a car floating in the middle of the street!

If we can project that the precursor events (for Mark's yard and Wahoo River) have likely reached flooding level by this time (23:00), we can conclude that there is an emerging "flood" event with more confidence and earlier than having to wait for more direct indicator events to appear in the data stream.

3.5 Connecting the Dots and Bridging the Gaps

Identifying emerging events involves two inter-related tasks: (1) connecting sub-events to each other and to the relevant emerging events within the event models; and (2) estimating the likelihood of possible emerging events. Each sub-event can be part of a number of emerging events and supports various event models to varying degrees. Our approach is to build a coherent and compact model that probabilistically accommodates as many mini-theories invoked by the sub-events as possible. This globally constrains possible solutions for relating sub-events, including discourse management across messages. As discussed above, the combination of mini-theories and techniques, such as paraphrasing detection, provides the needed underlying structure for bridging gaps in the interpretation, inferring and making use of information not explicitly mentioned, and assigning probable cause and effect to the events.

3.6 Related Work

There is little work with comparable depth and coverage of the TRIPS system. Boxer (Bos, 2008) is a broad-coverage semantic parser, but does not perform word sense disambiguation, leaving its predicates as the lexical items. Furthermore, its semantic role processing is limited to the few thousand words covered in VerbNet (Kipper et al., 2008). There is much recent work in learning semantic parsers (e.g., Brabavan et al, 2010; Chen & Mooney, 2011; Matusek et al, 2012; Tellex et al, 2013), but these systems operate only in highly restrictive domains and cannot be considered broad coverage. The emergence detection approach above resembles processes developed for plan and intention recognition based on finding minimal explanations (Kautz & Allen, 1986), most recently within probabilistic frameworks (e.g., Cascading Hidden Markov Models in (Blaylock & Allen, 2014) and Temporal Markov Logic Networks in (Song et al, 2013a; 2013b)). It shares many intuitions of the abductive approaches (e.g., Hobbs et al, 1993), but is cast in the framework of probabilistic optimization of event models.

4 Evaluations and Future Work

This is an opportune time to investigate events in great detail, given the growth of Event-related workshops (e.g., the first Events Workshop last year at NAACL-2013) and international evaluations of events at SemEval and the new NIST Event evaluation currently underway with our team's participation. The evaluations are expected to run annually, with progressively more complex tests each year. The metric used is an F-score over detected events and roles of participants based on *s-precision* and *s-recall*, where *s* refers to three values: good, perfect, and semantically correct.

In the NIST evaluation, systems are evaluated on their ability to extract explicit arguments from events. For example, in our running *rain/flood* example, systems would be expected to detect that 7 people drowned. A deficiency of current event-oriented evaluations is that systems are not evaluated for their ability to infer implicit arguments or relationships between sub-events. That is, systems are not expected to infer that "rising water" is entailed by "flooding conditions", nor are they expected to infer *flood* as an emerging event.

Our future work will evaluate our progress in an iterative fashion, with the results of each iteration being used to improve the processing of the new resources in the next iteration. We will employ the metric defined above, but we will apply it to inferred arguments and sub-event relationships, adopting human-produced gold standards from NIST (distributed by the Linguistic Data Consortium at the University of Pennsylvania).

Acknowledgments

The research and investigations described above have been supported by the Florida Institute for Human and Machine Cognition.

References

Allen, James, N. Chambers, G. Ferguson, L. Galescu, H. Jung, W. Taysom, PLOW: A collaborative task learning agent. In *Proceedings of the 22nd national conference on Artificial intelligence*, pages 1514-1519, 2007.

Allen, James, de Beaumont, W., Galescu, L., Orfan, J., Swift, M. and Teng, Choh M., "Automatically deriving event ontologies for a commonsense knowledge base," *In Proceedings of the International Conference for Computational Semantics*, 2008.

Allen, James, de Beaumont, W., Blaylock, N., Ferguson, G., Orfan, J. and Swift, M., "Acquiring commonsense knowledge for a cognitive agent." *AAAI Fall Symposium Series: Advances in Cognitive Systems (ACS 2011)*, Arlington, VA, November 4-7, 2011.

Allen, James, W. de Beaumont, L. Galescu, J. Orfan, M. Swift, and Choh Man Teng, Automatically Deriving Event Ontologies for a CommonSense Knowledge Base. In *Proceedings of the International Conference for Computational Semantics*. 2013.

Allen, James, and Teng, Choh Man, "Becoming different: A language-driven formalism for commonsense knowledge." *In 11th International Symposium on Logical Formalizations of Commonsense Reasoning,* 2013.

Bannard, Colin and Chris Callison-Burch, "Paraphrasing with Bilingual Parallel Corpora," in *Proceedings of the Association for Computational Linguistics*, Atlanta, GA, pp. 597-604, 2005.

Blaylock, N and Allen, James .F. Hierarchical Goal Recognition, in Sukthankar et al (eds), *Plan, Activity and Intent recognition*, Morgan Kaufman, 2014

Bos,f Johan. "Wide-Coverage Semantic Analysis with Boxer". In: J. Bos, R. Delmonte (eds): Semantics in Text Processing. STEP 2008 Conference Proceedings, pp.277-286, Research in Computational Semantics, College Publications

Burcea, I., Milenko Petrovic, Jacobsen, H-A. I know what you mean: semantic issues in Internet-scale publish/subscribe systems. In Proceedings of the First International Workshop on Semantic Web and Databases. September 2003

Chang, Fay, et al. "Bigtable: A distributed storage system for structured data." *ACM Transactions on Computer Systems (TOCS)* 26.2 (2008): 4.

Chen D.L., and R. Mooney.(2011) Learning to interpret natural language navigation instructions from observations. Proc. AAAI (AAAI-2011), pages 859–865

Cugola, Gianpaolo, and Alessandro Margara. "Processing flows of information: From data stream to complex event processing." *ACM Computing Surveys (CSUR)* 44.3 (2012): 15.

Dagan, Ido, Oren Glickman, and Bernardo Magnini. "The PASCAL recognising textual entailment challenge." In *Machine Learning Challenges. Evaluating Predictive Uncertainty, Visual Object Classification, and Recognising Tectual Entailment*, pp. 177-190. Springer Berlin Heidelberg, 2006

Das, Dipanjan, et al. "Probabilistic frame-semantic parsing." *Human language technologies: The 2010 annual conference of the North American chapter of the association for computational linguistics*. Association for Computational Linguistics, 2010.

Dorr, Bonnie J., David Farwell, Rebecca Green, Nizar Habash, Stephen Helmreich, Eduard Hovy, Lori Levin, Keith J. Miller, Teruko Mitamura, Owen Rambow, Florence Reeder, Advaith Siddharthan. "Interlingual Annotation of Parallel Text Corpora: A New Framework for Annotation and Evaluation," *Journal of Natural Language Engineering*, 16:3, Cambridge University Press, pp. 197--243, 2010.

Dorr, Bonnie J., "Machine Translation Evaluation and Optimization," Part 5, in Joseph Olive, John McCary, and Caitlin Christianson (eds.), *Handbook of Natural Language Processing and Machine Translation*, 2010.

Eckert, Michael, et al. "A CEP Babelfish: Languages for complex event processing and querying surveyed." *Reasoning in Event-Based Distributed Systems*. Springer Berlin Heidelberg, 2011. 47-70.

Fellbaum, C. *WordNet: An Electronic Lexical Database*. Cambridge, MA: MIT Press, 1998.

Galitsky, Boris. "Machine learning of syntactic parse trees for search and classification of text." *Engineering Applications of Artificial Intelligence* 26.3 (2013): 1072-1091.

Habash, Nizar, Bonnie J. Dorr, and Christof Monz, "Symbolic-to-Statistical Hybridization: Extending Generation-Heavy Machine Translation," in *Machine Translation*, pp. 23-63, November, 2009.

Hobbs, J. R., Stickel, M. E., Appelt, D. E. and Martin, P. A. Interpretation as abduction. *Artificial Intelligence*, 63(1-2):69–142, 1993.

Kautz, H. A. and Allen, James. F. Generalized Plan Recognition. *Fifth National Conference on Artificial Intelligence (AAAI-86)*, Philadelphia, PA, 1986.

Kipper, K., Korhonen, A., Ryant, N. and Palmer, M. "A large-scale classification of english verbs." *Language Resources and Evaluation Journal*, vol. 42, pp. 21–40, 2008.

Lin, Jimmy, Nitin Madnani, and Bonnie J. Dorr, "Putting the User in the Loop: Interactive Maximal Marginal Relevance for Query-Focused Summarization," in *Proceedings of the Conference of the North American Chapter of the Association for Computational Linguistics - Human Language Technologies (NAACL HLT)*, Los Angeles, CA, pp. 305-308, 2010.

Liu, H., Milenko Petrovic, Jacobsen, H.-A., Efficient and Scalable Filtering of Graph-based Metadata. To appear in Journal of Web Semantics. January 2006

Liu, H., Z. Wu, Milenko Petrovic, H.-A. Jacobsen, Optimized cluster-based filtering algorithm for graph metadata*. Information Sciences*. 181(24): 5468-5484. January 2011.

Madnani, Nitin and Bonnie J. Dorr, "Generating Phrasal and Sentential Paraphrases: A Survey of Data-Driven Methods", *Computational Linguistics*, 36:3, pp. 341-387, 2010.

Matuszek, C., E Herbst, L Zettlemoyer, D. Fox (2012), Learning to Parse Natural Language Commands to a Robot Control System, Proc. of the 13th International Symposium on Experimental Robotics (ISER)

Manshadi, M. H., James Allen, and Mary Swift, "Towards a universal underspecified semantic representation." In *Proceedings of the 13th Conference on Formal Grammar*, Hamburg, Germany, 2008.

Manshadi, M. and James Allen, "A Universal Representation for Shallow and Deep Semantics," In *Proceedings of LREC Workshop on Semantic Annotation*, Instanbul, Turkey, 2012.

Marcus, Mitchell, et al. Treebank-3 LDC99T42. Web Download. Philadelphia: Linguistic Data Consortium, 1999.

Palmer, M., D. Gildea, D. and P. Kingsbury, "The Proposition Bank: A corpus annotated with semantic roles," *Computational Linguistics Journal*, 31:1, 2005.

Petrovic, Milenko, Burcea, I., Jacobsen, H-A, S-ToPSS: Semantic Toronto Publish/Subscribe System. In *Proceedings of 29th International Conference on Very Large Databases (VLDB)*. September 2003. (System demonstration)

Petrovic, Milenko, Liu, H., Jacobsen, H.-A., CMS-ToPSS: Efficient Dissemination of RSS Documents. In Proceedings of 31st International Conference on Very Large Data Bases (VLDB). September 2005a. (System demonstration)

Petrovic, Milenko, H. Liu, H.-A. Jacobsen, G-ToPSS: Fast Filtering of Graph-based Metadata. In *Proceedings of 14th International World Wide Web Conference (WWW)*, pp. 539-547. May 2005b. Best Paper Finalist.

Pradhan, Sameer, et al. "Support vector learning for semantic argument classification." *Machine Learning* 60.1-3 (2005): 11-39.

Pustejovsky, James, "The Role of Event-Based Representations and Reasoning in Language," in *Proceedings of NAACL Workshop on EVENTS: Definition, Detection, Coreference, and Representation*, Atlanta GA, 2013.

Richardson, M., and Domingos, P. Markov logic networks. *Machine Learning* 62(1-2):107–136, 2006.

Ruppenhover, Joseph, Caroline Sporleder, Roster Morante, Colin Baker, Martha Palmer, "SemEval-2010: Linking Events and their Participants in Discourse," in *Proceedings of the NAACL HLT Workshop on Semantic Evaluations: Recent Achievements and Future Directions*, pp. 106-111, Boulder Colorado, 2009.

Song, Y., Henry Kautz, Yuncheng Li, Jiebo Luo, "A General Framework for Recognizing Complex Events in Markov Logic," In *Proceedings of the Plan, Activity, and Intent Recognition Workshop and Statistical Relational AI Workshop, AAAI*, Bellevue, WA, USA, 2013a.

Song, Y., Henry Kautz, James Allen, Mary Swift, Yuncheng Li, Jiebo Luo, Ce Zhang, "A Markov Logic Framework for Recognizing Complex Events from Multimodal Data," In *Proceedings of the 15th ACM International Conference on Multimodal Interaction (ICMI)*, Sydney, Australia. 2013b.

UzZaman, Naushad, Hector Llorens, Leon Derczynski, Marc Verhagen, James Allen, and James Pustejovsky, "TempEval-3: Evaluating Time Expressions, Events, and Temporal Relations," in *Proceedings of the International Workshop on Semantic Evaluation*, pp. 1-9, Atlanta, GA, 2013.

Vlachos, Michail, et al. "Identifying similarities, periodicities and bursts for online search queries." *Proceedings of the 2004 ACM SIGMOD international conference on Management of data*. ACM, 2004.

Wu, Eugene, Yanlei Diao, and Shariq Rizvi. "High-performance complex event processing over streams." *Proceedings of the 2006 ACM SIGMOD international conference on Management of data*. ACM, 2006.

Wun, A., Milenko Petrovic, H.-A. Jacobsen, A system for semantic data fusion in sensor networks. In *Proceedings of International Conference on Distributed Event-Based Systems (DEBS)* pp. 75-79. June 2007.

Zajic, David M., Bonnie J. Dorr, Jimmy Lin, Richard Schwartz. "Multi-Candidate Reduction: Sentence Compression as a Tool for Document Summarization Tasks," *Information Processing and Management Special Issue on Summarization*, 43:6, pp. 1549--1570, 2007

Zajic, David, Bonnie J. Dorr, and Jimmy Lin, "Single-Document and Multi-Document Summarization Techniques for Email Threads Using Sentence Compression," Journal of Information Processing and Management, 44:4, pp. 1600--1610, 2008