

From Narrative Text to VerbNet-Based DRSES: System Text2DRS

MS Project Report

Gang Ling

University of Nebraska Omaha

gling@unomaha.edu

Advisor: Yuliya Lierler

May, 2018

Computational linguists have long studied various logic forms for capturing essential semantic information carried by narratives. Among these, discourse representation structure (DRS) form is designed to acquire the entities, entities' properties, events, and event properties. This project report describes a system called Text2DRS that takes English narrative as an input and outputs a DRSES output in Neo-Davidsonian style.

1 Introduction

In this work we understand a narrative text to be a sequence of sentences with action verbs so that the succession of events is given in chronological order. Linguistically, action verbs are words that convey physical or mental actions, such as *grab*, or *go*. From the narrative text, readers can reproduce the sequence of events with related information in their minds. This process is not hard for a person but is a challenge for an artificial intelligence (A.I.) system agent to generate the same results automatically. Consider a simple three sentences discourse:

- John grabbed the suitcase. (1)
- John travelled to the hallway. (2)
- Sandra journeyed to the hallway. (3)

From the first sentence, a state-of-the-art A.I. agent can retrieve the direct information such as

```
grab is an action verb
John is an agent entity of action grab
suitcase is a object entity of action grab
```

The same pattern applies to the second and third sentence:

```
travel is an action verb
John is an agent entity of action travel
hallway is a destination entity of action travel
```

```
journey is an action verb
Sandra is an agent entity of action journey
hallway is an entity denoting a destination of action journey
```

However, what about other implicit information that a human would retrieve from this discourse that is sufficient to answer the following questions:

- *What is the location of the suitcase at different time points?*
- *Were John and Sandra in the same location at the end of described scenario?*

Such questions are difficult for modern A.I. agents. Human can process knowledge about the world (world knowledge/commonsense knowledge) and can effectively utilize it to perform inferences. To enable an A.I. agent to perform the similar inferences as a human, a software system has to utilize tools appropriate for describing world knowledge and drawing proper entailment from that.

An area of A.I. concerned with developing such tools is knowledge representation and reasoning (KRR). Action languages are formal languages that are used for capturing world knowledge about the effects of actions. Michael Gelfond, Daniela Inlezan and Yuliya Lierler have proposed a methodology for designing Question Answering (QA) systems that utilize an action language ALM to allow inferences based on complex interactions of events described in text [Lierler et al., 2017]. This methodology assumes the extension of the VerbNet lexicon with interpretable semantic annotations in ALM. It also relies on the construction of a tool chain of modern NLP systems, such as the LTH [Johansson and Nugues, 2007] and coreNLP [Manning et al., 2014]. And, this is the place where system Text2DRS, described here, fits in. Even through Text2DRS doesn't contain question answering component from the methodology mentioned above, it generates a knowledge representation of a given narrative, which forms the essential data for an A.I. agent to perform question answering tasks. The system is available for download at <https://www.unomaha.edu/college-of-information-science-and-technology/natural-language-processing-and-knowledge-representation-lab/software/text2drs.php> .

Related Work STool Text2Drs is similar to system Boxer [Bos, 2008], which is an open-domain NLP tool for semantic analysis that produces a DRS for a given narrative. Text2DRS provides additional information in comparison to Boxer. For example, Text2DRS relies on lexical resource VerbNet [Kipper-Schuler, 2005, Palmer, 2006] for annotating the specific relations between relevant entities and events mentioned in the narrative using the verb classes and thematic roles of VerbNet.

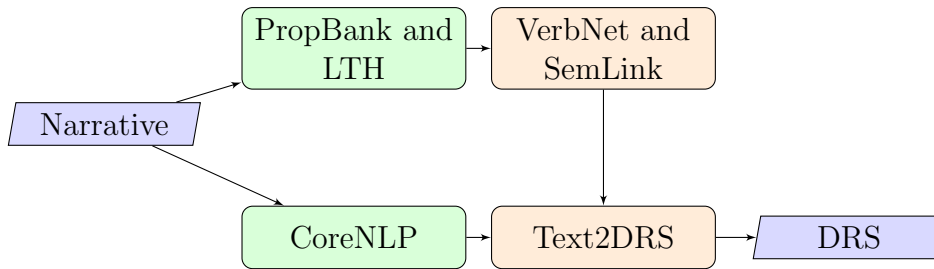


Figure 1: *NLP resources*

Report Outline Figure 1 presents a schematic architecture of the Text2Drs system. Note that lexical resources as PropBank [Palmer et al., 2005], VerbNet [Kipper-Schuler, 2005, Palmer, 2006], and SemLink [Bonial et al., 2013] form the important components of Text2DRS. Also, several modern natural language processing (NLP) tools such as LTH [Johansson and Nugues, 2007] and CoreNLP [Manning et al., 2014] are part of the system. Also such . We start this report by providing key details on these NLP tools and resources. We then discuss the intuitions behind the discourse representation structures [Kamp and Reyle, 1993] that are essential in understanding the output produced by the Text2DRS tool. In Section 3, we talk about the system architecture of Text2DRS as well as the implementation challenges and solutions. Then, we present the system evaluation results. At last, we conclude with future work discussion in Section 5.

2 Modern NLP Resources

2.1 PropBank

In English, a sentence contains three main components: subjects, verbs, and objects. More importantly, verbs stand for predicates that are used to describe events occurring in our world. These predicates take arguments which are participants of events. By processing verbs from a sentence, an A.I. agent can understand which events happen in a world around the agent.

PropBank [Palmer et al., 2005] is a Verb Lexical Resource that systematizes knowledge about verbs with respect to their predicate-argument or, in another word, event-participant structure. Consider “frame” for verb *grab*:

```

id: grab.01
Semantic roles: Arg0-PAG: grabber
                Arg1-PPT: entity grabbed
  
```

In this frame, Arg0 and Arg1 are called semantic roles, and *grabber* and *entity grabbed* are descriptions of these particular semantic roles. We can use the Propbank frames to annotate sentences with verb *grab* (and others) in a systematic fashion. We can identify event participants and their roles: this task is called “Semantic Role Labeling”. For example, a table below presents PropBank semantic role labeling for sentence (1).

| | | |
|------|---------|--------------|
| John | grabbed | the suitcase |
| Arg0 | grab.01 | Arg1 |

Similarly, we can generate a PropBank semantic role labeling table for sentences (2) and (3):

| | | |
|------|-----------|----------------|
| John | travelled | to the hallway |
| Arg0 | travel.01 | Arg4 |

| | | |
|--------|------------|----------------|
| Sandra | journeyed | to the hallway |
| Arg0 | journey.01 | Arg2 |

2.2 VerbNet

However, PropBank organizes each predicate individually without grouping similar meaning verbs together. Also, the semantic roles of the predicates are not always consistent across similar verbs. As a result, the A.I. agent needs to include more complex rule handling at each case separately. VerbNet [Palmer, 2006] is, like Propbank, a Verb Lexical Resource. Its frames adopt a richer system of thematic roles. (The term of “thematic roles” has the same meaning of “semantic role” in the PropBank.) More importantly, VerbNet groups relevant verbs into verb classes. For example: verb *travel* evokes VerbNet class:

```
run-51.3.2
Members: travel, journey, climb, step, zoom ...
Thematic roles: Agent
                  Theme
                  Location
```

As we see, verbs *travel* and *journey* are members of run-class. They have close meaning and share three thematic roles: Agent; Theme; Location. Now, we extend the PropBank semantic role labeling tables of the three sentences by adding verb class data from the VerbNet.

Sentence (1):

| | | |
|-------|--------------|---------------|
| John | grabbed | the suitcase. |
| Arg0 | grab.01 | Arg1 |
| Agent | steal-10.5-1 | Theme |

The predicate “grab.01” is associated with verb classes *steal-10.5*, *obtain-13.5.2*, and *hold-15.1* according to data entry in the PropBank. Verb class *obtain-13.5.2* may seem

to be more suitable in this case. But without the emotional meaning, these three verb classes represent a similar action. Since we are interested in verbs and events in this project, we can pick either one of three verb classes and use it in the final output. We will discuss this decision-making mechanism in the system architecture section. Right now, let us use the verb class *steal*. The same idea applies to other two sentences in our running example.

Sentence (2):

| | | |
|-------|------------|----------------|
| John | travelled | to the hallway |
| Arg0 | travel.01 | Arg4 |
| Theme | run-51.3.2 | Location |

Sentence (3):

| | | |
|--------|------------|----------------|
| Sandra | journeyed | to the hallway |
| Arg0 | journey.01 | Arg2 |
| Theme | run-51.3.2 | Location |

From the results for sentences (2) and (3), we find that distinct PropBank predicates *travel.01* and *journey.01* are converted into a single verb class *run-51.3.2*. Also, the words *hallway* in the two sentences are labeled by the same thematic role of *Location*, even though they have different semantic roles (*Arg4*, *Arg2*) in PropBank. Therefore, Verbnet Semantic Role Labeling allows us to simplify collected data but still maintain the sentences' information.

2.3 SemLink

SemLink[Bonial et al., 2013] contains mappings between PropBank and VerbNet.

Here we present sample entries in the SemLink resource for verbs *travel*, *journey*, and *grab*.

```
<predicate lemma="travel">
  <argmap pb-roleset="travel.01" vn-class="51.3.2-1">
    <role pb-arg="0" vn-theta="Theme" />
    <role pb-arg="1" vn-theta="Location" />
  </argmap>
</predicate>
```

```
<predicate lemma="journey">
  <argmap pb-roleset="journey.01" vn-class="51.3.2-1">
    <role pb-arg="0" vn-theta="Theme" />
  </argmap>
</predicate>
```

```

        <role pb-arg="2" vn-theta="Location" />
    </argmap>
</predicate>

</predicate>
<predicate lemma="grab">
  <argmap pb-roleset="grab.01" vn-class="10.5-1">
    <role pb-arg="0" vn-theta="Agent" />
    <role pb-arg="1" vn-theta="Theme" />
  </argmap>
  <argmap pb-roleset="grab.01" vn-class="13.5.2">
    <role pb-arg="0" vn-theta="Agent" />
    <role pb-arg="1" vn-theta="Theme" />
  </argmap>
  <argmap pb-roleset="grab.01" vn-class="15.1-1">
    <role pb-arg="0" vn-theta="Agent" />
    <role pb-arg="1" vn-theta="Theme" />
  </argmap>
</predicate>

```

Consider verb *travel*. In its mapping, the predicate *travel.01* is mapped to verb class 51.3.2-1 which is a subclass of *run-51.3.2*. The semantic role “arg0” is mapped to thematic role “Theme”, and semantic role “arg1” is mapped to the thematic role “Location”.

2.4 Discourse Representation Structures

In order to let the A.I. agent to access the data with an efficient knowledge representation form, we use discourse representation structures (DRSs) [Kamp and Reyle, 1993], as the system outputs format. A DRS is a structure in spirit of a first order logic formula used by linguists to encode key information carried by a narrative. A DRS consists of two parts:

- a universe that consists of so-called “discourse referents”, which represent the objects/entities/events/participants under discussion
- a set of DRS-conditions which encode the information that has accumulated on these discourse referents: event descriptors and participants’ roles information

Consider the narrative composed of sentences (1-3), its sample DRS is presented in Table [reftable:drs](#). The first block displays the entities and events introduced by the narrative. The entities are represented as *r1*, *r2*, *r3*, *r4* (“r” stands for a referent), and the events are represented as *e1*, *e2*, *e3*. The second block shows descriptive details about the entities and events of the narrative. The property is a mapping of an entity referent and its original text in the narrative. The eventType is a relation between an

event referent and its corresponding VerbNet class. In this example, both verb *travel* and *journey* are mapped to the verb class *run-51.3.2-1*, and verb *grab* is mapped to the verb class *steal-10.5-1*. An eventArgument relation presents information about events. For instance, eventArgument(e2,location,r3) says that entity *r3* (that has a property of being a “hallway”) plays a thematic role “location” of event *e2* (that belongs to VerbNet class *run-51.3.2-1*). The eventTime indicates the time order of the occurred events in the narrative. In the action language, we consider event begin at time zero. Therefore, the first event *e1* (represent verb “grab”) starts at “0”.

| <i>r1</i> | <i>r2</i> | <i>r3</i> | <i>r4</i> | <i>e1</i> | <i>e2</i> | <i>e3</i> |
|------------------------------|---------------------------------|------------|------------|-----------|-----------|-----------|
| entity(r1) | entity(r2) | entity(r3) | entity(r4) | | | |
| property(r1, John) | property(r2, suitcase) | | | | | |
| property(r3, hallway) | property(r4, Sandra) | | | | | |
| event(e1) | event(e2) | event(e3) | | | | |
| eventType(e1, steal-10.5-1) | eventTime(e1, 0) | | | | | |
| eventArgument(e1, agent, r1) | eventArgument(e1, theme, r2) | | | | | |
| eventType(e2, run-51.3.2-1) | eventTime(e2, 1) | | | | | |
| eventArgument(e2, theme, r1) | eventArgument(e2, location, r3) | | | | | |
| eventType(e3, run-51.3.2-1) | eventTime(e3, 2) | | | | | |
| eventArgument(e3, theme, r4) | eventArgument(e3, location, r3) | | | | | |

Table 1: DRS for the given narrative

2.5 LTH

In the tool chain of our Text2DRS implementation, the LTH system and Stanford CoreNLP system are used as pre-processing system components. LTH [Johansson and Nugues, 2007] is a semantic role labeler for unrestricted text in English that uses predicates and semantic roles from PropBank. For sentences (1), (2), and (3), LTH produces the following outputs:

```
[A0 John][V (grab.01) grabbed ][A1 the suitcase]          (i)
[A0 John][V (travel.01) travelled][A4 to the hallway]     (ii)
[A0 Sandra][V (journey.01) journeyed][A2 to the hallway] (iii)
```

These outputs are annotated using the semantic roles of the predicates *grab.01*, *travel.01*, and *journey.01* as defined in their PropBank frames schemas. Recall the listed frame schema of the predicates grab.01 on Section 2.1. The frame schemas of the predicate travel.01 and journey.01 follow:

```
travel.01: the act of moving to
Arg0-PPT: traveller
Arg1-LOC: location or path
Arg2-DIR: start point
Arg4-GOL: destination
```

journey.01: the act of moving to
Arg0-PAG: traveller
Arg2-LOC: location or path

2.6 CoreNLP

The Stanford CoreNLP system [Manning et al., 2014] provides a set of NLP tools including a coreference resolution system. Text2DRS utilizes the coreference resolution function from CoreNLP system to process a given narrative.

CoreNLP detects four mentioned entities in the discourse: *John*, *suitcase*, *hallway*, and *Sandra*. The system also recognizes that the entity *John* in the sentence (2) is the same entity that appears in the sentence (1). Similarly, the noun entity *the hallway* in the sentence (3) is the same noun entity that is used in the sentence (2). This process is called coreference recognition. As a result, CoreNLP identifies four unique entities in this discourse.

3 System’s Architecture, Challenges, and Output

As the outcome of this project, we developed and implemented a DRS generating system, named Text2DRS. It contains four main parts: LTH, CoreNLP, SemLink, and itself. In a nutshell, first LTH and CoreNLP preprocess a given narrative. Then, Text2DRS reads the outputs from these two systems and generates DRS using the mapping data from SemLink.

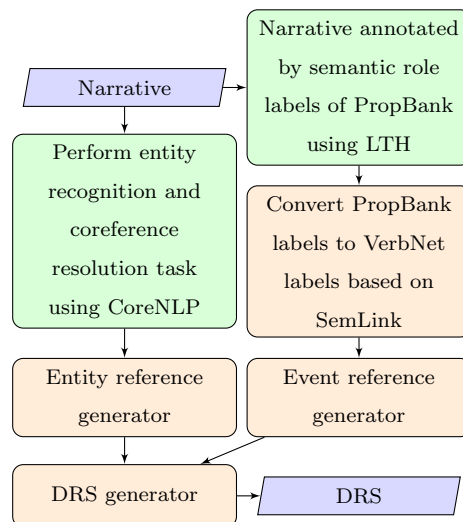


Figure 2: System architecture

Figure 2 presents a detailed Text2DRS architecture. The entity reference generator component creates an entity coreference map from the CoreNLP output and uses this map to assign a reference ID to each distinct entity. From the LTH output, Text2DRS looks up the related VerbNet class information in SemLink and returns the corresponding verb

class number along with the respective thematic roles. However, sometime SemLink maps one PropBank predicate into multiple VerbNet classes. In this case, we pick the first verb class from the data list and use it in the final output. After entity reference generator and event reference generator complete processing of the data, the DRS generator merges the data and outputs the DRS for the given narrative.

Challenges During the Text2DRS system implementation, there were three main challenges. First, LTH does not always label the semantic roles on the nouns (entities). For example, it may mark a preposition of a prepositional phrase by a semantic role. As a result, Text2DRS has to consider deeper syntactic structure of a sentence to properly assign roles to recognized entities.

The second issue that we have to overcome is missing mapping entries in the SemLink. Some predicates have partial mapping data in their entries, while others predicates do not have entries in SemLink.

The third challenge concerns generating DRSES for complex narratives as some sentences have multiple events (verbs) with multiple associated entities. An example sentence from “ROC-03” (see Section 6),

I stood up feeling confident and turned it in.

In this sentence, we can find three events as “stood”, “feeling”, and “ turned”, and two associated entities “I” and “it”. So, in a corresponding DRSES, we have to encode all three events with their carried information.

Current solutions In order to solve the second challenge, if an entity is assigned a semantic role by LTH, but Text2DRS cannot find the corresponding thematic role from SemLink, then Text2DRS replaces the semantic role with NONE-THEMEROLE in the output. If Text2DRS cannot find the verb class for the predicate in SemLink, then a NON-FOUND-IN-SEMLINK will be used as a label in the narrative’s DRSES.

To address the third challenge, since we have to include all event properties in the final output, we generate a byproduct in the Text2DRS, a text file called VerbNetsrl. It is an extension of the LTH’s output that adds a verb class data column for each predicate. And, based on the VerbNetsrl file, Text2DRS continues merging data and generates DRSES outputs.

System Output

Given the example narrative composed of sentences (1-3) as input, the Text2DRS generates the following output:

```
\% r1, r2, r3, r4, e1, e2, e3
\% =====
entity(r1). entity(r2). entity(r3). entity(r4).
```

```

property(r1, "John"). property(r2, "suitcase").
property(r3, "hallway"). property(r4, "Sandra").

event(e1). event(e2). event(e3).

eventType(e1, "10.5-1").
eventType(e2, "51.3.2-1").
eventType(e3, "51.3.2-1").

eventTime(e1, 0). eventTime(e2, 1). eventTime(e3, 2).

eventArgument(e1, "Agent", r1).
eventArgument(e1, "Theme", r2).
eventArgument(e2, "Theme", r1).
eventArgument(e2, "Location", r3).
eventArgument(e3, "Theme", r4).
eventArgument(e3, "Location", r3).

```

4 Evaluation

We evaluate system Text2DRS on sample narratives from two datasets (i) Facebook’s bAbl [Weston et al., 2015] collection and (ii) the ROCStories 2017 [Mostafazadeh et al., 2017] collection. We included all considered test cases in the appendix of this paper. Totally, we consider 20 test narratives (ten bAbI cases and ten ROCStories cases) that contain 177 sentences. Following is one of the ten bAbI test cases:

```

John travelled to the hallway.      Mary journeyed to the bathroom.
Daniel went back to the bathroom.  John moved to the bedroom.
John went to the hallway.          Sandra journeyed to the kitchen.
Sandra travelled to the hallway.   John went to the garden.
Sandra went back to the bathroom.  Sandra moved to the kitchen.

```

Here is an example of the ten ROCStories narrative test case:

```

Tom had a very short temper.
One day a guest made him very angry.
He punched a hole in the wall of his house.
Tom’s guest became afraid and left quickly.
Tom sat on his couch filled with regret about his actions.

```

The evaluation is divided into four parts

1. including entity recognition verification,

2. entity coreference correctness checking,
3. event recognition verification, and
4. event annotation correctness checking.

4.1 Entity Recognition Verification

In the entity recognition verification, we identify all the entities in the testing cases manually and compare them with entries in generated DRSEs.

| Entity recognition | Manually identified | Text2DRS identified |
|--------------------|---------------------|---------------------|
| Facebook bAbI | 264 | 264 |
| ROCStories 2017 | 161 | 161 |

4.2 Entity Coreference Correctness Checking

In the entity coreference correctness checking, we identify manually entities in each testing case and assign entity referent to the unique entities. Then we compare the numbers of unique entities (entity referents) from the manually created reports with Text2DRS’s outputs.

| Entity coreference correctness | Manually identified | Text2DRS generates |
|--------------------------------|---------------------|--------------------|
| Facebook bAbI | 79 | 80 |
| ROCStories 2017 | 103 | 132 |

These results show that CoreNLP has an excellent performance of entity coreference recognition when it processes bAbI narratives. The only incorrect case occurred in the Narrative 2 of bAbI (see Section 6). CoreNLP is unable to recognize that two entities *suitcase* from both sentences are the same. CoreNLP drops accuracy when it processes ROCStories narratives. Narrative ROC-04 (Section 6) is an example of a test case where CoreNLP is unable to resolve coreference properly.

4.3 Event Recognition Verification

In the event recognition verification, we repeat the same process as evaluating entity recognition.

| Event recognition verification | Manually identified | Text2DRS identified |
|--------------------------------|---------------------|---------------------|
| Facebook bAbI | 127 | 127 |
| ROCStories 2017 | 91 | 91 |

From the result table, we conclude that LTH locates all the events, and Text2DRS can assign event referents to them.

4.4 Event Annotation Correctness Checking

In the event annotation correctness checking, we collect used VerbNet verb-class data entries. We go through each eventArgument condition and compare the annotation with verb class data. The table below summarizes the results.

| Event annotation correctness | Text2DRS generates | Incorrect | Accuracy |
|------------------------------|--------------------|-----------|----------|
| Facebook bAbI | 264 | 27 | 89% |
| ROCStories 2017 | 199 | 67 | 66% |

If we find an incorrect annotation during our evaluation process, we take a further step by looking up the corresponding VerbNetsrl file, LTH outputs, and SemLink mapping entries. In this way, we can categories mistakes. For the case of bAbI narratives, all of the mistakes were due to a single fact: SemLink has a missing mapping entry for predicate *go.01* whereas the verb *go* occurred 27 times among the test cases. Similarly, SemLink has 67 missing mapping entries during the eventArgument generating processes within test cases in ROCStories.

5 Future Work

One of important future work directions is adding missing mappings into SemLink after we test more narrative cases. The correctness of event annotations heavily depends on the mapping data from the SemLink.

Another essential work is to improve system performance and extend output file format options. Currently, Text2DRS runs LTH and CoreNLP by Java command line, and both systems take time to load their modules. For example, CoreNLP takes 18 seconds on average to setup its pipeline, but it only uses two seconds on average to process an input narrative. We intend to implement server-client based architecture within Text2DRS to keep CoreNLP running as a process in the operating system. The idea applies to the LTH system.

Since LTH is no longer maintained, we intend to incorporate few other state-of-the-art semantic role labelers, such as Neural-dep-srl [Marcheggiani and Titov, 2017] or PathLSTM [Roth and Lapata, 2016] in the future. We anticipate that incorporating a different semantic role labeler will be a simple task. Indeed, Text2DRS utilizes the CONLL-2008 standard output of LTH. Provided that more modern mentioned systems produce the same output makes a transition to a new semantic role labeler seamless.

Last but not least, we plan to utilize the Text2DRS system in a larger framework capable of reasoning about events.

Acknowledgements

I would first like to thank my project advisor Dr. Yuliya Lierler of the Computer Science Department at the University of Nebraska Omaha. I can always find help from her when I had a question about my research or writing. I would also like to thank Dr. Parvathi Chundi and Dr. Ryan Schuetzler for serving on my project committees. It is a great honor to present my project report to them and gather feedback from them. Finally, I would like to thank the support from my family and my friends. The help from them is immeasurable through my studies and life.

References

- Claire Bonial, Kevin Stowe, and Martha Palmer. *SemLink*, 2013. <https://verbs.colorado.edu/semlink/> [Accessed: 2017].
- Johan Bos. Wide-coverage semantic analysis with boxer. In *Proceedings of the 2008 Conference on Semantics in Text Processing*, STEP '08, pages 277–286, Stroudsburg, PA, USA, 2008. Association for Computational Linguistics. URL <http://dl.acm.org/citation.cfm?id=1626481.1626503>.
- Richard Johansson and Pierre Nugues. *Language Technology at LTH*, 2007. <http://nlp.cs.lth.se/> [Accessed: 2017].
- Hans Kamp and Uwe Reyle. *From discourse to logic*, volume 1,2. Kluwer, 1993.
- Karin Kipper-Schuler. *VerbNet: A Broad-Coverage, Comprehensive Verb Lexicon*. PhD thesis, University of Pennsylvania, 2005.
- Yuliya Lierler, Daniela Incezan, and Michael Gelfond. Action languages and question answering. In *Proceeding of the 12th International Conference on Computational Semantics*, 2017. URL http://works.bepress.com/yuliya_lierler/70/.
- Christopher D. Manning, Mihai Surdeanu, John Bauer, Jenny Finkel, Steven J. Bethard, and David McClosky. The Stanford CoreNLP natural language processing toolkit. In *Association for Computational Linguistics (ACL) System Demonstrations*, pages 55–60, 2014. URL <http://www.aclweb.org/anthology/P/P14/P14-5010>.
- Diego Marcheggiani and Ivan Titov. Encoding sentences with graph convolutional networks for semantic role labeling. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, 2017. URL <http://aclweb.org/anthology/D17-1159>.
- Nasrin Mostafazadeh, Michael Roth, Annie Louis, Nathanael Chambers, and James F. Allen. *LSDSem 2017 Shared Task: The Story Cloze Test*, 2017. <http://cs.rochester.edu/nlp/rocstories/> [Accessed: 2018].
- Martha Palmer. *VerbNet*, 2006. <https://verbs.colorado.edu/~mpalmer/projects/verbnet.html> [Accessed: 2017].

Martha Palmer, Daniel Gildea, and Paul Kingsbury. The proposition bank: An annotated corpus of semantic roles. *Computational Linguistics*, 31(1):71–106, March 2005. ISSN 0891-2017. doi: 10.1162/0891201053630264. URL <http://dx.doi.org/10.1162/0891201053630264>.

Michael Roth and Mirella Lapata. Neural semantic role labelling with dependency path embeddings. In *In Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics*, 2016. URL <https://arxiv.org/pdf/1605.07515.pdf>.

Jason Weston, Antoine Bordes, Sumit Chopra, Alexander M. Rush, Bart van Merriënboer, Armand Joulin, and Tomas Mikolov. *Facebook Research bAbI*, 2015. <https://research.fb.com/downloads/babi/> [Accessed: 2017].

6 Narrative testing cases

1. Narrative one:
John grabbed the suitcase.
John travelled to the hallway.
Sandra journeyed to the hallway.
2. Narrative two:
John grabbed the suitcase.
John travelled to the hallway with suitcase.
3. Narrative three:
John grabbed the suitcase.
John travelled to the hallway by car.
4. Single Supporting Fact test:
John travelled to the hallway. Mary journeyed to the bathroom.
Daniel went back to the bathroom. John moved to the bedroom.
John went to the hallway. Sandra journeyed to the kitchen.
Sandra travelled to the hallway. John went to the garden.
Sandra went back to the bathroom. Sandra moved to the kitchen.
5. Two Supporting Facts test:
Mary got the milk there. John moved to the bedroom.
Sandra went back to the kitchen. Mary travelled to the hallway.
John got the football there. John went to the hallway.
John put down the football. Mary went to the garden.
John went to the kitchen. Sandra travelled to the hallway.
Daniel went to the hallway. Mary discarded the milk.
6. Three Supporting Facts test:
Mary got the milk. John moved to the bedroom.
Daniel journeyed to the office. John grabbed the apple there.
John got the football. John journeyed to the garden.
Mary left the milk. John left the football.

| | |
|-------------------------------------|---------------------------------|
| Daniel moved to the garden. | Daniel grabbed the football. |
| Mary moved to the hallway. | Mary went to the kitchen. |
| John put down the apple there. | John picked up the apple. |
| Sandra moved to the hallway. | Daniel left the football there. |
| Daniel took the football. | John travelled to the kitchen. |
| Daniel dropped the football. | John dropped the apple. |
| John grabbed the apple. | John went to the office. |
| Sandra went back to the bedroom. | Sandra took the milk. |
| John journeyed to the bathroom. | John travelled to the office. |
| Sandra left the milk. | Mary went to the bedroom. |
| Mary moved to the office. | John travelled to the hallway. |
| Sandra moved to the garden. | Mary moved to the kitchen. |
| Daniel took the football. | Mary journeyed to the bedroom. |
| Mary grabbed the milk there. | Mary discarded the milk. |
| John went to the garden. | John discarded the apple there. |
| Sandra travelled to the bedroom. | Daniel moved to the bathroom. |
| Sandra got the milk. | Daniel travelled to the garden. |
| Sandra went back to the bathroom. | Daniel took the apple there. |
| Mary went back to the hallway. | Daniel went to the hallway. |
| Sandra went to the kitchen. | Mary journeyed to the bedroom. |
| Sandra journeyed to the hallway. | Daniel put down the apple. |
| Daniel put down the football there. | Sandra journeyed to the garden. |
| Mary travelled to the office. | Sandra dropped the milk. |

7. Three Arg Relations test:

| | |
|------------------------------------|-----------------------------------|
| Fred picked up the football there. | Fred gave the football to Jeff. |
| Bill went back to the bathroom. | Jeff grabbed the milk there. |
| Jeff gave the football to Fred. | Fred handed the football to Jeff. |
| Jeff handed the football to Fred. | Fred gave the football to Jeff. |
| Jeff gave the football to Fred. | Jeff put down the milk. |

8. Yes No Questions test:

| | |
|-----------------------------------|----------------------------------|
| Mary got the milk there. | John moved to the bedroom. |
| Mary discarded the milk. | John went to the garden. |
| Daniel moved to the bedroom. | Daniel went to the garden. |
| Daniel travelled to the bathroom. | Sandra travelled to the bedroom. |
| Mary took the football there. | Sandra grabbed the milk there. |

9. Counting test:

| | |
|-----------------------------------|-------------------------------|
| Mary got the milk there. | John moved to the bedroom. |
| Sandra went back to the bathroom. | John got the football there. |
| Mary journeyed to the bathroom. | Mary gave the milk to Sandra. |
| John went back to the bathroom. | John left the football. |
| Sandra gave the milk to John. | John journeyed to the garden. |

10. Lists Sets test:

| | |
|------------------------------------|----------------------------------|
| Mary got the milk there. | John moved to the bedroom. |
| John picked up the football there. | John journeyed to the bathroom. |
| John went to the garden. | Daniel went back to the hallway. |
| John went back to the bathroom. | Mary went to the office. |
| Sandra went back to the kitchen. | Mary travelled to the hallway. |
| John travelled to the bedroom. | John picked up the apple there. |
| Sandra travelled to the bedroom. | Sandra journeyed to the kitchen. |

11. ROC_01:
Tom had a very short temper.
One day a guest made him very angry.
He punched a hole in the wall of his house.
Tom's guest became afraid and left quickly.
Tom sat on his couch filled with regret about his actions.

12. ROC_02:
Melody's parents surprised her with a trip to the big aquarium.
Melody took a nap during the two hour car ride to the aquarium.
When they arrived, Melody was energetic and excited.
At the aquarium Melody saw sharks, tropical fish and many others.
After five hours at the aquarium, Melody and her family drove home.

13. ROC_03:
The math teacher announced a pop quiz as class began.
While some students complained, he began passing out the quiz.
I took out my pencil and began to work.
About 5 minutes later, I finished.
I stood up feeling confident and turned it in.

14. ROC_04:
Robbie was competing in a cross country meet.
He was halfway through when his leg cramped up.
Robbie wasn't sure he could go on.
He stopped for a minute and stretched his bad leg.
Robbie began to run again and finished the race in second place.

15. ROC_05:
When I was 12 years old, my dad got angry and kicked me aggressively.
Afterward, I became very ill, and tasted something metallic.
I went to a doctor, and was informed that one of my kidneys was dead.
Ever since, I've had swelling and hypertension.
I started taking medications to combat the symptoms at 13.

16. ROC_06:
David noticed he had put on a lot of weight recently.
He examined his habits to try and figure out the reason.
He realized he'd been eating too much fast food lately.
He stopped going to burger places and started a vegetarian diet.
After a few weeks, he started to feel much better.

17. ROC_07:
Soren ran through the airport, pulling her bags behind her.
The female voice above her announced final boarding to Soren's flight.
She yelled for them to wait as she neared her gate, waving her arms.
The attendant at the desk gave Soren a sad, sympathetic look.
Nearly out of breath, Soren presented her pass and boarded the plane.

18. ROC_08:
Karl was a good baseball player in his youth.

As a middle-aged man, he joined an adult baseball team.
Karl is very competitive and will do anything to win.
One day, he slid into second base to beat a throw and hurt his knee.
The injury made Karl realize that he wasn't young anymore.

19. ROC_09:

Ken put a bottle of beer in the freezer.
He heard a popping noise.
He looked in the freezer and saw the bottle had burst.
He didn't want to wait for another beer to get cold.
He drank a warm beer instead.

20. ROC_10:

Justin was terrified of dogs.
His girlfriend had a dog and he wanted to feel comfortable with it.
Justin went to therapy to help him get over his fear.
After a few months of therapy Justin felt better around dogs.
Justin then moved in with his girlfriend and her dog.