Towards Deep Semantic Analysis with FrameNet

Daniel Bauer
Columbia University
bauer@cs.columbia.edu

FrameNet Workshop

April 19, 2013

This work is supported by the National Science Foundation under Grant No.IIS-0904361.
Language Understanding is Hard!

There was once a Wolf who saw a Lamb drinking at a river and wanted an excuse to eat it.
For that purpose, even though he himself was upstream, he accused the Lamb of stirring up the water and keeping him from drinking...

Minsky (1975)

- Need full representation of scene.
- Requires extra-linguistic world knowledge.
Owen is washing the floor.

The woman is sitting at the desk and using the computer. A dog is watching her. A cat is sleeping on the rug.
Text-to-scene generation: WordsEye

Coyne and Sproat (2001)

Shallow and Deep Semantics

primitives, relations, world knowledge grounding

how much is known about symbols?

$[[\text{life}]] = \text{life}'$

expressive power

unit of representation

words predicate/arguments sentences discourse
FrameNet Annotations are Shallow (I)

### Unit of Representation

<table>
<thead>
<tr>
<th>Developers</th>
<th>want</th>
<th>to</th>
<th>buy</th>
<th>the</th>
<th>land</th>
<th>for</th>
<th>a</th>
<th>Sainsbury</th>
<th>superstore</th>
</tr>
</thead>
<tbody>
<tr>
<td>NN2</td>
<td>VVB</td>
<td>TO0</td>
<td>VVI</td>
<td>AT0</td>
<td>NN1</td>
<td>PRP</td>
<td>AT0</td>
<td>NP0</td>
<td>NN1</td>
</tr>
<tr>
<td>Buyer</td>
<td>COM._BUY</td>
<td></td>
<td>Goods</td>
<td>NP</td>
<td>Obj</td>
<td>Recipient</td>
<td>PP</td>
<td>Dep</td>
<td></td>
</tr>
</tbody>
</table>

- predicate/argument structures
- frame-elements are filled with syntactic, not with semantic objects
- compositionality?
FrameNet Annotations are Shallow (II)

What do we know about a frame?

- Intentionally_Affect
  -inherits from Apply_heat
  -uses
  -Cooking_Creation

A **Cook** applies heat to **Food**, where the **Temperature_setting** of the heat and **Duration** of application may be specified. A **Heating_instrument**, generally indicated by a locative phrase, may also be expressed.

Lexical Units: *bake.v, barbecue.v, boil.v, braise.v, roast.n...*

Valence patterns:
- *John BAKED the aubergines in a preheated gas oven at 350°F for half an hour!*
- *The aubergines BAKE in the oven. N1*

- no formal specification of meaning of the frame?
- how do frame elements relate to each other?
Contents

1 Introduction

2 Sentence Meaning as Graphs

3 Defining Frame Meaning

4 Semantic Parsing

5 Conclusion
Contents

1 Introduction

2 Sentence Meaning as Graphs

3 Defining Frame Meaning

4 Semantic Parsing

5 Conclusion
Instantiating Frames: Types and Instances

\[
\text{Commerce\_buy}(\text{self: i6, buyer: i4, seller: i1, goods: i2, money: i3})
\]

' Mary bought an apple for $1.'
Graph Based Meaning Representation

Mary bought an apple for $1.
Why graphs?

A woman is sitting on a chair and using the computer.
Why graphs?

A woman is sitting on a chair and using the computer.

The man scratches his head.
Defining Frame Meaning

In particular, I thought of each case frame as characterizing a small abstract ‘scene’ or ‘situation’, so that to understand the semantic structure of the verb it was necessary to understand the properties of such schematized scenes.

Fillmore (1982)

- How to formally characterize scenes/situations?
- Probably depends on application.
- Two types of frames:
  - functional frames ← FrameNet
  - graphical frames ← VigNet
Two Levels of Visual Scene Description

Types of descriptions:

- **High-level:**
  
  The man washed the floor.

- **Low-level:**
  
  The man is on the floor. He is kneeling. He is holding the sponge. The bucket is near the man.
Coyne et al. (2011)

- Frames grounded in graphical primitives.
- Represent different visual realizations for functional frames.
- Vignettes extend frames by
  
  - optionally introducing new frame elements that participate in the visualization
  - describing the structure of decomposition into sub-frames using a special frame-to-frame relation

```
Commerce_counter( buyer, goods, money, seller)

(ISA) Commerce_buy
at_counter( partcpt1:buyer, partcpt2:seller, counter:c)
on(figure:goods, ground:c)
on(figure:money, ground:c)
```
Vignettes and FrameNet
Vignettes and FrameNet

- Removing (Agent, Cause, Source, Theme, ...)
- Inheritance
  - Wash_handheld_item
  - Wash_floor (Agent, Surface, Theme, SPONGE, BUCKET)
  - Wash_vehicle
  - Wash_body_part
- Subframe_parallel
  - Stance_at_location (Agent, Support_Surface, Pose = KNEEL)
  - Apply_instrument_to_target (Agent, Instrument, Target)
  - Next_to (Figure, Ground)
- Pose (Agent, Pose) On (Figure, Ground)
  - Reach (Agent, Target) Grasp (Agent, Theme)
Vignettes and FrameNet

- **Removing** (Agent, Cause, Source, Theme,...)
- **Inheritance**
- **Subframe_parallel**
- **Wash_handheld_item**
- **Wash_floor** (Agent, Surface, Theme, SPONGE, BUCKET)
- **Wash_vehicle**
- **Wash_body_part**
- **Stance_at_location** (Agent, Support_Surface, Pose = KNEEL)
- **Apply_instrument_to_target** (Agent, Instrument, Target)
- **Next_to** (Figure, Ground)

Diagram:

- Inheritance
- Subframe_parallel
- Removing(Agent, Cause, Source, Theme,...)
- Wash_handheld_item
- Wash_floor(Agent, Surface, Theme, SPONGE, BUCKET)
- Wash_vehicle
- Wash_body_part
- Stance_at_location(Agent, Support_Surface, Pose = KNEEL)
- Apply_instrument_to_target(Agent, Instrument, Target)
- Next_to(Figure, Ground)
Vignettes and FrameNet

Removing(Agent, Cause, Source, Theme,...)

Wash_handheld_item

Wash_floor(Agent, Surface, Theme, SPONGE, BUCKET)

Wash_vehicle

Wash_body_part

Stance_at_location(Agent, Support_Surface, Pose = KNEEL)

Apply_instrument_to_target(Agent, Instrument, Target)

Next_to(Figure, Ground)

On(Figure, Ground)

Pose(Agent, Pose)

Reach(Agent, Target)

Grasp(Agent, Theme)
Status of the VigNet Resource

VigNet currently contains:

- small set of primitive spatial relations (on, next-to, grasp ...)
- small set (about 30) ‘abstract‘ vignettes
  - holding/touching target or patient, using handheld instruments, using stationary machine, human poses...
- several hundred verbal vignettes using abstract vignettes
- entries for 2000 nouns mapping to about 3000 3D models (with physical attributes, parts, affordances).
- about 80 location vignettes (all rooms, including fixtures/affordances).
Semantic Parsing

The woman is sitting on a chair and using the computer

Using

agent

location

Posture

Chair

Computer

agent

machine

at-fixture

ground

Desk

Figure

on

Seat

agent

Sit on seat

on

Desk

Use machine

agent

Machine

agent

Desk

on
Synchronous Hyperedge Replacement Grammars

Jones et al. (2012)
Synchronous Hyperedge Replacement Grammars

Jones et al. (2012)
Synchronous Hyperedge Replacement Grammars

Jones et al. (2012)

Sit:Agent is sitting Sit:Seat and S

Sit:Agent is sitting Sit:Seat and S

Sit:Agent is sitting Sit:Seat and S

Sit:Agent is sitting Sit:Seat and S

Sit:Agent is sitting Sit:Seat and S
Synchronous Hyperedge Replacement Grammars

Jones et al. (2012)

The woman is sitting \textbf{Sit:Seat} and \textbf{S}
Synchronous Hyperedge Replacement Grammars

Jones et al. (2012)

The woman is sitting on a chair and S
Synchronous Hyperedge Replacement Grammars

Jones et al. (2012)

The woman is sitting on a chair and using the **Use:Instrument**
Synchronous Hyperedge Replacement Grammars

Jones et al. (2012)

The woman is sitting on a chair and using the computer
Synchronous Hyperedge Replacement Grammars

Jones et al. (2012)

The woman is sitting on a chair and using the computer
Synchronous Hyperedge Replacement Grammars

Jones et al. (2012)

The woman is sitting on a chair and using the computer
Inference

- Inference is hard: Choose correct rule based on broader context, violates context-freeness.
- Need to perform SHRG parsing and inference jointly.

Approaches:
- Encode some constraints in non-terminal alphabet.
- Beam search with Integer Linear Programming.
- Dual Decomposition.
Open Questions

- Inference
- Can we learn grammars automatically from (scene, description) pairs?
- Evaluation: Create full-text annotations by hand using new annotation tool.
FrameNet annotations are too ‘shallow’ for some applications.

Extension to FrameNet

- Graph-based meaning representation for full scenes.
- Vignettes to ground frames in semantic primitives.

Semantic parsing based on synchronous graph grammars.
Thank You!

Special thanks to Bob Coyne and Owen Rambow.


