Broad Coverage Spatial Language Understanding

James Allen, University of Rochester and IHMC
Outline

- Context and Disclaimers
- The TRIPS Language Understanding System
- Scales
- Spatial Ontology
- Examples of use
Context and Disclaimers
What is Deep understanding?

Students develop **deep** understanding when they grasp the relatively complex relationships between the central concepts of a topic or discipline. Instead of being able to recite only fragmented pieces of information, they understand the topic in a relatively systematic, integrated or holistic way. As a result of their deep understanding, they can produce new knowledge by discovering relationships, solving problems, constructing explanations and drawing conclusions.

Students have only **shallow** understanding when they do not or cannot use knowledge to make clear distinctions, present arguments, solve problems or develop more complex understanding of other related phenomena.

**DEPT. OF EDUCATION, QUEENSLAND**

**IN OTHER WORDS,**
**CONNECTING LANGUAGE TO OTHER COGNITIVE ABILITIES:**
**KNOWLEDGE, REASONING, ACTION, LEARNING, ...**

**SAME WITH MACHINES - DEEP UNDERSTANDING PRODUCES MEANING THAT IS USABLE FOR MULTIPLE TASKS, INCLUDING REASONING & EXPLANATION**
The Goal of The TRIPS Parser

- Broad-coverage parsers are inevitably shallow
  - essentially syntax (possibly with superficial predicate-argument structure)

- Deep semantic parsers are inevitably narrow
  - produce “deep” semantics for the domain they are trained on
  - but little transfer to new domains

<table>
<thead>
<tr>
<th></th>
<th>Broad Coverage</th>
<th>Narrow Coverage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shallow Representation</td>
<td>structural parsers</td>
<td></td>
</tr>
<tr>
<td>Deep Representation</td>
<td>?</td>
<td>semantic parsers</td>
</tr>
</tbody>
</table>

Can we achieve broad AND deep semantic parsing?
Understanding Requires Context

At a grocery store ... Customer: *black beans*? clerk: *aisle 3.*

*But in a home environment...*

When arriving home ... Spouse: *black beans?* You: *Oh, sorry, I forget to get them.*

When cooking ... Spouse: *black beans?* You: *in the cupboard.*

When exploring nutrition options ... Spouse: *black beans?* You: *227 calories in a cup*

When cooking (adding black beans to a pot) ... Spouse: *black beans?* You: *don’t you like them.*

To understand an utterance, we need to understand why someone is speaking to us, i.e., intention recognition.
The Dilemma

• Language technology is heavily based on interpreting structure
• But full understanding requires reasoning in context

Our Approach

A PRACTICAL MIDDLE GROUND
Requirements for the Logical Form

“universal vocabulary”
- there is one set of words and senses drawn from a generic ontology for all domains (except domain-specific technical vocabulary)

“no word left behind”
- we don’t know what may be critical in contextual interpretation later

“meaning for everyone”
- all words should map into an ontology used for reasoning

“preserve all detail and subtleties of phrasing”

“retain ambiguity whenever possible”
- quantifier scoping
- abstract word senses

“prefer compositional structures over idiosyncratic meanings”
- especially with multi-words
How are spatial concepts used in language?

MIGHT BE EASIER TO ANSWER

“What in language is NOT couched in spatial concepts!”

ENGLISH IS STRUCTURED AROUND WORDS THAT HAVE SPATIAL INTERPRETATIONS:
Space invades every part of speech

❖ PREPOSITIONS: in, on, out, by, beside, ...
❖ ADJECTIVES: near, close, adjacent, high, tall, ...
❖ VERBS: touching, supporting, covering, ...
❖ NOUNS: height, width, size, area, ...

How are all these related to each other?
The TRIPS Logical Form
The TRIPS Meaning Representation

- predicates are the common senses of the words, organized into a commonsense ontology capturing the underlying semantic notions of natural language (TRIPS ontology has about 4000 core upper-level concepts)

```
ONT::NATURAL-OBJECT
  ONT::ORGANISM
    ONT::ANIMAL
      ONT::MAMMAL
        [THE ONT::NONHUMAN-ANIMAL / dog]

ONT::EVENT-OF-STATE
  ONT::EVENT-OF-EXPERIENCE
    ONT::PERCEPTION
      [F ONT::ACTIVE-PERCEPTION / see]
        :experiencer

ONT::PHYS-OBJECT
  ONT::SUBSTANCE
    ONT::FOOD
      ONT::PREPARED-FOOD
        [THE ONT::FAST-FOOD / pizza]
          :neutral

[THE ONT::FAST-FOOD / pizza]
```

“The dog saw the pizza”
Formally, this is a constraint-based underspecified representation that subsumes Hole Semantics and MRS, 
Manshadi, Gildea & Allen, Computational Linguistics
A fragment of the event ontology

<table>
<thead>
<tr>
<th>ONTOLOGY TYPE</th>
<th>ROLES (INHERITED, NEW)</th>
<th>EXAMPLE VERBS</th>
</tr>
</thead>
<tbody>
<tr>
<td>SITUATION-ROOT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EVENT-OF-CHANGE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EVENT-OF-ACTION</td>
<td>AGENT</td>
<td></td>
</tr>
<tr>
<td>EVENT-OF-AGENT-INTERACTION</td>
<td>AGENT, AGENT1</td>
<td>meet, collaborate, ...</td>
</tr>
<tr>
<td>AGREEMENT</td>
<td>AGENT, AGENT1, FORMAL</td>
<td>agree, confirm, ...</td>
</tr>
<tr>
<td>EVENT-OF-CREATION</td>
<td>AGENT, AFFECTED</td>
<td>bake, establish, ...</td>
</tr>
<tr>
<td>EVENT-OF-CAUSATION</td>
<td>AGENT, AFFECTED</td>
<td>push, control, ...</td>
</tr>
<tr>
<td>MOTION</td>
<td>AGENT, AFFECTED, RESULT</td>
<td>go, disperse, ...</td>
</tr>
<tr>
<td>ACQUIRE</td>
<td>AGENT, AFFECTED, SOURCE</td>
<td>adopt, buy, ...</td>
</tr>
<tr>
<td>EVENT-OF-UNDERGOING-ACTION</td>
<td>AFFECTED</td>
<td>die, inherit, ...</td>
</tr>
<tr>
<td>EVENT-OF-STATE</td>
<td>NEUTRAL</td>
<td></td>
</tr>
<tr>
<td>POSITION</td>
<td>NEUTRAL, NEUTRAL1</td>
<td>contain, surround, ...</td>
</tr>
<tr>
<td>EVENT-OF-EXPERIENCE</td>
<td>NEUTRAL, EXPERIENCER</td>
<td>see, like, ...</td>
</tr>
<tr>
<td>AWARENESS</td>
<td>NEUTRAL, EXPERIENCER, FORMAL</td>
<td>believe, suspect, ...</td>
</tr>
</tbody>
</table>
Ontology Types, Roles & Restrictions

ONT::CONSUME
SEM: [Situation aspect=dynamic, time-span=extended, …]
ROLES:
  AGENT {required} [Phys-obj origin=living, …]
  AFFECTED {required} [Phys-obj comestible=+, …]
WordNet: consume%2:34:00, have%2:34:00, …

ONT::ANIMAL
SEM: [Phys-obj origin=living, …]

ONT::DEVICE
SEM: [Phys-obj origin=artifact, …]
Arguments vs Relational Roles

- Argument roles identify arguments in a predicate:
  - e.g., \( \text{PUSH}(e) \land \text{agent}(e, ag) \land \text{affected}(e, aff) \) in a Davidsonian-style representation

- Relational roles are causal/temporal relations between predicates

\[
\text{Ev}(e) \land \text{agent}(e, ag) \land \text{result}(e, p) \land \text{Occurs}(e,t) \Rightarrow \\
\text{Meets}(t, t') \land \text{Holds}(p,t') \land \text{figure}(p, ag)
\]

- e.g., “I walked into the store” in a picture ....

[Walk :agent I] [In :figure I :ground Store] :result
### TRIPS CORE SEMANTIC ROLES

<table>
<thead>
<tr>
<th>Role</th>
<th>Distinguishing Properties</th>
<th>Definition</th>
<th>Examples</th>
</tr>
</thead>
</table>
| AGENT     | +CAUSAL                    | Entity that plays a causal or initiating role as part of the event meaning  | The boy told a story  
The hammer broke the window  
The storm destroyed the house |
| AFFECTED  | -CAUSAL +CHANGED           | (non-causing) Entity that is changed as part of the meaning the event       | He carried the package  
The ice melted  
The ball hit the wall |
| NEUTRAL   | -CAUSAL -CHANGED +EXISTENT | Acausal argument, neither causing nor changed by the event, but which has existence | I saw him  
I want a pizza  
I told him a story |
| EXPERIENCER | - CAUSAL - CHANGED +COGNITION +EXISTENT | An entity undergoing a cognitive or perceptual state | The man knows the plan  
The dog saw the cat |
| FORMAL    | -CAUSAL -CHANGED -EXISTENT | Acausal argument with no temporal existence | He believes that the money’s gone  
I want to go  
He seems crazy |

### Diagram

- **causal?**
  - **AGENT**
  - **changed?**
    - **AFFECTED**
    - **existent?**
      - **cognition?**
        - **FORMAL**
        - **EXPERIENCER**
        - **NEUTRAL**
### Semantic properties of some relational roles

<table>
<thead>
<tr>
<th>Relational Role</th>
<th>Verb arguments</th>
<th>Figure of role prop’n</th>
<th>Temp. Relation between e &amp; r</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>RESULT</td>
<td>agent only</td>
<td>agent</td>
<td>te meets t_result</td>
<td>I walked into the store</td>
</tr>
<tr>
<td>RESULT</td>
<td>agent + affected</td>
<td>affected</td>
<td>te meets t_result</td>
<td>I pushed the box in the corner</td>
</tr>
<tr>
<td>SOURCE*</td>
<td>agent + affected</td>
<td>affected</td>
<td>ts_{source} overlaps te</td>
<td>I pushed the box from the shelf</td>
</tr>
<tr>
<td>TRANSIENT -RESULT*</td>
<td>agent</td>
<td>agent</td>
<td>t_{result} during te</td>
<td>I walked by the tree</td>
</tr>
<tr>
<td>METHOD</td>
<td>agent (+ others)</td>
<td>agent</td>
<td>te equals t_{method}</td>
<td>I moved the box by pushing it</td>
</tr>
<tr>
<td>LOCATION</td>
<td>any</td>
<td>event</td>
<td>n/a</td>
<td>I ran at the gym</td>
</tr>
<tr>
<td>MANNER</td>
<td>any</td>
<td>event</td>
<td>n/a</td>
<td>I ran quickly</td>
</tr>
</tbody>
</table>

* also has the second variant as with RESULT
The Resultative Construction(s)

What about “The man pushed the box in the room”

via Lexicon & grammar

FONT::PUSH

:agent

THE ONT::MALE-PERSON

:affected

THE ONT::BOX

:figure

ONT::IN-LOC

:figure

??

:ground

THE ONT::ROOM
Lexical Approach: the lexical entry contains the entire set of subcategorization frames

e.g., VerbNet entries for “push”

CARRY-11.4 (11 frames)
FORCE-59 (4 frames)
FUNNEL-9.3 (4 frames)
HOLD-15-1 (2 frames)
PUSH-12 (4 frames)
SPLIT-23.2 (6 frames)

TRIPS: two senses + a few templates

ONT::PUSH
agent-affected-templ
- “We pushed the cat”

ONT::PROVOKE
agent-formal-objectcontrol
- “We pushed him to do it”

All the other VerbNet senses correspond to one of these two + a spatial result
Other Resultative Constructions

❖ Resultative with Transitive Verbs
  ❖ They wiped the desk clean
  ❖ Sweep the dust into the bin
  ❖ He pushed it flat
❖ Resultative is intransitive
  ❖ The water froze solid
  ❖ I walked in the store
❖ Resultative with particles
  ❖ Lay the box down
  ❖ Lay down the box in the corner (2 results)
  ❖ Lay the box down in the corner (2 results)
  ❖ Lay the box in the corner down (1 result)
❖ Intransitive to transitive+result (our favorite!)
  ❖ The dog barked the cat up the tree

NOTE: Selecting of these rules over other interpretations requires commonsense knowledge about what states events typically cause:
  e.g., *They wiped the table clean* vs
  *They wiped the table happy*
Scales
(the cartoon version)

How do the different spatial concepts conveyed by the different parts of speech relate to each other?

for formulas, see

In the beginning, there was a scale ....

Scales are a conceptual organization of a set of values that can be compared (e.g., taller) sometimes quantifiable (e.g., 5 feet high)
Scales

Scales have a characteristic function that maps objects to the scale

CHAIR2 is taller than CHAIR1
properties (i.e., adjectives) are associated with a range of values ....
properties may overlap …

consider “short but not tiny”
Scales

properties may be relative to a reference class of objects ...

consider “tall for a chair but not for a stool”
Scales

scales may also be defined in terms of another object (GROUND)

e.g., “the chair is close to the door”
Scales

Scales may be quantifiable using measure phrases

e.g., “the chair is 15 inches tall”
Scales

scales support comparison operators: a objected is compared with with another (the COMPAR object)

e.g., “CHAIR2 is taller than CHAIR1 by 9 inches”

COMPARE = CHAIR1
Scales

Scales also enable selection of an object from a set (the REFSET).

e.g., “the tallest chair”
“The tallest of the chairs”

REFSET = {CHAIR1, CHAIR2, CHAIR3}
# Summary: Roles associated with scale-based properties

<table>
<thead>
<tr>
<th>Role</th>
<th>Definition</th>
<th>Example (argument is underlined)</th>
</tr>
</thead>
</table>
| FIGURE   | the argument that is being characterized with respect to other objects (the GROUND), a scale, or an relative sub scale (the STANDARD), | The red block  
The block is red.  
The larger dog  
The tallest building |
| GROUND   | the argument related to the FIGURE                                         | The building closer to the river                                                                  |
| COMPAR   | An explicit object with which the FIGURE is being compared                 | My dog is larger than your dog  
The building closer to the river than that                   |
| REFSET   | A explicit set of objects of which the FIGURE belongs                      | She is the tallest of the girls in the class  
The larger of the animals died.                  |
| SCALE    | The scale on which a predication is based (typically implicit in the predicate) | It is hotter in temperature  
It is hot spice-wise                                            |
| STANDARD | a relative subscale defined by a predicate, ranging from fairly simple (e.g., tall for a dog, the standard is the height subscale associated with dogs) to complex (e.g., short to reach the shelf) defined a standard that is a subscale of heights where someone could reach the shelf. | It is hot enough for taking a walk  
The shelf is too high to reach  
The ladder is a bit short to reach the shelf  
He is large for a dog  
He is old to be in third grade |
| EXTENT   | The amount by which the figure differs from the ground in a comparison operation | It is 6 inches longer than the shelf                                                             |
| DEGREE   | A qualitative measure of value on a scale                                  | He is very tall                                                                                  |
Sample parse involving most of the roles ...
SO FAR I’VE FOCUSED ON STRUCTURAL ASPECTS OF SPATIAL RELATIONS

NOW I TRY TO CLASSIFY THEM BY THEIR CONTENT
The Spatial Relations

Primarily adverbs (aka prepositions) and adjectives
The Spatial Objects

Primarily concrete nouns (types of locations)
The Spatial Verbs

Primarily stative verbs
The Spatial Scales

Primarily abstract nouns
Spatial Relations relate an object (the **FIGURE**)
to a reference object (the **GROUND**), sometimes implicit,
with respect to some measure function (the **SCALE**), typically implicit

*The box is near the door*

**ONT::NEAR-RELN**

**ONT::DISTANCE-SCALE**
The Spatial Relations

**ONT::POSITION-AS-POINT-RELN**
Both FIGURE and GROUND are viewed as point-like objects.

**ONT::ORIENTED-LOC—RELN**
SCALE involves a directional orientation.

- **ONT::DIRECTIONAL-VERT—RELN**  *above the counter*
  SCALE involves a vertical orientation.

- **ONT::NAVIGATIONAL—RELN**  *north of the barn*
  SCALE involves a vertical orientation.

**ONT::POS-DISTANCE**  *near the factory*
SCALE is distance between two points.

**ONT::AT-LOC**  *at the store*
SCALE is degree of co-location of FIGURE and GROUND.
Spatial Relations

ONT::POSITION-WRT-AREA-RELN
GROUND is an extended space

ONT::ACROSS    the path across the field
FIGURE bisects the GROUND

ONT::THROUGH   the road through the tunnel
FIGURE bisects the GROUND and is conceptually IN it

ONT::POS-AS-OPPOSITE the house across the street
FIGURE is on opposite side of GROUND from a reference object

ONT::DISTRIBUTED-POS the flowers throughout the field
FIGURE is a set of objects in or on the GROUND

ONT::AROUND    The path around the city
FIGURE is a space that surrounds the GROUND, or goes through the GROUND
Spatial Relations

ONT::POSITION-CONTAINMENT-RELN
GROUND is viewed as a container (physical or abstract), and may be an extended space

ONT::IN-LOC  *the dog in the box, the idea in my mind*
FIGURE is contained in the GROUND

ONT::CONTAINS  *the park contains a fountain*
FIGURE contains the GROUND

ONT::OUTSIDE  *the trees outside the park*
FIGURE is nor contained in the GROUND

*at*
Spatial Relations

ONT::DIRECTION

The prototypical sense of direction relates an object (the FIGURE) to itself at another time along some orientation.

*the dog moved away from the house*

\[ \text{DISTANCE}(\text{dog}, \text{house}, t2) > \text{DISTANCE}(\text{dog}, \text{house}, t1) \]

i.e., the RESULT of the move event is that the dog is further from the house than when it started (e.g., \text{AWAY}(\text{dog}, \text{house}, t1, t2))
Spatial Relations

**ONT::DIRECTION**
FIGURE is changing location relative to some GROUND

**ONT::DIRECTION-WRT-ENTITY-RELN**
Change is relative to some orientation based on an entity

**ONT::DIRECTION-FORWARD**
the dog moved forward

**ONT::TOWARDS**
It rolled towards the house

**ONT::DIRECTION-ROTATION**
It turned clockwise

**ONT::DIRECTION-WRT-VERTICALITY**
We pushed it up, they ran down the mountain

**ONT::CARDINAL-DIRECTION**
they moved east/eastward

FIGURE changes

FIGURE changes in a vertical direction
The Spatial Objects

Primarily concrete nouns (types of locations)
The Spatial Verbs

Primarily stative verbs
Spatial Relations Realized by Verbs

- While spatial relations are most commonly realized as preposition/adverbs and adjectives,
- Some spatial relationships can only be described using verbs.

  - The shovel is leaning against the fence.
  - The dog is touching the door.

  The ball is in the box == the box contains the ball.
Spatial Relations as Verbs

**ONT::POSITION**
NEUTRAL is in a spatial relationship with NEUTRAL1

**ONT::BE-AT**
NEUTRAL is at NEUTRAL in some postural position
*The cup is sitting on the table/hanging from the shelf/leaning against the wall*

**ONT::SURROUND**  *The sense surrounds the field*

**ONT::SUPPORT**  *The legs can support the table*

**ONT::COVER**  *The blanket covered the bed*

**ONT::CONNECTED**  *The two countries connect at the river.*

**ONT::INTERSECT**  *The roads cross near there*
Pondering

Where do spatial concepts end and non-spatial begin?

The blanket covered the bed

The car fits five people

He is holding a pizza!

He carried the backpack

He bought a backpack

does it matter?
Our Approach

(CONTEXTUALLY-INFLUENCED) GENERIC SEMANTIC PARSING LANGUAGE

A PRACTICAL MIDDLE GROUND

LOGICAL FORM

CONTEXTUAL INTERPRETATION

INTENDED MEANING IN CONTEXT
Examples of Applications
Building and Learning Structures in a Situated Blocks World Through Deep Language Understanding

Ian Perera, See paper this workshop
Constraint Interpretation

Referring Expression

Predicate

Feature

Operator

Constraint: 

(F (:* COLUMN-FORMATION COLUMN))

MOD:FIGURE

(F (:* LEFT LEFTMOST))

Predicate

Referring Expression

Constraint: 

(F (:* COLUMN-FORMATION COLUMN))

MOD:FIGURE

(F (:* RIGHT RIGHTMOST))

Predicate

Referring Expression
Extraction of the spatial roles and relations in descriptions of images

Parisa Kordjamshidi, Tulane

About 20 kids in traditional clothing and hats waiting on stairs

Extract the following roles and relations:

- (About 20 Kids) is the **trajector**
- (On) is the **spatial indicator**
- (Stairs) is the **landmark**
- (About 20 kids, on, Stairs) is the **spatial triplet**.
Initial Results

- The results show a significant improvement on spatial relation extractions based on experiments on a subset of CLEF annotated benchmark.

<table>
<thead>
<tr>
<th></th>
<th>Without TRIPS</th>
<th>With TRIPS relation labels</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Precision</td>
<td>Recall</td>
</tr>
<tr>
<td>Trajectory</td>
<td>61.111</td>
<td>64.706</td>
</tr>
<tr>
<td>Landmark</td>
<td>84.615</td>
<td>78.571</td>
</tr>
<tr>
<td>Indicator</td>
<td>100.000</td>
<td>85.714</td>
</tr>
<tr>
<td>Triplets</td>
<td>55.556</td>
<td>58.824</td>
</tr>
</tbody>
</table>

- Result 1: The most challenging part is the relation extraction, and previous models have good performance on Roles. It seems TRIPS has a significant impact on relation (triplet extraction). Particularly, the relation labels improve the recall and consequently improve the overall F1.

- Result 2: The TRIPS phrases types still improved the relation extraction but dropped the results of role classification more. We assume the phrase types vary a lot and generate sparse features. We have not used TRIPS ontology still to get the more generic types of phrases but we assume by integrating the ontology we can obtain better results.
Domains Using TRIPS Parser

Big Mechanism: reading biology research papers
"We hypothesized that MEK inhibition activates AKT by inhibiting ERK activity, which blocks an inhibitory threonine phosphorylation of EGFR and HER2, thereby increasing ERBB3 phosphorylation."

CWC Blocks World: Collaborative action
"Why don’t we place the large blue block behind that wall"

CWC Biocuration: exploring Bio pathways
"Is the amount of MAP2K1-MAPK1 complex sustained at a high level if we increase the total amount of MAPK1 and DUSP6 by 5 fold?"

Crop Modeling: Building causal models from text
"The government promotes high-yielding and drought/flood-tolerant rice varieties with policy to encourage the application of organic fertilizers, decreasing the cost on inorganic fertilizers"

Using the same grammar, lexicon, and ontology for every domain! and no training corpus required!

CWC Story Understanding
“Sandra was walking to the store. She passed a little girl who was crying on her front steps. Sandra asked her what was wrong. The little girl said she was locked out of her house. Sandra sat down and waited with her for her parents to come home”

Music Composition
“Move these notes up a step”

ASMA
(texting with teens)
“K”
Try it out and explore!

- Pointers to a TRIPS parsers customized to different domains at
  - www.trips.ihmc/parser
- also provides web services for programmatic access
- Browse the Lexicon and Ontology at
  - www.cs.rochester.edu/research/trips/lexicon/browse-ont-lex.html


