Inference by Coincidence and the Extraction of Relational Information

Simon Dennis
University of Colorado
Current Approaches to Extracting Propositional Information from Text

Parse

```
S
  /\   
NP  VP
  /\   /
 N  V  N
 Sampras defeated Agassi
```

Assign Roles

Verb (defeated)
Head Noun (Sampras)
Structural Properties (two up two down)

(c.f. Gildea & Jurafsky 2002 Blaheta & Charniak 2000)
Problems

- Require large sets of parsed sentences for training
- Must predefine semantic roles
- Require large set of role labeled sentences for training
- Still don’t work well especially across corpora
The SP Model in a Nutshell

- Assumes that people store a large number of sentence instances (sequential traces).
- When trying to interpret a new sentence they retrieve similar sentences from memory and align these with the new sentence.
- The set of alignments is an interpretation of the sentence.
- Training involves adding new traces to memory and inducing word-to-word correspondences that are used to choose the optimal alignments.
The set of words that aligns with each word from the target sentence represents the role that that word plays in the sentence.

{Ellen, Sue, Pat} represents the lovee role and {George, Michael, Joe} the lover role.

The model assumes that two sentences convey similar factual content to the extent that they contain similar words aligned with similar sets of words.

Interpretations are entered into memory as relational traces and relational retrieval and resolution affect subsequent alignment.
The SP Architecture

### Working Memory

<table>
<thead>
<tr>
<th>Who</th>
<th>Who</th>
</tr>
</thead>
<tbody>
<tr>
<td>did</td>
<td>did</td>
</tr>
<tr>
<td>Sampras</td>
<td>Sampras</td>
</tr>
<tr>
<td>beat</td>
<td>beat</td>
</tr>
<tr>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>#</td>
<td>#</td>
</tr>
</tbody>
</table>

### Sequential Long-Term Memory

- Sampras defeated Agassi
- Kuerten defeated Roddick
- Hewitt defeated Costa
- Who did Kuerten beat? Roddick
- Who did Hewitt beat? Costa

### Relational Long-Term Memory

- **Sampras**: Kuerten, Hewitt
- **Agassi**: Roddick, Costa
- **Kuerten**: Sampras, Hewitt
- **Roddick**: Agassi, Costa
- **Hewitt**: Sampras, Kuerten
- **Costa**: Agassi, Roddick
- **Kuerten**: Hewitt
- **Roddick**: Costa
- **Hewitt**: Kuerten
- **Costa**: Roddick
Sequential Retrieval

### Working Memory

<table>
<thead>
<tr>
<th>Who</th>
<th>Who</th>
</tr>
</thead>
<tbody>
<tr>
<td>did</td>
<td>did</td>
</tr>
<tr>
<td>Sampras</td>
<td>Sampras</td>
</tr>
<tr>
<td>beat</td>
<td>beat</td>
</tr>
<tr>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>#</td>
<td>#</td>
</tr>
</tbody>
</table>

### Sequential Long-Term Memory

- Sampras defeated Agassi
- Kuerten defeated Roddick
- Hewitt defeated Costa

Who did Kuerten beat? Roddick
Who did Hewitt beat? Costa

### Relational Long-Term Memory

- **Sampras**: Kuerten, Hewitt
- **Agassi**: Roddick, Costa
- **Kuerten**: Sampras, Hewitt
- **Roddick**: Agassi, Costa
- **Hewitt**: Sampras, Kuerten
- **Costa**: Agassi, Roddick
- **Kuerten**: Hewitt
- **Roddick**: Costa
- **Hewitt**: Kuerten
- **Costa**: Roddick
### Sequential Resolution

**Working Memory**

<table>
<thead>
<tr>
<th>Who</th>
<th>Who</th>
</tr>
</thead>
<tbody>
<tr>
<td>did</td>
<td>did</td>
</tr>
<tr>
<td>Sampras</td>
<td>Kuerten</td>
</tr>
<tr>
<td>Hewitt</td>
<td></td>
</tr>
<tr>
<td>beat</td>
<td>beat</td>
</tr>
</tbody>
</table>

| ?         | ?         |
|          |           |

<table>
<thead>
<tr>
<th>#</th>
<th>Roddick</th>
</tr>
</thead>
</table>

**Sequential Long-Term Memory**

- Sampras defeated Agassi
- Kuerten defeated Roddick
- Hewitt defeated Costa

Who did Kuerten beat? Roddick
Who did Hewitt beat? Costa

**Relational Long-Term Memory**

- **Sampras**: Kuerten, Hewitt
  - **Agassi**: Roddick, Costa
- **Kuerten**: Sampras, Hewitt
  - **Roddick**: Agassi, Costa
- **Hewitt**: Sampras, Kuerten
  - **Costa**: Agassi, Roddick
- **Kuerten**: Hewitt
  - **Roddick**: Costa
- **Hewitt**: Kuerten
  - **Costa**: Roddick
## Relational Retrieval

### Working Memory

<table>
<thead>
<tr>
<th>Who</th>
<th>Who</th>
</tr>
</thead>
<tbody>
<tr>
<td>did</td>
<td>did</td>
</tr>
<tr>
<td>Sampras</td>
<td>Kuerten, Hewitt</td>
</tr>
<tr>
<td>beat</td>
<td>beat</td>
</tr>
<tr>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>#</td>
<td>Roddick, Costa</td>
</tr>
</tbody>
</table>

### Sequential Long-Term Memory
- Sampras defeated Agassi
- Kuerten defeated Roddick
- Hewitt defeated Costa
- Who did Kuerten beat? Roddick
- Who did Hewitt beat? Costa

### Relational Long-Term Memory
- **Sampras**: Kuerten, Hewitt
- **Agassi**: Roddick, Costa
- **Kuerten**: Sampras, Hewitt
- **Roddick**: Agassi, Costa
- **Hewitt**: Sampras, Kuerten
- **Costa**: Agassi, Roddick
- **Kuerten**: Hewitt
- **Roddick**: Costa
- **Hewitt**: Kuerten
- **Costa**: Roddick
Relational Resolution

### Working Memory

<table>
<thead>
<tr>
<th>Who</th>
<th>Who</th>
</tr>
</thead>
<tbody>
<tr>
<td>did</td>
<td>did</td>
</tr>
<tr>
<td>Sampras</td>
<td>Sampras, Kuerten, Hewitt</td>
</tr>
<tr>
<td>beat</td>
<td>beat</td>
</tr>
<tr>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>#</td>
<td>Agassi, Roddick, Costa</td>
</tr>
</tbody>
</table>

### Sequential Long-Term Memory

- Sampras defeated Agassi
- Kuerten defeated Roddick
- Hewitt defeated Costa
- Who did Kuerten beat? Roddick
- Who did Hewitt beat? Costa

### Relational Long-Term Memory

- **Sampras**: Kuerten, Hewitt
- **Agassi**: Roddick, Costa
- **Kuerten**: Sampras, Hewitt
- **Roddick**: Agassi, Costa
- **Hewitt**: Sampras, Kuerten
- **Costa**: Agassi, Roddick
- **Kuerten**: Hewitt
- **Roddick**: Costa
- **Hewitt**: Kuerten
- **Costa**: Roddick
String Edit Theory

- Characterize the similarity and alignment of strings by sequence of edit operations required to turn one string into another (matches, changes, insertions and deletions).

  John loves Mary
  |   |   |
  Bert loves Ellen

- Deletes allow alignment of sentences of different lengths:

  Little John loves Mary
  |   |   |
  - Bert loves Ellen

  or

  John loves Mary -
  |   |   |
  - Bert loves Ellen

- 63 ways to align sentences not all equally likely
- Edit probabilities trained using a version of EM algorithm
The Bayesian Framework

- Calculate expected probability of substitution given the sentential context:

\[
E_k [P(<W_m,T_i>|T)] = \sum_{k=1}^{N} \frac{P(S_k \mapsto T|T)}{\sum_{i=1}^{N} P(S_i \mapsto T|T)} P(<W_m,T_i>|S_k \mapsto T,T)
\]

- Employs Bayesian likelihood framework:

\[
\frac{P(S_k \mapsto T|T)}{P(S_k \mapsto T|T)} = \frac{P(T|S_k \mapsto T)P(S_k \mapsto T)}{P(T|S_k \mapsto T)P(S_k \mapsto T)} = \frac{P(T|S_k \mapsto T)}{P(T|S_k \mapsto T)}
\]
Relational Retrieval and Resolution

- Similar model to sequential case – assuming relational representation of target was generated by editing one of the relational traces in memory.

\[
E_k[P(<W_m,T_i>|RT)] = \sum_{k=1}^{N} \frac{P(R_k \leftrightarrow RT|RT)}{\sum_{i=1}^{N} P(R_i \leftrightarrow RT|RT)} P(<W_m,T_i>|R_k \leftrightarrow RT,RT)
\]
Intentional vs Extensional Semantics

- **Intentional Semantics**
  - Defined by intended use

  Defeated(Sampras, Agassi)

- **Extensional Semantics**
  - Defined by enumerating instances

  Sampras: Kuerten, Hewitt
  Agassi: Roddick, Costa

- Most systems employ intentional semantics
- SP uses extensional semantics:
  - Easier to make mapping from text to meaning representation
  - Allows system to take advantage of **inference by coincidence**
Inference by Coincidence

- Simple inference is an emergent property of the model as there is coherence between the fillers of different roles.

  e.g. Sampras claims 14th Grand Slam title.
  Sampras defeated Agassi at Flushing Meadows

- If Sampras won the tournament there is a good chance he won the match.

- There is overlap between the winner-of-match role and the winner-of-tournament role.

- No inference process needed – occurs as a consequence of the causality being expressed by the corpus.
Question answering

- 69 articles were taken from the Association of Tennis Professionals (ATP) website at http://www.atptennis.com/.
- 377 questions of the form:

  Who won the match between X and Y? X

- Choice of Domain
  - Naturally occurring text available
  - Questions not susceptible to type heuristics
  - Opportunity for inference by coincidence
Results

Correct, 251
Loser, 97
Other Player, 18
Other Symbol, 11
Sequential Only

- Other Player: 342
- Correct: 30
- Other Symbol: 0
- Loser: 5
Inference by Coincidence

Results

![Bar chart showing accuracy of responses. The chart includes categories for Correct, Wrong, Literal, Inference, and Other. The data points are as follows:

- Correct:
  - Literal: 79
  - Inference: 113
  - Other: 78

- Wrong:
  - Literal: 31
  - Inference: 27
  - Other: 49

Accuracy of Response vs. Number of Responses.]}
Conclusions

- No grammar
- No predefined semantic roles
- No conceptual schema
- Yet questions that require role information can be answered
- Automatic “inference by coincidence” can occur as an emergent property of retrieval without any inference process