Natural and Formal Languages – Current Activity

Henrik Björklund
with thanks to the group

Umeå University

March 26, 2014
Media in context (MICO)
Subdivide the content into logical units and do, e.g.,

- face detection and recognition,
- speech and music discrimination,
- **speech transcription**, 
- logotype recognition,
- quality assessment,
- **sentiment analysis**, 
- character recognition, and
- mood detection.
Extractors

Subdivide the content into logical units and do, e.g.,

- face detection and recognition,
- speech and music discrimination,
- **speech transcription**,  
- logotype recognition,
- quality assessment,
- **sentiment analysis**,  
- character recognition, and
- mood detection.

Interactive learning of extractors
Mildly context-sensitive languages
Mildly context-sensitive languages
Syntax-based text classification

Into love out again

and thus I went thus I go
Syntax-based text classification

Due to the graph representation of the sentence, the natural text can be reconstructed as follows:

Into love and out again, thus I went and thus I go.
What does the following little program do?

```ruby
s = "aaababccc"
if s =~ /(a*b)*/ then
  if $' =~ /a.*c/ then
    return "yes"
  end
end
return "no"
```
What does the following little program do?

```plaintext
s = "aaababccc"
if s =~ /(a*b)*/ then
    if $' =~ /a.*c/ then
        return "yes"
    end
end
return "no"
```

Well, importantly it is not \((a*b)^*(a.*c)\)
Millstream systems

Figure 3: A sample configuration that relates a syntactic and a semantic tree.

the domain state, namely the state that the first argument (commonly called the agent) loves the second (the patient). The links establish correspondences between nodes in the two trees showing, for example, that the verb of the sentence corresponds to the function loving whose two arguments correspond to the two noun phrases of the sentence. We note here that this correspondence must respect the different thematic roles (agent and patient) of the noun phrases, thus making sure that the agent is really the first argument of the semantic function. Otherwise, the state object obtained by evaluating the semantic tree would express the wrong thing.

In realistic settings, one would use more elaborate term graphs on both the syntactic and the semantic sides. For example, one might decompose the verb into its stem love and the inflection s indicating present tense. Semantically, one would at least add a node above the current root. This node would be a function taking a state as its input and turning it into a situation in the present, that is, it would reflect the temporal information provided by the inflection. That node would then be linked with the inflection that gives rise to it (see, e.g., [26]). This slightly more elaborate configuration is shown in Figure 4. However, since we primarily want to convey the idea behind our way of constructing configurations, we will stick to the more simplified type of configurations shown in Figure 3 for the examples discussed throughout the rest of this paper.

3. Building Configurations Incrementally by Graph Transformation

We are given a sentence. We want to "understand" it. Technically, more precisely and more modestly, we want to turn this sentence into a "correct" Millstream configuration, if possible. What is a correct Millstream configuration for a given sentence and how can we find it? These questions are addressed in the present section of this paper. As described above, a Millstream system contains k modules, one for each of the k trees in a configuration. Furthermore, we are...
Further projects

- Parsing of contextual graph grammars
Further projects

- Parsing of contextual graph grammars
- Weighted automata: $n$ best trees
Further projects

- Parsing of contextual graph grammars
- Weighted automata: \( n \) best trees
- Survey of grammatical inference
Further projects

- Parsing of contextual graph grammars
- Weighted automata: $n$ best trees
- Survey of grammatical inference
- And the Berglund conjecture again...