Statistical Transfer Updates

Michael W Goodman
9th DELPH-IN Summit
31 July 2013

Semantics in MT showing active interest

- AMR as interlingua
- Several papers upcoming at ACL 2013
  - “Semantic Parsing as Machine Translation” by Jacob Andreas, Andreas Vlachos, Stephen Clark
  - “Semantic Roles for String to Tree Machine Translation” by Marzieh Bazrafshan and Daniel Gildea

- Last year I proposed statistical transfer
  - For deep MT, parse/generate as usual, but transfer MRSs as an SMT task
  - Train a transfer model from MRS bitexts
  - Decode by mapping MRS subgraphs
  - Semantic language model to help with reassembling a complete MRS
- Unfortunately I don’t have results to present, but I’ll share some developments

Training

First challenge is getting useful alignments

- Graph-based alignment is hard (any takers?)
- If we can linearize the MRS, we can use standard methods (e.g. GIZA++)
  - By character position (CFROM)
  - By graph traversal; e.g., mrs-traverse(INDEX)

```
mrs-traverse(x, done):
if done[x] == True: return;
posit x;
done[x] = True;
for y sharing x's label:
   mrs-traverse(y, done);
for z in x's arguments:
   mrs-traverse(z, done);
```

Training

Extracting subgraphs

- Can the alignments be used to find useful subgraphs from the MRSs?
- Hierarchical subgraphs?

Decoding

- Rather than mapping to MRS fragments and reassembling later, map subgraphs in-place, relinking in- and out-arguments as necessary
  - Something like synchronous CFGs, but not a tree
  - Or perhaps hyperedge replacement grammars
- Layered decoding
  - Preds and arguments
  - QEQs
  - Variable properties

Questions

- Who has an active interest in deep MT (i.e. will work on it in the next year)?
- What sub-paradigms will you be using?
  - Traditional, hand-written transfer rules
  - Inferred transfer rules
  - SMT with semantics (e.g. statistical transfer)
  - other?
- What challenges do you anticipate?
- What resources will you use?
- What language pairs?