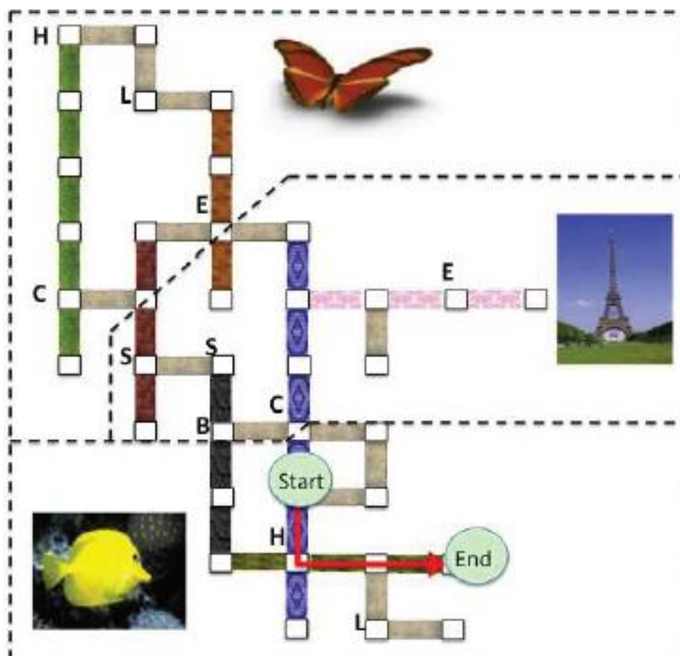
Unsupervised PCFG Induction for Grounded Language Learning with Highly Ambiguous Supervision - Kim and Mooney '12

Presented by Vempati Anurag Sai

SE367 – Cognitive Science: HW3

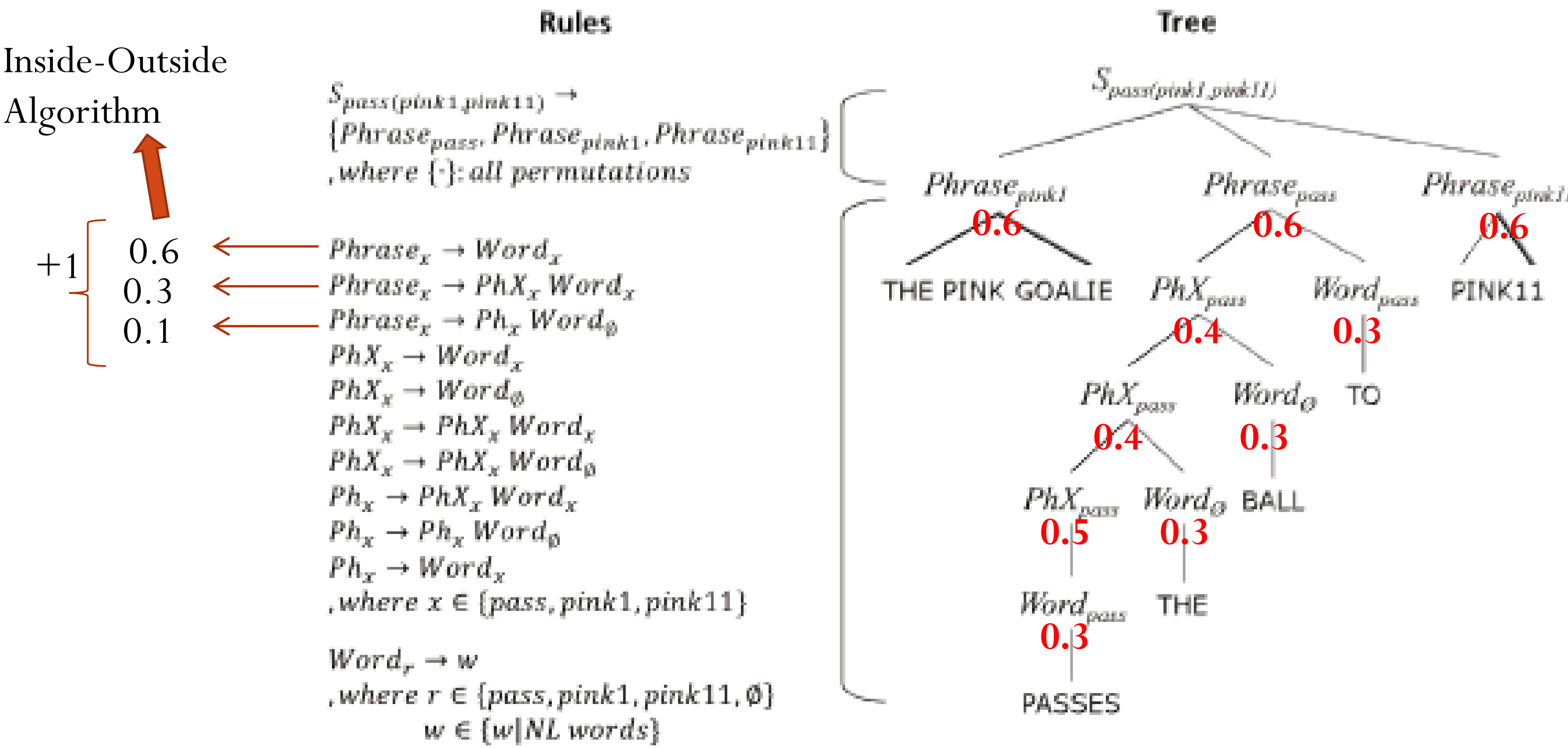
Introduction

- "Grounded" language learning
  - Given sentences in NL paired with relevant but ambiguous perceptual context, being able to interpret and generate language describing world events. Eg. Sports casting problem (Chen & Mooney (CM), '08), navigation problem (Chen & Mooney, '11) etc.
- **Navigation Problem**: Formally, given training data of the form $\{(e_1, a_1, w_1), \ldots, (e_N, a_N, w_N)\}$, where e_i is an NL instruction, a_i is an observed action sequence, and w_i is the current world state (patterns of floors and walls, positions of landmarks, etc.), we want to produce the correct actions a_j for a novel (e_j, w_j).



Related Work

- Borschinger et al. ('11) introduced grounded language learning based on PCFG (Probability Context Free Grammar) which did well in low level ambiguity scenarios like sports casting but, fails to scale to tasks where each instruction can refer to a large set of meanings as in Navigation problem.



Related Work

- There are combinatorial number of possible meanings for a given instruction which again grows exponential in number of objects and world-states that occur when the instruction is followed.
  - CM'11 avoid enumerating all the meanings and build a semantic lexicon that maps words/phrases to formal representations of actions
  - This lexicon is used for obtaining MR (Meaning representation) for an observed instruction.
  - These MRs are used to train a sematic parser capable of mapping instructions to formal meanings

Instruction: "at the easel, go left and then take a right onto the blue path at the corner"

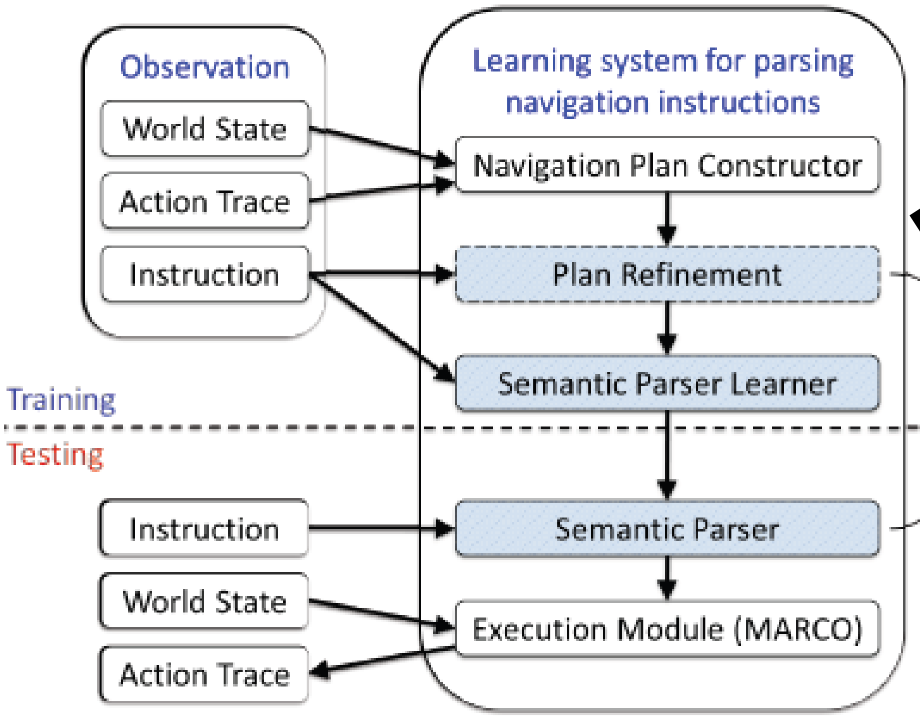
Landmarks plan:
Travel (steps: 1) ,
Verify (at: EASEL , side: CONCRETE HALLWAY) ,
Turn (LEFT) ,
Verify (front: CONCRETE HALLWAY) ,
Travel (steps: 1) ,
Verify (side: BLUE HALLWAY , **front: WALL**) ,
Turn (RIGHT) ,
Verify (back: WALL , **front: BLUE HALLWAY** , front: CHAIR ,
front: HATRACK , left: WALL , right: EASEL)

Proposed Method

- Our Method:

- For each action a_i, let c_i be the landmark plan representing context of each action and landmarks encountered. Now a particular plan p_i, as suggested by the instruction would be a subset of c_i. As we can see, there are many possible plans that could be MR of an instruction.
  - Combinatorial matching problem between e_i and c_i
- Given: Training set with (e_i, c_i) pairs.
- Lexicon is learnt by evaluating pairs of words/phrases w_j , and MR graphs, m_j , and scoring them based on how much more likely m_j is a subgraph of the context c_i when w_j occurs in the corresponding instruction e_i.

Changes to CM'11



Lexicon learnt by scoring (w_j, m_j) pairs

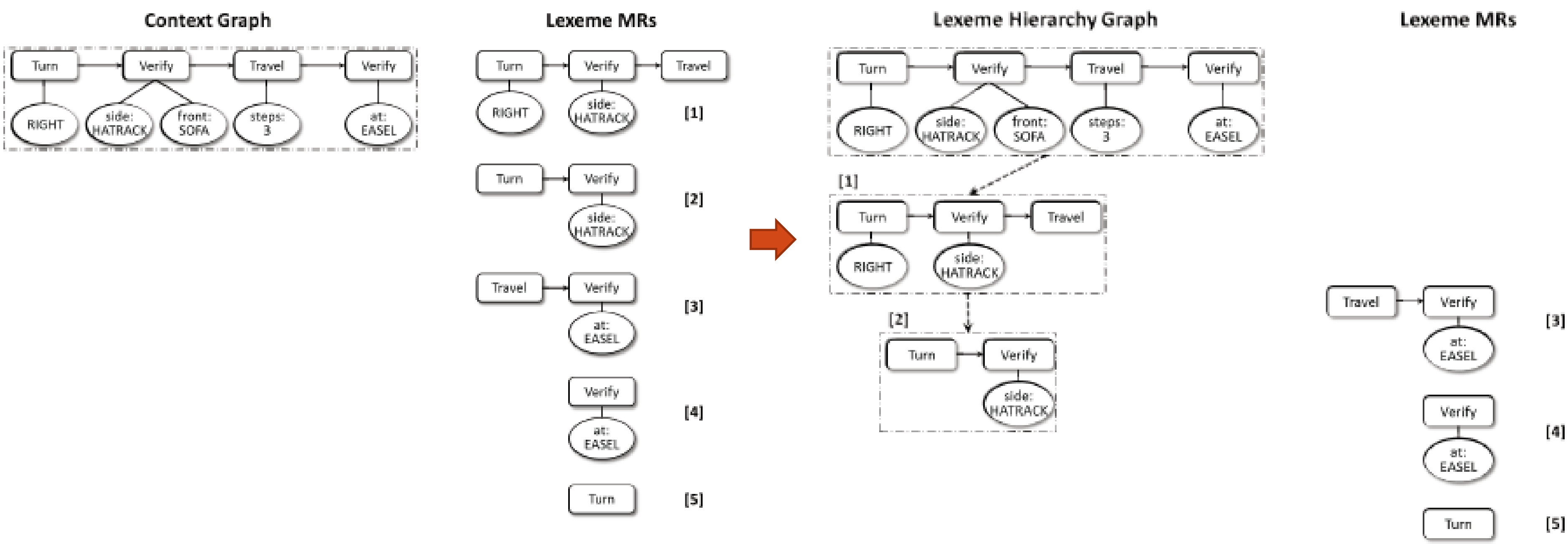
$p_i = \arg\max_j S(w_j, m_j)$ such that, w_j belongs to e_i

(e_i, p_i) pairs used as training inputs for semantic parser learner

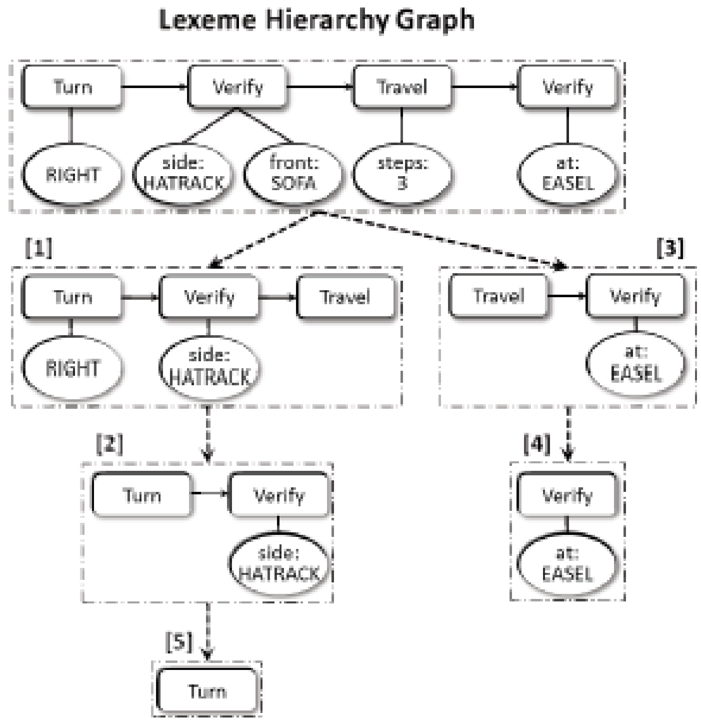
The chunk that is changed

PCFG Framework

- Lexeme Hierarchy Graph (LHG)
  - Since lexeme MRs are analogous to syntactic categories in that complex lexeme MRs represent complicated semantic concepts whereas simple MRs represent simple concepts, it is natural to construct hierarchy amongst them.
  - Hierarchical sub graph relationships between the lexeme MRs in the learned semantic lexicon to produce a smaller, more focused set of PCFG rules.
  - Analogous to hierarchical relations between non-terminals in syntactic parsing



Continued…



Completely built LHG

- Pseudo Lexems
- LHGs of all the training examples are used to generate production rules for PCFG.
- Instead of generating NL words from each atomic MR, words are generated from Lexeme MRs and small Lexeme MRs are generated from complex ones.
  - No Combinatorial explosion!!!!

Continued...

$$Root \rightarrow S_c, \qquad \forall c \in contexts$$

$$\begin{aligned}
&\forall non\text{-}leaf\ node\ and\ its\ MR\ m \\
&S_m \rightarrow \{S_{m_1}, \dots, S_{m_n}\}, \\
&\quad where\ m_1, \dots, m_n: children\ lexeme\ MR\ of\ m, \\
&\qquad\qquad \{\cdot\}: all\ k\text{-}permutations\ for\ k = 1, \dots, n
\end{aligned}$$

k-permutations of child MRs for every Lexeme MR node

Production rules generated from LHGs

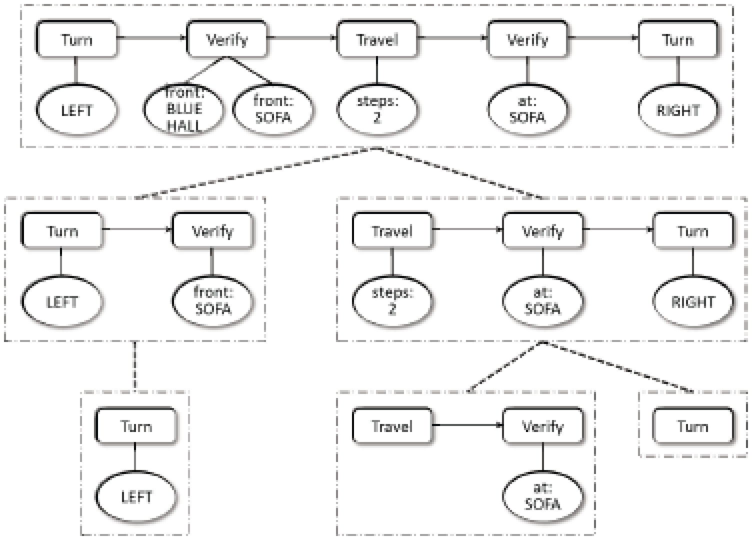
$$\begin{aligned}
&\forall lexeme\ MR\ m \\
&S_m \rightarrow Phrase_m \\
&Phrase_m \rightarrow Word_m && PhX_m \rightarrow PhX_m\ Word_m \\
&Phrase_m \rightarrow PhX_m\ Word_m && PhX_m \rightarrow PhX_m\ Word_\emptyset \\
&Phrase_m \rightarrow Ph_m\ Word_\emptyset && Ph_m \rightarrow PhX_m\ Word_m \\
&PhX_m \rightarrow Word_m && Ph_m \rightarrow Ph_m\ Word_\emptyset \\
&PhX_m \rightarrow Word_\emptyset && Ph_m \rightarrow Word_m \\
&Word_m \rightarrow s, \quad \forall s\ \text{s.t.}\ (s, m) \in lexicon\ L \\
&Word_m \rightarrow w, \quad \forall word\ w \in s\ \text{s.t.}\ (s, m) \in lexicon\ L
\end{aligned}$$

$$Word_\emptyset \rightarrow w, \qquad \forall word\ w \in NLs$$

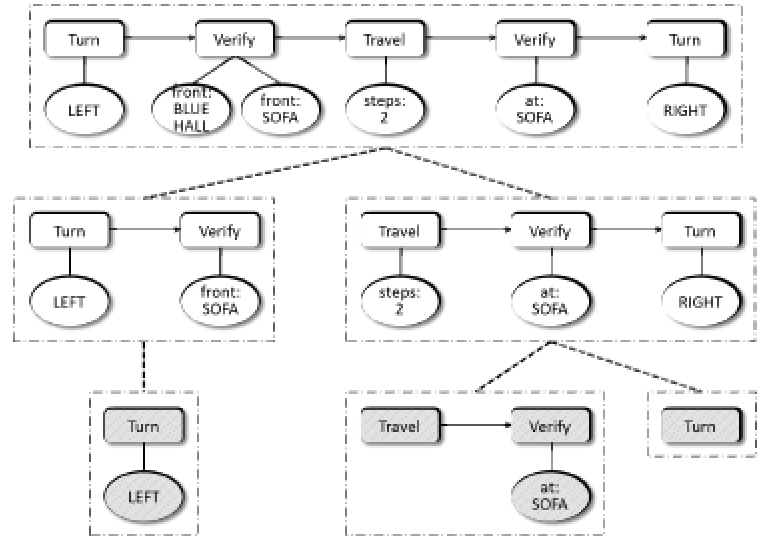
- Including k-permutations of child MRs for every Lexeme MR node makes the rule book more rich. This results in producing MRs that weren't present in the Training set which wasn't possible in Borshinger et al.

Parsing Novel NL Sentences

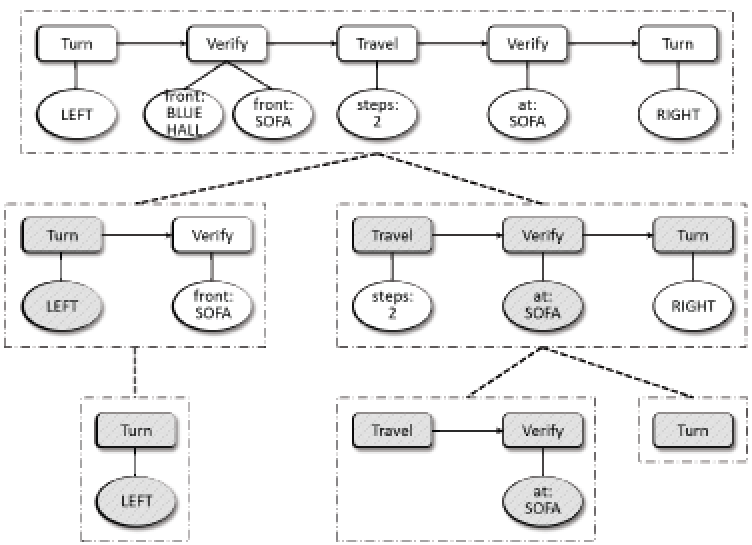
- To learn the parameters of the resulting PCFG, the Inside-Outside algorithm is used. Then, the standard probabilistic CKY algorithm is used to produce the most probable parse for novel NL sentences (Jurafsky and Martin, 2000).
- Borschinger et al. simply read the MR, m, for a sentence off the top nonterminal of the most probable parse tree. However, in this paper, the correct MR is constructed by properly composing the appropriate subset of lexeme MRs from the most-probable parse tree.



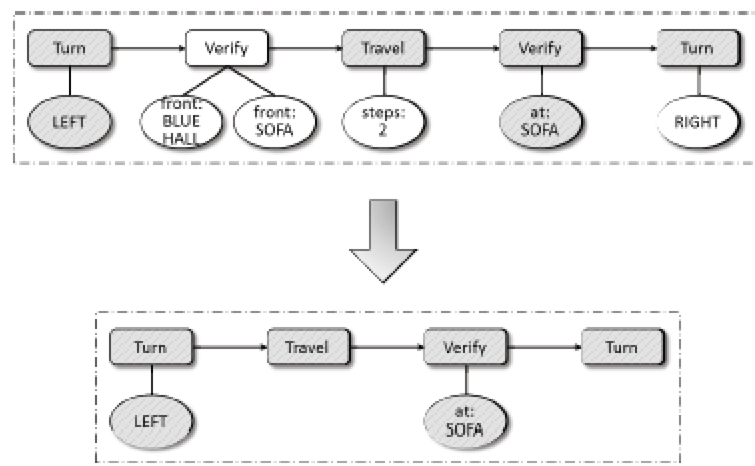
(a) Pruned parse tree showing only MRs for S_m nodes



(b) Leaf nodes have all their elements marked



(c) Upper level nodes are marked according to leaf-node markings



(d) Removing all unmarked elements for the root node leads to the final MR output

Results

- Measure of how good the system is able to convert NL sentences into correct MRs in a new test environment:

	Precision	Recall	F1
Our system	87.58	∗65.41	∗**74.81**
CM	∗90.22	55.10	68.37

- Efficiency in executing novel test instructions:

	Single-sentence	Paragraph
Our system	∗**57.22%**	∗**20.17%**
CM	54.40%	16.18%

References

- Joohyun Kim and Raymond J. 2012. Mooney, "Unsupervised PCFG Induction for Grounded Language Learning with Highly Ambiguous Supervision"
- Benjamin Borschinger, Bevan K. Jones, and Mark Johnson. 2011. "Reducing grounded learning tasks to grammatical inference"
- David L. Chen and Raymond J. Mooney. 2011. "Learning to interpret natural language navigation instructions from observations"

QUESTIONS???