Learning to parse natural language commands to a robot control system

By,

Cynthia Matuszek, Evan Herbst, Luke Zettlemoyer, Dieter Fox

Presentated by,

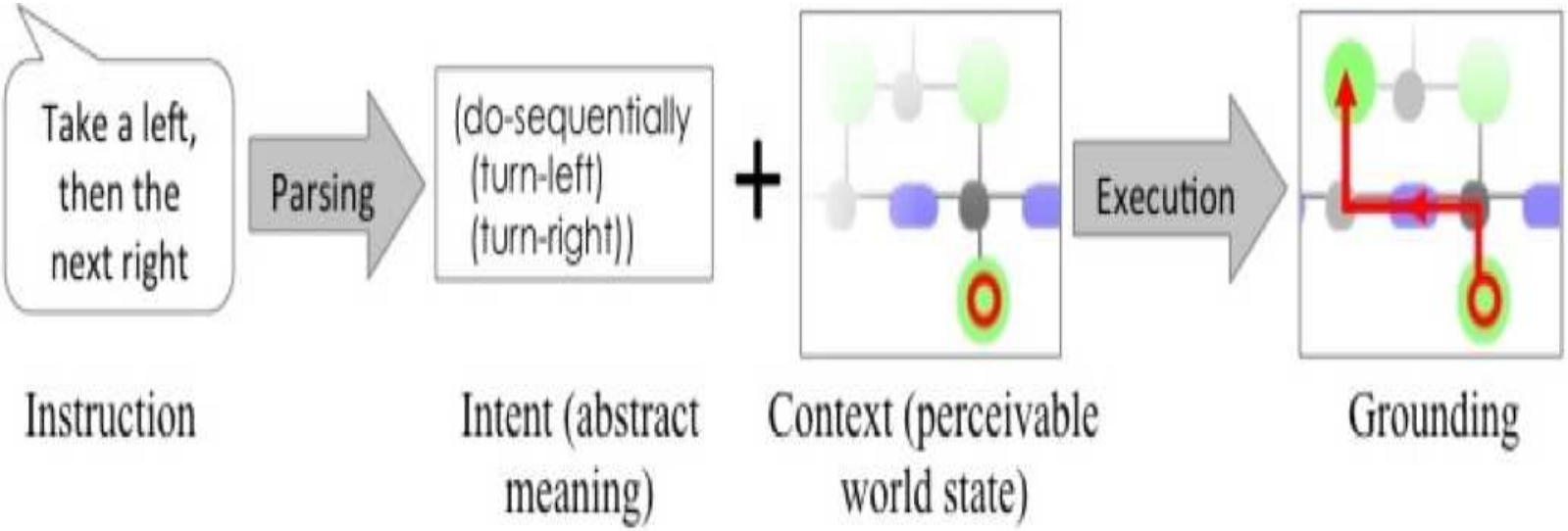
Praveen Dhinwa

Roll no 10526

Objective of the paper:

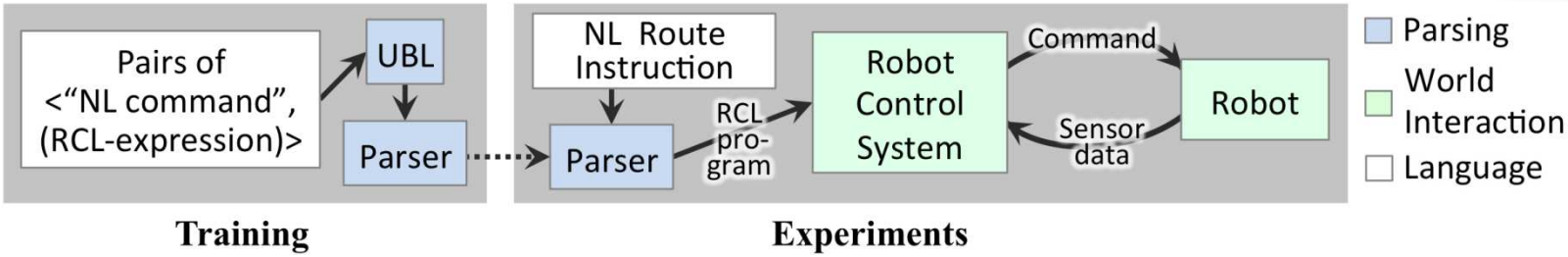
- Recently natural language interaction with robots has increased tremendously.
- This approach is parser based approach based on English commands and corresponding control language expressions.
- This experiment will deal with processing of route instructions (in natural language) in a indoor environment by a robot.
- language–interpreting human language into semantically informed structures in the context of robotic perception and actuation.
- Our goal is to investigate whether it is possible to learn a parser that produces correct, robot-executable commands for such instructions commands.

The task


Take a left, then the next right
Parsing
(do-sequentially (turn-left) (turn-right))
+
Execution
Instruction
Intent (abstract meaning)
Context (perceivable world state)
Grounding

- The task: Going from NL to robot control. First, the natural language command is parsed into a formal, procedural description representing the intent of the person. The robot control commands are then used by the executor, along with the local state of the world, to control the robot, thereby grounding the NL commands into actions while exploring the environment.
- Parsing natural language to expressive formal representations such as lambda-calculus
- most approaches rely on a manually constructed parser to map from NL commands to lambda-calculus, rather than learning grounding relations from data.

Technical Details:



The high-level architecture of the end-to-end system. Training is performed by learning a

parser from English instructions to RCL. In the experimental phase, the learned parser maps NL

instructions to an RCL program that is executed by the robot

RCL (Robot Control Language)

- **Locations**
  - **Current-loc:loc => robot's current location**
  - **forward-loc:loc => location ahead of robot**
  - **left-loc:loc => to robot's left**
  - **right-loc:loc => to robot's right**
  - **exists:t [loc] => does [loc] exist?**
- **Logic:**
  - **and:t [t] [t] boolean 'and'**
  - **or:t [t] [t] boolean 'or'**
  - **not:t [t] boolean 'not'**
- **Movement**
  - **move-to:t [loc] move to [loc]**
  - **turn-left:t take next available left**
  - **turn-right:t take next available right**

Similarly for querying the type of node, eg is node a room , hall , 3-way junction, hallway etc. and also for turning left and right

Sample RCL code

- **“Go left to the end of the hall.”**

Corresponding RCL Code:

```
(do-sequentially
   (turn-left
   (do-until
   (or
         (not
                  (exists forward-loc))
         (room forward-loc))
   (move-to forward-loc)))
```

Parsing

- For this work, parsing is performed using an extended version of the Unification- Based Learner, UBL
- The grammatical formalism used by UBL is a probabilistic

version of combinatory categorial grammars

- UBL can learn a parser solely from training data of the form

f(x_i , z_i) , where x_i is a natural-language sentence and z_i is a corresponding semantic language sentence.

Eg.

go to the second junction

S/NP NP/NP NP/N N

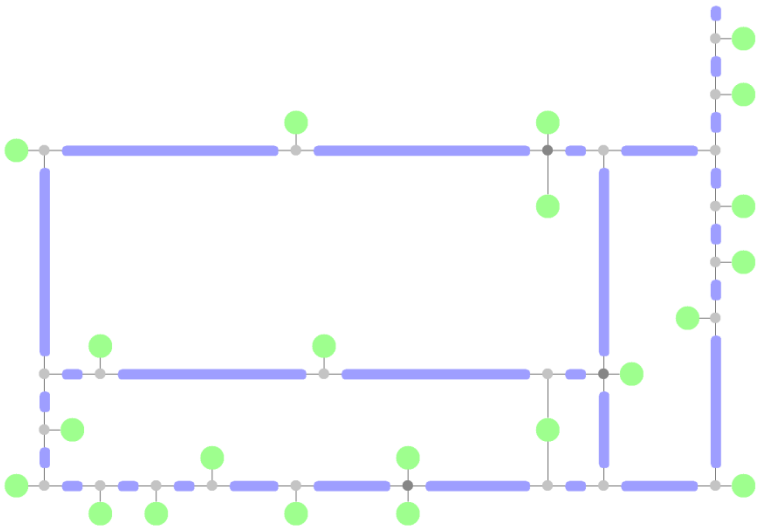
(move-to forward) [null] (do-n-time 2 x) (until(junction current-loc) y)

So overall:

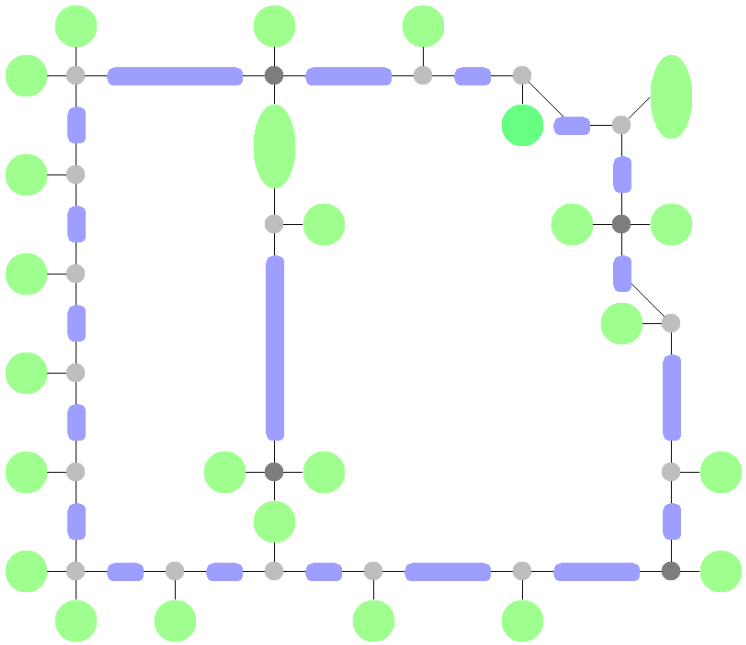
do-seq(do-n-times 2(until junction current_loc) (move-to-forward)

Dataset & Maps

Map A

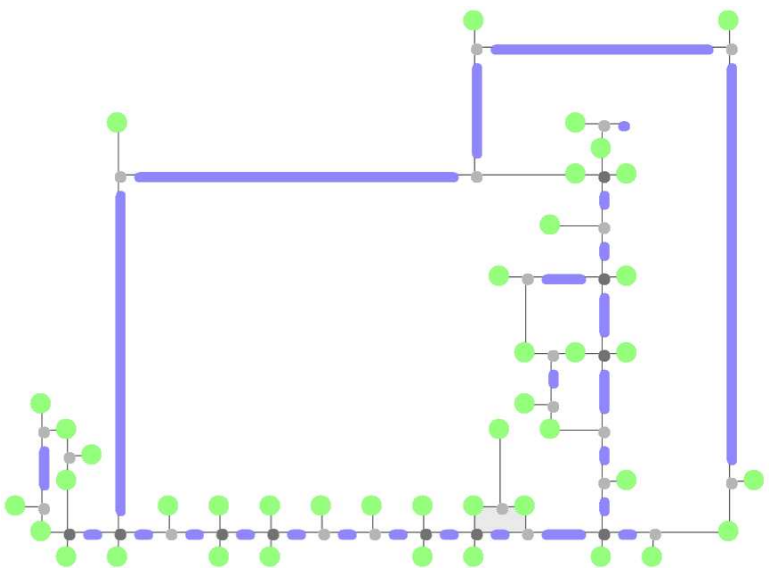


Map B

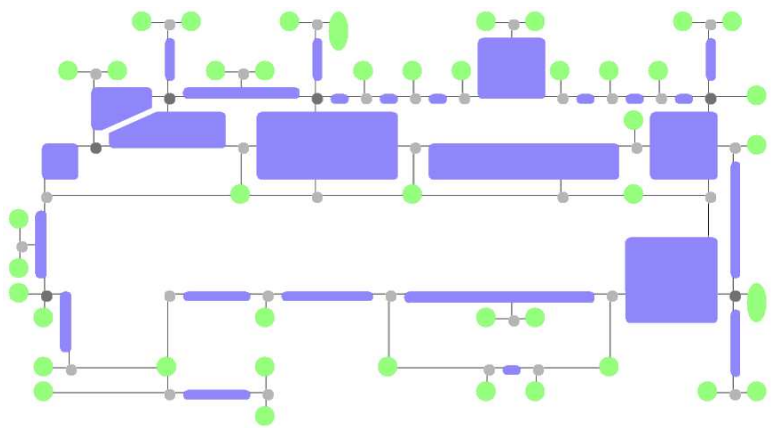


Dataset and maps

Map C



Map D



Datasets

Training :

First two maps shown in the previous slides for testing and the second two for testing. For training and test purposes, each English route description is also paired with an associated RCL annotation.

Map A and B were automatically generated using Voronoi Random.

Map C and D were manually constructed maps.

As A and B were relatively simpler ,So these were

constructed.

For English instruction 189 unique sentences generated by non-experts.

Further added route instructions for more rigorous testing. Finally enriched data set contained 418 NL route instructions with corresponding RCL commands

Testing

- using 189 simple instructions collected from training, data was tested against map B. 71.5 % success.
- Using various combinations of previous enriched data set 1200 unique paths through Map D were constructed. 1000 paths described by single NL sentences and 200 paths having on average 5 NL sentences.
- For 1000 simple paths , success was 66.3 +- 10.7 %
- For 200 complicated paths , success was 48.9 %

Conclusion

- It demonstrates that it is possible to have a parser able to handle complex , procedural natural language commands for robot instruction.
- It is also demonstrated that it is possible to combine advanced natural language processing with robotic perception and control systems. Eg. Go to and go left .
- Local error recovery will increase the results notably.
- Since the parser is probabilistic in nature, it can also generate a ranked list of possible RCL programs, each of which could generate a joint model for grounding

References

- 1. Y. Artzi and L.S. Zettlemoyer. Bootstrapping semantic parsers from conversations. In Proc.
- of the Conf. on Empirical Methods in Natural Language Processing, 2011.
- 2. M. Beetz, T. Arbuckle, T. Belker, M. Bennewitz, W. Burgard, A.B. Cremers, D. Fox,
- H. Grosskreutz, D. H¨ahnel, and D. Schulz. Integrated plan-based control of autonomous service
- robots in human environments. IEEE Intelligent Systems, 16(5), 2001.
- C. Boutilier, R. Reiter, M. Soutchanski, and S. Thrun. Decision-theoretic, high-level agent
- programming in the situation calculus. In Proc. of the National Conference on Artificial Intelligence
- (AAAI), 2000.
- 4. S.R.K. Branavan, L. Zettlemoyer, and R. Barzilay. Reading between the lines: Learning to
- map high-level instructions to commands. In ACL, pages 1268–1277, 2010.
- 5. W. Burgard, A. B. Cremers, D. Fox, D. H¨ahnel, G. Lakemeyer, D. Schulz, W. Steiner, and
- S. Thrun. Experiences with an interactive museum tour-guide robot. Artificial Intelligence,
- 114(1-2):3–55, 1999.
- 6. D.L. Chen. Fast online lexicon learning for grounded language acquisition. In Proc. of the
- Annual Meetings of the Association for Computational Linguistics (ACL), 2012.
- 7. D.L. Chen and R. Mooney. Learning to interpret natural language navigation instructions
- from observations. In Proceedings of the 25th AAAI Conference on Artificial Intelligence
- (AAAI-2011), pages 859–865, 2011.
- 8. J. Dzifcak, M. Scheutz, C. Baral, and P. Schermerhorn. What to do and how to do it: Translating
- natural language directives into temporal and dynamic logic representation for goal
- management and action execution. In Proc. of the 2009 IEEE Int'l Conf. on Robotics and
- Automation (ICRA '09), 2009.
- 9. A. Ferrein and G. Lakemeyer. Logic-based robot control in highly dynamic domains. Robotics
- and Autonomous Systems, 56(11), 2008.
- 10. S. Friedman, H. Pasula, and D. Fox. Voronoi random fields: Extracting topological structure
- of indoor environments via place labeling. In IJCAI, pages 2109–2114, 2007.
- 11. D. H¨ahnel, W. Burgard, and G. Lakemeyer. GOLEX - bridging the gap between logic
- (GOLOG) and a real robot. In Proc. of the German Conference on Artificial Intelligence
- (KI), Germany, 1998.
- 12. F. Ingrand, R. Chatila, R. Alami, and F. Robert. PRS: A high level supervision and control
- language for autonomous mobile robots. In Proc. of the IEEE Internation association

- **Thank You**
- Questions ?