Towards Generating Route Instructions Under Uncertainty: A Corpus Study

Verena Rieser and Amanda Cercas Curry
Interaction Lab
Heriot-Watt University
Edinburgh, UK
v.t.rieser@hw.ac.uk

Abstract

The overall aim of this work is to develop a principled data-driven approach for generating route instructions in spatial domains when faced with various types of uncertainty, e.g. user location, view shed, distance to target etc. As a first step, we conduct a corpus study investigating how humans give instructions in different scenarios. We find that human instruction givers produce different route instructions based on the information available to them. This motivates a context-adaptive approach for generating route instructions.

1 Introduction

Systems that generate route instructions have recently attracted a lot of attention from the dialogue and Natural Language Generation (NLG) communities, e.g. (Koller et al., 2007; Dethlefs and Cuayáhuitl, 2011; Dethlefs et al., 2011; Janarthanam et al., 2012; Dräger and Koller, 2012) etc. In this research we investigate how to generate route instructions when faced with uncertainty, e.g. about the user's location, view shed, distance to target etc. As a first step, we conduct a corpus study to empirically investigate how humans give instructions in different scenarios. In particular, we compare object references and quantitative descriptions. Previous research seem to suggest that landmark-based route instructions ("*Walk towards the Castle*") are easier to understand than distance-based ones ("*Walk 300 meters*") (Lovelace et al., 1999; Dräger and Koller, 2012). Here, we investigate the choices human Instruction Givers make when confronted with different types of uncertainty. We draw conclusions based on the different distributions of observed surface forms across three different corpora.

2 Corpus Annotation

We manually annotate navigation instructions in two Wizard-of-Oz corpora collected as part of the SpaceBook project (Janarthanam et al., 2014). We follow an annotation scheme by (Levit and Roy, 2007) developed for the HCRC MapTask corpus (Thompson et al., 1993), which we modify to account for situated dialogues. We also utilise the original annotations from MapTask. These three corpora are collected in different setups and thus introduce different types of uncertainty between Instruction Giver (IG) and Instruction Follower (IF):

MapTask (MT): IF and IG, share the same spatial representation in form of a paper map, i.e. distances and landmarks are known to both. The location of the IF is hidden to the IG.

SpaceBook1 (SB1): The IG follows the IF through the city of Edinburgh while communicating on the phone. That is, the IG knows location and view shed of the IF.

SpaceBook2 (SB2): The IG tracks the IF on Google Maps and also has access to Google StreetView. The exact location of the IF is unknown due to a noisy GPS signal.

The annotation scheme decomposes an utterance into navigational information units (NIUs). These NIUs are then further specified according to various aspects of instruction giving, e.g. actions, path descriptions etc. Here we only report on aspects relevant to generation under uncertainty:

Verification Actions aim to clarify uncertainty about position or orientation of the IF.

Reference Objects serve as anchors for identifying directions or positions.

Quantitative Aspect encode how far the traveler should move.

Label	SpaceBook Example	SpaceBook1	SpaceBook2	MapTask
Total NUIs		316	414	2132
Verif.:Position	<POSITION verifier="WIZARD" reference="LANDMARK">Are you standing outside Informatics in Edinburgh?</POSITION>	17.7%	8.2%	11.02%
Verif.:Orientation	<ORIENTATION verifier="WIZARD" reference="LANDMARK">can you see the National Museum of Scotland in front of you?</ORIENTATION>	4.1%	3.1%	0.8%
Object:Landmark	see above.	43.7%	21.7%	33.1%
Object:Streetname	<POSITION verifier="WIZARD" reference="STREETNAME"> This is West Nicolson Street. </POSITION>	7.6%	22.9%	N/A
Object:Proximity	<TURN descriptor="NIL" reference="NIL" quantitative="PROXIMITY">Turn at the next crossing</TURN>	8.2%	7.0%	N/A
Object:UserCentric	<MOVE descriptor="STRAIGHT" reference="USER" quantitative="FALSE"> Just keep walking in the direction you are going.</MOVE>	17.4%	16.4%	5.6%
Object:Cardinal	<MOVE descriptor="STRAIGHT" reference="CARDINAL" quantitative="FALSE">Please continue walking South. </MOVE>	0.3%	0.2%	31.8%
Object:NIL	<TURN descriptor="LEFT" reference="FALSE" quantitative="FALSE">you wish to turn left</TURN>	17.4%	19.3%	13.3%
Object:Other		5.4	12.5	5.1%
Quantitative:Time	<MOVE descriptor="STRAIGHT" reference="OTHER" quantitative="TIME"> About one minute down the road.</MOVE>	0%	1.2%	N/A
Quantitative:Distance	<MOVE descriptor="STRAIGHT" reference="NIL" quantitative="DISTANCE">follow for three hundred meters. </MOVE>	0%	0.2%	64.8%

Table 1: NUIs label distributions: Frequencies within corpora.

3 Results

We highlight and discuss the main differences observed between the three corpora, based on their frequency of occurrence as summarised in Table 1.

- **Verification actions** make up between 11-22% of all possible actions. They are almost twice as frequent in SpaceBook2 than in SpaceBook1 and MapTask. We hypothesise that this occurred when the IG lost sight of the IF. In general, the IG tends to verify the IF's position, but less so the orientation/ view shed.
- **Landmarks** are the most common reference object in SpaceBook1 and MapTask, but SpaceBook2 uses **Steetnames** more often. This can be attributed to the fact that in this scenario the IG tracks the IF on a digital map where street names are indicated. Whereas in SpaceBook1 Landmarks are more prominent due to the shared view shed.
- **Proximity** is only used in SpaceBook1 and SpaceBook2 since the IG had an estimate of the IF's position, whereas in MapTask the IF's location is hidden. Similarly, **UserCentric** instructions are generated relative to the IF's position and thus only occur in SpaceBook1 and SpaceBook2.
- **Cardinal** directions hardly occur in the SpaceBook scenarios, but in MapTask this information is relative to the paper map, and thus a shared point of reference which is used in over one third of the cases.
- The main difference between SpaceBook and MapTask is the occurrence of **quantitative** descriptions. In the SpaceBook scenarios quantitative descriptions hardly ever occur, whereas in MapTask about 65% of instructions are quantified. We attribute this difference to the fact that distances can be easily estimated from a paper map, whereas distances from a digital map or while walking down the street are harder to judge.

4 Discussion and Future Work

In summary, we find that human instruction givers produce different route instructions based on the information available to them: Landmarks are preferred if the view shed of the instruction follower is known; User centric and instructions based on proximity only occur if the location is known; Cardinal directions occur if the orientation is known; And quantitative descriptions are limited to cases where the scale is known. We therefore conclude that in contrast to claims by previous work, it is not always preferred to generate instructions based on landmarks, but good route instructions depend on the contextual information available. In future work, we will generate context-dependent instructions based on a framework for generation under uncertainty (Rieser and Lemon, 2009; Rieser et al., 2014) and test their effectiveness with real users (Janarthanam et al., 2012). We will also measure inter-annotator agreement and run significant tests for the above annotations.

Acknowledgments

The research leading to this work has received funding from the Engineering and Physical Science Research council, UK (EPSRC) under project no. EP/L026775/1. We would also like to thank Srinivasan Janarthanam and Robin Hill for help and discussion.

References

Nina Dethlefs and Heriberto Cuayáhuitl. 2011. Combining hierarchical reinforcement learning and bayesian networks for natural language generation in situated dialogue. In *Proceedings of the 13th European Workshop on Natural Language Generation (ENLG)*.

Nina Dethlefs, Heriberto Cuayáhuitl, and Jette Viethen. 2011. Optimising natural language generation decision making for situated dialogue. In *Proc. of the Annual SIGDIAL Conference on Discourse and Dialogue*.

Markus Dräger and Alexander Koller. 2012. Generation of landmark-based navigation instructions from open-source data. In *Proceedings of the Thirteenth Conference of the European Chapter of the ACL (EACL)*, Avignon.

Srinivasan Janarthanam, Xingkun Liu, and Oliver Lemon. 2012. A web-based evaluation framework for spatial instruction-giving systems. In *Proc. of Annual Meeting of the Association for Computational Linguistics (ACL)*.

Srinivasan Janarthanam, Robin Hill, Anna Dickinson, and Morgan Fredriksson. 2014. Proceedings of the eacl 2014 workshop on dialogue in motion. pages 19–27. Association for Computational Linguistics.

Alexander Koller, Johanna Moore, Barbara di Eugenio, James Lester, Laura Stoia, Donna Byron, Jon Oberlander, and Kristina Striegnitz. 2007. Shared task proposal: Instruction giving in virtual worlds. In Michael White and Robert Dale, editors, *Working group reports of the Workshop on Shared Tasks and Comparative Evaluation in Natural Language Generation*.

Michael Levit and Deb Roy. 2007. Interpretation of spatial language in a map navigation task. *IEEE Trans. Syst., Man, Cybern. B, Cybern*, 37(3).

Kristin L. Lovelace, Mary Hegarty, and Daniel R. Montello. 1999. Elements of good route directions in familiar and unfamiliar environments. In *Proceedings of the International Conference on Spatial Information Theory: Cognitive and Computational Foundations of Geographic Information Science*, COSIT '99, pages 65–82, London, UK, UK. Springer-Verlag.

Verena Rieser and Oliver Lemon. 2009. Natural Language Generation as Planning Under Uncertainty for Spoken Dialogue Systems. In *Proc. of the Conference of European Chapter of the Association for Computational Linguistics (EACL)*.

V. Rieser, O. Lemon, and S. Keizer. 2014. Natural language generation as incremental planning under uncertainty: Adaptive information presentation for statistical dialogue systems. *Audio, Speech, and Language Processing, IEEE/ACM Transactions on*, 22(5):979–994, May.

Henry S. Thompson, Anne Anderson, Ellen Gurman Bard, Gwyneth Doherty-Sneddon, Alison Newlands, and Cathy Sotillo. 1993. The hcrc map task corpus: Natural dialogue for speech recognition. In *Proceedings of the Workshop on Human Language Technology*, HLT '93, pages 25–30, Stroudsburg, PA, USA. Association for Computational Linguistics.