The GRUVE Challenge: Generating Routes under Uncertainty in Virtual Environments

Srini Janarthanam and Oliver Lemon

Interaction Lab, MACS

Heriot-Watt University

Edinburgh, United Kingdom

`sc445,o.lemon@hw.ac.uk`

www.macs.hw.ac.uk/InteractionLab

Abstract

We propose a shared task based around generation of instructions for pedestrian users navigating in open-world virtual environments. An important variant of this task involves handling uncertainty about the user's location (as would happen in the real world with a standard GPS system). We motivate and explain the task, propose metrics for evaluation of systems, describe the planned software setup, and propose a timeline for the challenge.

1 Introduction

Providing route instructions and descriptions for users is an interesting and a challenging task. Route-giving tasks have recently attracted active research in the NLG and dialogue systems communities (Dale et al., 2002; Cheng et al., 2004; Richter et al., 2008; Cuayhuitl et al., 2010; Dethlefs and Cuayáhuitl, 2011; Dethlefs et al., 2011). Route-giving (whether in virtual or real environments) involves many decisions, including when to instruct, what instructions and/or descriptions to give, and how to verbalize them. Research has shown that inclusion of landmarks in route instructions is highly effective (May et al., 2003; Schroder et al., 2011). In order to include landmarks in instructions, decisions such as which landmarks to include, how best to refer to them, and so on, must be taken. Another interesting issue for real-world route-giving is that it is not always possible to know where the user is, or where they are looking. Even when using tools like GPS trackers, there is an element of uncertainty in the pedestrian user's location, so generation under uncertainty becomes important (Lemon et al., 2010). Finally, instructions and referring expressions should also take into account the pedestrian's field of view or "viewshed", which is not directly

observable but may be inferred from uncertain information about location and orientation.

Virtual environments provide an important development and test infrastructure for real-world systems. They avoid the need for costly and time-consuming real-world experiments and data-collections, while allowing manipulation of the spatial environment to investigate specific issues and contexts for NLG systems.

There is therefore an interesting and practical shared task in which research teams can collaborate using a shared infrastructure to investigate NLG issues in route giving tasks to pedestrians in outdoor virtual environments, where different types and degrees of uncertainty can be manipulated experimentally. The GRUVE challenge targets these tasks, with the expectation that its results will be informative for real-world pedestrian navigation systems.

2 Related work

The "Generating Instructions in Virtual Environments" (GIVE) challenge has been running successful shared tasks since 2009 (Koller et al., 2007; Byron et al., 2007). In this task, human users log into a virtual world over the Internet in which they are free to walk around inside building-like environments with several rooms and corridors. The objective (for users) of these tasks is to follow the instructions given to them (in text), navigate around, push buttons to disable or enable alarms, open or close doors, and finally recover a hidden trophy. Several teams participated in this shared task to build systems that will generate instructions online to the users. The generation systems were provided with the user's location and viewshed (i.e. what objects in the world are in the user's view). In the first version of the challenge, the users moved "block by block" in a grid-based virtual environment. Therefore it was possible to give instructions such as "move 3 steps forward".

Figure 1: An outdoor virtual environment from SecondLife

However in the latest version of this challenge, the users move continuously and not discretely (Gargett et al., 2010). This challenge examined the issues concerning generating instructions and referring expressions in situated contexts. Our proposed challenge is similar to the GIVE challenge in the sense that it involves systems generating instructions for navigation, and generating referring expressions to refer to entities in the world. But in contrast, in this challenge, we propose to use an outdoor virtual environment where route instruction giving and referring to outdoor entities would be for pedestrian navigation in city-like environments, involving issues such as uncertainty in user's location and viewshed.

3 GRUVE Shared tasks

We propose a collection of shared tasks or challenges which will allow exploration of a number of issues in NLG:

- NLG under uncertainty

- the generation of instructions and route descriptions

- generation of referring expressions

- situated NLG

- optimisation of NLG

- adaptive NLG for different users

- NLG in interactive systems.

The proposed tasks will take place in an outdoor virtual environment and will be variants of route giving tasks. The basic task will be to get the user (who sees a first-person perspective, pedestrian view of the environment, see e.g. Figure 1) from location A to location B. The task can vary along the following dimensions of system knowledge:

- precision of user location information

- precision of user gaze direction / contents or user viewshed

- previous knowledge of user behaviour

- amount of feedback from user and its reliability/ noise.

We propose to evaluate NLG systems developed by the participating teams in route instruction giving tasks under various conditions discussed above. In the simplest case, the system has total information about the user location, heading/gaze direction, history of interaction/behaviour, and a clear and detailed set of feedback signals (e.g. "I am lost", "I am confused", "repeat", "rephrase", etc) as on-screen buttons. The challenge will be a case of constructing optimal messages for the user based on complete knowledge of their situation, which is akin to generating instructions for players of video games, as in an interactive version of the original GIVE task. (We discuss notions of optimality shortly). At the other end of the spectrum we may have to generate instructions for un-

known users whose location we are very insure of, where feedback signals are very noisy, and where we don't have much idea what direction they are facing. This latter set of conditions is similar to real-world city navigation problems. There is a wide range of possibilities across this spectrum. For example, one task would be to generate instructions to users whose location is uncertain. In such situations, it becomes necessary to not only instruct the users but also question them in order to reduce uncertainty arising due to their location. Therefore, the NLG systems should be able to generate both instructions and questions in order to successfully complete the task. The NLG system should also be able to decide when to question the user and when to instruct him.

Instructions can also be generated in two formats: *a priori* or *in situ*. In the *a priori* format, the entire set of instructions to follow from source to destination are given to the user at the starting point. On the other hand, in the *in situ* format, a sequence of instructions are given to the user on the fly one by one as he walks along the route at appropriate times. We believe that all these generation tasks involve subtasks like content planning, referring expression generation, aggregation, realization, timing, and so on, and therefore this challenge would be of interest to many. We invite the community to discuss the range of tasks in detail.

4   Software infrastructure

As in the GIVE challenge, we will ask users to log in to a virtual environment, running on a server, and they will then encounter different NLG systems in a variety of tasks. We propose to reuse and modify the existing GIVE infrastructure for building a 3D interactive outdoor pedestrian environment. However, if it is not suitable, we propose to build the infrastructure using one of the following tools:

- OpenWonderLand[1]

- OpenSimulator[2]

- OpenSceneGraph[3]

- Unreal engine[4]

- Google Sketch Up[5]

- jMonkeyEngine[6]

- X3D[7]

One of these tools will be chosen and additional "feedback" buttons will be added to the user's GUI. These buttons will allow the user to "say" things like: 'Yes', 'No','OK', 'I'm lost', 'repeat', 'I'm confused', 'quit' and so on. Speech will be delivered to the user's browser via a TTS engine, or wizard voice, or prerecorded prompts can be used. A route planner will be a part of this infrastructure that will generate plans for navigation. This route plan will contain route directions from source to destination along with information on landmarks on the way. This route plan along with information specific to the user (i.e. location, viewshed, confidence scores, and button requests) will form the inputs to the NLG system. This infrastructure will then be made available to the teams for developing their own NLG system. Since this is the first time the challenge is organised, there will be no data available. All collected data will be released for future versions of the challenge.

5   Evaluation metrics

We propose to evaluate the participating systems based on objective metrics such as task completion, time taken, and so on. We also propose to obtain ratings from the users based on the quality of the interaction they had with the system. They will be asked to rate the features of the system based on how confusing or easy it was to follow the instructions, and so on.

6   Proposed Schedule

1. Software infrastructure in place: Oct 2011

2. Tasks and metrics defined: Nov 2011

3. Entrants collected: Dec 2011-Jan 2012

4. System development: Jan 2012- April 2012

5. Collect users/subjects via MTurk or other crowdsourcing method: April 2012

6. Run the challenge: April-May 2012

---

[1]http://openwonderland.org
[2]http://opensimulator.org/
[3]http://www.openscenegraph.org/
[4]http://www.unrealengine.com/

[5]http://sketchup.google.com/
[6]http://jmonkeyengine.org/
[7]http://www.web3d.org/

7. Report results: INLG 2012

8. Release data via web: post INLG 2012

7 Future work

In the future, we hope to extend this infrastructure so that users can interact with the system using text or speech input instead of propositional inputs using buttons. This will introduce an element of uncertainty in speech/text input as well in terms of ambient noise, underspecified referring expressions and so on.

8 Conclusion

In this paper, we presented a shared task for research teams to collaborate and investigate the issues and challenges in giving instructions for outdoor pedestrian navigation. We briefly presented a set of interesting new problems in this task. The GRUVE challenge targets route giving tasks to pedestrians in outdoor virtual environments, where different types and degrees of uncertainty can be manipulated experimentally, with the expectation that its results will be informative for real-world pedestrian navigation systems.

We hope to discuss with members of the NLG community how to modify the existing GIVE infrastructure for this task and how we can best collaborate with other researchers in developing and refining the challenge.

Acknowledgments

References

D. Byron, A. Koller, J. Oberlander, L. Stoia, and K. Striegnitz. 2007. Generating Instructions in Virtual Environments (GIVE): A challenge and evaluation testbed for NLG. In *Proceedings of the Workshop on Shared Tasks and Comparative Evaluation in Natural Language Generation*.

H. Cheng, L. Cavedon, and R. Dale. 2004. Generating Navigation Information Based on the Driver's Route Knowledge. In *Proceedings of the Coling 2004 Workshop on Robust and Adaptive Information Processing for Mobile Speech Interfaces*.

H. Cuayhuitl, N. Dethlefs, L. Frommberger, K.-F. Richter, and J Bateman. 2010. Generating Adaptive Route Instructions Using Hierarchical Reinforcement Learning. In *Proceedings of the International Conference on Spatial Cognition (Spatial Cognition VII), Portland, OR, USA*.

R. Dale, S. Geldof, and J. P. Prost. 2002. Generating more natural route descriptions. In *Proceedings of the 2002 Australasian Natural Language Processing Workshop*.

Nina Dethlefs and Heriberto Cuayáhuitl. 2011. Hierarchical reinforcement learning and hidden markov models for task-oriented natural language generation. In *Proc. of ACL*.

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

A. Gargett, K. Garoufi, A. Koller, and K. Striegnitz. 2010. The GIVE-2 Corpus of Giving Instructions in Virtual Environments. In *Proceedings of the 7th Conference on International Language Resources and Evaluation (LREC), Valletta, Malta*.

A. Koller, J. Moore, B. Eugenio, J. Lester, L. Stoia, D. Byron, J. Oberlander, and K. Striegnitz. 2007. Shared Task Proposal: Instruction Giving in Virtual Worlds. In *Workshop on Shared Tasks and Comparative Evaluation in Natural Language Generation*.

Oliver Lemon, Srini Janarthanam, and Verena Rieser. 2010. Generation under uncertainty. In *Proceedings of INLG / Generation Challenges*.

A. J. May, T. Ross, S. H. Bayer, and M. J. Tarkiainen. 2003. Pedestrian navigation aids: information requirements and design implications. *Personal and Ubiquitous Computing*, 7(6):331–338.

K.-F. Richter, M. Tomko, and S. Winter. 2008. A dialog-driven process of generating route directions. *Computers, Environment and Urban Systems*, 32(3):233245.

C. J. Schroder, W. A. Mackaness, and B. M. Gittings. 2011. Giving the 'Right' Route Directions: The Requirements for Pedestrian Navigation Systems. *Transactions in GIS*, 15(3):419–438.