Navigational Instruction Generation as Inverse Reinforcement Learning with Neural Machine Translation

Andrea F. Daniele
TTI-Chicago

Mohit Bansal
UNC - Chapel Hill

Matthew R. Walter
TTI-Chicago

Modern robotics applications require robots to be able to communicate with humans seamlessly and effectively. Natural language provides a flexible and efficient means of communication. Significant advancements have been made in developing robots capable of interpreting free-form instructions, but less attention has been devoted to endowing robots with the ability to generate natural language.

We propose a method that enables robots to generate natural language instructions that allow their human partners to navigate a priori unknown environments. This problem is a surrogate for the general domain of instruction generation, which is important for human-robot collaborative tasks, such as search-and-rescue, exploration, and teaching. This work presents what is, to the best of our knowledge, the first framework that generates natural language route instructions that reflect the preferences and rhetorical style of humans.

There are two primary challenges to generating effective, natural language route instructions, which are characteristic of the more general problem of free-form generation.

The first challenge is *content selection*, the problem of deciding what and how much information to convey to the user. Existing navigational solutions typically rely on a set of hand-crafted rules that consider only street names and metric distances as valid candidates. In contrast, studies show that people prefer instructions that reference physical, salient landmarks. We propose a method (Fig. 1) that models this content selection problem as a Markov decision process (MDP) with a learned policy that decides what and how much to include in a formal language specification $\mathcal{C}$ of the task (path) p. We learn this policy via inverse reinforcement learning (IRL) from demonstrations of route instructions provided by humans. This avoids the need for hand-crafted selection rules, and allows our method to automatically adapt to the preferences and communication style of the target populations, and to simultaneously choose information that minimizes ambiguity while avoiding verbosity.

The second challenge is *surface realization*, which is the task of synthesizing a natural language sentence that refers to the selected content. Existing solutions rely on sentence templates, generating sentences by populating manually defined fields (e.g., "turn ⟨direction⟩"). As expected, the use of such templates reduces coherence across sentences and limits domain ability. We address the surface realization problem through a neural sequence-to-sequence model that "translates" a formal language specification of the selected command $\mathcal{C}$ into a natural language sentence Λ. Our model (Fig. 1) takes the form of an encoder-aligner-decoder architecture that first encodes the formal task specification with a recurrent neural network using long short-term memory (LSTM-RNN). The model then decodes (translates) the resulting abstraction of the input into a natural language sentence (word sequence), using an alignment mechanism to further refine the selected information and associate output words with the corresponding elements in the input formal specification. The use of LSTMs as the hidden units enables our model to capture the long-term dependencies that exist among the selected information and among the words in the resulting instruction.

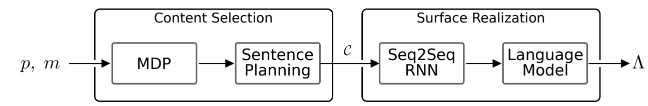

Figure 1: Our method generates free-form instructions through inverse RL and sequence-to-sequence learning.

We train and test our content selection and surface realization models on the benchmark SAIL dataset consisting of human-generated route instructions and the corresponding paths that human participants followed. Instructions generated with our method achieve a sentence-level BLEU score of 72.18%, indicating their similarity with the reference set of human-provided instructions. We additionally conduct human evaluation experiments in which 42 participants were asked to follow instructions generated by our method as well as by humans (from the SAIL dataset) in a 3D virtual world. The experiments reveal that our method produces instructions that are as accurate, effective, and usable as those generated by humans. In particular, the participants reached the intended destination 4% more often and did so in less time while following our instructions. They also rated our instructions as being more informative than those generated by humans (by 19%), and easier to follow (by 20%).

In summary, we briefly described a method for natural language generation in the context of navigation that produces instructions that are unambiguous, easy to follow, and that adhere to the preferences and rhetorical style of humans.