While much research exists on resolving spatial natural language references to known locations, little work deals with handling references to unknown locations. In this paper we introduce and evaluate algorithms integrated into a cognitive architecture which allow an agent to learn about its environment while resolving references to both known and unknown locations. We also describe how multiple components in the architecture jointly facilitate these capabilities.

All previous approaches to spatial reference resolution have used a static environmental map which cannot be changed once reference resolution begins, and have only dealt with natural language commands, and not interrogative or declarative utterances.

We present algorithms for spatial reference resolution integrated into a cognitive robotic architecture that significantly improve previous proposals by: (1) systematically adding unknown places to the map, which allows robots to meaningfully communicate about unknown places without having to first discover their exact location, (2) updating the map as the agent discovers unknown environments, which allows robots to have natural language interactions about new environmental features discovered while navigating to an unknown place, and (3) generating action sequences only when they are actually needed to visit the referenced location (information is stored in a location-independent form, and not interrogative or declarative utterances).

We describe how two unknown locations are reconciled when the robot is meaningfully communicating about its environment while resolving references to both known and unknown locations. We also describe how multiple components in the architecture jointly facilitate these capabilities.

SPEX finds a suitable identifier for “the hall” but not for “the room” as it knows of no rooms at the end of the nearby hallway. SPEX thus creates a representation for a new room and notes that it is connected at a topological level to the nearby hallway.

At a metric level, SPEX can only plan as far ahead as he other side of the doorway, but can use the fact that the target location is at the end of the hall and is to the right of the current location to create an action script which should bring the robot to the target location.

Our approach has a number of advantages. Storing the information gleaned from natural language and through exploration in a location-independent format allows the robot to (1) travel to previously described locations, (2) describe how two unknown locations are positioned relative to each other, (3) pause an action sequence and then later resume it from another location, and (4) return to a known location after visiting an unknown one.

Finally, augmenting the robot’s world model based only on descriptions allows a robot to learn a map purely through dialogue if it is able to extract sufficiently accurate semantic representations; previous approaches would not be able to learn a map of their environment from dialogue alone.

This work was in part funded by NSF grant 111323 and in part funded by ONR grant N00014-11-1-0493.