A Cognitively Motivated Route-Interface for Mobile Robot Navigation

Mohammed Elmogy, Christopher Habel, and Jianwei Zhang

Department of Informatics, University of Hamburg,

Vogt-Kölln-straße 30, D-22527 Hamburg, Germany.

{ elmogy, habel, zhang}@informatik.uni-hamburg.de}

Abstract A more natural interaction between humans and mobile robots can be achieved by bridging the gap between the format of spatial knowledge used by robots and the format of languages used by humans. This enables both sides to communicate by using shared knowledge. Spatial knowledge can be (re)presented in various ways to increase the interaction between humans and mobile robots. One effective way is to describe the route verbally to the robot. This method can permit computer language-naive users to instruct mobile robots, which understand spatial descriptions, to naturally perform complex tasks using succinct and intuitive commands. We present a spatial language to describe route-based navigation tasks for a mobile robot. The instructions of this spatial language are implemented to provide an intuitive interface with which novice users can easily and naturally describe a navigation task to a mobile robot in a miniature city or in any other indoor environment. In our system, the instructions of the processed route are analyzed to generate a symbolic representation via the instruction interpreter. The resulting symbolic representation is supplied to the robot motion planning stage as an initial path estimation of route description and it is also used to generate a topological map of the route's environment.

1 Introduction

A more natural interaction between humans and mobile robots – with the least collective effort – can be achieved if there is a common ground of understanding [1, 2]. A natural language interface supports more natural styles of interaction between robots and their users. Route descriptions are considered as one of the more important natural language interfaces between humans and mobile robots for applying an effective human-robot interaction.

To describe a navigation task to a mobile robot, route instructions are used to specify the spatial information about the route environment and the temporal information about the move and turn actions which will be executed by the robot [3]. Good route instructions should contain adequate information on these two aspects by considering the spatial environment of the robot and the relevant navigation and perception actions. To express the route in an effective way, the rules and sequence of commands should be expressed vey concisely. Natural language uses symbols and syntactic rules to interact with the robots which dispose of represented knowledge at the symbolic level.

On the other hand, spatial reasoning on the natural language route is essential for both humans and mobile robots. Spatial reasoning gives robots the ability to use human-like spatial language and provides the human user with an intuitive interface that is consistent with his innate spatial cognition [4]. It can also accelerate learning by using symbolic communication, which has been shown in [5].

This paper is organized as follows. Section 2 discusses some current implementations of natural language interfaces for both mobile robots and simulated artificial agents. In section 3, the structure of our route instruction language (RIL), which is used to describe the route for the mobile robot, is presented. Section 4 discusses the creation of the symbolic representation of the route. The grounding of the symbolic representation with the perceptual data in the physical environment is illustrated in section 5. Finally, the conclusion is presented in section 6.

2 Related Work

In the last three decades, there has been considerable research on spatial language and spatial reasoning. This motivates the research interest of using spatial language for interacting with artificial navigational agents. Many researchers [4, 6, 7, 8] have proposed frameworks using natural language commands in simulated or real-world environments to guide their artificial agents during navigation. In this section, some implementations of natural language interfaces for mobile robots and simulated agents will be discussed.

In our group, Tschander et al. [6] proposed the idea of a cognitive-oriented Geometric Agent (GA) which simulates instructed navigation in a virtual planar environment. This geometric agent can navigate on routes in its virtual planer environment according to natural-language instructions presented in advance. In their approach, Conceptual Route Instruction Language (CRIL) is used to represent the meaning of natural language route instructions. It combines the latter with the spatial information gained from perception to execute the desired route. Tellex and Roy [7] implemented spatial routines to control the robot in a simulator. They defined a lexicon of words in terms of spatial routines and used that lexicon to build a speech-controlled robot in a simulator. Their system is unified by a high-level module that receives the output from the speech recognition system and simulated sensor data, creates a script using the lexicon and the parse structure of the command, and then sends appropriate commands to the simulated robot to execute that command. However, their current implementation acts only on the current snapshot of sensor readings which leads to errors in the robot's behavior.

On the other hand, there are considerable research efforts in developing various command sets for mobile robots and robotic wheelchairs [9-12]. The mobile robot community has created systems that can understand natural language commands. Many research efforts [6, 8, 13, 14] focus on using spatial language to control the robot's position and behavior, or to enable it to answer questions about what it senses. In general, previous work in this area has focused on developing various command sets for mobile robots and robotic wheelchairs, without directly addressing aspects of language that are context-sensitive. Torrance [14] implemented a system that is capable of mediating between an unmodified reactive mobile robot architecture and domain-restricted natural language. He introduced reactive-odometric plans (ROPs) and demonstrates their use in plan recognition. The communication component of this architecture supports a typewritten natural language discourse with people. This system was brittle due to place recognition from odometric data and the use of IR sensors for reactive motion control. The resulting ROPs do not contain error-reducing stopping conditions, and this has caused problems in some parts of the tested environment where hallways do not sufficiently constrain the reactive navigation system.

Skubic et al. [8] implemented robot spatial relationships combined with a multimodal robot interface that provides the context for the human-robot dialog. They showed how linguistic spatial descriptions and other spatial information can be extracted from an evidence grid map and how this information can be used in a natural human-robot dialog. With this spatial information and linguistic descriptions, they established a dialog of spatial language. To overcome the object recognition problem (the system does not support vision-based object recognition), they have defined a class of persistent objects that are recognized and named by the user.

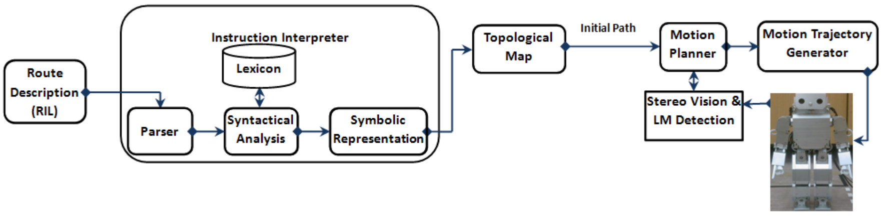

Fig. 1. System architecture.

3 Route Instruction Language (RIL)

In our system, we present a spatial language – called Route Instruction Language (RIL) [15] – to describe route-based navigation tasks for a mobile robot. This language is implemented to present an intuitive interface that will enable novice users to easily and naturally describe a route to a mobile robot in indoor and miniature city environments. We proposed this language to avoid ambiguity and misunderstanding during route description. Therefore, a non-expert user can describe the route for the mobile robot by using simple and easy to understand instructions. Fig. 1 shows an overview structure of our system.

The RIL is developed to describe the route between the start and end points to a mobile robot. It is intended as a semi-formal language for instructing robots, to be used via a structured graphical user interface. RIL provides elementary instruction statements which are processed to supply the robot with a sequence of motion actions. During navigation, this sequence of actions is processed by the motion planner to determine the footstep placements which will be effected by the humanoid robot to execute the route. Each statement in the RIL constitutes a spatial instruction which relates verbally coded motion concepts to one or more landmarks by use of a suitable spatial relationship.

Table 1. RIL command set and their syntax.

Command Type	Command Name	Syntax
Position	$START()	$START ([Pre1|Direction], Landmark1, [Pre2], [Landmark2])
	$STOP()	$STOP (Pre1|Direction, Landmark1, [Pre2], [Landmark2])
	$BE()	$BE (Pre1|Direction, Landmark1, [Pre2], [Landmark2])
Locomotion	$GO()	$GO([Count], [Direction]| [Pre1], [Landmark1], [Pre2], [Landmark2])
	$CROSS()	$CROSS ([Pre1], Landmark1, [Pre2], [Landmark2])
	$PASS()	$PASS ([Pre1], Landmark, direction, [Pre2], [Landmarket2])
	$FOLLOW()	$FOLLOW ([Landmark1], Pre, Landmark2)
Change of Orientation	$ROTATE()	$ROTATE (Direction, Pre, Landmark)
	$TURN()	$TURN ([Count], [Pre1], Direction, [Pre2], [Landmark])

The commands of the RIL and their syntax are shown in Table 1. Each instruction of the RIL specifies motion verbs, directions, destinations, and landmarks. The RIL commands are divided into three basic types: position, locomotion, and change of orientation commands. The position commands are used to indicate the current position of the robot during navigation. They are also used to determine the start and end points of the route. The Locomotion commands are used to instruct the robot to move in the spatial environment in a specific direction or to follow a certain path. The last category is the change of orientation commands, which are used to rotate around a landmark or turn in a certain direction.

The command syntax consists of a command word and an arbitrary number of arguments as shown in Table 1. The command word indicates the action which will be taken by the mobile robot and is represented in the imperative form of the verb, e.g., GO, TURN, BE, etc. Each argument is a place holder for a specific group of words such as prepositions, directions, the number of turns, and landmarks. To add more flexibility to the command syntax, multiple kinds of command syntax have been defined. Mandatory arguments are typed without any brackets, whereas optional arguments are placed between rectangular brackets '[]'. The pipe symbol '|' indicates an OR operator. Fig. 2 shows an example of a route description from the railway station to the McDonald's restaurant in our miniature city using RIL.

(a) Miniature City Map

```
$START( RailwayStation, right)
$GO( forward, into, Street)
$PASS( Saturn, left)
$PASS( BurgerKing, right)
$PASS( C&A, left)
$BE( at, CrossRoads)
$GO( forward, into, Street)
$PASS( Church, left)
$PASS( KarStadt, right)
$BE( at, CrossRoads)
$TURN( right)
$GO( forward, into, Street)
$PASS( TownHall, left)
$STOP( McDonalds, left)
```

(b) Route Description

Fig. 2. A route description from the railway station to the McDonald's restaurant in our miniature city using RIL.

We carried out an experiment to test the suitability of RIL for communicating a route description to a robot. 18 participants took part in the experiment (age 22 to 35 years). None of the participants had any background knowledge on route instructions and robotics. First, we gave them a description of the RIL syntax, a map of the miniature city, and an example of a suitable route description. We asked them to describe a route between the railway station and the McDonald's restaurant in the miniature city as depicted in Fig. 2(a). 89% of the participants described the route correctly, but the rest are confused about how to use some commands and parameters. 83% of the participants stated that the RIL is simple and easy to learn, but the rest of them preferred to use a controlled natural language without any specific syntax for the instructions. 78% of the participants agreed that it is better to provide the commands of RIL with many optional parameters than to restrict them to a single syntax.

4 Instruction Interpreter

The instruction interpreter is used to discriminate, identify, and categorize the motion actions of the processed route description. It combines definitions from the lexicon according to the parse structure of the instruction, creating a symbolic script that describes the navigation process. The generated symbolic representation is used to create a topological map for the route environment. It is also supplied to the motion planner as an initial path estimation of the navigation task to help in generating the footstep placements for the humanoid robot. This symbolic script is based on CRIL representation which was developed by our group [6].

The instruction interpreter contains a simple parser, a lexical analysis, and a syntactic analysis. The parser is supplied by the route description text. It separates the text into individual instructions. Each instruction is split into sequence of words using space and punctuation characters as delimiters. The resulting list is entered at the syntactical analysis stage to identify the structure of instructions by comparing their structure with a list of all kinds of instruction syntax which are understandable by the robot. Each word is looked up in the lexicon to obtain its type and features. The available types of words in the lexicon are command verbs, directions, prepositions, numbers of turns, and landmarks. Each verb entry in the lexicon consists of an action verb and an associated script composed from the set of its primitives and depends on the specified arguments passed to its instruction. It is defined as a script of primitive operations that run on data extracted from the analyzed instruction.

After analyzing the route instructions syntactically and connecting each resulting verb with its motion procedure, the symbolic representation of the route is generated. The resulting symbolic script consists of three basic components: motion actions, spatial relationships, and landmarks. The motion actions are classified into the following four different actions:

- BE_AT Action: It presents the position of the robot during navigation. It identifies the start, current, and end positions of the robot during navigation.
- GO Action: It indicates the motion actions which should be taken by the mobile robot.
- VIEW Action: It is used to notice a landmark in a certain direction or region during navigation.
- CH_ORIENT Action: It is used to indicate a change in the current orientation of the mobile robot motion during navigation based on a specific direction or landmark.

The spatial relationships are classified into two types. First, relations represent a location with respect to a landmark. Second, relations specify a direction with respect to one or two landmarks. Finally, the landmark features are retrieved from the knowledge base. They contain data about their shape, color or color histogram, and recognition method values. In addition to the retrieved features, the relationship feature is extracted from the processed route to describe the relation between the current processed landmark and other landmarks in the same path segment. It is used to handle uncertainty and missing information during the robot navigation. Landmarks in our miniature city are classified into definite and indefinite landmarks depending on their features. Definite landmarks have unique characteristics which single them out from among the other landmarks in the miniature city, such as the Burger king restaurant, the Saturn store, and the town hall. On the other hand, indefinite landmarks have a number of properties that are not unique such as buildings, crossroads, and streets.

After creating the symbolic representation of the route, the robot requires an adequate representation of the route environment. This representation should be abstract enough to facilitate higher-level reasoning tasks like strategic planning or situation assessment, and still be detailed enough to allow the robot to perform lower-level tasks like path planning/navigation or self-localization [16]. The topological map representation is used to describe relationships among features of the environment in a more abstract representation without any absolute reference system [17]. Our implementation of the topological map represents the robot's workspace in a qualitative description. It presents a graph-like description of the route where nodes correspond to significant, easy-to-distinguish landmarks, and arrows correspond to actions or action sequences which will be executed by the mobile robot [18].

5 Symbol Grounding

After building the topological map, the resulting symbolic representation is supplied to the motion planner as initial path estimation. The motion planner uses both the symbolic representation and the output of the stereo vision and landmark recognition stage to calculate the desired footstep placements of the humanoid robot to execute the processed route. The motion planner grounds the landmark symbols to their corresponding physical objects in the environment. Therefore, the symbolic and physical presentations of the landmarks should be integrated. Many researchers [19-21] have worked on the symbol grounding problem to solve the problem of incorporating the high-level cognitive processes with sensory-motoric processes in robotics. Cognitive processes perform abstract reasoning and generate plans for actions. They typically use symbols to denote objects. On the other hand, sensory-motoric processes typically operate from sensor data that originate from observing these objects. The researchers tried to maintain coherence between representations that reflect actions and events, and the produced stream of sensory information from the environment. Accordingly, mobile robots need learning abilities that constrain abstract reasoning in relation to dynamically changing external events and the results of their own actions [4].

Harnard [19] considered perceptual anchoring as an important special case of symbol grounding. The anchoring is defined as the process of creating and maintaining the correspondence between symbols and sensor data that refer to the same physical objects [22]. We used a perceptual anchor to incorporate the symbols of the landmarks represented in the symbol system (Σ) and the physical landmarks retrieved from the perceptual system (Π). The predicate grounding relation (g) is used to encode the correspondence between predicate symbols and admissible values of observable attributes. It contains the values of the landmark properties, such as color histogram values, shape, area range, and recognition method values. We used a color histogram, scale invariant features transform (SIFT), and the bag of features methods to recognize the landmarks. As shown in Fig. 3, the perceptual anchoring (α) for a landmark contains a pointer to its symbol (δ), a pointer to its physical object (π), and its signature (γ).

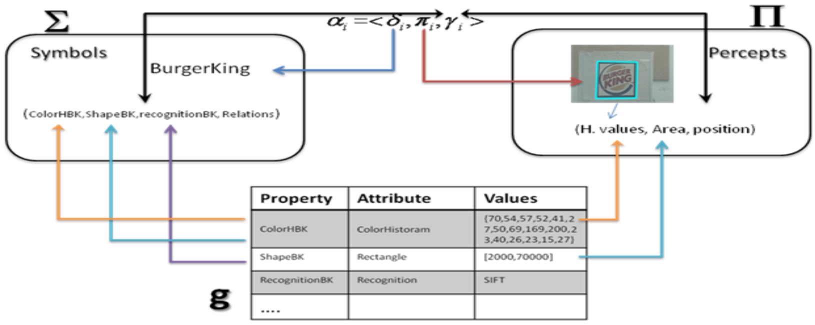

Fig. 3. Anchoring process between the symbolic and perceptual systems

6 Conclusion

We presented RIL, a semi-formal language to be used by non-expert users to instruct mobile robots. Based on RIL, we designed and realized an intuitive interface to mobile robots preventing misunderstanding and ambiguities in route descriptions. Starting from a set of commands, the instruction interpreter stage performs the analysis of route instructions and its lexicon relates the internal procedures to perceptual objects and specifies actions that can be carried out by the mobile robot. The instruction interpreter analyzes the route to generate its equivalent symbolic representation which is supplied to the motion planner as initial path estimation.

The resulting symbolic representation of the route is used to generate a graphical representation of the route to supply the robot with global route information and to prevent it from getting trapped in local loops or dead-ends in unknown environments. Finally, the symbolic representation is supplied to the motion planner to ground the landmark symbols to their equivalent physical objects by using perceptual anchoring.

References

1. Kiesler,S.: Fostering common ground in human-robot interaction. In proceedings of the IEEE international workshop on robot and human interactive communication (ROMAN), pp. 729-734 (2005)
2. Brennan, S. E.: The grounding problem in conversations with and through computers. In Social and cognitive psychological approaches to interpersonal communication. Fussell SR & Kreuz R J (Eds.), pp. 201-225 (1991)

3. Habel,C.: Incremental generation of multimodal route instructions. In natural language generation in spoken and written dialogue, AAAI spring symposium 2003. palo alto, CA., pp. 44-51 (2003)
4. Lauria, S., Bugmann, G., Kyriacou,T., Bos, J., Klein, E.: Training personal robots using natural language instruction. IEEE Intelligent Systems, vol.16, pp. 38-45 (2001).
5. Cangelosi, A., Harnad, S.: The adaptive advantage of symbolic theft over sensorimotor toil: grounding language in perceptual categories. Evolution of Communication, vol.4, pp. 117-142 (2000)
6. Tschander, L. B., Schmidtke, H., Habel, C., Eschenbach, C., Kulik L.: A geometric agent following route instructions. In Spatial cognition III, pp. 89-111 (2003)
7. Tellex,S., Roy, D.: Spatial routines for a simulated speech-controlled vehicle. In proceedings of the 1st ACM sigchi/sigart conference on human-robot interaction, Salt Lake City, Utah, USA, pp. 156-163 (2006)
8. Skubic,M., Perzanowski, D., Blisard, S., Schultz, A., Adams, W., Bugajska, M., Brock, D.: Spatial language for human-robot dialogs. Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions, vol.34, no. 2, May, pp. 154-167 (2004)
9. Schulz, R., Stockwell, P., Wakabayashi, M., Wiles, J.: Towards a spatial language for mobile robots. In proceedings of the 6th international conference on the evolution of language pp. 291-298 (2006)
10. Tellex, S., Roy, D.: Grounding language in spatial routines. In proceedings of AAAI Spring Symp. on control mechanisms for spatial knowledge processing in cognitive / intelligent systems (2007)
11. Pires, G., Nunes, U.: A wheelchair steered through voice commands and assisted by a reactive fuzzy-logic controller. Journal of Intelligent and Robotic Systems, vol.34, pp. 301-314 (2002)
12. Simpson, R. C., Levine, S. P.: Adaptive shared control of a smart wheelchair operated by voice control. In proceedings of the 1997 IEEE/RSJ international conference on intelligent robots and systems (IROS '97), pp. 622-626 (1997)
13. Bischoff, R., Jain, T.: Natural communication and interaction with humanoid robots. In second international symposium on humanoid robots. Tokyo (1999)
14. Torrance, M. C.: Natural communication with mobile robots. MIT Department of Electrical Engineering and Computer Science. (1994)
15. Elmogy, M., Habel, C., Zhang, J.: Spatial Language for Route-Based Humanoid Robot Navigation. In proceedings of the 4th international conference on spatial cognition (ICSC2009), Roma, Italy, to be published (2009)
16. MacMahon, M.: Marco: a modular architecture for following route instructions. In proceedings of the AAAI workshop on modular construction of human-like intelligence, Pittsburgh, PA, pp. 48-55 (2005)
17. Zavlangas, P. G., Tzafestas, S. G.: Integration of topological and metric maps for indoor mobile robot path planning and navigation. Methods and applications of artificial intelligence, pp. 121-130 (2002)
18. Elmogy, M., Habel, C., Zhang, J.: Robot topological map generation from formal route instructions. In proceedings of the 6th international cognitive robotics workshop at 18th european conference on artificial intelligence (ECAI), Patras, Greece, pp. 60-67 (2008)
19. Harnad, S.: The symbol grounding problem. Physica D. Nonlinear phenomena, vol.42, no. 1-3, June, pp. 335 – 346 (1990)
20. Chella, A., Coradeschi, S., Frixione, M., Saffiotti, A.: Perceptual anchoring via conceptual spaces. In proceedings of the AAAI-04 workshop on anchoring symbols to sensor data (2004)
21. Karlsson, L., Bouguerra, A., Broxvall, M., Coradeschi, S., Saffiotti, A.: To secure an anchor - a recovery planning approach to ambiguity in perceptual anchoring. AI Communications, vol.21, no. 1, pp. 1-14 (2008)
22. Coradeschi, S., Saffiotti, A.: An introduction to the anchoring problem. Robotics and Autonomous Systems, vol.43, pp. 85-96 (2003)