Common Sense Data Acquisition for Indoor Mobile Robots

Mykel J. Kochenderfer
Honda Research Institute USA, Inc.
800 California St., Suite 300
Mountain View, CA 94041
mkochenderfer@hra.com

Rakesh Gupta
Honda Research Institute USA, Inc.
800 California St., Suite 300
Mountain View, CA 94041
rgupta@hra.com

ABSTRACT
The objective of this research is to enhance the intelligence of mobile robots that work in home and office environments with common sense. Since indoor common sense knowledge is extremely vast, it may be collected in a similar fashion to the Open Mind family of distributed knowledge capture projects over the Internet. This paper describes the collection of data through the Open Mind Indoor Common Sense (OMICS) website. The knowledge was collected through sentence templates that were dynamically generated based on user input. This was converted into relations and saved into a database. We discuss the results of this online collaborative effort and mention various applications of the collected data.

Categories and Subject Descriptors
I.2.4 [**Artificial Intelligence**]: Knowledge Representation Formalisms and Methods—*Frames and scripts*; I.2.6 [**Artificial Intelligence**]: Learning—*Knowledge acquisition*

1. INTRODUCTION
To accomplish more than extremely narrow tasks in a home or office environment, a mobile robot must posses some amount of common sense. Such common sense includes knowledge about human desires, objects and their locations, and causality. Although our domain is restricted to indoor home and office environments, the amount of knowledge required for a robot to function is still quite vast. Since common sense does not require expert knowledge, the data may be collected as part of a public online collaborative effort over the Internet.

Distributed online knowledge acquisition projects, such as those associated with the *Open Mind Initiative*[7], have been quite popular. The Open Mind Common Sense project, led by Push Singh at the MIT Media Lab, has accumulated a corpus of 600,000 pieces of knowledge from 11,500 users (as of August 2003) over the past three years[1]. Other projects such as Open Mind Word Expert[2] and Open Mind 1001 Questions[3], have also been successful.

This paper describes the Open Mind Indoor Common Sense (OMICS) project that became publicly available at the beginning of August 2003[4]. We report on our results and experiences with working with online contributors. The final section of the paper discusses how we will use the data we collected.

2. METHODS
In this section we first describe the framework of the knowledge base and the relations necessary to capture those types of common sense most useful to an indoor robot. We then describe how we built a website to capture this data in a way that is friendly to non-expert users and how we converted the data into machine-understandable relations.

2.1 Knowledge Base Framework
The framework of this project is *object-centric*. It is assumed that the desires of the users and the actions taken by the robot are grounded in the properties of objects in the world. The robot is useful to humans because it can perform operations on objects in such a way that a desired effect is achieved.

The robot can observe properties of objects in its vicinity and it can perform actions that change the properties of objects. In this system, a statement is a pair $\phi = (o, p)$ where o is some object and p is an adjective describing the property. Statements may be thought of as assertions about the property of an object in the world or actions to be taken on an object to

[1]http://commonsense.media.mit.edu
[2]http://www.teach-computers.org/word-expert.html
[3]http://www.teach-computers.org/learner.html
[4]http://openmind.hri-us.com

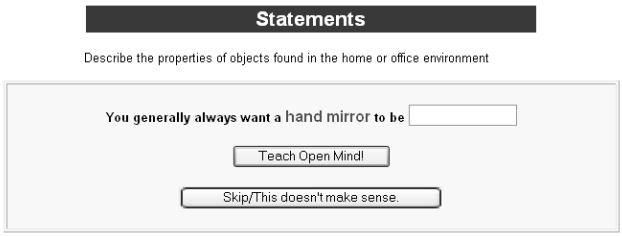

Figure 1: A template sentence.

achieve a particular effect. For example, the statement (*cup-of-coffee*, *hot*) can mean "a cup of coffee is hot" or represent the action "make a cup of coffee hot."

Since we want our robot to understand the consequences of its actions, we wish to capture such common sense knowledge as: (o_1, p_1) causes (o_2, p_2). For example, the statement (*fan*, *on*) causes (*room*, *cool*). We also wish to capture knowledge such as (o_1, p_1), which would indicate that you would want (o_2, p_2). For example, the statement (*cup-of-coffee*, *cold*) would make you want the statement (*cup-of-coffee*, *hot*) fulfilled.

At any given point in time, the robot observes a set of statements that are true and the robot is able to execute a set of statements. The general problem is to decide which statements to execute in order to achieve perceived goals.

It is important that the robot knows where objects are generally located (e.g. a pillow may be found in the bedroom). The robot should also know what various objects are used for (e.g. a microwave is used to heat food).

2.2 Sentence Templates

We need some way to convert the common sense in the minds of non-expert users into relations in a knowledge base. Following the style of the Open Mind Common Sense website, we decided to use sentence templates. Users are prompted to fill in the blanks of sentences with words, as shown in Fig. 1.

Different activities capture different types of knowledge. Below is a summary of some of the activities:

Objects

In this activity, the user is asked to identify the types of objects commonly found in a home or office. The user may be prompted to type in an object name that comes to mind or simply label an image of an object. It is important to allow a user to simply type in any indoor object that comes to mind because we want to include all relevant objects in the database even if we do not have their picture.

Image labelling can link multiple labels to the same object (e.g. the labels "sofa" and "couch" might be both associated with "h0141.gif"). The images themselves can be used for training the object recognition system of the robot. When the website became public, the database initially contained a set of over 400 images of indoor objects selected by hand from a collection of stock photography.

Statements

In the 'statements' activity, the user is prompted with a question such as, "You often want a fan to be *blank*." This activity pairs objects with properties in the knowledge base. The objects that appear in these sentence templates come from the objects entered by users in the other activities, such as the 'objects' activity.

Uses

This activity associates objects with their uses. For example, the user might be prompted with the form, "A hanger is used to _____." Again the objects come from user input.

Causes

This activity captures causality. For example, a form might ask, "A computer is off when a _____ is _____." If the user enters a new object or a new object-property pair, it will be entered into the object or statement table. The object and property that makes up the first part of the sentence is formed dynamically by selecting a random object from the knowledge base.

Desires

This activity helps the robot determine what needs to be done in various situations. A template form might ask, "You might want a fan to be blowing if you notice that your _____ has become _____."

Locations

This activity associates objects with the rooms where they are typically found. For example, the user might be prompted with, "A room where you generally find a dinner table is the _____."

Proximity

This activity associates objects with each other based on proximity. A sample form would be, "You generally find a frying pan in the same room as a _____."

Senses

This activity disambiguates the intended sense of various objects entered into the database by other users. Fig. 2 shows a sample form. The objects to disambiguate are selected from previous user entries and the senses are from WordNet [3].

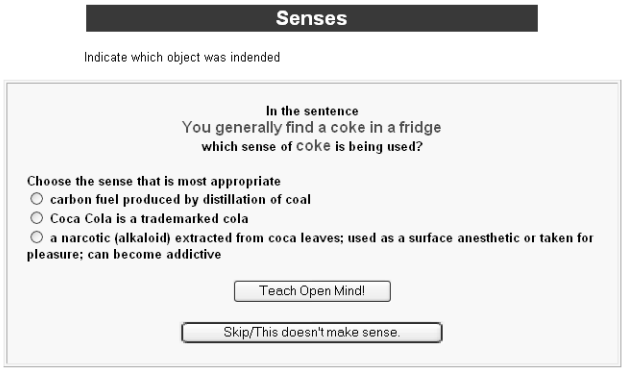

Figure 2: A word sense disambiguation form.

Freeform

This activity allows users to type in any form of common sense knowledge that might not be captured by any of the other activities. Although it is quite difficult to convert freeform sentences into useful relations, it provides us with a sense of the types of knowledge the general public would like an indoor robot to understand. Analysis of freeform sentences will later lead to the creation of new activities.

2.3 Data Collection

Once a user logs on with their account, they are presented with a random activity. After a random number of entries for a particular activity, the system prompts them with a new activity. Users may also manually switch between the activities.

2.4 Data Quality Review

The completed sentence templates are stored in the database as raw sentences. These sentences are not parsed into relations until after an administrator "approves" of them (see Fig. 3). It is important that there be some way of ensuring data quality since the data (such as names of objects and their properties) are used to generate new sentence templates. Sentence templates containing misspelt objects or objects that do not appear in a home or office environment would propagate errors in the knowledge base.

Once an administrator approves of an entry, it is parsed into relations immediately. There is currently no need for part-of-speech tagging or lemmatization as with the Open Mind Common Sense project since the sentence templates are structured and designed in such a way that they implicitly cue the user as to what part-of-speech and tense they should use.

3. RESULTS

To advertise the OMICS site, a message was sent to the Open Mind Initiative mailing list at the beginning of August. With no other direct advertising, within three weeks we had 171 users and 18,000 entries with about

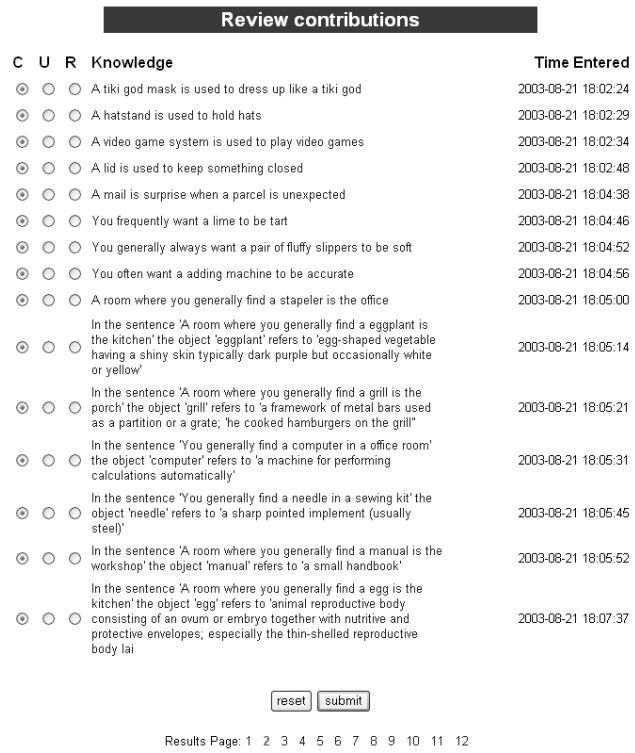

Figure 3: The review form used for administrators to approve of user entries. Administrators may either commit, uncommit, or reject entries.

17,000 of them accepted. We had a weekly contest (lasting four weeks) where the top contributor was awarded an Open Mind t-shirt. Other Open Mind projects have used similar contests to help motivate submissions. Winners had their names and hometowns listed on the title page of the site.

3.1 Observations

The OMICS site was publicly announced on August 5th and the first t-shirt prize was awarded August 6th. A significant portion of the entries were submitted within two days of the announcement. Prizes were also awarded on August 12, August 19, and August 26. The first few weeks of data collection is plotted in Fig. 4.

The quality of the data was actually quite good. Less than 10% of the submissions were rejected. Entries that were rejected tended to fall within one of the following categories:

- Misspelling: e.g. "A room where you generally find a exercise bike is the bym."
- Grammatically incorrect: e.g. "A magazine rack is used to stored magazines."
- Unclear or loose statements: e.g. "People cry when they can't get it."

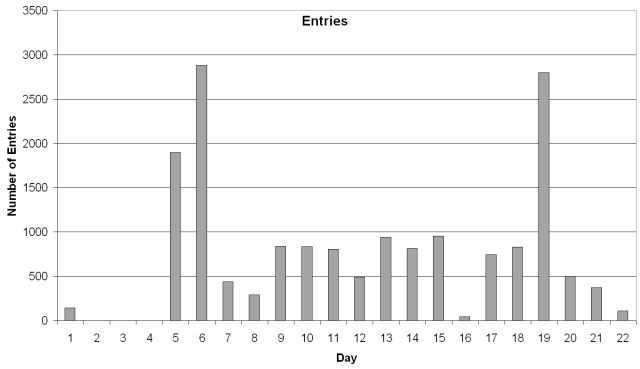

Figure 4: The number of entries submitted plotted against the day in August during data collection.

- Outside the scope of home and office environments: e.g. "A trap is set when a predator is hunting."
- Nonsense: e.g. "You generally want a green light to be addressed to you."
- Inappropriate: e.g. suggestive or obscene

The 'causes' activity had the highest rejection rate. Deciding how one object affects another object proved to be difficult for some users. Interestingly, almost all word sense activities were answered correctly. Even users that entered "rogue" data in other activities generally entered the correct sense of the words in the 'sense' activity.

Users that appeared to have hostile intentions, indicated by sentence completions that were of a crude or sexual nature, also submitted useful data. Surprisingly, a few users that might be classified as malicious were among the top contributors of good data.

Fig. 5 shows the ranking of the various activities and the number of submissions. Not surprisingly, users spent significantly more time on the 'objects' activity. This is probably because the 'objects' activity requires the least amount of thought and because it was the only activity that involved images. Although users were allowed to submit their own images of indoor objects, very few users actually did.

3.2 Feedback

Although we have received little feedback on the OMICS website, comments thus far have been largely positive. One of the weekly winners entered data with her seven-year-old son. She had the following to say about the site:

> As a teacher I think it is really great and put my son on it with me here at home—It was a great mom and kid project for several days. My little one who is 5 will be working on it too this week. It really forces us to do some critical thinking—the relationships and location section were great for him as were the free questions that he could come up with on his own.

Activity	Count
Objects	3717
Uses	2517
Locations	2458
Statements	2448
Proximity	1702
Freeform	1213
People	1086
Desires	1054
Causes	1025
Senses	646
Images	23

Figure 5: The number of submissions for each activity.

Some users were concerned about their spelling errors and the spelling errors that were part of their sentence templates. Most grievous spelling errors were filtered out by the administrators, but some minor spelling errors were allowed into the database. Other users were concerned that the data they entered already existed in the database. One user commented:

> I still don't really know what to put in the "freeform" category, or what needs to be worked through to be the most help in the project (I'm guessing something like "computer & mouse" gets overused, while other objects go ignored?).

The general tone of the e-mails we received were positive and indicated a desire by the users to see the project succeed in making robots more intelligent.

4. DATA APPLICATIONS

The data collected from the Open Mind family of projects have been applied to a wide variety of problems. Open Mind Word Expert has been applied to Word Sense Disambiguation [1] and Open Mind Common Sense has been applied to textual affect sensing [2]. In this section, we describe some ways in which the data collected as part of OMICS is being used.

4.1 Action Selection

Since the OMICS database contains information about desires, the robot is able to deduce what actions it should take based on its observations. The robot would

want to take actions that satisfy either these goals directly or accomplish something that causes a goal to be satisfied.

The following example comes from the OMICS database. Suppose that the robot observes (*coke-can*, *warm*). The OMICS database contains the piece of knowledge that if you observe (*coke-can*, *warm*) then you would want (*coke-can*, *chilled*). The robot cannot directly execute the statement (*coke-can*, *chilled*), but it can execute the statement (*coke-can*, *refrigerated*). The robot would then decide to execute (*coke-can*, *refrigerated*) because it knows from the knowledge base that (*coke-can*, *refrigerated*) causes (*coke-can*, *chilled*).

4.2 Inference

Since the relations are structured and interconnected, rule-based inference may be applied to produce sentences that were not explicitly entered into the database. For example, if the database contains the facts that a spoon is found near a bowl and a bowl is generally found in the dining room, the inference engine would be able to tell you that a spoon may be found in the dining room with a certain degree of confidence. Any sort of theorem proving environment, such as SRI SNARK [6], may be used to do inference over the database.

The database has been used in conjunction with WordNet to infer the location of objects that were not found in the database. For example, the object "patchwork" was never entered into the OMICS database. However, our inference engine was able to observe that a patchwork is a hyponym of "quilt." The OMICS database contained the fact that quilts are typically found in bedrooms and living rooms, and so the system was able to infer that patchworks are found in bedrooms and living rooms.

5. FURTHER WORK

Although the robot can use the common sense knowledge at a very high level to determine which tasks to pursue, it is not yet intelligent enough to actually execute the mid-level actions that accomplish these goals. It remains to be seen how common sense can be used in the actual execution of various tasks, such as cleaning a bathtub.

One might use the teleo-reactive program framework, as proposed by Nils Nilsson [4, 5], to accomplish such basic tasks as "make a cup of coffee hot" or "pop some popcorn." Although it does not require an expert to accomplish such tasks, it is not clear how to best capture this kind of knowledge through a web interface.

6. CONCLUSIONS

The Open Mind Indoor Common Sense project has successfully captured thousands of pieces of common sense knowledge about home and office environments. The indoor focus of the website and the structured activities have lead to data that is useful to a robot in anticipating desires, understanding causality, and interacting with objects.

7. ACKNOWLEDGMENTS

The authors would like to thank David Stork and Push Singh for their helpful comments and suggestions and the users of the Open Mind Indoor Common Sense website for their data and feedback.

8. REFERENCES

[1] Timothy Chklovski (2003). Using Analogy to Acquire Commonsense Knowledge from Human Contributors. MIT Artificial Intelligence Laboratory technical report AITR-2003-002, February 2003.

[2] Hugo Liu, Henry Lieberman, and Ted Selker (2003). A Model of Textual Affect Sensing using Real-World Knowledge. *Proceedings of the Seventh International Conference on Intelligent User Interfaces (IUI 2003)*, pp. 125-132. Miami, Florida.

[3] George A. Miller, et. al. (2003). WordNet: A Lexical database for the English Language. `http://www.cogsci.princeton.edu/~wn`

[4] Nils J. Nilsson (1992). Toward agent programs with circuit semantics. Technical Report STAN-CS-92-1412, Department of Computer Science, Stanford University.

[5] Nils J. Nilsson (1994). Teleo-reactive programs for agent control. *Journal of Artificial Intelligence Research*, 1:139158.

[6] Mark E. Stickel, Richard J. Waldinger, and Vinay K. Chaudhri (2001). A Guide to SNARK. `www.ai.sri.com/snark`.

[7] David G. Stork (1999). "The Open Mind Initiative," *IEEE Expert Systems and Their Applications* pp. 16–20, May/June 1999.