Human-Centered Adaptive Mobile Robot based on On-line Dialogue and Stochastic Experience Representation

Tetsunari INAMURA*1,2  Ken NAKA*2  Masayuki INABA*2  Hirochika INOUE*2
*1 Japan Science and Technology Corporation (JST), CREST
*2 Dept. of Mechano-Informatics, Faculty of Engineering,
The University of Tokyo. 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656 JAPAN

Abstract
In this paper, we propose a novel method for personal robots to acquire autonomous behaviors gradually based on interaction between human and robots. In this method, behavior decision models and dialogue control models are integrated using Bayesian Networks. This model can treat interaction experiences using statistical processes, and surelyness of decision making is represented by certainty factors using stochastic reasoning. The robots not only decide behavior, but also make suggestions to and ask questions of the user using the certainty factors. Consequently, the certainty factors enable the behavior acquisition to be more effective. We investigate the feasibility of this method with obstacle avoidance tasks for mobile robots. Through experiments on a real mobile robot, we have confirmed that the mobile robot acquires robust behavior decision models against changes of environment and uncertainties of sensors, with only a few teaching and learning sessions.

1 Introduction
Recently, research on personal robots which interact with human in daily life, is gaining a great deal of attention. These human-centered robots are claimed to achieve intelligent behavior such as submision to users and observation of users. However, it is difficult for personal robots to predict in advance what kind of situation will occur. Therefore, an approach in which environmental information and users’ preference are embedded in robots by developers, can encounter difficulties.

To approach this problem, teaching and learning are often used for behavior acquisition. There is a trade-off between the two approaches. On the one hand, when the teaching method is used, the robot cannot cope with dynamic changes of environment. Users must teach the robot again when the environment changes. On the other hand, when the learning method is used, there is no guarantee that the robot acquires behavior which is suitable for the users’ preferences. Users must prepare appropriate teaching signals or rewards in order to let the robot acquire the desired behavior.

One of the examples which reveals the issue of teaching approaches is a mobile robot developed by Shibata et al.[1]. In a real-time interactive system on the mobile robot, a user boards the robot and teaches the robot how to move in an unfamiliar corridor environment. With this system it is easy to alter the behavior because several basic behaviors have been embedded and the user determines the combination of these behaviors. Nevertheless the system cannot cope with changes in the environment. If a change occurs, the robot must be taught all over again. This problem is caused by the fixed representation of the relation between visual information and behaviors.

One of the examples which reveals the issue of learning approaches is Asada’s reinforcement learning framework[2]. The aim of their research is autonomous behavior growth; it doesn’t consider the interaction between users and robots. Therefore, the problem is mentioned that there is no guarantee that the acquired behavior meets users’ preference. Ishiguro et al have proposed a method in which mobile robots use users’ instruction in the beginning of learning[3]. However their approach did not consider communication from the robot to the user.

A learning method which can reflect the users’ preference is needed, as the learning mechanism for personal robots. Therefore, it is impossible to acquire suitable behavior unless only learning mechanisms implemented by the developer are prepared. It is desired that the robot acquires behavior through two-way interaction between users and robots. In this paper, we focus on interaction between human and robots in order to integrate teaching and learning approaches. In our approach, the robot makes every effort to decide behavior by itself, and makes questions and suggestions in case of need.

In section 2, we propose a stochastic experience representation in order to realize this hybrid behavior acquisition framework. In section 3, two applications of the framework for a real personal mobile robot are proposed. In section 4, we show results of the personal robot application, and we conclude in section 5.
2 Adaptive Interaction based on Stochastic Experience Representation

It is necessary for the robot to describe the relation between sensor and behavior in order to decide the autonomous behavior. Additionally, in this paper, it is necessary that the robot describe its interaction experience and decide a strategy of dialogue.

Considering these conditions, we have adopted a Bayesian Network[4] which is a stochastic inference model. Although Bayesian Networks are intended primarily for diagnosis and reasoning systems, recently they have also been used for interaction frameworks between human and intelligent systems[5]. In the model, relation between sensory information, user's instructions and recognition result of sensory pattern are represented as network of concept nodes, and degree of causal relation are represented as link with probability. Using the model, the robot can infer an unobserved phenomenon based on observed phenomenon.

The reasons for adoption of stochastic representation are the following:

1. On-line learning is available using simple cumulative frequency calculation from interaction experience.
2. Certainty factors of behavior decision can be represented by probabilities. It is also useful for the robots to control dialogue strategy.
3. Stochastic processes can reduce the influence of noise such as users' mistakes.

2.1 Definition and Notation

This network model consists of two kinds of node: evidence nodes and hypothetical nodes. Each node has some independent propositions, which correspond to status, situation, behavior and so on. A set of these propositions have propositional symbols, represented by random variables. The relations between each proposition are represented by conditional probabilities. When the model is used for the robots, the evidence nodes become sensor nodes, and hypothetical nodes become behavior nodes, as shown in Fig. 1.

Let \( S = \{s_1, s_2, \ldots, s_n\} \) be the random variable of a sensor node, and let \( B = \{b_1, b_2, \ldots, b_m\} \) be the random variable of a behavior node. A causal relationship between sensor node \( S \) and behavior node \( B \) is described by the following matrix:

\[
CPT_{SB} \stackrel{\text{def}}{=} P(S|B)
\]

\[
= \begin{pmatrix}
P(s_1 | b_1) & P(s_2 | b_1) & \cdots & P(s_n | b_1) \\
P(s_1 | b_2) & P(s_2 | b_2) & \cdots & P(s_n | b_2) \\
\vdots & \vdots & \ddots & \vdots \\
P(s_1 | b_m) & P(s_2 | b_m) & \cdots & P(s_n | b_m)
\end{pmatrix}
\]

This matrix is called the Conditional Probabilities Table (CPT).

2.2 Learning Mechanism

The learning mechanism using Bayesian Networks is divided into three phases. First is the users' instruction phase. The robots act following the users' instruction, and store the instruction and sensory data into the experience database. Second is the learning phase. The system transfers the experience database into a probabilistic causal relation between each proposition. Third is the reasoning phase. The robots infer the users' instructions based on observed sensory information, and make a behavior decision, questions, and suggestions to the users.

Observation phase

Basically, the robots act following the users' instructions. While acting, the robots observe what kind of situation they encounter, what kind of behavior the user instruct, what kind of sensor information is input, and so on. These observation results are translated into propositional symbols, and stored as a personal database for learning and adaptation.

Learning phase

Let \( N \) be the number of observed experiences, and \( N_j \) be the number of propositional symbols \( b_j \), which are instructed by the user across the \( N \) experiences. Also, let \( n_{ij} \) be the number of propositional symbols \( s_i \) observed by the sensors in the \( N_j \) instructions. Learning is executed for each link and each node, namely conditional probabilities and posterior probabilities. The most simple calculations are the cumulative frequency as following:

\[
P(s_i | b_j) = P(S = s_i | B = b_j) = \frac{n_{ij}}{N_j} \quad (2)
\]

\[
P(b_j) = P(B = b_j) = \frac{N_j}{N} \quad (3)
\]

Reasoning phase

Conversely when the robot acts autonomously, and must decide behavior by itself, the probability of the plausible behavior \( P(B | S) \) is reasoned using sensor information \( S \) and \( P(S | B) \).
Before reasoning, the robot observes sensor information and current situation for the evidence nodes, and inputs them to the Bayesian Network. After that, probabilities of the random variable of the target behavior node is calculated. Using Baye’s rule, the calculation is expressed as following:

\[
BEL(B) \triangleq P(B|S) = \frac{P(S|B) \cdot P(B)}{P(S)} = \frac{CPT_{S|B} \cdot P(B)}{P(S)} \tag{4}
\]

In the case that the number of nodes is two, the probabilities are calculated using the above equation, however, in the case that the number of nodes is more than two, it becomes more complicated. An algorithm for calculation of the probabilities of unobserved propositions based on observed propositions on the network structure is proposed by Pearl[4]. In this method, probabilities of unobserved propositions are calculated based on adjoined observed propositions. Information about observed propositions is propagated for all nodes in the network, with a bound on the number of calculations proportional to the number of nodes, unless the network has a loop structure.

We adopted this algorithm for reasoning phase. As the propagation algorithm is used to compute the details of the algorithms in [4]. In the case of this paper, the calculation is the following for the Bayesian Network shown in Fig. 2:

\[
BEL(B) = \alpha \prod_{i} \lambda_{i}(B) \pi(B) \tag{5}
\]

where \(\alpha\) is a coefficient for normalization, \(\pi(B)\) is prior probability of the behavior node, and \(\lambda_{i}(B)\) is the likelihood vector which is calculated by

\[
\lambda_{i}(B) = CPT_{S|B} \lambda(S_{i}) \tag{6}
\]

where \(\lambda(S_{i})\) is sensor input of \(i\)-th sensor node.

This probability \(BEL(B)\) is called the certainty factor, which is utilized for behavior decision and dialogue strategy.

**Dialogue Control Strategy**

The certainty factors are also utilized for question generating and suggestion making. Four levels of autonomous behavior are defined based on the certainty factors. Two cases are defined: whether the users’ instruction and robots’ decision is the same, or different. As a consequence, eight situations are defined by the combination of these four levels and two cases, a certain dialogue strategy corresponds to each situation. There are four main strategies: pointing out the mistakes, offering an alternative plan, making questions, and confirmation of the decisions. For example, in the case where the certainty factor is fairly high and the user’s instruction differs from the inference result, the robot points out the user’s mistake. At the utterance generation, the robot uses proposition symbols which correspond to each node and proposition in order to make natural sentence. The details of the concrete examples of dialogue are explained in section 3.

3 Applications and Experiments

3.1 Obstacle avoidance based on uncertain instructions

As an example application, learning of obstacle avoidance tasks for mobile robots are adopted[6]. In the task, no strategy for avoidance is given to the robot. Only the Bayesian Network shown in Fig. 7 is given.

Considering obstacle avoidance tasks, we have adopted distance sensors for the sensor nodes. In the network, a behavior node has a random variable \(B = \{\text{Forward, Stop, Left, Right}\}\), and eight sensor nodes have a random variable \(S = \{\text{Near, Middle, Far}\}\). A task given to the robot is to run the corridor environment shown in Fig. 3, while avoiding obstacles.

**Observation and Learning phase**

While the learning has not converged, the robot observes the users’ instructions and operations. A direct instruction via joystick lever corresponds to the behavior node \(B\). Sensory information is observed at the same time. After the measurement, the results are translated into propositional symbols \(\{\text{Near, Middle, Far}\}\), and inputted into the Bayesian Network.

**Reasoning and Dialogue phase**

Using the manner described in section 2.2, the robot infers plausible behavior and outputs certainty factors for each behavior propositional symbol, \(\{\text{Forward, Stop, Left, Right}\}\). Strategies of the behavior decision and dialogue control are decided based on these certainty factors. The observation and reasoning processes are repeated about every tenth of a second.
For example, the certainty factors for two proposition are the same because of the robot is located in front of an obstacle. The robot tries to ask a question; “I can’t decide which is the better selection, turn to the right? or to the left?” In another case in which the robot reasoned the behavior ‘go forward’, and the user instructs it to turn right, the robot tries to point out the user’s mistakes; “Is this a correct behavior, really?”

This kind of dialogue is often generated in the beginning of the learning, however, the frequency decreases as the learning goes on.

3.2 Person searching based on statistical dialogue experience

Human-centered mobile robots are required to move toward a target person, not only move to a target place. Many navigation methods are proposed, however, it is difficult to navigate toward target persons, because each person moves about fields of daily life. Therefore a behavior is required that the robot estimates the locations where the target person stays roughly and makes communications with surrounding persons when evidence information is short.

Here, we designed a Bayesian Network shown in Fig.4 in order to describe the relation between location and person. Propositions on the person node indicate each person. Propositions on the room node indicate each place tag in the environmental map which acquired by the mobile robot. Propositions on the time node indicate time band, such as Morning, Noon and Night.

Observation and Learning phase

The robot observes and records “when, where, and who does the robot encounter?” in the filed of daily life. Location information is gained using a navigation method mentioned later. Personal identification is executed using visual information [7] and interviews with encountered persons. These information are translated into propositional symbols with time information; stored at the experience database.

After the storage, stochastic relation between room and person and time are calculated according to the Bayesian Network model shown in Fig.4.

Reasoning and Dialogue phase

When users order to the robot like “Please deliver this baggage to Mr. A”, the robot infers the location where the Mr. A stays. In this process, current time is inputted into time node; personal propositional symbol is inputted into person node. The robot takes out the certainty factors of the room node as the result of inference. When certainty factor has ambiguities the robot try to make questions to users in order to decide the target location.

3.3 System configuration of a personal mobile robot

A mobile robot has two binocular cameras for obstacle avoidance and an omnidirectional camera for navigation as shown in Fig.5. The robot can measure distance from it to obstacles using the binocular cameras. Distance information which observed along with horizontal line as shown in Fig.3 are correspond to each sensor node.

For navigation, we adopted a visual navigation method based on omnidirectional view sequence[8]. In the method, the robot can make environmental map and recognize some places in which the robot had been.

Some Bayesian Networks can be applied at the same time. Here we explain an experiment which uses a example explained in Sec.3.1 and Sec.3.2. A task given to the robot is “delivery tasks” in which the robot delivers some baggage to target persons, with avoiding obstacles and making communications with surrounding people.

4 Result of delivery tasks

We have experimented the delivery tasks at an office environment shown in Fig.7. At the experiment, we set up an experiment condition on the assumption that the robot have a perfect environmental map for this field, and person B have not encountered the robot so that the inference for person B is impossible for the robot.

Fig.6 shows the experiment in which the mobile robot makes communications between users and achieves a delivery task.

At first, the robot was ordered to deliver a baggage to person B. However, the robot had not made communication with person B so that the robot try to make communication in order to search person B (In scene from 1 to 3). During navigation to the target
5 Summary and Conclusions

In this paper, we described the framework of learning mechanism and dialogue control strategy using Bayesian Networks. Through a "delivery task", we show the feasibility of the method for real application on mobile robots. On the obstacle avoidance function, the robot can learn a sufficient avoidance strategy within few interactions. Observed on the person searching function, the robot can cope with unfixed persons' behavior using on-line communication between surrounding persons.

There are some mobile robots which can interact with humans in daily life environment\[^{[9]}\]\[^{[10]}\]. However, these robots cannot reuse the experience of the interaction for development and learning for the autonomous behavior. In our approach, the robots are in learning situations and dialogue situations at the same time. Personal robots ought to be developed through interaction with humans in daily life which includes learning, teaching and converging. It is not desirable that learning and dialogue be separated into phases. We think that the Bayesian Network based interaction framework has some advantages for modeling the intelligence of personal robots.

Acknowledgments

This research was partially supported by the "Robot Brain Project" of Core Research for Evolutional Science and Technology (CREST) of the Japan Science and Technology Corporation (JST).

References


Figure 6: A deliver task

Figure 7: An office environment for the deliver task

Table 1: Dialogue between the mobile robot and users

| User A  | “Please deliver for Mr.B”                        | (Pict. 1) |
| Robot   | “Which does Mr.B stay at the resting room or the conference room?” |         |
| User A  | “I think that Mr.B stays in resting room.”       |         |
| Robot   | “I see. Please put the baggage on me.”            | (Pict. 2) |
| Robot   | “I’m going toward the resting room.”              | (Pict. 3) |
| Robot   | “Some obstacles are detected!”                    | (Pict. 7) |
| Robot   | “I’m avoiding the obstacles.”                     | (Pict. 8) |
| Robot   | “The avoiding is over.”                           | (Pict. 9) |
| Robot   | “Please open the door.”                           | (Pict. 11) |
| Robot   | “I’m entering toward the resting room.”           | (Pict. 13) |
| Robot   | “There is a baggage for Mr.B.”                    | (Pict. 14) |
| User B  | “Thanks.”                                         | (Pict. 15) |

66