Synthetic Character with Bayesian Network and Behavior Network for Intelligent Smartphone

Sang-Jun Han and Sung-Bae Cho

Dept. of Computer Science, Yonsei University,
134 Shinchon-dong, Sudaemoon-ku, Seoul 120-749, Korea
{sjhan,sbcho}@sclab.yonsei.ac.kr

Abstract. As more people get using mobile phones, smartphone which is a new generation of mobile phone with computing capability earns world-wide reputation as new personal business assistant and entertainment equipment. This paper presents a synthetic character which acts as a user assistant and an entertainer in smartphone. It collects low-level information from various information sources available in smartphone, such as personal information, call log, and user input. Collected information is used for inferring high-level information with Bayesian network and determining agent's goal. This character takes a behavior appropriate to current situation using behavior selection mechanism. The proposed agent is toward a framework for integrating reactive and deliberative behavior.

1 Introduction

Recently mobile phones have become an essential tool for human communication. As more people use mobile phones, various services based on mobile phone networks and high-end devices have been developed. Smartphone which integrates the functions of personal digital assistant (PDA) and mobile phone earns world-wide reputation as new personal business assistant and entertainment equipment because it is all-in-one device: many technologies such as wireless voice/data communication, digital camera, and multi-media player are converged into one device. Current smartphone is just high-end mobile phone, but, they have the potential of being utilized to many novel services by developing innovative applications. However, current mobile devices have constraints of limited processing power, and awkward interaction devices. We are in need of AI techniques specialized in smartphone to cope with these constraints and make intelligent services in real.

There are three major issues in implementing intelligent service in constrained environment.

- To gather information which provides meaningful features for user's state: it is never an intelligent technique to require directly explicit information from user. Desired technique should be able to provide sufficient information while not bothering the user and not invading the user's privacy.

R. Khosla et al. (Eds.): KES 2005, LNAI 3681, pp. 737–743, 2005.

- To infer and predict user's state from collected data: predicting user's state from data can be formulated as conventional classification task. Many AI methods have been successfully applied to this problem.
- Service selection or composition: we can select one service from pre-defined service library or compose novel services appropriate to inferred user's state dynamically.

Here we used personal information, communication log, and the state of smartphone device as the information source, which provides the low-level clues used for inference and action selection. Bayesian network is used for inferring the high-level states from low-level information. In constructing the Bayesian network, we employ commonsense knowledge actively to deal with incomplete, imprecise data. It helps out inferring process by magnifying implicit and hidden clues in that kind of data. In order to generate services appropriate to user's state, we employ the concept of synthetic character based on behavior selection network. A character plays the role of a good medium for entertainment as well as a user assistant with intimate user interface. Its behavior reflects not only user's states such as how he/she feels and how busy he/she is, but device's states such as remaining battery power and things displayed on screen.

2 Intelligent Agent for Smartphone

The proposed agent consists of four components: perception system, emotion system, motivation system, and action selection system. The perception system checks the changes of personal and device information by monitoring address book, call log, user's input, currently running programs, etc. It determines the low-level states such as the category of ongoing scheduled event, the time elapsed after last input, and the number of missed call. The emotion system infers the high-level states such as affect and how busy the user from the low-level state. It uses Bayesian network and the low-level states is used as the evidence for inference. The motivation system determines the goal of the agent using low-level and high-level states. The action selection system chooses agent's action appropriate to the current state. Figure 1 shows an overview of the proposed agent.

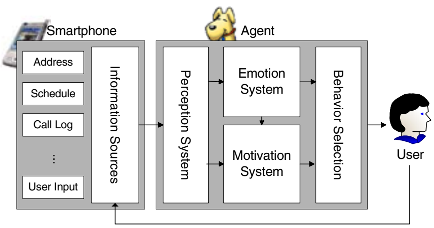


Fig. 1. An overview of the proposed agent.

2.1 Collecting Useful Information

Perception system collects basic information (low-level state) which is used to infer high-level state, determine agent's goals, or select appropriate action. The design methodology of perception system is as follows. First, we discover some events that seem to be related with user's state or agent's goal and define them as low-level state. Next, we specify the possible values that each state has and make concrete conditions for each value of state. For instance, to infer how he/she busy is, we use 'the number of missed calls' as related variable and define 'many', and 'few' as possible states of this variable. The state of 'the number of missed calls' is "many" if the number of calls which were not answered in the last two hours is more than five and 'few' if less than five. It monitors user's input, call/text message log, and changes in device's state or personal information continuously and update the value of the states using defined rules.

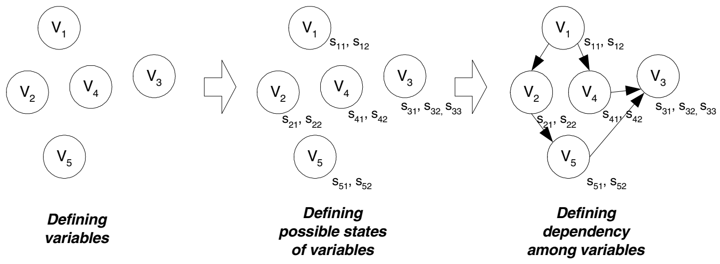
2.2 Inferring High-Level States

The low-level state is not enough to recognize user's unspoken needs and respond to it properly. It is needed to mine higher-order information from low-level information. However, there is much uncertainty in attempts to recognize high-level state such as user's object and affect. Bayesian probabilistic inference is one of the famous models for inference and representation of the environment with insufficient information.

The design methodology of Bayesian network is as follows. First, we discover some events that seem to be related user's context and define them as variables in Bayesian network. Next, specify the states that each variable has and make concrete conditions for each state of variable. For instance, to infer how he/she busy is, we use "the number of missed calls" as related variable and define "many", "few" as possible state of this variable. The state of "the number of missed calls" is "many" if the number of calls which did not be answered in the last two hours is over five and "few" if under 5. The accuracy of user modeling hinges upon how we define the variables in Bayesian network and the state of each variable.

After defining variables and its states, we construct the structure of Bayesian network considering the dependency among variables. In this stage, we have to specify the topology and the probability distributions. There are two ways to do this: constructing it automatically from data with learning algorithm and manually using expert's domain knowledge. In this work, we have constructed it manually because it is difficult to gather much amount of personal information data. Figure 2 illustrates the design process of Bayesian network.

The emotion system infers user's high-level state using probabilistic inference technique. We have used user's affect, how busy he/she is, and how close he/she is with someone as the high-level states. Bayesian network is used for inferring high-level states. The low-level states are used as variables in Bayesian network. Perception system traces the changes of personal information and communication log. Then, it sets the values of the variables in Bayesian network.



	s_{51}	s_{52}
s_{21}	.2	.4
s_{22}	.8	.6

Fig. 2. The design process of Bayesian network.

The structure of Bayesian network is constructed considering the dependency among variables. We have to specify the topology and the probability distributions. There are two ways to do this: constructing it automatically from data with learning algorithm and manually from expert's domain knowledge. In this work, we have constructed it manually because it is difficult to gather enough personal information data.

Bayesian network for this problem has 33 observable variables whose state can be specified by observing user's behavior or personal information and 19 unobservable variables whose state is specified from relationship with other variables. All variables have two states and there are 56 dependencies defined among the variables.

- **Inferring user's affect.** Valence-Arousal (V-A) space has been applied to infer user's affect. V-A space is a simple model which represents affect as the position in the two-dimensional space. It has commonly used in previous studies on affect recognition [2]. It uses two eigenmoods: valence and arousal. Every affect can be described in terms of these eigenmoods. The valence axis ranges from negative to positive. The arousal axis ranges from calm to excited. For instance, "anger" is considered low in valence while high in arousal. In this work, we have used just four simple emotions: joyful, anger, sad, and relaxed depending on the sign of value in each axis.
  General commonsense knowledge has been utilized to define relationships among variables. For example, if there are many business schedules, user's emotion is likely to be negative. Many positive words or emoticons in user's incoming message box imply that user's emotion is likely to be positive.
- **Inferring how the busy user is.** The possible clues for inferring how busy the user is may be the amount of today's schedules, whether it is a holiday, and how frequently he uses the phone. Many business schedules would indicate that user is very busy with his work. No schedule in the evening implies that user has a rest. Busy user is likely to cause many missed calls or unanswered messages. With these commonsense rules, we define causal relationships among the variables.

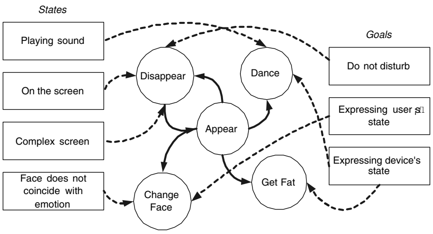


Fig. 3. A part of behavior network (Dotted: External link Solid: Internal link).

- **Inferring how close the user with someone.** In order to infer how close the user with people who are registered in address book, we have discovered the clues such as the content of incoming text messages, the name of group which he/she is added to, the frequency and duration of phone call, etc. The commonsense knowledge has been used to construct Bayesian network as follows. People who frequently make contact with or are registered at "friends" group in address book can be regarded as close. People who never got contact with in long time can be regarded as not close.

2.3 Determining Goal of Agent and Selecting Appropriate Action

The ultimate aims of our agents are assisting and entertaining the user in the way which is appropriate to current situation. Therefore, it is needed to keep it to be suitable for current situation by changing it from time to time. For this reason, we determine the goal of the agent considering information about the user and the device which is collected by perception system. Situation-goal rules are used for determining current goal. The situation part consists of the low-level states observed by perception system and high-level states inferred by Bayesian network. Things which are supposed to be need by the used in particular situation is set as agent's goal. Action that is appropriate to current situation is selected using behavior network. Whether the agent should make action and what action should be activated is determined by selection mechanism of behavior network. Behavior network has the advantage of flexibility with which human can deal in comparison to simple rule-based selection mechanism. It continuously changes the activation level of behaviors through selection procedure is adopted for flexibility [3]. The behavior network used in this paper has 26 states, 6 goals, 15 behavior nodes. Table 1 illustrates the basic behaviors of the synthetic character and figure 3 shows a part of our behavior network.

3 A Working Scenario

To develop the agent prototype, we have used the Microsoft Pocket PC phone edition SDK 2003. It provides a good implementation of real smartphone device containing PIMS, voice conversation, text message service, web browser, and

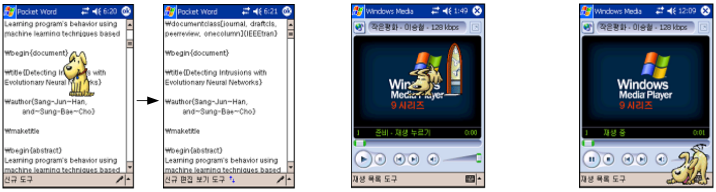
Table 1. Basic behaviors of character.

Behavior	Description
Disappear	Disappears from the screen
Appear	Appears on the screen
Move	Moves to an appropriate position
Sleep	Lies down and sleeps
Turns over	Turns over while sleeping
Wake up	Wake up from sleeping
Get fat	Gets fat (the size of character increases)
Get slim	Gets slim (the size of character decreases)
Change facial expression	Express affect through facial expression
Notice something	Raises a notice board and informs something
Take a rest	Sits down on a chair and takes a rest
Dance	Dances shaking its tail
Vibrate	Vibrates for a while (use vibration motor)
Move to center	Moves to the center of the screen
Happy	Smiles and gets delighted

audio/video player. We have observed the agent's behavior according to user's input in order to demonstrate the feasibility of the agent.

User executes a word processor to read an e-book. Then the screen is filled up with many characters. The value of low-level state, 'complex screen', becomes true and 'do not disturb' is set as the agent's goal. 'disappear' behavior gets additional activation because all of its precondition: 'complex screen', 'awake', and 'on the screen' is true and current goal: 'do not disturb' is in the add list of 'disappear'. As time goes by, the global threshold gets lowered and the activation value of 'disappear' gets higher. After several iteration of activation update procedure, 'disappear' behavior is selected and character is disappeared from the screen as shown in Figure 4(a). User can read the e-book comfortably without manual turning off.

User plays music after reading an e-book. Then, 'complex screen' state becomes false and 'playing sound' becomes true. The motivation system set agent's goal as 'expressing device's state' because 'playing sound' changes from false to true. At first, there are no behaviors of which all preconditions are met because almost behaviors has 'on the screen' as precondition, but they give activation to the 'appear' behavior because 'on the screen' is in the add list of 'appear'. Consequently, 'appear' is selected and the character enters the screen. After that, many behaviors start to compete and the 'dance' behavior has highest activation because all of its preconditions such as 'playing sound', 'awake' and 'on the screen' are met and current goal: 'expressing device's state' is in its add-list. Then, the character begins to dance and entertains the user who is listening to music as shown in Fig. 4(c). Figure 5 shows the agents behavior according to the changes of the low-level and high-level states.



(a) 'disappear' behavior (b) 'appear' behavior (c) 'dance' behavior

Fig. 4. Agent's behavior.

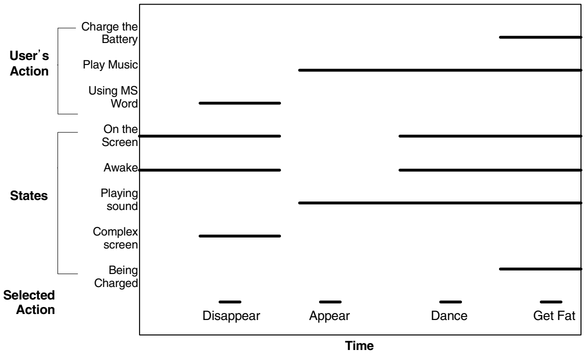


Fig. 5. Changes of states and goals according to user's behavior.

4 Conclusion

In this paper, we proposed a synthetic character as a user assistant and an entertainer. It collects low-level information from various information sources available in smartphone. They are used as clues of high-level information which is inferred by Bayesian network and the basis of determining agent's goal. High/low-level states and agent's goal are inputted into behavior selection mechanism in order to take a behavior appropriate to current situation.

Acknowledgement

This research was performed for the Ubiquitous Computing Research Program funded by the Ministry of Information and Communication of Korea

References

1. E. Charniak, "Bayesian Networks without Tears," *AI Magazine,* vol. 12, no. 4, pp. 50-63, 1991.
2. R. Picard, "Affective Computing," *Media Laboratory Perceptual Computing TR 321,* MIT Media Laboratory, 1995
3. P. Maes, "How to Do the Right Thing," *Connection Science Journal,* vol. 1, no. 3, pp. 291-323, 1989.