Generating Dialogue

Naoyuki OKADA, Jun-ichi NAKAMURA, Tsukasa NOMA and Masato TOKUHISA
Department of Artificial Intelligence
Kyushu Institute of Technology
Iizuka 820, Japan
e-mail: okada@pluto.ai.kyutech.ac.jp

1 Introduction

The history of Artificial Intelligence(AI) has been iterating diversification and specialization in its research area,and each sub-area has made great advances by efforts of many researchers. To make the current level of AI techniques much closer to the human intelligent processing, integrated approaches to AI are inevitable as a reaction to the diversification and specialization. We are promoting a project named AESOPWORLD, by which we try to make a computer simulation of a human-like agent with intellect and emotion at a lower-grade of elementary school. The AESOPWORLD project aims at the following integrations:

(1) Language association with mind Non-linguistic mental activities are associated with the deep structures of language.

(2) Symbol grounding in perception or motion Abstract symbols for representation of mental activities ground in concrete analogue signals for control of sensors or actuators.

(3) Integration of Multi-media Vision, speech, and even motion are integrated through mind.

(4) Integration of intellect and emotion Subjective knowledge processing is connected with objective emotion arousal.

A human-like agent, or "k-agent," simulating a protagonist of an Aesop fable was constructed: He has a desire,makes plans to satisfy it,takes action to execute them, recognizes his surroundings,gets emotional and utters those processes in natural language.

2 Purpose, Progress, Features of the Research

This research work was motivated by term (1) in Section 1, aiming at dialogue generation deeply based on the agent's mental activities. The research plan was made as follows:

1) Modeling language generation from mental activities,

2) Implementation of another agent, or "p-agent,"

3) Modeling dialogue structures,

4) Dialogue simulation between k- and p-agents.

The annual subjects were scheduled as follows:

- 1st year Execution of 1)

  The mind of k-agent is composed of a lot of units or "modules" like Minsky's agent. Note that our "agent" is different from that of Minsky. The key idea of the modeling is that deep case-frames are embedded in some modules to produce deep structures when they are activated. Thus, their non-linguistic activities are described by the deep structures.

- 2nd year Execution of 2) and 3)

  Another agent starts with the copy of k-agent. Then,its knowledge data about everyday experiences or parameters of emotion arousals are changed to form another personal. Next, "invitation" is selected as a topic of dialogue. Its discourse structure is modeled by two semantic networks: one is for the inviting side and the other for the invited side. Each network shows dependencies among dialogue units on one side.

- 3rd year Execution of 4)

  First,knowledge about invitation is investigated from both sides. Then, the two kinds of networks are implemented on k- and p- agents with the investigated pieces of knowledge. Thus, a dialogue simulation is performed between the two agents.

This dialogue generation has the following features:

- Dialogue structures are modeled by semantic networks.
- The arrows of the network don't show transition information from unit to unit, but "dependencies" among dialogue units. Transitions are considered as "performance" of dialogue,and controlled by other information.
- Both speakers, or inviting and invited sides, have not only one's own network but the other's to guess the other's mind.
- The dialogue deeply depends on the non-linguistic mental activities.

3 AESOPWORLD System

This research work is supported by the AESOPWORD system constructed previously. The system has two big components: k-agent and an environment. The mind of k-agent comprises seven domains, i.e. instinct, emotion, recognition, planning, expression, memory, and language. To the mind are added two domains, perception and action as sensor and actuator, respectively.

The architecture of the agent follows the so-called multi-agent model. Each domain contains a lot of modules each of which has a simple function. When a module is activated by some stimulus in some domain, it activates its "adjacent" modules. Then, the adjacent modules activate their adjacent ones. And thus, module activations propagate inside or outside of the domain. This is called "chain activation" of modules. Mental activities such as recognition or planning are captured by this chain activation.

Seven domains of the k-agent are implemented on workstations. Unfortunately, our techniques of sensor and actuator are currently too poor for the k-agent to move around in a real environment,and so the second component,or environment,supports the functions of sensor and actuator, and provides an environment for k-agent's moving around on computer graphics.

4 Results of the Research

The AESOPWORLD system has been augmented to two human-like agents. They are able to have dialogue about invitation.

4.1 Modeling of Language Generation

Modules are classified from several points of view:

- Behavior Some are static, e.g. concepts of substance, whereas others are dynamic, e.g. planning.
- Concept formation Some are formed concepts, and others are tentative knowledge or episodic data.
- Intellect and emotion Some are objective knowledge about the real world, and others are subjective emotions.

We pay attention to dynamic modules, which are implemented with a frame-structure. When they are activated, they begin to run by stored programs and produce some output as the result. We added a slot to the frame-structure, and embedded two kinds of deep case-frames or predicates: one describes the behavior and the other the result. The domain of language watches activations of dynamic modules, and picks up the predicates to describe the activation of the module.

4.2 Construction of P-agent

Implementation of p-agent starts with the copy of k-agent. Its basic architecture and high-level concept data were kept as they were. But low-level knowledge and episodic data were changed. In particular,episodic data is important, because both agents often make example-based judgements. Furthermore, many parameters for computation,e.g. evaluation of planning or acceptance of emotion stimuli, were changed. Thus, two human-like agents with different personal were constructed.

4.3 Modeling Dialogue

If p-agent want to invite a plan to k-agent, he has to "introduce" it to k-agent first of all. Then, he has to try "detailed explanation" or "strong persuasion" according to the p-agent's positive or negative response, respectively. Here are three dialogue units on the inviting side: plan introduction, detailed explanation and strong persuasion. The latter two depend on the first one according to the attitude of k- agent. On the other hand, there are dialogue units on the invited side : listening-to-plan, detailed comprehension and plan rejection. These structures show the framework of the dialogue about invitation on both sides.

Each unit has a detailed structure, or sub-network. For example, the strong persuasion contains "analysis of rejection" and several strategies for persuasion, e.g."emphasis of the merits of the plan" or "compromise of the plan to the other side." Finally, the unit of sub-network performs several kinds of functions :

- Communications with the domain of language for linguistic comprehension and generation of utterances,

- Interpretation of linguistically analyzed utterances,
- Production of responses or reactions to the input utterances,
- Communications with related modules,e.g.planning,recognition expres sion, or knowledge control.

4.4 Implementation

The main parts of the results obtained in 4.1 through 4.3 were implemented on the original AESOPWORLD system. The following is a rough scenario for the experiment: The k- and p-agents are fox and raccoon,respectively, and live in some natural environment. The k-agent is very thirsty, and is going to a mansion to drink water in the pond. On the way,he happens to meet with p-agent at a cross. He is invited to come to a river to drink water with p-agent, and persuaded with several invitation strategies. It is suggested that the mansion may be crowded with people and dangerous now. So, he changes his mind, and agrees to come to the river.

The two agents move around in a natural environment on a graphic display with speech and music generators. They make the dialogue when they meet at the cross. The mental states of both sides can be uttered in natural language as monologue. In particular, the emotional states are automatically converted to background music of the screens.

5 Conclusions

This research work was motivated by "language association with mind" in AESOPWORLD project. It aims at dialogue generation which is deeply based on the mental activities of human-like agents, and has obtained the following results:

1) Modeling language generation from mental activities Two kinds of predicates were embedded in the case-frame slot of dynam- ic modules, and the language domain watches their activations.
2) Implementation of p-agent The copy of k-agent was improved to have another personal.
3) Modeling dialogue structures The structure of invitation was represented by the two semantic net- works on inviting and invited sides.
4) Dialogue simulation between k- and p-agents. Experiments were done in the world of Aesop fables.

In conclusions, this augmented AESOPWORLD system could be expected to be a good workbench for the development of mental-activities-based dialogue systems.

Main Publications

[1] Okada,N.:Integrating Vision, Motion, and Language through Mind, AI Review(to appear).

[2] Okada,N.: Semantique du Langage Naturel—-Vers une Approche Tech- nologique, Traitment Automatique(to appear).

[3] Noma,T. et al.:Translating from Natural Language Story to Computer Animation, Proc. SPICIS '92, pp.475-480(1992).

[4] Nakamura,J. et al. : Automatic Back-ground-music Generation Based on Actor's Mood and Motions,The Journal of Visualization and Comp- uter Animation, Vol.5, No.4, pp.247-264(1994).

Keywords: dialogue generation, dialogue structure, mental activity, invitation