1998 International Computer Symposium
Workshop on Computer Graphics and Virtual Reality
December 17-19, 1998, N.C.K.U., Tainan, Taiwan, R.O.C.

# Action Hierarchy with Interactive Motions

*Shiao-Ting Sun, Yu-Wei Jeng, Min-Lung Lee and Jia-Sheng Heh*
*Dept. of Info and Computer Eng., Chung Yuan Christian University*
*E-mail.* jsheh@ice.cycu.edu.tw

**Abstract**

The motion architecture of agents with simulated behaviors and interactive motions is proposed in this paper. Through such architecture, agents not only can play their own roles autonomously but also can act in response to others' motion.

For the purpose of analyzing the behaviors and interactive motions of the human beings, an *action hierarchy* is developed. An action hierarchy of an agent is composed of five layers: script units, action units, basic motions, body motions and coordinate units. Through such hierarchy, the motion of a single agent can be completed. For representing the interaction of agents, an *interactive motion model* is also developed. As the interaction between human beings, agents can accept commands from other agents and give appropriate response.

An experiment system is presented with a script that a father agent and a child agent go to school together. Similar action hierarchies are built within both agents and interactive motion model provides their communication. With such interaction, when the father agent walks quicker, the child agent will get aware and walk faster.

**Keywords:** *action hierarchy, virtual agent, motion model, interactive motion*

## 1.Introduction

Although there are cartoon and animated characters, most of the actors and actresses in movies are real human beings. By the progress of computer technology, *virtual movie* generation is not a dream any more [1]. It is possible to design a virtual human cast of a movie in a computer. To act as a player in a virtual movie, a *virtual agent* has to possess abundant facial expressions and vivid motions. Moreover, human can behave or response by receiving messages from outside world. This agent has also to accept commands and communicate with other (agents or human beings). These messages may include symbolic language or body language.

The facial expression, behavior, emotion, and gesture of an agent can be produced through computer graphics. Many researchers make effect on analyzing the characteristics of basic facial expressions to recognize human expression [2][3][4]. On the other hand, there are some researches, such as the agent architecture Tok [5], simulating the behavior, emotion, and motion of human beings. Also, human gesture interaction, a kind of body language, can take advantage of the natural and intuitive communication mode [6].

For human behavior, Shallice (1978) has postulated the existence of what he calls the *action system.* Action system is one component of human cognitive processing. The result from the action system is outputs to motor system [7]. In the adaptation of Shallice's action system, Martindale (1980) proposed a four-layer structure: action units, script units, disposition units, and subself units. The *action unit* is the top layer, which contains every basic act performing by a person. For example, "go to school", "have breakfast", "open book", and "kiss" all belong to the action units. At lower levels of the action system there are nodes controlling more and more general types of actions. The second level includes nodes coding what Schank and Abelson (1977) call *scripts*, such as "locomotion from home to school", "restaurant going", and "study" belong to the script units. Another lower two layers, *disposition* and *subself*, are beyond our investigation.

A tree structure, called *identity hierarchy*, is proposed in the study of human behavior [8]. The notion of *identity level* is represented clearly in the relational terms people use to signify connections among identities (Danto, 1963; Goldman, 1970). Using *by*-relationship as a tool for uncovering the organization of a set of identities, these identities are arranged hierarchically.

Stretched from Martindale's action unit and script unit of four-layer structure, an *action hierarchy* is developed. With the script unit and the action unit kept, three more layers, basic motions, body motions and coordinate units, are proposed to describe the detailed motions.

This paper uses action hierarchy as the motion architecture of agents to simulate human behaviors and interactive motions. A formal description of action hierarchy will be introduced in Section 2. Section 3 utilizes action hierarchy to analyze the interactive motions among agents. The design of such motion system is investigated in Section 4. Section 5 implements an experiment system of two agents: father agent and child agent. Section 6 is a conclusion.

## 2. Action Hierarchy

For the purpose of analyzing these behaviors and interactive motions, an *action hierarchy* is developed. The action hierarchy is composed of five layers: script units, action units, basic motions, body motions and coordinate units. For the analysis of agent motion, a three-dimensional *coordinate system*, *global* (GCS), *body* (BCS), and *floor* (FCS) [9], is adopted as agents' fundamental motion model. 15 characteristic points are chosen to control the motions of an agent. By the setting of characteristic points, agents can present different behaviors and interactive motions.

In real world, every one lives in different place (scene) at different time. The whole life of human contains many scenes. In every scene, the behaviors of a person are limited. And any kind of behavior can be decomposed to many detailed motions. In the viewpoint of virtual person, it is the same with human. The motion system can be divided into five layers. The first layer in motion system is called *script unit*, which includes personages, a place and all the behaviors that can be done in a scene. For example, "A student sits in the restaurant and eats the breakfast". The personage of the example is the student, the place is the park and the behaviors are 'sit' and 'eat'.

An agent acts by following a *script* in a scene. A *script* in a specific scene can be divided into several *actions* in the second layer, such as a student "gets up" in a morning, "eat breakfast", and then "goes to school". Therefore, the second layer we called is *action unit*. The third layer includes many *basic motions* decomposed from the action units. For example, the action of "getting up from the bed" can be decomposed to several basic motions: "lie", "sit on the bed", and "stand up". The *action unit* includes more than one *basic motion*. Then we have *body motions* as the fourth layer. For example, the basic motion "lie" can be decomposed to several body motions: "left hand lie", "right hand lie", "left leg lie", "right leg lie" and "head lie". Finally, the fifth layer, *coordinate unit*, records all coordinates of body motions. These coordinates depict the state of the characteristic points of a virtual person.

These five layers in motion system constitute a tree-like structure, called *action hierarchy*, as shown in Figure 1. The heavy arrows, dotted arrows and light arrows represent script-action, action-basic motion and basic motion-body motion relationships, respectively. The body motion-coordinate relationships are not shown in the figure. The former two arrows (script-action and action-basic motion) are cognitive relationships and are usually accomplished in the episodic and semantic memory of human brain [7]. The latter two relationships (basic motion-body motion and body-motion-coordinate) involve human motor system and are always executed through our muscles.

In a arrow of Figure 1, its upper (or lower, respectively) unit is the *predecessor* (or *successor*, respectively) of the corresponding relationship. A predecessor may have more than one successor and conversely, a successor may possess more than one predecessor. Therefore, action hierarchy is not a tree but a lattice. The top-down processing of action hierarchy may activate successive actions. On the other hand, given a series of successor actions, their upper predecessor may turn on, then make motion recognition.

### 3. Analysis of Agent Motion

Some properties can be found out in action hierarchy. For a predecessor, many of its successors may be processed at the same time. This makes *action synchronization*. On the contrary, the successor motions usually need a (non-unique) execution order. The *sequencing* of motions can be indicated by a set of attached values on the corresponding arrows. The successor motion with smaller value will be executed earlier. When the values of successors in one level are the same, it means these successors are synchronized.

Every agent has its (or his/her) own action hierarchy. When one agent, says agent A, intends to present some motion, it makes top-down processing and the corresponding coordinates for this motion is transmitted to other agents. This is just a *single agent motion*. Each single agent motion can accept two kinds of commands, one is outside command and the other is inside. *Outside command* is a set of actions, which belong to one level of the hierarchy and sent by the user. Therefore the user can command the agent in his (her) way, such as "right left hand" in *body motion*, "stand up" in *basic motion*, "get up" in *action unit*, etc. *Inside command* is the command in the interactive motion model, shown as Figure 2. Suppose that there is more than one agent in the action system. As another agent, says agent B, senses the coordinates (or coordinate changes) of agent A from *the sensor*, and *the sensor* will pass these into its (agent B's) own bottom-up processing action system, then interprets these motions. This makes an *interactive motion of agents*. Note that each agent interprets coordinate through its own action system, so misunderstanding is possible.

**4. Designing Motion System**

The motion system of an agent includes two parts: the sensor and the motor system. The *sensor* accept the coordinates of the coordinate systems in other agents, and then goes on the bottom-up processing in the action recognition system. On the other hand, the *motor system* possesses just the above-mentioned action hierarchy, which recognizes the corresponding inside commands (scripts, actions, basic motions or body motions) sent from the sensor or accepts the outside commands assigned by the user, and then process top-down to form the corresponding motions. All these processing modules are shown in Figure 3.

In the interactive motion, agents interact with each other by the sensors. For example, there are two agents: agent A and agent B. The sensor in agent A can detect the change of coordinate when agent B takes actions. According to the change of coordinate, the sensor A decides the response which agent A needs to do. Then it sends the inside command to the motor system of agent A. The motor system A will send inside commands to the suitable level, then agent a will response accordingly.

**5.Experiment System**

An experimental system is built up on IBM-Compatible Personal Computer. An agent with an action system may receive messages from a Control Panel or other agents. Commands such as "get up" or "sit down" can be set on Control Panel. The agent will accept these commands and act accordingly. Lower level actions like "stretch out your right hand" is allowable also. If the message comes from another agent, they may have interactive motion. And the agent can communicate with another through KQML (Knowledge Query and Manipulation Language), which is a standard language designed for agents.

For example, a father agent and a child agent go to school together. Both of the their script units are "Father and child go to school together". The father agent's action units is "go to school" and the child agent's is "follow". Both of their basic motions are "walking". These two agents have KQML communication, which means they know each other's position. Since the steps of the child agent are smaller, the distance between the father agent and the child agent will become larger. When the child agent become aware of larger distance, he will speed up. The experiment result is shown in Figure 4.

The virtual person in our system is 2-dimensional originally, but now we try to display it in 3 dimension. This virtual person is quite different from the traditional agent. It can communicate with other agents, act like humans, and present as human body. In the future, this agent will be translated from a passive agent to a smart agent. The ability of learning will be designed by neural network. And the human motions and expressions can be learned much more easily and automatically.

## 6.Conculsion

It's our dream to realize the virtual person in the virtual world. In the motion system, the initiative step was realized. The motion system gives agents the ability of acting, moving, and being aware of another agent in the same environment.

The learning ability of agent will be developed in the future. The learning mechanism can be designed by neural network. And the human motions and expressions can be learned much more easily and automatically.

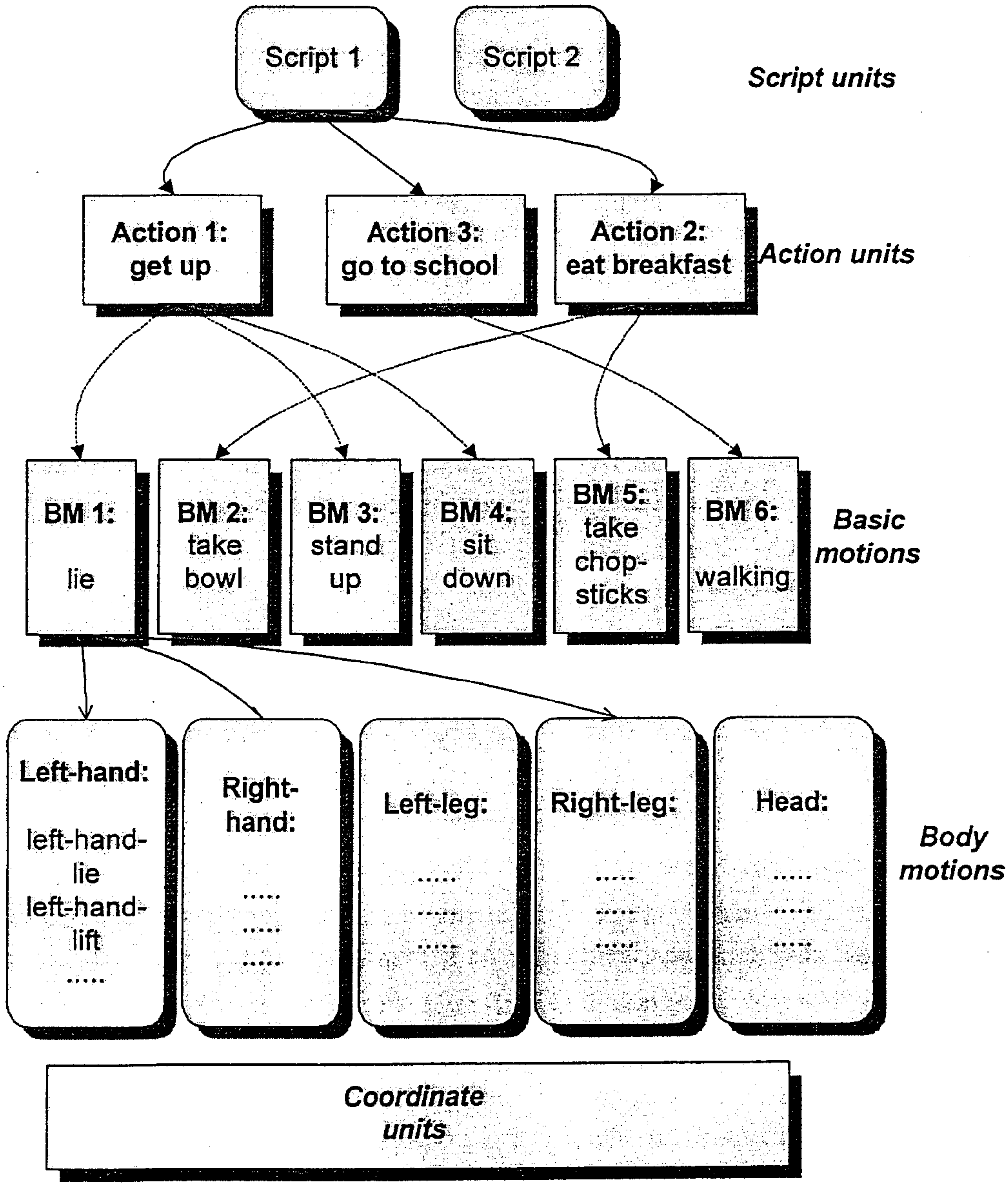


Figure 1 An action hierarchy

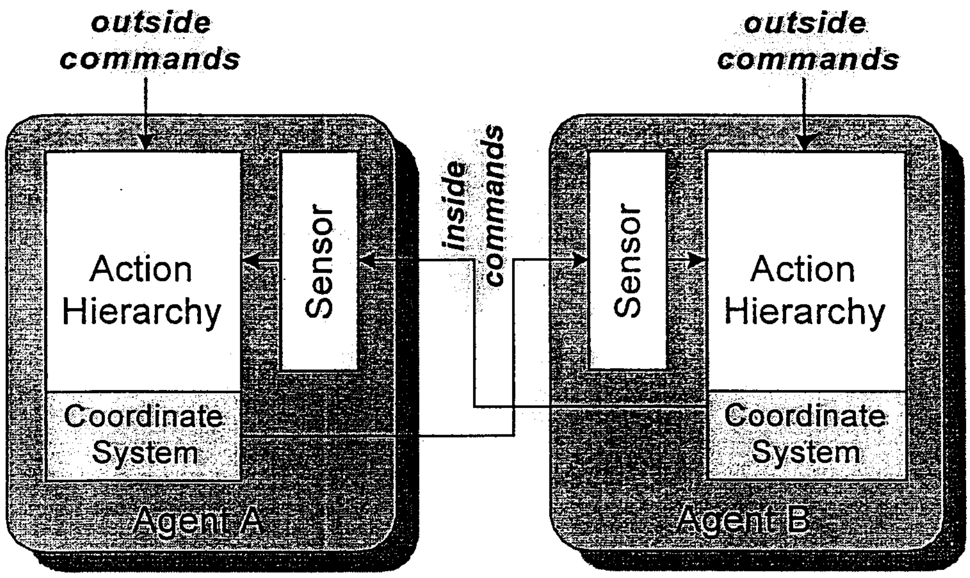


Figure 2 Interactive Motion Model

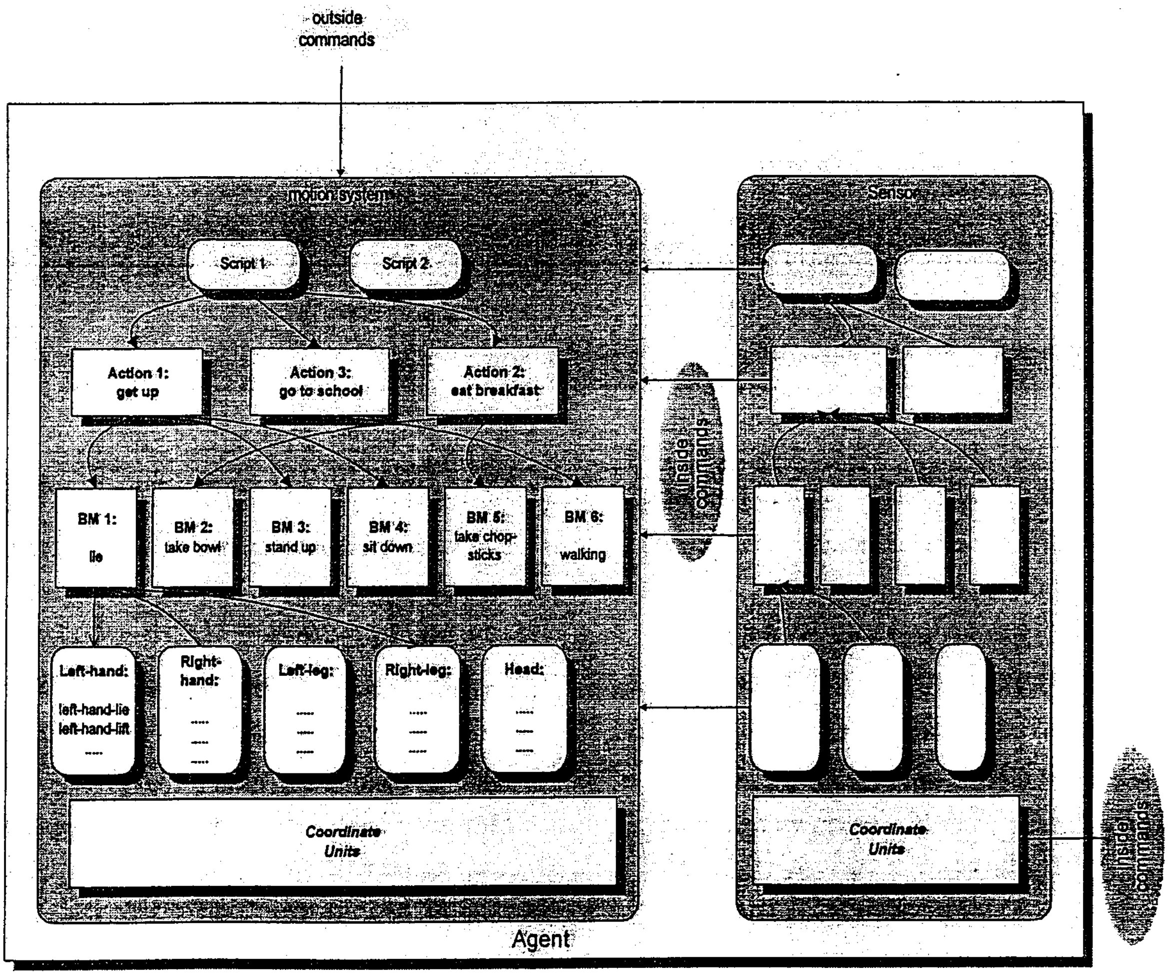


Figure 3 Motion system of an agent

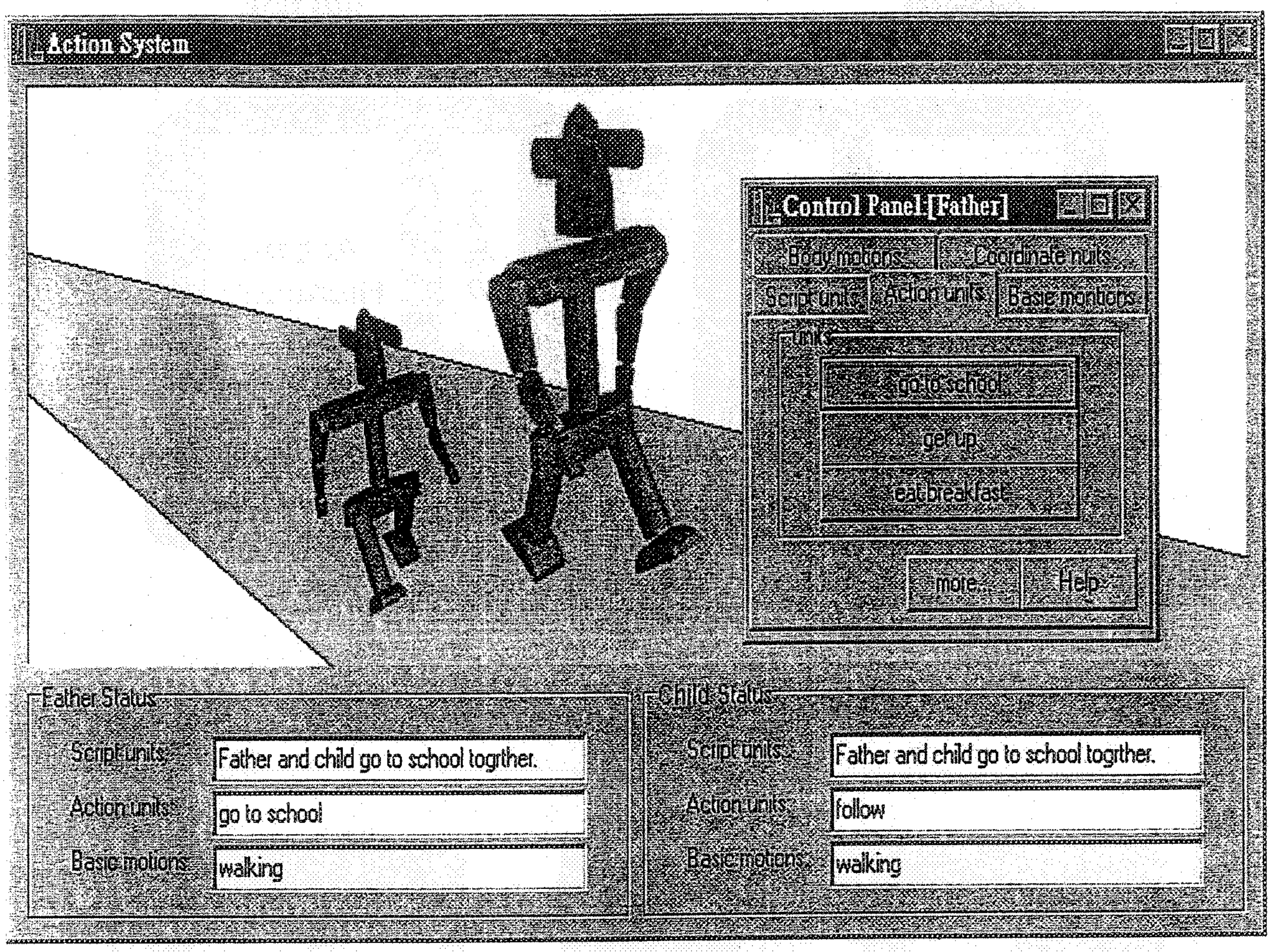


Figure 4 The experimental system

References

[1] Yasuo Matsuyama , "Multiple Descent Cost Competition: Restorable Self-Organization and Multimedia Information Processing", *IEEE Transactions on Neural Networks*, Vol.9, No.1, January 1998, pp.160-122

[2] Hiroshi KOBAYASHI and Fumio Hara, "Analysis of the Neural Network Recognition Characteristics of 6 Basic Facial Expressions", *JSME International Journal*, Series C, Vol.39, No.2, 1996, pp.323-331

[3] Kin Choong Yow and Roberto Cipolla, "Feature-based human face detection", *Image and Vision Computing 15*, 1997, pp.713-735

[4] Mark Rosenblum, Yaser Yacoob, and Larry S. Davis, "Human Expression Recognition form Motion Using a Radial Basis Function Network Architecture", *IEEE Transactions on Neural Networks,* Vol.7, No.5, September 1996, pp.1121-1138

[5] Joseph Bates, A. Bryan Loyall, and W. Scott Reilly, "An Architecture for Action, Emotion, and Social Behavior", *Artificial Social Systems, 4th European Workshop on MAAMAW'92*, July 29-31, 1992, pp.55-68

[6] Francis K.H. Quek, "Unencumbered Gestural Interaction", *IEEE Multimedia*, Winter 1996, Vol.4, Num:3, pp.36-47

[7] Colin Martindale, *Cognitive Psychology: A*