Curtis Roads

The Tsukuba Musical Robot

The 1985 International Exposition inaugurated the Science City of Tsukuba, Japan as a world technological center. Some 47 foreign countries and 28 Japanese corporations participated in the fair (Japan Association for the International Exposition 1985). Visitors to the Japanese theme pavilion saw an astounding technological demonstration. Musical scores could be placed in front of a talking robot seated at a Yamaha digital organ. The 90-kg robot, equipped with a video camera "eye," studied the score for several seconds. Then it launched into a performance of the score using its ten fingers and two feet (Fig. 1). When a performance was complete, the robot could take verbal requests for songs from the audience. It could also play along with a singer, tracking the singer's pitch and tempo and adjusting its own keyboard performance to match the singer.

The robot was designed to play an organ in much the same way as human beings do. Equipped with arms and legs shaped like a human being's, the robot had a total of 50 joints. Its fingers simulated those of a live musician. Capable of striking keys at a rate of 15 times per second, the robot could play a variety of music from simple Bach chorales to popular music.

At Tsukuba, the body of the robot was set up on stage in the middle of an exhibition room, while the computer control systems were installed in a separate room 20 m away. Thirteen Sumitomo Electric fibreoptic data links between the robot and its control systems transmitted score image data, performance data, and control signals. The robot control system consisted of 17 Z8001 and Z8002 16-bit computers and 50 Z8094 8-bit computers interconnected by means of fibreoptic data links. The maximum transmission capacity of the links was 32 Mbytes/sec.

The prototype robot, called WABOT-2 (short for Waseda Robot), was developed at Waseda University, Tokyo (Mu Research Group 1985). Research and development efforts at Waseda University were broken into several tasks: the vision system (team led by Professor Sadamu Ohteru), the conversation system (team led by Professor Katsuhiko Shirai), a singing voice tracking system (team led by Professor Seinosuke Narita), and limb control (team led by Professor Ichiro Kato). WABOT-2 was completed in 1984.

A second version of the robot, similar in appearance but with increased speed and reliability was built by a team of engineers at Sumitomo Electric, Ltd. This second version was exhibited in 1985 at Tsukuba.

The Vision System

The robot recognizes all symbols in a score in 10 to 15 sec. Figure 2 shows the symbols it recognizes. During its score reading phase it also plans all its subsequent performance movements. Two seconds after it finishes reading the score it has memorized the score and begins playing. Thus, the robot does not sightread as it plays.

The robot's camera eye reads the score from 1 m away. The charge-coupled device (CCD) camera and frame buffer have a resolution of 2000 by 3000 pixels (6 million pixels total).

In order to sort out the fine lines of musical staff notation, the robot's eye renders each frame into a 64-by-64-line grid. A powerful light illuminates the score, but no lamp can be assumed to produce uniform illumination. Thus, at all times the robot's vision system controls a parameter representing the lightness of each square in the grid. This system compensates for stray light and illumination irregularities.

A special computer subsystem was developed in order to recognize note heads. The robot can recognize note heads that vary in size from 8-by-8 pixels to 16-by-16 pixels. Eight standard note-head patterns are stored to accommodate the reading of various shapes of note heads. A typical score recognized and played by the system is shown in Fig. 3.

Computer Music Journal, Vol. 10, No. 2, Summer 1986,

Fig. 1. The Tsukuba musical robot.

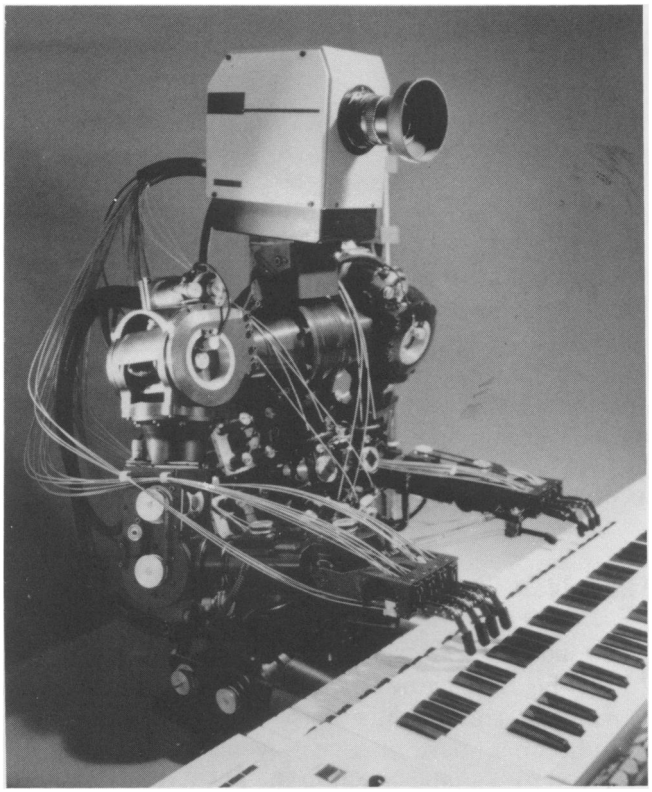
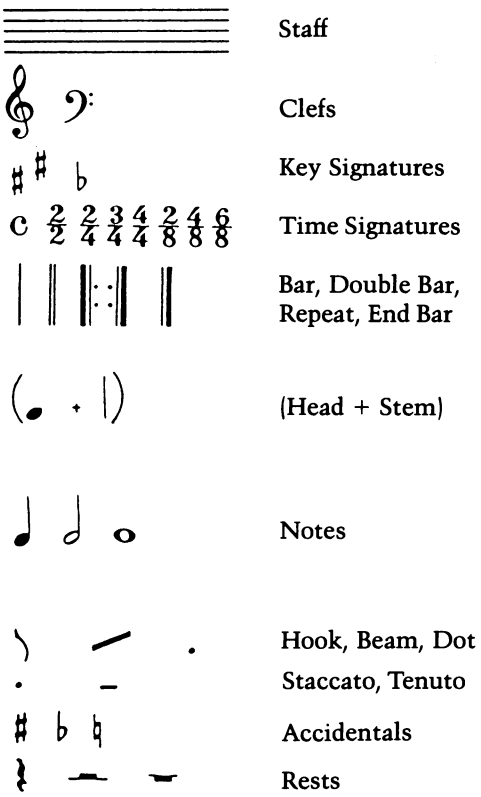


Fig. 2. Music symbols recognized by the robot.



Conversation System

The robot incorporates a speech synthesis and analysis subsystem for conversation with human beings. When a human being speaks (in Japanese) to the robot, its phoneme-based speech understanding system performs linear predictive and cepstrum analysis to determine the spectrum and pitch of the spoken utterance. The robot understands a limited number of sentences. The robot can respond with speech of its own, since it has a vocabulary of 180 utterances, mostly having to do with handling requests from the audience. A conversation between the robot and a human being is shown in Fig. 4.

Singing Voice Tracking System

The robot is designed to track a human singer who sings into a microphone connected to the system. The robot plays the organ along with the human, with tempo regulated in part by the vocalist. For pitch analysis, a system of five narrowly tuned bandpass filters is used to derive a fundamental frequency every 30 msec. If the singer is out of tune, the organ player can adjust the tuning of the organ for a more euphonious ensemble effect. Ironically, the robot musician cannot hear its own performance.

Limb Control System

According to the designers, the aim of the limb control system was to realize dexterity and speed in the robot's gestures. The design of the limbs was anthropomorphic, i.e., modeled after a human being's two hands, five fingers, and two feet.

The fingers of the robot consist of a two-joint thumb and four three-joint fingers (Fig. 5). The hands can move across the keyboard at a speed of 1.5 m/sec. The legs have four joints each. The left

Fig. 3. Fox and Goose, *a typical score recognized and played by the robot.*

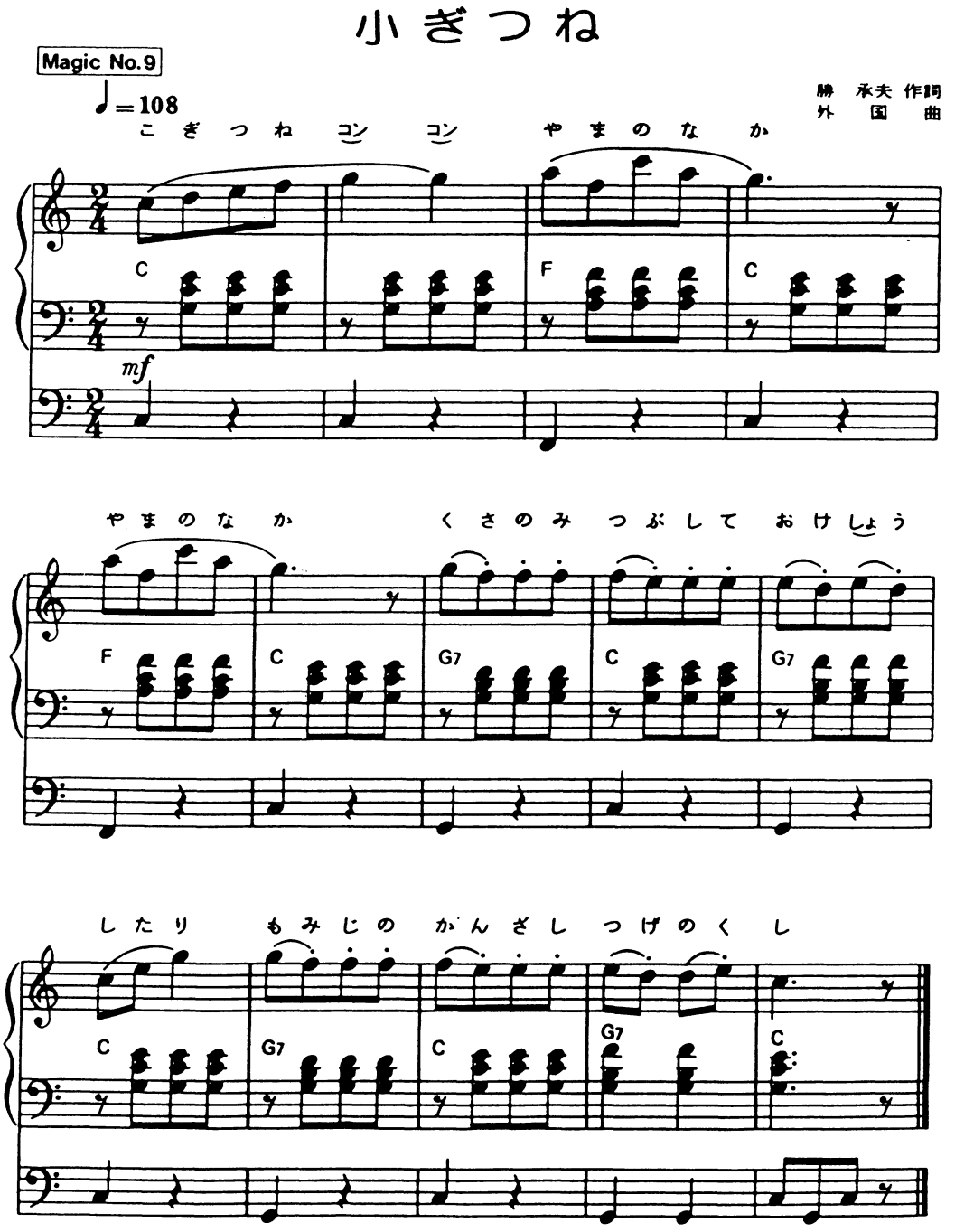


Fig. 4. Transcript of conversation between a man and the robot.

Fig. 5. Posture of the robot's hand showing movement of the individual fingers on the keyboard.

Man:	Kon'nichiwa. (Hello.)
Robot:	Kon'nichiwa. Boku wa *musician robot* desu. Gakufu wo yonde denshi piano wo enso shitari, uta ni awasete banso shitari dekirundesuyo. (Hello. I am a musician robot. I can read music scores, play the electric piano, and accompany the song.) Nanika hiitemimashoka? (Shall I play for you?)
Man:	Onegaishimasu. (Yes, please.)
Robot:	Soredewa "Tsuioku" wo kiitekudasai. (I will play *The Way We Were.*) —Plays—
Robot:	Ikagadeshitaka? (How was that?) Hokani nanika shimashouka? (Shall I play something else?)
Man:	Banso wo shitekudasai. (Accompany me, please.)
Robot:	Nani wo utaimasuka? (What do you want to sing?)
Man:	Nani ga hikeruno? (What can you play?)
Robot:	*Zosan* . . . sorekara *Quma no oyako* ga hikemasu. (I can play *Elephant Family*, . . . and *Mother Horse and Her Foal.*) Soredewa kyokumei wo itte kudasai. (Which do you want?)
Man:	*Zosan* wo onegaishimasu. (*Elephant Family*, please.)
Robot:	Soredewa *Zosan* wo hikimasu node piano ni awasete utatte kudasai. (I will play *Elephant Family.* Will you sing along with me?) —Plays—
Robot:	Taihen yokatta desuyo. (You sang very well.) Hokani nanika shimashouka? (Shall I play something else?)
Man:	Iie. Arigatou gozaimashita. (No, thank you.)
Robot:	Douitashimashite, Soredewa sayonara. (You're welcome, Good-bye.)

Fig. 5

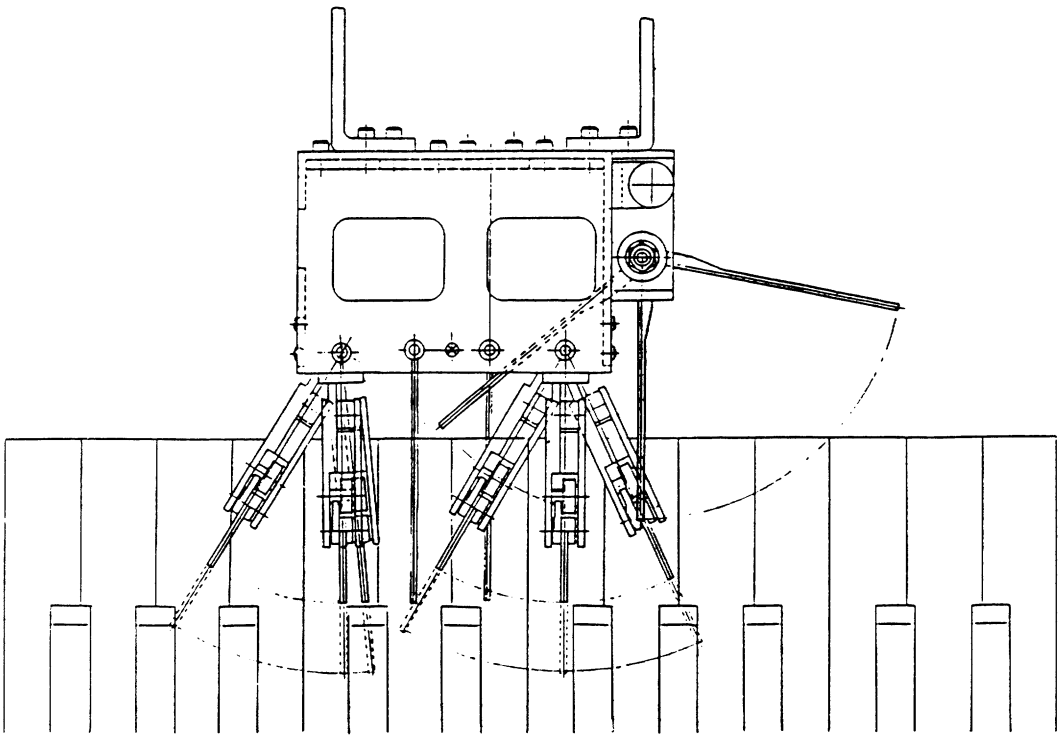
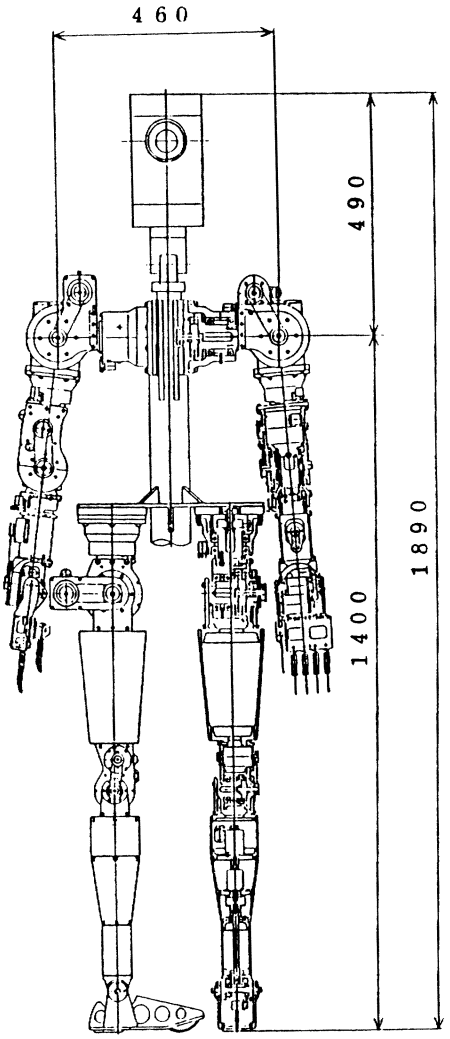


Fig. 6. Construction of the robot's limbs.



foot of the robot plays the pedal bass clavier, while the right foot plays the "expression pedal" or amplitude control. Each of the robot's arms has seven degrees of freedom, deriving from its shoulder, elbow, and wrist movements (Fig. 6). Overall control of each arm is controlled by a separate computer, and each degree of freedom in the arm is controlled by a separate computer. Thus, 16 microprocessors are devoted to arm movements.

All movements are planned in advance, when the score is recognized. The sequence of movements is planned by a program that carries out various searches to determine the most efficient arm and finger movements for a particular score.

Scope of the Project

The development project at Waseda University involved some 50 students over a period of 3.5 years. The development cost of the robot was approximately 500,000,000 yen, or about $2,000,000, although the actual cash spent was only a portion of this, or several hundred thousand dollars, according to Professor Ohteru. Sumitomo Electric, Nippon Gakki (Yamaha), Namco, and Nippon Electric all contributed materials and services to the project.

According to Professor Ohteru, one of the developers of WABOT-2, a plan to develop another robot with different (nonmusical) capabilities is currently in the planning stages (Ohteru 1985).

References

Japan Association for the International Exposition. 1985. *Tsukuba Expo'85: Official Guide Book.* Tsukuba: Japan Association for the International Exposition.

Mu Research Group. 1985. "Wabot-2." *Bulletin of the Science and Engineering Research Laboratory.* 112. Tokyo: Waseda University.

Ohteru, S. 1985. "WABOT-2." Lecture presented 6 August 1985 at the M.I.T. Artificial Intelligence Laboratory, Cambridge, Massachusetts.