Interface 2 (1973) pp. 159-162.

Some Programs for Real-Time Computer Synthesis and Composition

BARRY D. TRUAX

(received 13 July 1973)

1. INTRODUCTION AND SUMMARY

Programs for sound synthesis and composition that operate in real time have been developed by the author during the past year and used on the PDP-15 computer at the Institute of Sonology in Utrecht. They have been designed to operate within the 12K core memory capacity of the 18-bit wordlength machine, using between one and five of the 12-bit D-A converters, the real-time clock, two dectape units, and other peripheral devices. Because of the emphasis on exploiting the real-time potential of the machine, the user is able to work interactively and receive the sound result of his specifications and modifications very quickly.

The two major synthesis methods used are fixed waveform synthesis and frequency modulation which produces time-dependent spectra. Both methods are characterized by sampling rates that are variable within a limited range. The compositional method distributes user-defined sound "objects" as events within a delimited frequency-time field that is determined by a Poisson Distribution. The two major versions in use at present for compositional, pedagogical and theoretical purposes are POD5 and POD6, which can only be summarized here.[1]

2.1 GENERAL CHARACTERISTICS OF THE PROGRAMS

Given the case of discrete, independent sound events, the Poisson distribution is mathematically the most appropriate method of syntactically organizing such events randomly in time. Its use has the advantage of allowing one to concentrate on larger formal structures that concern entire sections, and yet have the occurrence of individual events conditioned by a parameter such as density that characterizes the whole distribution. Thus, one is encouraged to look at groups of events and ongoing tendencies from the events of one section to those of another.

Density is the major variable of the Poisson Distribution; the user specifies its initial value in sounds per second and indicates whether it will generally increase or

1 An extensive account of these programs, and the introductory version POD4 is printed by the Institute of Sonology as "The Computer Composition – Sound Synthesis Programs POD4, POD5 & POD6", *Sonological Reports No. 2,* 1973, which is invailable on request from the Institute.

decrease throughout a section. As well, the user delimits the frequency-time area by means of a mask which declares certain areas available at any given time, with zero area of a segment of a mask generating a rest. It is implied that recognition of the occurrence of an event is the most basic perceptual determinant, followed by that of the density of such occurrences in time. Both occurrence and density are organized by the Poisson distribution, which is made to fit within a general frequency area. The nature of the event is further specified by the user in terms of sound objects, which gather together certain synthesis variables into a unit describing a potential event, the occurrence of which is determined by various selection procedures assigning objects to a given frequency-time event as calculated by the Poisson method.

For each section, the user specifies the number and identification of objects available, as well as the selection principle to be used. Those available using statistical procedures are:
1. Alea: equally weighted probability
2. Ratio: variably weighted probability
3. Tendency: variably weighted probability in time

In addition, selection according to a fixed sequence may be requested. The use of these selection procedures represents a simple approach to sonological coding of larger groups of sound events, the downward progression from one statistical procedure to the other involving successive simplifications of representation.

The realization of such a distribution of events is conditioned by various performance variables which may be adjusted by the user on perceptual grounds without the need of recalculating the distribution. These include:
1. Speed of performance.
2. Object data (dependent on synthesis method: see below)
3. Time mode:
   i) Dependent mode - The next entry occurs only after the entire duration of the previous event, as well as the time delay derived from the Poisson distribution have elapsed.
   ii) Independent mode - The clock time, representing the duration between successive time co-ordinates calculated by the Poisson procedure, has precedence and the duration of the event is allowed to continue only as long as the clock is counting that duration. This mode represents a strict interpretation of the calculated time structure.

The user directs the flow of the program from the teletype by means of a choice code, each option being identified by an alphabetic character and an integer from 1 to 5. Groups of related options are linked by having the same mnemonic code letter. Other information may be requested from the program on teletype, paper tape or graphic output. The co-ordinates needed for the Poisson distribution are temporarily stored on a small dectape unit, and are retrieved and read into core memory when a performance is requested. Changes and corrections may then be made with certain options and, if recalculation is required, the co-ordinates on dectape are then re-written, prior to synthesis.

Ten compositional sections may be created in succession, each with different performance variables, selection method and Poisson characteristics. Other programs allow completed sections to be stored as permanent files and retrieved later for further work. At the end of each section, all the data used for its construction is printed on teletype for future reference. The teletype protocols generated in the use of the program provide a unique record of compositional activity, and their systematic study should provide theoretical insight into the mental processes involved in real-time composition.

3. SOUND SYNTHESIS METHODS

3.1 Fixed Waveform Synthesis (POD5)

Sounds of fixed periodicity and constant timbre may be produced by repeatedly generating a given waveform described by a stored set of numbers representing instantaneous pressure levels during one period. Although this represents a severe approximation of real sounds, it still is useful to begin with such a simplification, as one is familiar with it from its usage in most analogue and digital oscillators, and the reduction of acoustic complexity at this point allows one to concentrate on higher level processes as well, within a limited machine facility.

The present program, used in POD5, employs an array of 500 data samples describing one period, and produces frequencies within an accuracy of about a quarter tone in the range 60 to 8000 Hz. with sampling rates from 30 to 47 KHz., but usually above 40 KHz. Sixty amplitude factors are used to produce smooth attack, steady state and decay amplitude sections in the envelope of a sound. The basic waveform data may be obtained in the following ways:
1. From a library of common waveform formulas within the program.
2. Input of paper tape from other programs, including samples of live sound read onto the computer's disk and scanned to obtain a single period.
3. Harmonic construction (with up to 50 partials) based on a single stored waveform.
4. Hybrid construction from sections of other waveforms.

These methods have been kept as simple as possible to allow the user to make experiments quickly and yet accurately without a great deal of technical data. The waveform data is stored on a separate dectape unit and loaded into core memory when needed for synthesis. In the compositional sections, five different waveforms may be used at a time, but 30 sound objects may be defined by the user, each characterized by:

1. waveform number
2. envelope data describing attack, steady state and decay times.
3. amplitude modulation frequency.

Amplitude modulation of the signal and spatial positioning between two channels are additional options with this synthesis method. In the case of amplitude modulation, the frequency of the modulation may be expressed either in absolute Hertz, or else as a constant fraction of the carrier frequency. Since two sidebands are produced at a distance equal to the modulating frequency above and below the carrier, the fractional specification assures a constant timbre over the entire frequency range. As for the spatial positioning, each event may be given one of seven positions between the left and right channels according to various selection principles, such as aleatoric, time or frequency correlation. During the attack portion, the sound appears to move from the centre to the position assigned, and then back during the decay; therefore, longer sounds give the impression of movement as well as identifiable position. The basic sampling rate for the two-channel output is in the range of 20 to 27 KHz.

3.2 Frequency Modulation Synthesis (POD6)

The present program is a real-time realization of a method of time-dependent sound production developed by John Chowning of Stanford University. The basic principle of synthesis involves a carrier sine wave, frequency c, modulated by another sine wave, frequency m, amplitude d. The modulation index I is defined as the ratio d/m, and its value controls the strength of the pairs of upper and lower sidebands produced about the carrier frequency, each separated by a frequency distance equal to m. The basic spectrum generated is of the form:

$$/ c \pm n \cdot m / \quad n = 1,2,3,\ldots \text{ for nth order sidebands}$$

It should be noted that both harmonic and inharmonic spectra are produced in this way, depending on whether the c-m ratio is in a simple integer proportion (e.g. 1:1, 2:1, etc.) or else involves rational values (e.g. 5:7, 4:13, etc.). The absolute value signs signify that if the lower sidebands are negative, they are reflected back into the positive domain with a phase inversion. The amplitude of the nth order sideband is given by the nth order Bessel function $J_n(I)$, where I is the modulation index.

If I is made to vary during the course of the sound, then the amplitudes of the sidebands will also change and the resultant spectrum will be time-dependent. Large values of I produce strong high-frequency components, whereas low values result in small deviations from the basic sine wave. The correlation of the modulation index I with the amplitude envelope is crucial; the present program uses a basic pattern of growth of the modulation index from zero to the maximum value during the attack, followed by a decay of I during the steady amplitude state, and finally I remaining constant during the amplitude decay, acting as a kind of reverberation.

Again, thirty sound objects may be defined by this method, having as variables:
1. envelope values (similar to POD5)
2. c-m ratio (carrier to modulating frequency)
3. maximum modulation index

As in the other synthesis method, sixty amplitude factors are used, and the sampling rate is optimized by increasing the number of samples per period used at lower frequencies, from four or five per period at the highest to the full array size at the lowest frequencies. In this method, a permanent array of size 512 describing a sine wave is used as the basic data from which both the sound material and some parts of the real-time calculation are derived. The sampling rate is variable, but generally in the range of 13 to 17 KHz. The need for synthesis methods capable of generating dynamic spectra, such as one is accustomed to in instrumental and environmental sounds, seems evident, if the computer's potential for compositional use is to be realized. The present program represents a very simple, yet effective method by which this can be achieved.

4. CONCLUSION

Further developments in both the synthesis and compositional capability of the programs are constantly being implemented, as the experience derived in their use over the last year has suggested that they must continue to grow as our insight into the complexities of programmed musical intelligence develops. It is of great importance, then, that such programs are as accessible as possible to the user, allowing him to realize his ideas quickly, and at the same time deepen his insight into the process in which he is involved. Running them on the smaller machines, such as the PDP-15, that are now increasingly available, with ample time for interactive compositional, pedagogical or theoretical work, has proved quite successful in creating a first-hand learning situation for the user, from whose experience suggestions arise leading to further developments in the programs.

Barry D. Truax,
Dept. of Communications,
Simon Fraser University,
Burnaby 2, B.C.,
Canada