

Modelling musical cognitive processes with the use of supercomputers

Francesco Carreras
CNUCE/CNR via S.Maria 36, 56126 Pisa, Italy
e-mail: F.Carreras@cnuce.cnr.it

Abstract

The development of realistic cognitive models requires computing resources that the progress of technology and the use of parallel processing techniques made available in the last few years in many research institutions. In this paper the results of the modelling of two different musical cognitive environments are described.

Introduction

In recent years the modelling of cognitive processes has become a research goal in many institutions around the world. From one side this interest has been stimulated by the new developments in neurology, psychology, brain and cognitive sciences and by a renewed attention to the phenomena connected to music perception; e.g. see [Kru90][La92][La97][Sei91]. On the other side a tremendous progress in computer hardware technology made cheap computing power available in many research laboratories. A realistic model for the simulation of perceptive and cognitive processes requires large amounts of computer resources both in terms of memory capacity, computing power and large disk storage space. The availability of supercomputers and parallel computers with usable software development environments made it possible to design, implement and use, also without the assistance of computer specialists, of large simulation models in the musical domain. Various parallelizing tools and primitives, such as for instance Express, MPI, MPL, allow the transformation of programs written in languages like Fortran or C from a batch version into a parallel one with limited effort and skill. Application programs analysis and decomposition techniques are becoming of common use in the scientific community.

Given these developments a shift of attention was possible from the difficulties of implementation to the modelling process in itself. Large models with high memory and computing requirements, which can be decomposed and distributed over many different

processors of the same parallel system, can be designed without particular attention to the computing environment constraints. The latter have indeed become less and less restrictive during the past decade. A consequence is that the implementation of larger and more realistic models is more feasible than until a few years ago. The possibility of controlling many different consoles from one single screen allows more productive ways of conducting a project by having several applications running at the same time on different processors of the same computer. In the following paragraphs two different cognitive models that were implemented on a parallel computer will be presented, with special focus on their technological side.

Cognitive models for simulating music perception

A theory of music cognition has recently been developed by M. Leman [Le95] and a computer model was realised with the aim of verifying the basic assumptions of the theory. A central point of the theory is that the tonal relationships imbedded in the western tonal music are learnt by repeated exposition to that type of music so that internal memory representations of those relationships develop by adaptation to the musical environment. The model employs sampled real music which is filtered and processed by an auditory model, that simulates the perceptive transformations of the human auditory system. A self-organizing neural network, SOM [Ko95], is then used to extract high-level implied structural information out of these images in terms of a mapping onto a two-dimensional grid. In particular the SOM is trained with so-called context images, which are obtained by integration over a fixed time window of a sequence of contiguous stable structures called completion images and fed into the network in random order. Through a process of self-organization schemata are developed that represent tonal information in terms of a topology that represents the circle of the fifths. This approach is called "ecological" for the reason that the role of the programmer is limited to the specification of the interactions between the musical environment and a model of perceptual learning [Le96][Le97]. A second model has been designed and implemented [Ca98A,B] that makes use of some of the components used in the previous model. Here again an ecological approach is followed. This means that the data are real sampled music that no score-based or symbolic information is used and that the results develop by self-organization. The aim of this model is to extract of the harmonic information embedded in a musical piece. The classification of the possible chord configurations used for training a neural network follows a recent research proposal that is part of an original harmonic theory by L.Balsach [Ba97]. For each the 12 notes of an octave 11 so-called convergent chords are assigned according to precise criteria stated in the theory [Ba94]. In our modelling approach these chords are sampled and processed by the auditory model previously mentioned. The resulting completion images are then passed through an onset detection module [Mo97] that is able to precisely define the perceived onset of all the musical events of the musical piece. A choice is made of a representative part (40 samples of 4 ms each) of every musical event which, in

this case represents one of the 132 chords. A SOM is then trained and a map is designed that shows the responding regions for each of the training chords (being there represented by 40 points). The process of chord classification is then carried on with a piece of real music that is to be passed through the trained neural network acting as an analyser of the input vectors configurations. The final outcome is a table that contains, for each musical event, the chords that respond with a value above a fixed threshold. Each of the responding chords has a weight that represents the degree of fit with the given event. The analysis of these results can be interpreted as a harmonic analysis of the musical piece.

The use of parallel computers in the simulation of large cognitive models

The first model mentioned above was gradually developed over several years. The first version was designed and implemented by M.Leman on a four T800 transputer system using 3L-C and then Express [Le89]. The hardware limitations of the system allowed only small scale simulations. In 1994 the model was ported on a nCUBE2 parallel system hosted at CNUCE, Pisa. Different configurations were tested and monitored with the aim of finding the best trade-off between computing and communication load. A final configuration of eight 16 Mbyte processors was selected to perform a simulation with a 20 by 20 grid and 2000 vectors of 56 components each, for a total of 360000 vectors because the simulation process was run for several epochs. The execution required almost 24 hours and the balance between communication and execution time was 51,5% and 48,5% respectively. The third set of simulations was then performed on an IBM SP2 parallel system of CNUCE. The model was completely rewritten by F.Carreras and M.Leman [Ca96] in order to exploit the memory capacity and the power of the new system. The new program was parallelized with MPL and was designed so as to be run in different processor and neural grid configurations. The goal this time was ambitious. The simulation was to use as input data the sampled complete first book of the Das Wohltemperierte Klavier by J.S.Bach. The images (56 data points each) employed were over 380000 to be randomly used for several epochs. Also the employed grid was now larger to obtain a finer resolution. A 30 by 30 and a 100 by 100 grid were used. The idea at the base of this implementation was to assign each processor part of the grid. This would then communicate the value of the most responding neuron, along with its coordinates, to all other nodes. The absolute most responding neuron was then chosen and each node operated the due adaptation for the neurons that were into its assigned part of the grid, having in mind that the grid was conceived as a torus shape, i.e. each edge is connected to its opposite edge. Several simulations were performed with the objective of finding the set of system parameters that would produce the highest speed-up. The SP system can work in IP (Internet protocol mode) whereby the communication between nodes uses an Internet like protocol. This mode is slower than the US (User Space) mode but allows several users to share the same node. The latter protocol makes use of a High Performance Switch that considerably increases the

internode communication speed, but requires dedicated nodes. A table reports the data of different parameters configurations for the same simulation.

Table 1

PROCS	CMODE	TELAP	TCOMM%	SPEEDUP	
1	-	633.92	0	1	1
2	IP	440.28	24.65	1.43	
2	US	363.76	9.78	1.75	
4	IP	381.07	53.56	1.66	
4	US	205.49	18.17	3.08	
8	IP	172.71	47.08	3.67	
8	US	118.12	16.07	5.37	

Where PROCS is the number of processors; CMODE is the User Space or Internet Protocol; TELAP is the clock elapsed time (with dedicated system); TCOMM% is the percentage of interprocessor communication time and SPEEDUP is a measure of the degree of parallel processing.

The configuration that assured a balance between speed-up and shared use of the system was chosen and most of the simulations were performed with 4 processors in US mode. Each epoch, with 380.000 vectors, required about 5 hours of elapsed time. The results of the simulations, of great interest from a theoretical point of view, were reported in several presentations [Le95][Le96][Ca96].

The use of supercomputers for speeding up musical modelling research

In the second project the supercomputer was used in a different way. Now the volume of the data involved in the simulations is low and the issue is not in system speed but rather in the necessity of performing many different tasks in a short time and if possible at the same time. A chain of programs, that represents the perceptive and cognitive processes, must be repeatedly performed. The chain is made of an auditory model, a periodicity analysis autocorrelation step, an on-set analysis, a neural network training or testing. A critical problem was the determination of the type of musical data to be used for training the network. Synthetic Shepard tones [She64] were used first, and real music data, recorded with different techniques, were then tried. Each simulation runs for about seven hours on one SP2 node. A second chain of computations was tied to the creation of 495 four-note chord configurations to be then tested with the trained network, whose results for each chord were subsequently processed to produce a table with chord decompositions. The

work plan encompassed the preparation of many different sound configurations, in order to compare the chord decomposition results, and choose the highest scored one for the analysis of musical examples. The two chains of programs had to be performed many times, and data acquisition and preparation, programs adjustments and data analysis were some of the many tasks required.

The solution that was adopted to increase the productivity of the research process was to choose a remote workstation equipped with an X Window System simulator. In the specific case eXodusPowerPC™ was utilized. Short-cut connections to different nodes of the supercomputer were defined. The workstation allowed for the simultaneous use of tens of remote consoles and edit screens. With a setup of this kind one console was used to start and follow a simulation process; an edit screen could then show partial results or visualization of map layouts [Kra92]; at the same time development work was going on from another console that could utilize other edit screens. Data acquisition from an external musical keyboard was performed directly from the workstation as well as resampling to 20 Khz. The results were then transferred to the supercomputer and there processed. With a precise planning of all the work that could be done in parallel the otherwise much time-consuming development activities of this project could be carried out in a reasonable time a few weeks. The musical results of this project will be presented in [Ca98A][Ca98B].

Conclusion

The high computational requirements that the modelling of cognitive musical processes imply can be met by the use of supercomputing and parallel technology. Simulations that were inconceivable until a few years ago are now possible if parallel computers are used. Large scale simulations are generally possible if neural networks models are used. For the SOM, the feed-forward model and other neural paradigms decomposition and parallelization techniques now exist that make the modelling of large cognitive problems possible. On the other side the simultaneous performance of different activities on the same supercomputer and from the same workstation allows a sensible increase in the productivity of the research development cycle. Two examples of distinct approaches to the development of musical cognitive models were presented from the viewpoint of the use of computer technology. The former made use of the parallel computing facilities of the system for solving a highly computational demanding simulation while the latter facilitated the development work by allowing the simultaneous execution of many tasks related to different problems.

Bibliography

[Le89] Leman M., Van Renterghem P., *Transputer implementation of the Kohonen feature map for a music recognition task*, Proc. of the 2nd International Transputer Conf.: Transputers for Industrial Applications II (BIRA, Antwerpen Belgium) 1989.

[She64] Shepard R., *Circularity in judgments of relative pitch*, The Journal of the Acoustical Society of America, 36, 2346-2353, 1964.

[Kru90] Krumhansl C., *Cognitive foundations of musical pitch*, New York, NY: Oxford Univ. Press.

[Sei91] Seifert U., *The schema concept – a critical review of its development and current use in cognitive science and research on music perception*. In A. Camurri and C. Canepa (Ed.), IX CIM, Genova 1991

[Kra92] Kraaijveld M., Mao J., & Jain A., *A non-linear projection method based on Kohonen's topology preserving maps*. In Proceedings of the 11th ICPR. Los Alamitos, CA: IEEE Comput. Soc. Press 1992.

[La92] Langner G., *Periodicity coding in the auditory system*, Hearing Research, 60, 115-142, 1992.

[VIm92] Van Immerseel L., Martens J., *Pitch and voiced/unvoiced determination with an auditory model*. The Journal of the Acoustical Society of America, 91, 3511-3526, 1992.

[Ba94] Balsach L., *Harmonic Convergence*, Translated from La convergencia Harmonica, Barcelona 1994.

[Ko95] Kohonen T., *Self-Organizing Maps*, 30-Springer Series of Information Sciences, Springer-Verlag, 1995.

[Le95] Leman M., *Music and Schema Theory*, 31-Springer Series in Information Sciences, Springer-Verlag, 1995.

[Ca96] Carreras F., Leman M., *Distributed parallel architectures for the simulation of cognitive models in a realistic environment*. In E. D'Hollander, G. Joubert, F. Peters, & D. Trystram (Eds), *Parallel computing: State-of-the-art perspective*. Amsterdam Elsevier, 1996.

[Ca96] Carreras F., Leman M., *Schema and Gestalt: Explorations of perceptual Learning*. In Proceedings of JIC96-Brugge, M. Leman (Ed.), Brugge 1996

[Le96] Leman M., Carreras F., *The self-organization of stable perceptual maps in a realistic musical environment*. In G. Assyah (Ed.), *Proceedings of the Journées d'Informatique Musicale 1996*. Caen: Univ. De Caen – IRCAM.

[Le96] Leman M., Carreras F., *Psychoneural Isomorphism by Computer Simulation*. In *Music, Gestalt and Computing*, M. Leman (Ed.), Springer LNAI State-of-the-Art Survey 1997

[Ba97] Balsach L., *Application of Virtual Pitch Theory in Music Analysis*, Journal of new Music Research, 26-3, 1997.

[La97] Langner G., *Temporal processing of pitch in the auditory system*. Journal of new Music Research, 26, 1997.

[Le97] Leman M., *The convergence paradigm, ecological modelling and context-dependent pitch perception*. Journal of new Music Research, 26-2, 1997.

[Mo97] Moelants D., Rampazzo C., *A computer system for the automatic detection of perceptual onsets in a musical signal*, In *Kansei: the Technology of Emotion*, AIMI International Workshop, A. Camurri (Ed.), Genova 1997

[Ca98A] Carreras F., Leman M., Petrolino D., *Towards the convergence of the symbolic and subsymbolic approaches to virtual pitch theory*, submitted to *Symposium on Musical Cognition and Behavior: Relevance for Music Composing*, Rome May 1998.

[Ca98B] Carreras F., Leman M., Petrolino D., *Content-based extraction of music harmonic information*, submitted to XII Colloquio di Informatica Musicale, Udine Sept. 1998