Journal of Applied Sciences

ISSN 1812-5654

science alert

ANSInet
an open access publisher
http://ansinet.com

Journal of Applied Sciences 6 (4): 775-778, 2006
ISSN 1812-5654

Sequential and Parallel Genetic Algorithms for the Hybrid Flow Shop Scheduling Problem

[1,2]K. Belkadi , [2]M. Gourgand ,[1]M. Benyettou and [1]A. Aribi
[1]Department of Computer Science, University Mohamed Boudiaf,
USTO, BP 1505 Oran M'Naouer, 31000 Oran, Algeria
[2]University Blaise Pascal, Clermont Ferrand, LIMOS, UMR CNRS 6158,
Campus des Cézeaux, 63177 Aubière Cedex, France

Abstract: In this study we treat the scheduling problem in Hybrid Flow Shop production systems (HFS) and that using the sequential Genetic Algorithms (GA) and Massively Parallel Genetic Algorithms (MPGA). A comparison between the sequential and parallel version of Genetic Algorithm is established. This comparison relates to the quality of the solution. The simulation of parallelization according to the massively parallel model by using a MIMD architecture could improve considerably the quality of the solutions obtained, compared with sequential GA. This success of this model of parallelism is closely linked to its parameters and in particular the number of threads implied in the search and the replacement strategy of the individuals. The replacement of the individuals by application of the temperature principle of simulated annealing could give results clearly better than the MPGA with competition.

Key words: Hybrid flow shop, massively parallel genetic algorithm, optimization, scheduling, hybridization, simulated annealing

INTRODUCTION

Sequential Genetic Algorithms (GA) are stochastic research techniques introduced in the beginnings of the Seventies (Holland, 1975; Goldberg, 1994). They are inspired by the mechanisms of the biological evolution of the species and very much used for combinative optimization.

Sequential GA suffers from a major problem which is premature convergence. In order to cure this disadvantage, several selection and replacement strategies were introduced to create a genetic diversity in the population. And that gave rise to parallel genetic algorithms and among them the Massively Parallel Genetic Algorithm model (MPGA) (Heijligers, 1996; Cantü Paz, 2000).

The MPGA model is intended to be applied to a massively parallel architecture (SIMD: Single Instruction Multiple Data). This model present at itself a new inspiration of the origin biological model, which is the evolution of the species, by introducing the notion of local selection managed by each individual in an autonomous way, as well as the notion of distance which separates an individual from his neighbours and who determines really his local vision or its local neighbouring. The genetic operators will be strongly modified to support this model of parallelization.

In the purpose of exploiting this model of, we simulated this model on a MIMD architecture (Multiple Instruction Multiple Data) with shared memory.

For the resolution of Hybrid Flow Shop (HFS) scheduling problem, a comparison between the sequential and parallel version of GA is established. This comparison relates to the quality of the solution.

MATERIALS AND METHODS

Presentation of hybrid flow shop: An organization in Hybrid Flow Shop (HFS) (Fig. 1) is constituted of a set of stages, where e ach stage is composed of one or more parallel machines. The various jobs visit the stages in the same order. On each stage, a job is treated by only one machine. Between each stage, the jobs can wait or not in limited or unlimited stocks (Billaut *et al*., 2000; Vignier *et al*., 1999).

The problem treated is the scheduling of n jobs in entry of the system and the assignment of these jobs on the machines. We seek a scheduling and an assignment which optimize a criterion of performance. The criterion to be optimized (to minimize here) is C_{max} (the time of total

Corresponding Author: Dr. K. Belkadi, Department of Computer Science, University Mohamed Boudiaf, USTO, BP 1505 Oran M'Naouer, 31000 Oran, Algeria

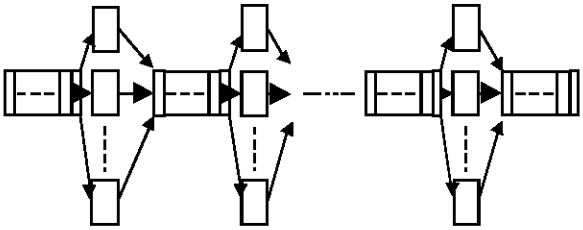

Fig.1: Presentation of FSH

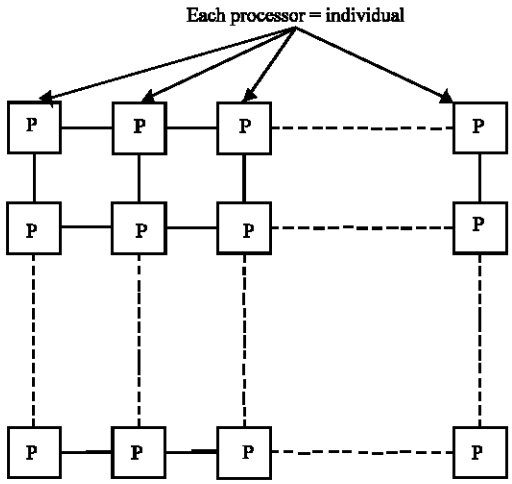

Fig. 2: Architecture SIMD

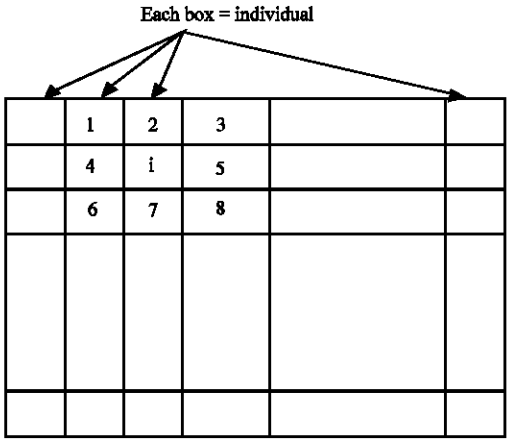

Fig. 3 : Matrix 2D of simulation

completion of works). By using the notation of Vignier (1997), we can represent it by FH2 (P3, P2)||C_{max}.

Presentation of the MPGA parallel model: This model introduces the notion of local selection managed by each individual as well as the notion of distance which determines its neighbouring.

The distribution of the population with an individual per processor is represented in a parallel algorithm with fine grain said data parallel. In this schema, the major element is the topology of the communication network which conditions the diffusion and the distribution of the individuals as well as the parameter of distance which defines the local subset of neighbours with which the individual enters in competition. This implicitly determines the degree "of isolation" of the individual and, consequently, the diversity of the individuals resulting from the application of the genetic operators. Because of the local interactions, the risk of premature convergence can be reduced, because the characteristics of the individuals take much more time to be propagated in the network. It is only after one certain significant number of generations that we see appearing groups of individuals having similar characteristics called zones of convergence (Mahfoud and Goldberg, 1995; Prabhu, 1996).

Simulation of MPGA algorithm on MIMD architecture: The importance of the MPGA lies in the selection and replacement operations which become distributed and take account of the notion of neighbouring. The MPGA is a strongly synchronous algorithm thus its simulation must respect this behaviour.

The simulation of SIMD architecture (distribution of the individuals) (Fig. 2) can be done using an adequate data structure which is a matrix (Fig. 3). Each box comprises a number of the individual in the initial population. Each individual i located by his number, can select his partner among his 8 direct neighbours in the matrix (their numbers are also located in the matrix).

The programming model chosen for the simulation of the MPGA is Master/Slaves. Indeed, MIMD architecture offers a restricted number of processors and cannot for this fact to distribute the individuals at a rate of an individual per processor. For that, each thread of execution deals with treating a number of individuals. Threads of execution share the matrix 2D. Each thread has a number of lines of the matrix 2D and is taken care off to evaluate the individuals whose numbers appear in these lines.

The Master thread initializes the treatment by the random generation of the population. Then it fills the matrix 2D by the numbers of the individuals in the initial population and launches the different research threads (slaves). Each thread (slave) executes the genetic operators for each individual who is assigned to him and that for a given number of generations. For each individual, thread selects a neighbour according to a

selection strategy (Roulette wheel). The selection phase can introduce access conflicts on the matrix 2D into the zones shared by threads. These latter must be synchronized themselves before the selection.

The crossover and mutation operators remain unchanged of those of sequential GA. For the replacement phase, the children individuals generated by the crossover and the mutation enter in competition with the current individual. The competition is governed by a probabilistic choice which depends on the fitness of the individual and those of his two children.

RESULTS AND DISCUSSION

The algorithm of the MPGA was implemented on a Biprocesseur architecture and it was tested on Hybrid Flow Shop of the same type as that studied for sequential GA.

For the experimental tests, the choice was fixed on Hybrid Flow Shop constituted of 2 stages with 3 machines on the first stage and 2 machines on the second stage. The storage taken into account is stock by stage for all stages.

Sequential GA is based on a direct adaptation of the basic operators of GA for HFS problem (Vignier, 1997) in particular in the process of the coding/decoding of the chromosome which is strongly linked to the problematic of HFS scheduling. The variables of decisions linked to the HFS are of two types. One is linked to the assignment problem and the other to the scheduling problem. The genome constitutes of two chromosomes, one for the assignment and the other for scheduling (sequencing). For the assignment, the chromosome is represented by a vector A (a vector for each stage) whose components materialize each work. Thus the A(i) number indicates the number of the machine (in the stage) on which is affected the operation of work i. In the same way, for the sequencing problem, a matrix P contains the set of the precedence constraints between the operations of works at a given stage (Vignier, 1997).

The crossover and mutation operators apply to the two chromosomes (assignment and scheduling) of the genome of any individual. The crossover used is a uniform classical crossover. The mutation operator is very near to the crossover operator but it operates only on the sequencing chromosome by choosing a random machine and changing the order of a couple of the work assigned to this machine (Vignier, 1997).

Concerning the own parameters for sequential GAs which must be defined, their values are fixed and given as follows (Sahraoui, 2002): the population size is 200, the probability of crossover is 0.5, the probability of mutation is 0.005 and the number of iterations is 100.

Sequential GA applied to the HFS scheduling problem gave the results presented in Table 1 (the cost values of average Cmax and execution time). These values will be used as reference for a comparison between the sequential and parallel implementation of GA.

The MPGA depends on several parameters which we have studied to decide convergence of the algorithm (Aribi, 2004). These parameters are the number of threads implied in research, the selection strategy as well as the replacement strategy of the individual children.

The results concerning the MPGA are taken for a number of threads = 4, the selection strategy by roulette wheel and the replacement strategy by competition and by Simulated Annealing. In the selection strategy by roulette wheel (Goldberg, 1994), an individual i selects his partner among his eight direct neighbours according to a skewed wheel based on the fitness (objective function) of the eight neighbours.

We introduced a hybridization between the MPGA and Simulated Annealing by introducing the temperature parameter which regularizes the replacement schema (Aribi and Belkadi, 2003; Bachelet *et al.*, 1998).

Comparison between the MPGA and sequential GA: To study the contribution in quality which the parallelization strategy brings, a comparison between the MPGA and

Table 1: Average C_{max} and average CPUs for sequential GA

No. of jobs	Average C_{max}	Average CPUs (sec)
5	145.0	0.5
10	260.4	2.0
15	447.5	3.0
20	526.1	6.0

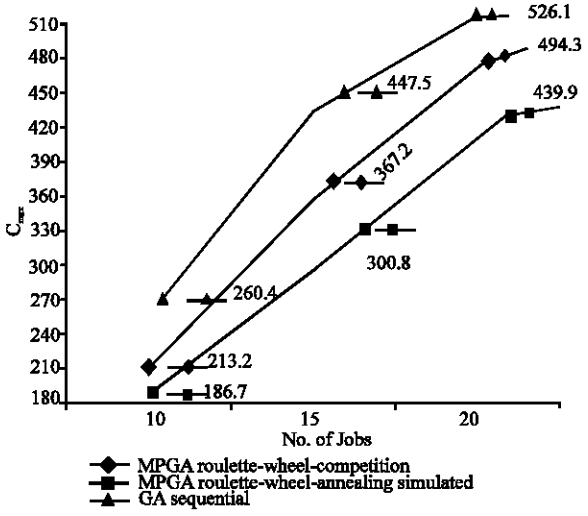

Fig. 4 : Comparison between the MPGA (competition and SA) and sequential GA

the sequential GA in term of quality of solution (Fig. 4). The graph represents the variation of average C_{max} with the number of jobs (N = 10, 15, 20) and it comprises three series of data: the series of the sequential GA, the series of MPGA by competition and the series of the MPGA by Simulated Annealing (SA).

In this Fig. 4 we notice the contribution of the distribution of the selection and the decentralization of the replacement strategy in the improvement of the results obtained by sequential GA. The MPGA is that which gave the best results. There is a significant gap between the curves which represent the sequential GA and the MPGA results. This gap remained constant compared to the increase of the HFS complexity (variation of the number of jobs). The contribution of the replacement strategy by SA is significant and among these three series of data it is the series of the MPGA-SA which gave the best results.

CONCLUSIONS

The simulation of parallelization according to the massively parallel model by using a MIMD architecture could improve considerably the quality of the solutions obtained, compared with sequential GA for the resolution of the Hybrid Flow Shop scheduling problem (HFS). The application of the MPGA depends on two essential parameters which condition the improvement of the results. These parameters are the number of threads implied in the search and the replacement strategy of the individuals. The replacement of the individuals by application of the temperature principle of simulated annealing could give results clearly better than the MPGA with competition. Moreover, we can say that this type of parallel genetics algorithmic presents a biological model of inspiration near to nature by underling the notion of local selection and the notion of neighbouring. This aspect was completely neglected in the sequential GA model.

REFERENCES

Aribi, A. and K. Belkadi, 2003. Parallel hybridization of parallel simulated annealing and genetic algorithms for hybrid flow shop. 2nd International Conference on the Productique, CIP'2003, Algiers (Algeria). CD-Rom, Session A4, Soft computing 2, Art. 57, pp: 8.

Aribi, A., 2004. Parallel meta-heuristics applied to the hybrid flow-shop scheduling production systems. Thesis of Magister, USTO, Oran (Algeria).

Bachelet, V., Z. Hafidi, P. Preux and E.G. Talbi, 1998. Towards the co-operation of the meta-heuristics. Parallel calculators, networks and systems distributed. Hermes, 10: 211-223.

Billaut, J.C., A.Vignier, C. Proust and M.C. Portmann, 2000. A survey of hybrid flow shop problems. Deterministic scheduling for computer science and production (ODIP), School of Summer, Aussois, France.

Cantü Paz, E., 2000. Markov chain models of parallel genetic algorithms. IEEE Trans. Evolutionary Computation, 4: 216-226

Goldberg, D.E., 1994. Genetic Algorithms: Exploration, Optimization and Machine Learning. Addison Wesley Edition, France.

Heijligers, M., 1996. The application of genetic algorithms to high-level synthesis. Ph.D Thesis, Eindhoven University of Technology.

Holland, J.H., 1975. Adaptation in Natural and Artificial System. The University of Michigan Press, Ann. Arbor.

Mahfoud, S. and D.E. Goldberg, 1995. Parallel recombinative simulated annealing: A genetic algorithm. Parallel Computing, 21: 1-28.

Prabhu, D., 1996, A study of massively parallel genetic algorithms with application to image interpretation. Ph.D Thesis, Tulane University

Sahraoui, M., 2002. Optimization methods applied to the hybrid flow shop production systems. Thesis of Magister, Oran (Algeria).

Vignier, A., 1997. Contribution to the resolution of the monogam multi-machines scheduling problems (HFS). Ph.D Thesis, University of Tours.

Vignier, A., J.C. Billaut and C. Proust, 1999. Hybrid flow-shop scheduling problems: State of the art. RAIRO, Operational Res., 33: 117-183.