JOB SEQUENCING PROBLEM IN A SEMI-AUTOMATED PRODUCTION PROCESS

Roberto Mosca
Filippo Queirolo
Flavio Tonelli
Department of Production Engineering
University of Genoa
Via all'Opera Pia 15, I-16145
Genoa, Italy
e-mail: tonelli@itim.unige.it

KEYWORDS
Sequencing, Simulators, Optimization, Parallel Processing

ABSTRACT

In this work the authors addresses the problem of sequencing a set of jobs on a single machine using a genetic algorithm and simulation. The goal is to find the schedule that minimizes the total earliness and tardiness penalties of all jobs, under the assumptions that no pre-emption of jobs is allowed and all jobs are available at time zero. In order to accelerate the search process, the Authors also implemented a procedure for genetic algorithm initialization. Simulation has been used for the fitness evaluation of the population's members: in this way, one of the most critical issues related to evolutionary computation has been successfully addressed. This hybrid approach led to an effective tool adopted for the scheduling in a real production plant, where three bottling lines are used and several kind of product are commercialized.

INTRODUCTION

In the paper the problem of single machine scheduling (SMS) is addressed. SMS represents the simplest of all possible scheduling problem. Nevertheless SMS provides important basis for heuristics approaches that are adopted in the cases of more complex production processes [Pinedo 2002].
Indeed flow shop, flexible flow shop, job shop and open shop scheduling problem are often addressed decomposing the original planning process into many sub-problem that can be solved by using single machine techniques.
Since 1955 several approaches have been proposed for this class of scheduling problems: EDD rule [Jackson 1955]; WSPT rule [Smith 1956; Lawler 1978; Sidney and Steiner 1986]; branch and bound methods [Nowicki and Zdrzalka 1986]; number of late jobs minimization [Moore 1968]; the scheduling problem as a knapsack problem [Gens and Levner 1981; Potts and Van Wassenhove 1988]; due date assignment problem [Cheng and Gupta 1989]; dynamic programming approaches [Potts and Van Wassenhove 1982, 1987]; PTAS [Chekuri et al. 1997; Schuurman and Woeginger 1999]; earliness and tardiness penalties [Baker and Scudder 1990]; SMS with multiple objectives [Chen and Bulfin 1993; Hoogeveen and Van Der Velde 1995]; the SMS as a TSP [Gilmore and Gomory 1964; Bianco et al. 1988; Wittrock 1990].
This work adopts an evolutionary approach for SMS greatly integrated by simulation: in the studied case, the SMS looks as a sequencing problem and the sequencer has been realized by using an hybrid solution, which joins genetic algorithm (GA) and simulation.
The authors deal with the proposed approach starting from a real industrial case study, concerning an Italian SME operating in the large consumer good market.
The remainder of the paper is organized as follow: after a brief introduction to the studied case, the proposed approach is followed up and the genetic paradigm and operator are described; then the contribute of simulation is highlighted and the conclusions are drawn.

THE REAL INDUSTRIAL CASE

The studied production process is a flowshop process: three semi-automated production lines are sited in the plant and the equipment is connected by rolling tapes. Two lines request the presence of eight employs, while the third needs of ten persons. According to products' technological cycles long setup times are required (from 2 to 6 hours). They are due to the substitution of some tools, equipment calibration for the specific size of the produced bottles and testing.

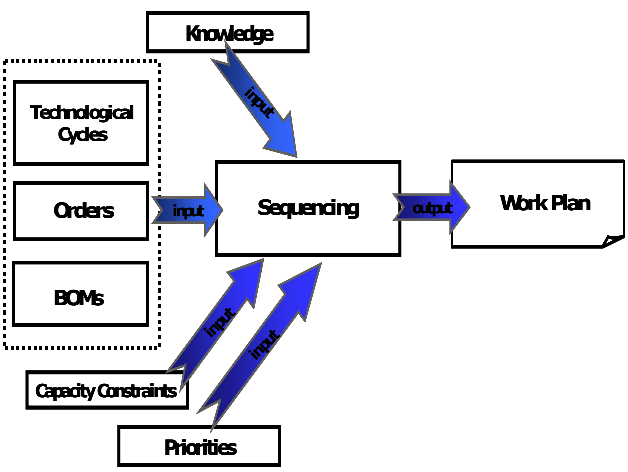

Figure 1: Information Flow Diagram in the Studied Company

The commercialized items

Twenty-one items are commercialized in France and in Northern and Middle Italy. Production is generally organized according to make to stock policy. Nevertheless a significant part of the produced items can be manufactured just when unforecasted orders come: in this case a mixed (make-to-stock/ make-to-order) policy is followed and the rescheduling of the activities is required.

Hidden set-up

When production is realized by using one of the two 8-employs lines, two persons execute PET bottles' preforms insuffling operations or the setup activities on another production line: in this way it is possible that the setup time is hidden and the total completion time is reduced.

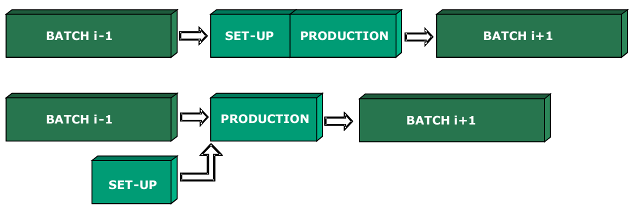

Figure 2: Hiding Setup Significantly Reduces Total Completion Time

Lot sizing solutions

The batches are determined as a function of the production rates: a batch is equal to the produced quantity during a turn. In some particular cases (generally during summer) the minimum batch size is equal to the production realized in an half of a turn.

ONE SINGLE MACHINE SCHEDULING

Because of the features of the production processes, the scheduling problem is reduced to a job sequencing problem. It means that the goal is determining the sequence, which optimizes the performance indices. This approach is consistent because the three production lines cannot work in a parallel way. Thus one can think to each lines as a single machine that transforms raw materials in finished goods (each flowshop is modeled by observing the production bottleneck of the lines). Sequencing is an integer optimization problem on bounded domain. The Authors faced this problem by using an hybrid approach, which joins a genetic algorithm and a simulator.

GAS FOR SEQUENCING

In order to find a solution for job sequencing, searching for an optimum (or sub-optimum) order of a list of objects is required. This problem is generally classified as a constrained optimization problem (COP). GAs have been successfully applied in the case of continuous function optimization [Chambers 1995] on bounded domain. Even if binary codes of the problem are generally adopted, several real cases concern multidimensional search. Many solutions have been provided for coding problem (see for example [Mitchell 1995]). In this work, the Authors adopted a code based on a alphabet of n symbols.
Traditional genetic operators are named crossover and mutation: crossover is responsible for mixing the portions of the sequences selected as parents; mutation introduces stochastic ness in the research process [Mitchell 1995].
However, in several industrial cases, developers should design specific genetic operators, based on the characteristics of the treated problem.

An OR viewpoint for sequencing problem

It's well known that COPs belong to the class of NP-hard problem [Mackworth 1977; Freuder 1978]. Therefore no classic research algorithm is available which is able to find the optimal solution (of every COP) in polynomial time.
Several Authors [Dechter and Pearl 1988, 1989; Fox et al. 1989; Freuder 1982, 1985] deeply investigated the performance of different heuristic algorithms for COPs. In spite of these significant efforts, a need for an efficient method is required by industry. Since 1983, theoretical bases have been drawn for the selection of the best heuristic for the assignment of the value to the variables [Nudei 1983]. However that and following results are characterized by the probabilistic approach: no effective solution of the faced problem is guaranteed.
Some algorithms are especially suitable for simple COPs; this is the case of problem where no cycles and/or implicit constraints are. Nevertheless the COPs concerning real problem are characterized by high dimensional research space, many cycles and implicit constraints. As a result determining optimal solution of a COP is generally performed by trial-and-error search algorithm.
For these reasons determining robust general purpose algorithms for this class of problem is a relevant issue both from theoretical and applicative viewpoint.

Heuristic Genetic Algorithms

During last ten years evolutionary computation has represented an interesting and important method approaching COPs.
Contrary to traditional search methodology, the aim of the heuristic genetic algorithms (HGAs) are not the exhaustive exploration (performed in some way) of all the possible solutions: GAs improve the fitness of the current set of individuals (chromosomes) by the weighted random selection of one or more parents and create a certain number of children from these parents by the genetic operators.
The main advantage deriving from evolutionary approach is related to the fact that, during each loop of the algorithm (i.e. a generation) a whole set of potential solutions (i.e. the population) is evaluated. In this way next individuals are generated by the simultaneous evaluation of many

production sequences. On the other hand, traditional methods examine just one solution during a loop.
In the literature, authors generally refer to this property as the parallel computation or the parallel research process. Parents are selected with a method by which selection probability is proportional to parent's fitness value.
Traditional GAs are based on selection and crossover operators that use uniform random number and are suitable for problems, characterized by not correlated independent variable [Davis 1985; Davidor 1991]. As a result, classic GAs should be adapted to the specific problem in order to be applied to constraint satisfaction problems (CSPs) and to COPs. The most important modifies concern the code and genetic operators.
The proposed approach substitutes uniform random mechanisms with new operators, which allow to preserve (almost partially) the feasibility of the candidate solutions.
The adopted genetic operators are partially derived from heuristic approach found in literature [Chambers 1995] and, when requested, they have been modified and integrated with problem specific knowledge. On the other hand, new operators have been designed in order to improve the GA's performance.
The main goal has been to integrate stochastic method (proper of traditional GA) with heuristic approaches. In this way, the best feature of each class of techniques have been followed up. As a result, hybrid approach minimizes the lackness of each class of mechanisms and new promising sequences could be introduced during the search process.
However, it's important to highlight the limitations of evolutionary computation: while exhaustive exploration of the research domain ever leads to a feasible solution, GAs lead to suboptimal solutions and it's impossible to state that no solution can be find for the studied problem.

Genetic Operators

The genetic algorithm is developed by employing problem specific reproduction, crossover and mutation operators.
In particular, the authors focused on crossover operator because it determines the role of exchanging information during evolution and this process often cause redundant or lost features in some solutions.
In this study the uniform order-based crossover is adopted because it is considered to best fit to the job sequencing problem, since it allows both the absolute and relative positions of parents' job sequences [Chambers 1995].
In order to accelerate the converge of the genetic algorithm, the authors also implemented a procedure for genetic algorithm initialization: the initial sequences of jobs are partially constructed generating sequence with EDD and SPT rules.

Unit and relationship of the proposed model

Sequencing problem has been faced by a multi-objective function composed by two components: the goal was to find the schedule that minimizes the sum of the total setup time and the earliness and tardiness penalties of all jobs, under the assumptions that no pre-emption of jobs is allowed and all jobs are available at time zero. Each job has its own due date, earliness and tardiness penalty weights.

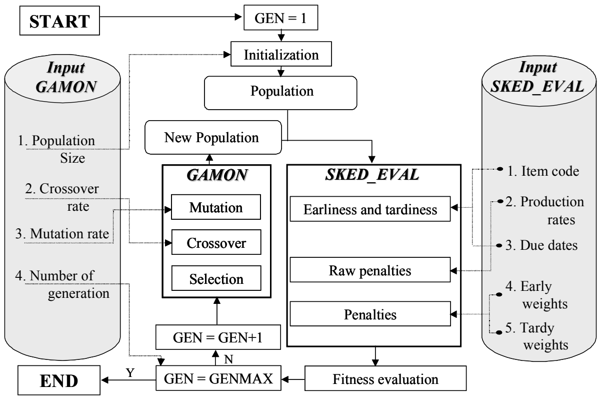

Figure 3: Schematic Representation of the Proposed Model

The first component optimizes the impact of setup times on the total completion time. In presence of production sequences dependent of the set-up times, a formal model for the problem is required because of strongly NP-hard nature of the sequencing. In accordance with [Pinedo 2002], the authors considered the scheduling problem as a Traveling Salesman Problem (TSP). This approach is easily understandable if one considers that the cities of the TSP represent the jobs and the distances model the setup times. This approach is suitable for problems characterized by setup times that depend only on the adjacent jobs of the production sequence. Since in this case production is realized on three different automated lines (modeled as three machines), setup times depend on the all production sequence. As a result modifying the fitness of a sequence by a specific routine is requested. The following pseudo-code describes this procedure:

```
If last_item_produced_on (prod_line_i) ==
  actual_scheduled_item_on (prod_line_i)
  then
    Setup_time = 0
Else
    Setup_time = TSP_distance_between
      (last_item_produced_on (prod_line_i),
      actual_scheduled_item_on (prod_line_i))
End
```

The second component of the fitness function is derived from a traditional heuristic function, commonly adopted in the case of operative scheduling problem: the total earliness and tardiness [Baker 1995]. In accordance with this approach the (partial) fitness value is calculated as the weighted sum of the earliness and tardiness.:

$$TotalPenalty = \sum_{i=1}^{\text{N of Items}} ItemPenalty(i)$$

where:

$$ItemPenalty(i) = \sum_{j=1}^{\text{N of items in advance}} Earliness(j) \cdot Quantity(j) \cdot EarlyWeight(j) +$$

$$+ \sum_{k=1}^{\textit{N of items in delay}} Tardiness(k) \cdot Quantity(k) \cdot TardiWeight(k)$$

SIMULATION FOR FITNESS EVALUATION AND SCHEDULING ROBUSTNESS ANALYSIS

In adopting evolutionary computation the definition of a fitness function is required. Nevertheless in the major part of the real cases no explicit fitness function is available. So genetic approaches are strictly related to all the research areas which can lead to an approximation of the real fitness function as accurate as possible. Obviously simulation represents a suitable technique for solving this critical issue. In this work simulation has been used for the evaluation of the earliness and tardiness penalties (ETP) of the population's members. In this way it was also possible to validate the deterministic scheduling: simulation allows production sequences' robustness evaluation.

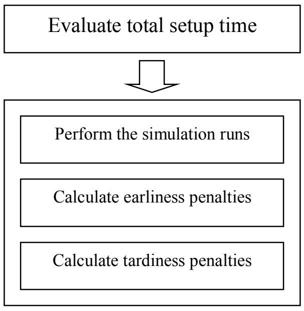

Figure 5: Flow Chart of the Fitness Evaluation Process

ETP estimation

Three simulators of the production processes have been designed and realized in order to evaluate the fitness of the production sequences. In accordance with [Mosca 1982], the experimental error has been calculated as the Mean Square Pure Error (MSpe):

$$SSpe = \sum_{i=1}^{n} \sum_{i=1}^{m} (Y_{ij} - \bar{y}_i)^2$$

$$MSpe = \frac{SSpe}{Gdl}$$

where: $Gdl = N^{\circ}_{RUN} - 1$

Y_{ij} is the production rate every τ hours

$\bar{y}_i$ is the average (on a run, with a fixed τ)

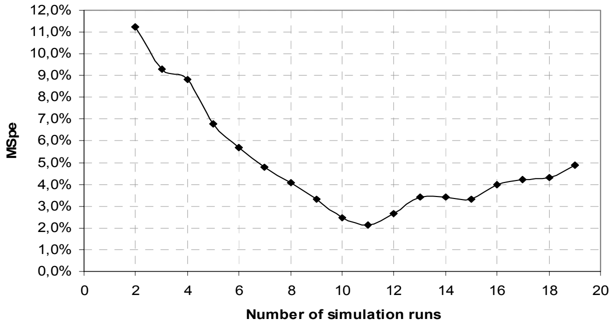

Figure 6: MSpe versus the Number of Run

For each sequence several runs are performed and the sequence's fitness value is calculated as the average of the earliness and tardiness penalties observed at the end of each run. In order to determine the optimal number of runs, the relation between the MSpe and the number of runs has been investigated.
According to the results of the MSpe analysis on the number of runs, the authors calculated the fitness as the average of the 11 simulation runs. In this way it was possible to bound the experimental error under the 3.15%.

Stochastic simulation for sequencing validation

The proposed approach allows also to the validation of the production sequences in a stochastic framework. Indeed the result of the deterministic scheduling (i.e. sequencing) process is obtained by using a fitness function which integrates earliness and tardiness penalties calculated on the basis of the performance observed in a stochastic environment.

APPLY THE PROPOSED METHODOLOGY

In order to enabled further application of the proposed methodology, a brief list of the required activities is provided:

1. search or design a traditional GA;
2. determine a code for the chromosome (number of gene is equal to the planning horizon; symbols of the alphabet are as many as the class of products);
3. extend crossover and mutation operator in accordance with the specific constraints;
4. evaluate if the designed operators finally preserve children feasibility;
5. in the case that feasibility is not completely preserved, design a repairing procedure;
6. build the simulators of the production processes;
7. perform verification and validation of the simulators;
8. determine the optimal number of simulation runs;
9. design the procedure for the evaluation of total setup time (TSP is a useful formal model);
10. integrate simulators and setup evaluation procedure in the GA so that fitness is calculated as the sum of earliness and tardiness penalties and total setup time.

CONCLUSIONS

A scheduling problem with distinct due date in a single machine is considered in this work. The adopted hybrid (genetic and simulative) approach led to a tool effectively adopted for the scheduling in a real production plant, where three bottling lines are used and several kind of product are commercialized. The authors approached the scheduling problem in a semi-automated production process (i.e. flowshop) as a sequencing problem. This is one of the well-known very hard combinatorial optimization problems. GAs are a good tool for COP, even if specific genetic operators should be designed in order to preserve the feasibility of the candidate solutions.

Adopting evolutionary computation a fitness function is required. Even if in the real industrial cases no fitness function is available, the genetic approach can be greatly supported by simulation. As known, simulation leads also to a validation of the deterministic scheduling in a stochastic environment.

Future work concerns the extensions of the proposed approach by using a multi-genetic-agent system, where each agent searches for a local solution and a supervisor manage the evolution of the GAs.

ACKNOWLEDGEMENTS

The authors wish to thank the referees for suggestions that have proved very useful in revising the submitted version of this paper.

REFERENCES

Baker, K.R. 1995. *Elements of Sequencing and Scheduling*. K. Baker. Amos Tuck School of Business Administration. Dartmouth College, Hanover, New Hampshire.

Baker, K.R. and G.D. Scudder. 1990. “Sequencing with Earliness and Tardiness Penalties: A Review”. *Operations Research*, Vol. 38, 22-36.

Bianco, L.; S. Ricciarelli; G. Rinaldi; and A. Sassano. 1988. “Scheduling Tasks with Sequence Dependent Processing Times”. *Naval Research Logistic Quarterly*, Vol. 35, 177-184.

Chambers, L. 1995. Practical Handbook of Genetic Algorithm, CRC Press, New York.

Chekuri, C.; R Motwani; B. Natarajan, and C. Stein. 1997. “Approximation Techniques for Average Completion Time Scheduling“. In *Proceedings of the Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 609-617.

Chen, C.-L. and R.L. Bulfin. 1993. “Complexity of Single Machine, Multi Criteria Scheduling Problems”. *European Journal of Operational Research*, Vol. 70, 115-125.

Cheng, T.C.E. and M.C. Gupta. 1989. “Survey of Scheduling Research Involving Due Dates Determination Decision”. *European Journal of Operational Research*, Vol. 38, 156-166.

Davidor, Y. 1991. “Epistasis Variance: A View-point on GA-Hardness”. In *Proceedings of FOGA-90*. Morgan Kaufmann, 23-35.

Davis, L. 1985. “Applying Adaptive Algorithms to Epistatie Domains”. In *Proceedings of IJCA-85*. Lawrence Erlbaum Associates. 162-164.

Dechter, R. and J. Pearl 1988. “Network-Based Heuristic for Constraint - Satisfaction Problems”, *Artificial Intelligence*, Vol. 34, 1-38.

Dechter, R. and J. Pearl. 1989. “Tree clustering for constraints networks”. *Artificial Intelligence*, Vol. 38, 353-366.

Freuder, E.C. 1978. “Synthesizing Constraint Expressions”. *Communications of the ACM,* Vol. 21, N. 11, 958-966.

Freuder, E.C. 1982 “A Sufficient Condition for Backtrack-Free Search”. *Journal of the ACM*, Vol. 29, 24-32.

Freuder, E.C. 1985. “A Sufficient Condition for Backtrack-Bounded Search”. *Journal of the ACM,* Vol. 32, 775-761.

Fox, M. S.; N. Sadeh; and C. Baykan. 1989. “Constrained Heuristic Search”. In *Proceedings of 11th IJCAI*. Detroit. 309-315.

Gens, G.V. and E.V. Levner. 1981. “Fast Approximation Algorithm for Job Sequencing with Dead-lines”. *Discrete Applied Mathematics*, Vol. 3, 313-318.

Gilmore P.C. and R.E. Gomory. 1964. “Sequencing a One-State Variable Machine: A Solvable Case of the Travelling Salesman Problem”. *Operations Research*, Vol. 12, 655-679.

Hoogeven, J.A. and S.L. Van de Velde. 1995. “Minimizing Total Completion Time and Maximum Cost Simultaneously Is Solvable in Polynomial Time”. *Operation Research Letters*, Vol. 17, 205-208.

Jackson, J.R. 1955. “Scheduling a Production Line to Minimize Maximum Tardiness”. Research report 43. Management Science Research Project, University of California, Los Angeles.

Lawler, E.L. 1978. “Sequencing Jobs to Minimize Total Weighted Completion Time Subject to Precedence Constraints”. Annals of Discrete Mathematics, Vol. 2, 75-90.

Mackworth, A.K. 1977. “Concistency in Networks of Relations”. *Artificial Intelligence*. Vol. 8, 99-118.

Mitchell., M. 1998. *An Introduction to Genetic Algorithms*. First MIT Press Paperback Edition, Massachusetts.

Moore, J.M. 1968. “An n Job, One Machine Sequencing Algortihm for Minimizing the Number of Late Jobs”. *Management Science*, Vol. 15, 102-109.

Mosca, R.; P. Giribone; G. Guglielmone. 1982. “*Optimal Lenght in O.R. Simulation Experiment of Large Scale Production System*”. In *Proceeding IASTED - AMS’82*. 78-81.

Nowicki E. and S. Zdrzalka. 1986. “A Note on Minimizing Maximu Lateness in a One-Machine Sequencing Problem with Release Dates”. *European Journal of Operational Research*, Vol. 23, 266-267.

Nudei, B. 1983 “Consistent-Labeling Problems and their Algorithms: Expected Complexities and Theory Based Heuristics”. *Artificial Intelligence*. Vol. 21, 135-178.

Pinedo, M. 2002. *Scheduling – Theory, Algorithms, and Systems. Second Edition*. Prentice Hall, Englewood Cliff, N. J.

Potts, C.N. and L.N. Van Wassanhove. 1982. “A Decomposition Algorithm for the Single Machine Total Tardiness Problem”. *Operation Research Letters*, Vol. 1, 177-181.

Potts, C.N. and L.N. Van Wassanhove. 1987. “Dynamic Programming and Decomposition Approaches for the Single Machine Total Tardiness Problem”. *European Journal of Operational Research*, Vol. 32, 404-414.

Potts, C.N. and L.N. Van Wassanhove. 1988. “Algorithms for Scheduling a Single Machine to Minimize the Weighted Number of Late Jobs”. *Management Science*, Vol. 34, 843-858.

Schuurman, P. and G. Woeginger. 1999. “Polynomial Time Approximation Algorithms for Machine Scheduling: Ten Open Problems”. *Journal of Scheduling*, Vol. 2, 203-214.

Sidney, J.B. and G. Steiner. 1986. “Optimal Sequencing by Modular Decomposition: Polynomial Algorithms”. *Operations Research*, Vol. 34. 606-612.

Smith, W. E. 1956. “Various Optimizers for Single Stage Production”. *Naval Research Logistic Quarterly*, Vol. 3, 59-66.

Wittrock, R.J. 1990. “Scheduling Parallel Machines with Major and Minor Setup Times”. *International Journal of Flexible Manufacturing Systems*, Vol. 2, 329-341.

AUTHOR BIOGRAPHY

ROBERTO MOSCA is full professor of "Industrial Plants Management" and "Economy and Business Organization" and he is currently the head of Department of Production Engineering at University of Genoa. He has worked in the simulation field since 1969 developing interesting enhancements in the application of DOE. His research work focuses on the application of AI for industrial plant management and the evaluation of new modeling techniques and simulation languages and architectures.

FILIPPO QUEIROLO is a member of DIP Research Group at University of Genoa. He’s research interest include Multi Agent Systems, Artificial Computation, Operation Management, Simulation, Pattern Recognition and Signal Processing. He is a member of SCS, Liophant Simulation Club and vice-president of SACS.

FLAVIO TONELLI is a complex systems management researcher in the Department of Production Engineering at University of Genoa. He earned a PhD in Production systems and industrial plant management from University of Parma and degrees from Genoa University. His research interests include planning production and control and logistics systems, simulation modeling, object oriented software design and prototyping,. He is a member of ANIMP, Liophant Simulation Club and Society for Computer Simulation.