AI TECHNIQUES FOR THE IMPLEMENTATION OF NEW ORGANIZATIONAL STRUCTURES IN THE RETAIL INDUSTRY

Alessandra Orsoni
Kingston University
Kingston upon Thames, Surrey KT1 1LQ, UK
E-mail: aorsoni@alum.mit.edu

KEY WORDS: Island-based Organizations, Decision Support Systems, Genetic Algorithms, Logistic Management, Optimization Techniques.

ABSTRACT

This paper introduces a hybrid decision support system (DSS) for the logistic management of island-based organizational structures. The main benefits expected from the implementation of this innovative organizational structure are improved workers satisfaction, productivity, and quality of service to the customer, and reduced risks of manpower shortage. The DSS is based on a hybrid architecture combining the propositive capabilities of genetic algorithms (GAs) and the scenario-testing strength of simulation for the iterative optimization of resources allocation to dynamic workloads. The design and the customization of the DSS are discussed in the paper with reference to an example application relevant to a large Italian retail chain. The application consists of a feasibility study for the implementation of the island-based concepts in a prototype retail store. In this context the DSS optimizes the allocation of the current pool of resources, in terms of quality of service to the customers, workers' job satisfaction and resources cost. In addition, the DSS provides guidelines and recommendations for future training and hiring.

INTRODUCTION

The concept of island-based organizational structures is a new approach to business organization and management developed and first introduced in France by a small consulting firm, Consilium 2000, operating in the field of human resources management. Islands are work groups built on the basis of complementary workers' skills and schedule requirements.

Some of the readers may be familiar with other types of organizational structures based on workgroups. Cell-based manufacturing systems (Prickett, 1994; Slack, 1988; Schonberger, 1986), for instance, are well-known exmples of organsations based on workgroups. What set manufacturing cells and islands apart are the purposes of their implementation and, consequently, their grouping criteria. Manufacturing cells are typically introduced to achieve production related objectives such as reducing lead time and Work In Progress (WIP) by increasing flexibility (Slack, 1988) and streamlining production (Schonberger, 1986; Prickett, 1994). Islands, instead, address the needs of labour intensive business environments with the objectives of reducing the cost of resources overtime, increasing resources productivity, and reducing the risks of manpower shortage, by better meeting workers schedule and job preferences. To a large extent, islands are self managed in that workers get to choose their working hours and duties on a bi-weekly basis as long as the overall work-plan for the island is feasible and covers the entire workload allocated to the island for that (the 2-week) period.

While the concept of islands is fully developed in theory and a prototype island-based hyper-market is now operating in France, the implementation of this business structure in existing organizations represents a totally different challenge as it raises the issues of defining the islands around rather specialized resources and of deciding the most effective way to expand their skills and competences in order to meet the poly-functional needs of each island. The major difference between the two situations is that the prototype French hyper-market was created around the concept of islands and, thus, specifically designed to operate according to this innovative business structure, meaning that the personnel was hired and immediately cross-trained for these purposes.The optimization of resource allocation to islands in a traditional work environment is a complex example of logistic management problem, which involves the definition of the islands in terms of types and amounts of workload, and mix of allocated resources, as dynamic variables following the patterns of customers' flow in the store. Several are the performance drivers, which are to be decided upon in the formulation of a feasible and cost-effective allocation plan. At the whole service level the relevant performance measures include resources cost, productivity, and customer service, which are driven by other parameters including workers job satisfaction, average customer queuing time for different services, and product display/availability for purchase. The first and most important step to be competed prior to the design and implementation of the islands, is a thorough assessment of the workloads in the different areas of competence, with their seasonal and daily fluctuations. Second it is necessary to clearly assess the current skills and competences of the available resources, their potential and willingness to undergo training and learn new skills, matching them with the identified types of workloads. Third it is necessary to understand workers

Proceedings 15th European Simulation Symposium
Alexander Verbraeck, Vlatka Hlupic (Eds.)

ISBN 3-936150-28-1 (book) / 3-936150-29-X (CD)

priorities in terms of work schedule and preferences in terms of duties. The efficient allocation of a specific resource to a job on a given work-shift is highly dependent not only on the expected performance of such an allocation, but also on the implications that this choice may have on the performance of all the remaining allocations. These interdependencies are mainly due to the fact that the number of available resources per skill/competence area is limited and, therefore, the allocation of a first-choice resource (e.g. the most efficient one) to a job may leave second or third choice resources available for allocation to other duties. In such a sense the optimization of the process at the individual job level is likely to produce a resource allocation plan which is sub-optimal from the perspective of the service as a whole (Orsoni, 2000). Simulation-based testing is the only approach by which each resource allocation scenario can be thoroughly tested taking into account primary, secondary, and tertiary impacts of each allocation choice on the performance of the whole service over the 2-week period. However, when looking at business organizations counting over 50 employees, exhaustive testing, on all possible combinations of resources allocations to jobs and workshifts, becomes computationally intensive and time consuming if handled by simulation alone. The approach proposed in this paper employs AI techniques based on Genetic Algorithms (GAs) for preliminary screening of the resource allocation scenarios to be subsequently tested through simulation. The iteration of the procedure feeding the simulation outputs back into the GAs effectively leads to the identification of an optimized solution for the entire service which is the result of best trade-offs among workloads rather than workload-based optimization

BACKGROUND

The integration of AI and simulation techniques constitutes a powerful approach that the authors have successfully implemented in the development of decision support systems (DSSs) for complex industrial applications (Mosca et al, 1998; Bruzzone et al, 2001). Relevant applications to supply chain management rely on the anticipatory capabilities of artificial neural networks (ANNs) for demand forecasting (Giribone et al, 1997; Giribone and Bruzzone, 1997), for resources/materials usage prediction (Giribone and Bruzzone, 1998), and for the definition of projected scenarios providing the boundary conditions for simulation-based testing of alternative management policies. Other applications (Bruzzone et al., 2002; Bruzzone and Signorile, 1998) use AI techniques, based on either ANNs or genetic algorithms (GAs), as DSS planning/optimization modules. Two major factors drive the choice between ANNs and GAs, these are: the nature of the decision context and the availability of historical data. ANNs are capable of establishing correlations between corresponding sets of input and output parameters even when explicit analytical or logical formulations may not be found (Padgett and Roppel, 1992; Hillis, 1989). This ability is based on “experience”, which in turn is provided to them by training on extensive sets of historical data (Anderson and Rosenfield, 1988). When applying ANNs to complex decisional problems, adequate precision is difficult to achieve in reasonable learning times (Padgett and Roppel, 1992; Anderson and Rosenfield, 1988): these objectives can only be achieved in situations involving a limited number of input and output parameters, if a large enough set of representative training data is available (Bruzzone et al, 2001). In contrast, GAs are well-suited to address complex combinatorial problems involving multiple variables and do not require preliminary training (Bruzzone et al, 2001; Goldberg, 1989). They iteratively seek performance improvement by trial and error (Bruzzone and Signorile, 1998; Goldberg, 1989; Koza, 1992). At each iteration an “evolutionary step” is made: the least efficient solutions are discarded, while attempts are made to improve the best performing ones by mixing (cross-over) and modification (mutation) (Goldberg, 1989). Both ANNs and GAs may be effectively combined with simulation to address complex optimization problems, however, for the particular application described in this paper, GAs appear better suited because of the lack of historical data on resource allocation, and because of the large number of parameters involved (including the number and types of resources with their specific skills and availability constraints, and the number and types of workloads with their seasonal and daily fluctuations.) The optimization problem examined in this paper is especially complex because a global optimum needs to be found based on multiple interdependent decisions, which cannot be separately optimized: in the island-based organizational structure, the efficiency of the allocation of a particular resource to a workload is reflected in the processing performance of other workloads and, through them, in the performance of the entire service.

STRUCTURE OF THE APPROACH

The DSS consists of a hybrid architecture, which combines discrete event stochastic simulation and Artificial Intelligence (AI) techniques based on Genetic Algorithms (GAs) in the form of two interactive modules (Figure 1). In this architecture the simulation module tests the feasibility and the effectiveness of alternative scenarios of resource allocation to the various jobs and workloads and to the scheduled work shifts with respect to the performance of the entire service on the 2 weeks. Starting from the measures of performance provided by the simulation module on a first set of scenarios, the GA-based optimization module generates new, improved scenarios for further testing. The procedure is iterated until an optimum combination of resources scheduling and job allocation is reached. The major advantage of using GAs in the optimization process is that the search for the optimum solution begins from an entire “population” of scenarios

(Goldberg, 1989) and, thus, from multiple points in the space of the possible solutions, which highly increases the chances of finding the actual optimum, rather than a sub-optimum. In addition, GAs are based on stochastic rather than derterministic rules, which further improves the effectiveness of the search (Goldberg, 1989). In order to define a suitable set of initial scenarios, the optimization tool is first run under the hypothesis of ideal conditions (independent workloads and unlimited number of resources) to determine the most suitable resource for each job-unit/workload (i.e. the resource of first choice based on the skills/competences available). Based on this information, the mix of resources and their allocation are adjusted until the feasibility of the resource schedule is reached in the real conditions (i.e. each workload unit is allocated to an available resource), as represented by the process defined in the simulation tool. The feasibility of the allocation plan is simply the ability to meet customer demand to a user-defined extent (e.g. 90% of customers serviced with wait time under 7 minutes). The performance measures extracted from the simulation tool for the feasible scenario are then used as input to the GAs to generate new scenarios that progressively move the performance of the solution from mere feasibility to maximum job satisfaction, in terms of schedule and duties, compatible with the expected workloads, and minimum cost, until the most suitable solution is found.

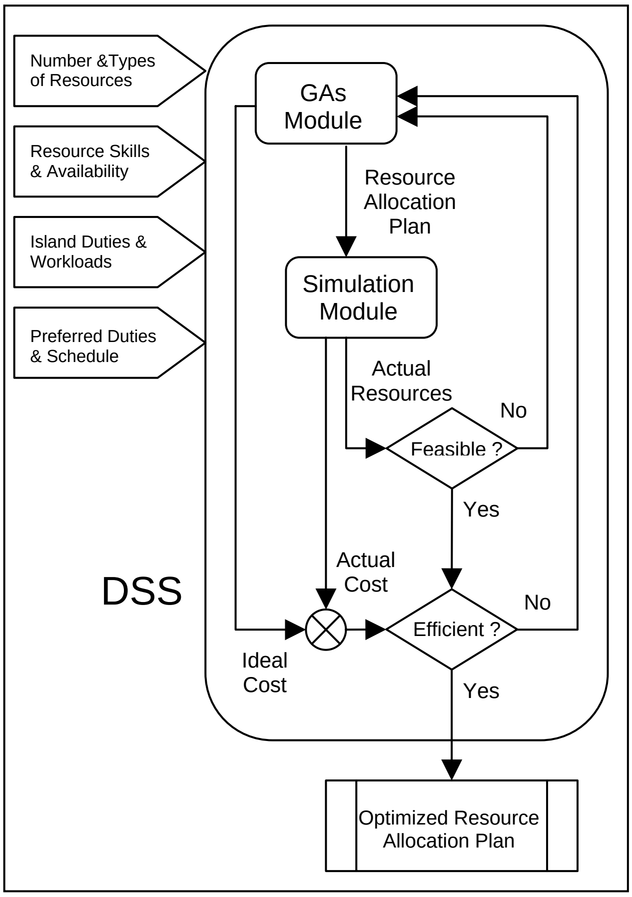

Figure 1: Functional Schematic of the Hybrid DSS

With reference to the case study, the measure of performance chosen as reference for the optimization process combines the percentage of unmet workers schedule and job preferences over the 2-week period and the extra cost to the company in terms of percentage resources overtime. The optimal solution will be the one capable of minimizing this performance measure. In the first instance this measure is taken as a weighed sum of the two performance drivers, assigning the same importance, namely weight, to each one of them.

The analysis conducted using the DSS leads to the identification of the most effective island design for the examined store, and to the formulation of guidelines and recommendations for future training and hiring of resources aimed at achieving the most effective mix.

THE IMPLEMENTED SIMULATION MODEL

The evaluation of a resources allocation plan with respect to its feasibility and efficiency requires a tool capable of simulating the various activities according to variable resource requirements, availability, and utilization. In order to support this testing phase a discrete-event simulation model was built in the C++ environment. The simulation model is quite simple in nature as it basically handles each service station as a queuing system subject to the stochastic variability of the number of servers available and of the respective service time. These aspects of the service could easily be modeled using standard process simulation packages, such as SIMPROCESS or ARENA, capable of handling queuing systems and stochastic variables. However it was preferred to implement the model in the C++ environment for flexibility and ease of integration with the GAs-based optimisation module.

In the model, the daily activities of the store are represented in terms of product handling and display, product preparation, customer service and cashier's duties across the main product categories of the store. Each of the available resources is characterized with a resource profile containing the suitability (yes/no) for each type of duty/product category and a corresponding skill level, namely a coefficient, capable of influencing the duration of each activity and, thereby, the number of customers serviced and/or the number of product units handled in the reference time unit. Specifically each activity is characterized by a reference triangular distribution representing its baseline time to complete, intended as an averaged distribution across the range of suitable resources. Whenever in the course of the simulation run a value of activity duration is extracted out of the baseline distribution, its value is scaled using the skill factor of the resource currently working, which in turn depends on the number of years of experience on the job as well as on the specific competences of such a resource as indicated in the resource's profile. This is the extent to which the individual resources are personalised in the simulation model: the model does not involve the detailed modeling of human resources in their multiple behavioural aspects, as the worker is only relevant to the testing of the resources allocation plan proposed by the GAs module in terms of his/her availability and service time. By talking to the store managers it was found that the optimal resolution in resource allocation testing is obtained considering a reference time unit of one hour, as it would not be efficient to consider resources re-allocation for shorter times. Finer scheduling resolution may be considered in the future for the management of store contingencies/emergencies, but these are not accounted for in the study at this stage. Prior studies conducted in the reference store, and additional data collection through their information systems database, enabled the re-construction of the dynamic patterns of customer flows across the different areas of the store. The main sources of information were the cashier's receipts of which both electronic and printed records are available in the store's database. The electronic records of cashier's receipts for the past two years were sorted to extract the number of customers accessing the store with its seasonal and daily fluctuations and the corresponding purchase patterns by product category. This data was then analysed and discussed with the store managers to re-construct reasonable patterns of workload by store area and duty to be used as reference in the simulation model. Seasonal, weekly, and daily patterns were extracted for two reference periods: Summer and Winter, broadly intended, as activities and customer flow patterns vary dramatically between the two periods in relation to tourism and seasonal changes in lifestyle. In the first instance the study did not address special periods such as major holidays, strikes, and other events, as it was mainly focused on the identification of generalized patterns for each one of the two broadly defined seasons. Special attention is paid to the handling of scheduling emergencies, intended as the unavailability of scheduled resources. For each resource an unavailability rate was defined based on cumulative store trends and estimates provided by the managers. For the purposes of the simulation model the resources unavailability rate was defined as:

$$1 - Av = \frac{MABS}{MTBE + MABS} \qquad (1)$$

where:

Av = resource availability rate
MABS = mean time of absence
MTBE = mean time between emergencies

At the beginning of each simulated work-shift the pool of scheduled resources actually available is updated using the Monte Carlo technique to extract punctual values from the corresponding MABS and MTBE distributions. For the purposes of the current work-shift the duties of the missing resources are shared and reassigned to the suitable resources actually present, with an impact on service efficiency (i.e. customers queuing time). For the purposes of the following work-shifts, for the entire duration of the resource MABS, a spare resource is introduced and the corresponding overtime costs are computed.

At the end of each simulation run a report file is generated, which records all the relevant simulation details. These include resource availability, percentage utilization and overtime costs incurred and quality of customer service as defined in the previous section (percentage of customers who waited less than 7 minutes for service). This set of information is fed back into the GAs module where it is used to compute a performance index for the given resource allocation plan.

EXAMPLE APPLICATION

The DSS was applied to perform a feasibility study for the implementation of the island-based organizational structure in the examined context (i.e. the reference store described in the previous sections). In particular the analysis was aimed at highlighting the potential benefits from the introduction of such a business structure, by comparing its performance to the average performance observed with the current organization. For the purposes of this analysis a reasonable extent of resources interchangeability was assumed across product categories within the three major types of activities: product handling and display, product preparation and customer service, while it was assumed, that cashier's duties would be shared among all the available resources who would dedicate approximately 20% of their time to cover cashier's workloads. Three product groups were defined aggregating product categories which have similar requirements in terms of resources skills. Based on such product grouping, six islands were defined as aggregation of duties and workloads, three pertaining to product handling and display (one for each product group) and three pertaining to customer service (again one for each product group). An additional island was defined grouping duties and workloads related to product preparation. This last island is treated separately from the others as only limited interchangeability is allowed for the resources involved in product preparation, because the activities are highly product-specific and require special resource training. In the future it may be possible that resources cross-training will lead to full interchangeability also in this area, but for the time being this option was not considered as it is not of immediate implementation, given the level of specialization of the current resources.

Based on the current pool of resources the GA-based module of the DSS was customized to manipulate binary input/output strings in which "1" bits indicate the presence of given resources for a given scheduling interval and "0" bits indicate the absence of the remaining resources for the same time interval. The objective function for the optimization process was set to be a percentage performance indicator combining the percentage workers satisfaction, in relation to their preferred duties and schedule, and the percentage of extra costs related to resources overtime. The number of strings for each island is equal to the number of scheduling intervals included in the scheduling horizon. For the purposes of this application the scheduling horizon was set to one day and the scheduling interval was fixed to one hour. The procedure was then iterated for each day in the 2-week period of analysis, as the store manager typically finalize their resource allocation schedules on a bi-weekly basis.

The logical steps for the optimization of the resource allocation schedule are explained in the following. The system is fed with the bi-weekly duties and workloads by island and with the corresponding preferences indicated by the resources. Based on this information the GAs-based module produces a first attempt resource allocation plan selecting the resource of first choice for each duty according to resources skills. Starting from this preliminary resource allocation plan, the GAs-based module proposes a schedule for the workloads of the different islands distributing them over the designated scheduling intervals. The scheduling constraints are chiefly related to the number of equally skilled resources available for island. If the resource of first choice is not available for all workloads on each scheduling interval, the original plan is not feasible and, thus, a resource of second choice is assigned to some of the given workloads. This process is iterated until a feasible solution of minimum cost is found. The resource allocation plan and schedule proposed by the GAs-based module are ideal ones because they are tested for feasibility and cost-effectiveness without considering the impact of resources skills on their productivity in relation to the actual customer flows in the store. The actual feasibility and cost of each solution can be verified through the discrete event stochastic simulation module, which explicitly accounts for these performance drivers. Realistic service times and costs are computed in the simulation model by associating a stochastic duration to each activity, which is further influenced by the specific resource's skill level, and availability.

APPLICATION RESULTS

The customization of the DSS, as described in the previous section, and its application to the analysis of the potential benefits from the implementation of an island-based organizational structure in the reference store, led to an interesting comparison between the current and the projected performance for the existing and the proposed structure, respectively. Three performance measures were defined for the purposes of this comparison. These include, quality of customer service, workers job satisfaction and overtime cost incurred. The quality of customer service is measured as the percentage of customers waiting less than 7 minutes for service. Worker's job satisfaction is measured in terms of percentage ability of the allocation plan to meet the workers' preferred work schedule and duties. Finally the impact of resources overtime cost is calculated as a percentage of the ideal cost of the service assuming that none of the resources worked overtime.

Prior to actually using the simulator to test alternative resources allocation plans, an estimate of the number of simulation replications needed per simulated scenario, was obtained using the classic methodology based on the analysis of the temporal evolution of the Mean Square pure Error (MspE). As mentioned in the previous paragraphs of this same section, for the examined application the length of the simulation run is fixed to fourteen days of activity as the actual planning horizon, which store managers typically refer to, is two weeks. The MSpE-based method was then applied to determine the number of replications needed to obtain sensible results out of the simulator or, in other words, the number of 2-week periods that need to be simulated for such purposes. Plotting of the temporal evolution of the MspE for the three designated output variables led to the identification of the optimal number of replications required, which was found to be 7. This may appear a relatively small number even considering the rather specialized and self-contained nature of the application context. However, it should be considered that 9 replications involve the simulation of 126 work-days, which in turn correspond to 2016 time units for scenarios involving two eight-hour work-shifts per day.

Figure 2 summarizes these results. As shown in the figure, the major benefits expected from the introduction of the optimized island-based structure in the examined store are a significant reduction in resource overtime cost (approximately 68 %), a relevant increase in workers job satisfaction (approximately 32%), and a marginal improvement in the quality of service (approximately 11%)

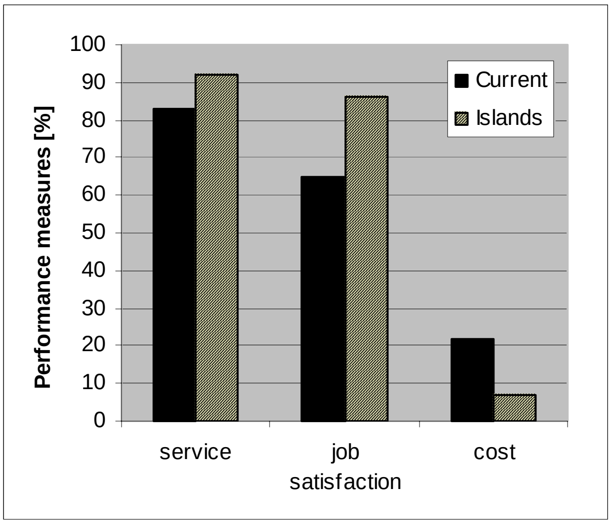

Figure 2: Performance Comparison between the Current and the Optimised Island-Based Structure

The output values shown in the figure correspond to the average value of each output calculated over the 9 simulated replications. The level of confidence associated to such estimates of the real output can be measured through the dispersion of the punctual data around each average value. For this scenario it was found that each individual output would fall within the range of ± 5.8 % of the corresponding average: specifically, ± 5.1 % for the quality of service, ± 4.6 % for job satisfaction, and ± 5.8 % for overtime cost.

It is important to observe that these results only represent the marginal improvements that may be achieved at zero cost, by simply revising the allocation criteria for the existing pool of resources. Far more interesting will be the benefits that may be obtained from the implementation of other, more efficient solutions proposed by the GAs module during the optimization process, which are currently not feasible due to the lack of skills/resources. Such solution may be implemented at the cost of resources cross-training and/or new hiring and proper cost-benefit analysis will be conducted in the course of further research to provide management with more effective recommendations for the successful implementation of the island-based structure.

CONCLUSION

The paper presented a hybrid decision support system (DSS) to optimize resources allocation in island-based organizational structures. The DSS architecture integrates GA-based optimization and stochastic discrete event simulation to assess the feasibility and the effectiveness of dynamic resource allocation plans. The paper also illustrated the application of the DSS to the conversion of the existing organization of a reference retail store to an island-based structure. The customization of the tool for the example application, led to an operative definition of islands, based on the actual types of workload and skill requirements, and to the formalization of an improved resource allocation methodology capable of best meeting workers preferred schedules, while minimizing service cost. The experimental results for this application demonstrate the robustness of the methodology and its applicability to a wide range of business organization and logistic management problems.

REFERENCES

Anderson, J. and E. Rosenfeled. 1988. *Neurocomputing - Foundation of Research*. MIT Press, Cambridge, MA.

Bruzzone, A.G.; R. Mosca; A. Orsoni,; and R. Revetria. 2001. “Forecasts Modelling in Industrial Applications Based on AI Techniques”, In *Proceedings of CASYS2001* (Liege, Belgium).

Bruzzone, A.G.; R. Mosca; R. Revetria; and A. Orsoni. 2002. “System Architecture for Integrated Fleet Management: Advanced Decision Support in the Logistics of Diversified and Geographically Distributed Chemical Processing”. In *Proceedings of the 2002 Conference on AI, Simulation and Planning in High Autonomy Systems* (Lisbon, Portugal).

Bruzzone, A.G. and R. Signorile. 1998. “Simulation and GAs for Ship Planning and Yard Layout.” *SIMULATION*, *71*(2), 74-83.

Giribone, P.; A.G. Bruzzone; M. Antonetti; and G. Siciliano. 1997. “Modelling Innovative Energy Management Techniques in Telecommunication Stations Through the Application of Neural Models”. In *Proceedings of the 1st World Congress on Systems Simulation* (Singapore).

Giribone, P. and A.G. Bruzzone.1997. “Artificial Neural Networks as Adaptive Support for the Thermal Control of Industrial Buildings.” *International Journal of Power and Energy Systems, 19*(1), 75-78.

Giribone, P. and A.G. Buzzone. 1998. “Artificial Neural Networks as Meta-Modelling Techniques for Car Component Industrial Production”. In *Proceedings of Neurap 98*. (Marseilles, France).

Goldberg, D.E. 1989. *Genetic algorithms in search, optimization, and machine learning.* Addison-Wesley, Reading, MA.

Hillis, D.W. 1989. *The connection machine*. MIT Press, Cambridge, MA.

Koza, J.R. 1992. *On the Programming of Computers by Means of Natural Selection.* MIT Press, Cambridge, MA.

Mosca, R., P.Giribone, A.G.Bruzzone, A.Orsoni, and S. Sadowski, 1998. “Evaluation and Analysis by Simulation of a Production Line Model Built with Back-Propagation Neural Networks, *International Journal of Modelling and Simulation, 17*(2), 72-77.

Orsoni, A. 2000. “Dynamic Process Simulation for the Design of Complex Large-Scale Systems with Respect to the Performance of Multiple Interdependent Production Processes”. Doctoral Thesis. MIT, Cambridge, MA.

Padgett, M.L. and T.A. Roppel. 1992. “Neural Networks and Simulation: Modeling for Applications” *Simulation 58*(5).

Prickett, P. 1994. “Cell-Based Manufacturing Systems: Design and Implementation.” *International Journal of Operations & Production Management 14*(2). 4-17.

Schonenberg, R. 1986. “World Class Manufacturing: the Lessons of Simplicity Applied.” Free Press, New York, NY.

Slack, N. 1988. “Manufacturing Systems Flexibility: an Assessment Procedure.” Journal of Computer-Integrated Manufacturing Systems. Vol 1. 25-31.