HIERARCHICAL CONTROL FOR CONTINUOUS FLOW SIMULATION OF MANUFACTURING SYSTEMS

Karim Tamani[1], Reda Boukezzoula[1], Georges Habchi[2]

[1]University of Savoie, Polytech'Savoie, [1]LISTIC Laboratory,
[2]SYMME Laboratory,
Domaine Universitaire – BP 80439, 74944 Annecy–Le–Vieux Cedex France

{karim.tamani, reda.boukezzoula, georges.habchi}@univ-savoie.fr

Abstract

This paper aims at developing hierarchical control architecture for continuous-flow simulation in order to regulate production in multiple-part-type flow shops. The approach uses a continuous-flow approximation to model the discrete flow of parts in a manufacturing system. The control strategy specifies how to allocate limited system capacity among all the part types to follow the solution of the continuous-flow model as closely as possible. The control objectives are to keep the actual production close to the demand, while maintaining the average work-in-process inventory and lead time to satisfactory levels. The control architecture is hierarchical. It allows combining different decisions into a unified model. This architecture is composed of basic-level distributed fuzzy logic controllers supervised by a higher level decision-maker. At the bottom level of the hierarchy, individual decisions are based on local information and expert's knowledge to adjust the machine's processing rate. At the top level of the hierarchy, the supervisory controller combines both local information and global performance indicators in order to tune the action of the lower distributed fuzzy controllers. The global performance indicator used in the supervisory level evolves in a tolerance interval defined by the normal operating conditions of the process. When a performance indicator value is outside of the predefined tolerance interval, an abnormal behaviour occurs. In this case, the supervisor allocates the production capacity or reduces the production throughput according to the aggregated global performance indicators. Simulation results through continuous-flow simulator of production network are presented to illustrate the feasibility of the proposed approach.

Keywords: Manufacturing system, fuzzy control, supervisory control, performance indicators, aggregation operator, continuous-flow simulation.

Presenting Author's biography

Karim Tamani is PhD candidate in Automatic and Electrical Engineering at the University of Savoie. The main subject of his thesis includes the development of fuzzy control methodology based on the fusion of performance indicators in the simulation of manufacturing systems.

1 Introduction

A manufacturing system characterized by its stochastic nature, is modelled by both qualitative and quantitative variables. Usually, the performance of manufacturing systems must be enhanced by optimising performance measures such as throughput, work-in-process (WIP) and cycle time with respect to some relevant decision variables. In this case, it is generally convenient to describe a manufacturing system in the form of an analytical model to get the accurate solutions as quickly as possible. However, as the complexity of the system increases, it gets more and more difficult to accommodate it in the analytical model due to the involved uncertainty. In such situations, the resort to simulation has been shown to be an effective alternative [11].

Simulation is an essential tool for the design and analysis of complex systems that cannot be easily described by analytical models. It can be discrete or continuous depending upon the way that the flow of parts is modelled in a manufacturing system. Pritchett et al. [] described the fundamental differences between discrete-event and continuous-time models. It has been proven possible to incorporate models of either type in the simulation software intended for the other. Indeed, in discrete-event simulation the state variable which reflects the discrete nature of the operations takes a finite set of alternatives, whereas in continuous-flow simulation the state variable includes real values which are an approximation of the discrete flow of parts in manufacturing systems [14]. In this paper, particular attention is paid on continuous-flow simulation. Its efficiency over discrete simulation results from the reduced number of the observed events by the simulator particularly in the case of high-volume manufacturing with unreliable resources and finite capacities. In addition, the continuous-flow simulation may provide a framework to develop an accurate control strategy based on the control theory which is traditionally used in the continuous processes.

However, a simulation model mainly acts as a tool for performance analysis. It is essentially a trial and error methodology, and does not directly provide explanations for the observed system behaviour. These limitations are due to the facts that the decision-making process and its impact on the production system are not taken into account during the simulation [11,20]. In this field, concepts have been proposed to integrate a fully distributed control process in the simulation model, i.e., individual decision is based only on the local performance indicators [5,15]. In this case, the control strategy is driven by crisp control rules in the following form [11]:

IF the control objective given in terms of threshold is not satisfied, THEN apply the adequate action according to the predefined program based on the cause and effect relation.

Although such strategy may lead to a substantial reduction of the number of simulation runs allowing system optimisation, it suffers from the lack of an automatic integration of the knowledge system in the control process and its inability to handle multiple and possibly conflicting objectives. Furthermore, this strategy may lead to “chattering” phenomena identified by oscillation around the threshold value. Another drawback of this control strategy is due to the fact that the information concerning the overall production system performance (necessary for the operating system optimisation) is not integrated in the decision mechanism of control.

In order to improve the control process in the simulation, it has been shown in [16,21] the usefulness of fuzzy control approach in comparison to the crisp control particularly in reducing the chattering phenomena. Indeed, fuzzy control approaches provide gradual action rather than an abrupt one. Furthermore, fuzzy logic offers a good tool allowing the integration of human expert knowledge in the control procedure and deals with the uncertainty of the production environment [1,4,12,19]. In addition, the synthesis of a global controller based on the fusion of the multiple and possibly conflicting objectives concerning the overall production-system performance is necessary to guarantee the satisfaction of the system’s control objective [12].

Consequently, it can be concluded that an efficient control architecture must satisfy the following properties:

- a modular representation of the production process to be controlled,
- a decentralized local and autonomous controller based on local information able to cope with uncertainties,
- a centralized supervisory controller based on fusion of global information able to cope with the multiple and possibly conflicting objectives.

Based on these statements, the method developed in this paper uses a hierarchical structure consisting of a supervisor at the higher level and fuzzy local controllers at the lower level. In manufacturing control literature, several hierarchical control algorithms have been developed, mainly for scheduling and planning problems [1,3,4,10,12]. In most cases, the actions provided by the higher level of the control hierarchy depend only on the information of the overall system. This may not be an effective way to adapt the global action with the local one.

In the proposed control architecture, the supervisor decision mechanism combines both local and global information for producing an additive control component to the local control action. This action aims at adjusting the resource's processing rates with regard to its maximum production capacity in order to regulate the production flow at each production stage. The supervisor action allocates the remaining production capacity according to the global performance indicators. In other words, the supervisor evaluates the system state and adjusts production rate for each local controller according to the remaining capacity of the corresponding resource (local information).

The control action of the supervisor is defined according to the behaviour of each global performance indicator given by its tolerance interval (which defines operating conditions). Our objective is to propose a way to combine the different actions related to each performance indicator in order to determine the supervisor action. For this goal, an aggregation mechanism based on the weighted sum operator is proposed. The key issue is to determine the weights associated for each global performance indicator. This latter is evaluated according to the dissatisfaction degree of each control objective given by the metric between the performance indicator measure and its tolerance interval bounds. In this case, the larger the value of the metric is, the worse the control objective is and thus more important is the assigned weight.

To summarize, the originality of this work lies in the answers given to the following questions:

- Given a tolerance interval (objectives) specified by the designer on the performance indicators of the system ($P_1,\ldots,P_L$):

  $$P_l \in P_l^{obj} = [P_l^{\min}, P_l^{\max}],\; l = 1,\ldots,L$$

  is it possible to synthesize a global control strategy able to maintain the performance indicators of the manufacturing system within this tolerance interval? What is the best control structure to be adopted?

- How the performance indicators are combined to achieve the global control performances?

- How robust the proposed strategy is in the case of complex manufacturing systems and against random disturbances such as machine failures and demand fluctuations?

The remainder of this paper is organized as follows. In section 2, statements and model assumptions are given for the considered problem. In section 3, a local control strategy is developed for a single machine, multiple-part-type, and thereafter extended to the general manufacturing system. Next, the supervisory control strategy is detailed in section 4. Simulation results are given in section 5 along with comparisons between the proposed method and the fully distributed and the surplus-based control approaches [12]. Finally, concluding remarks are provided in section 6.

2 System description

The production network under examination consists of a number of N machines M_i (i =1,…,N) with M intermediate buffers B of finite capacity and producing P different part types (j = 1,…,P). Each part type requires a number of K_j operations in a given sequence (route). Each machine M_i may perform P_i (j=1,…,P_i; with $P_i \le P$) different part type j, each of them may need K_{ij} (k=1,…, K_{ij}) different operations on machine i. The buffers are homogeneous. One buffer is defined for each part type at each stage such that B_{ijk} denotes the buffer of the part type j after the kth operation on machine i. These notations allow us to model re-entrant flow [8], that is the same part type can visit the same machine more than once if necessary. The possibility of assembly and disassembly operations on each machine is also allowed. In the assembly operation, a machine M_i obtains two or more parts from more than one upstream buffer, assemble them into a single part, and send this latter to a downstream buffer. The disassembly operation involves a machine M_i tacking single parts from one upstream buffer, separates them to V (v=1,…,V) parts, and sends them to downstream buffers $B_{ijk}^{v}, v = 1,\ldots,V$ (for an example, see Fig. 5).

Let us assume the following properties:

- The manufacturing system is controllable and there exists some indicators $P_1,\ldots,P_L$ able to quantify the system performances.
- The time to fail and the time to repair are modelled by exponentially distributed random variables.
- The processing time for each part type is given and deterministic.
- The setup and transportation times are negligible or are included in the processing times.
- The demand rate d_j for each part type j is constant and prespecified.
- Each part type j is prespecified by a route or a sequence.

3 Distributed control architecture

The problem arising in the production network described above is the control job flow under limited capacity. This is in part due to the combinatorial nature of the decision problem. Consequently, policies used in practice are often of the distributed type and based on the decomposition topologies. For this latter, according to the production floor modelling approach introduced and explained in [19], every manufacturing system may be divided into basic elementary sub-systems regarding the production flow, which include transfer, assembly and disassembly subsystems [19,9,13]. Each of the three elementary subsystems may be implemented in terms of fuzzy controllers [19].

To make clear how the distributed control strategy is designed, the basic idea is first illustrated through the elementary production resource: a single machine involving transformation process with one upstream buffer and one downstream buffer.

3.1 The continuous-flow model for a single machine, multiple-part-type

In production flow approaches, for an efficient implementation, the part movement is processed as a continuous-flow model so that the system dimension is reduced [14]. This approach is used in this paper where the time is discretized, i.e., t_n is defined as the *n*th discrete instant.

In the multiple-part-type systems, the distribution of machine operation times to the different part types, and consequently the decision of the capacity allocation for each part type, is a difficult task. In this case, in order to reduce the problem complexity, each machine M_i is "virtually" divided in as many sub-machines (or *partial machines* as defined in [3]) as the number of the operations of the different part type to be performed in the original machine.

In other words, the original multiple-part-type machine M_i which performs P_i different part type, each one involves K_{ij} operations, is divided into $\sum_{j=1}^{P_i} K_{ij}$ single-part-type sub-machines.

Let us denote by m_{ijk} the sub-machine of M_i which performs the *k*th operation on the part type *j*. Then:

- τ_{ijk} : the processing time of the *k*th operation of part type *j* on M_i.
- $u_{ijk}^{\max}$: the maximum rate at which the sub-machine m_{ijk} can process a *k*th operation of part type *j*.

Then, the production rate of the sub-machine m_{ijk} performing the *k*th operation of the part type *j*, is given by:

$$u_{ijk}(t_n) = r_{ijk}(t_n)u_{ijk}^{\max} \qquad (1)$$

with $r_{ijk}(t_n) \in [0,1]$, is the control variable weighting the maximum production rate $u_{ijk}^{\max}$ allocated for processing part's *j* operation *k* on machine *i*. The maximum possible production rate of submachine m_{ijk} is given as follow:

$$u_{ijk}^{\max} = \left(\frac{d_j \tau_{ijk}}{\sum_{j=1}^{P_i} \sum_{k=1}^{K_{ij}} d_j \tau_{ijk}} \right) \left(\frac{1}{\tau_{ijk}} \right) \qquad (2)$$

The product $d_j\tau_{ijk}$ gives the required production to satisfy the demand of part type *j*. The sum $\sum_{j=1}^{P_i} \sum_{k=1}^{K_{ij}} d_j \tau_{ijk}$ is the total required production to satisfy all part types demand on the original machine M_i. This sum should be less than or equal to the availability of M_i.

The time availability of M_i is given according to its repair and failure rate g_i and f_i respectively.

Let us define the state of the original machine M_i at time t_n by a binary variable $\alpha_i(t_n)$ such that: when the machine is down, $\alpha_i(t_n) = 0$, otherwise $\alpha_i(t_n) = 1$. The delay for M_i being down and up is defined according to an exponential distribution with the averages of $1/f_i$ and $1/g_i$, respectively. In this case, the availability of M_i when taken in isolation is given by:

$$e_i = \left(\frac{g_i}{g_i + f_i} \right) \qquad (3)$$

Then the total demand is feasible if and only if:

$$\sum_{j=1}^{P_i} \sum_{k=1}^{K_{ij}} d_j \tau_{ijk} \le e_i \,, i = 1, \ldots, N \qquad (4)$$

This is a convenient way to check the feasibility of the demand.

The sub-machine m_{ijk} must be up and down at the same time as M_i. Then according to the machine's state, the capacity constraint is given by:

$$\begin{aligned} &if\ \alpha_i(t_n) = 0, \quad u_{ijk}(t_n) = 0. \\ &if\ \alpha_i(t_n) = 1, \quad \sum_{j=1}^{P_i} \sum_{k=1}^{K_{ij}} u_{ijk}(t_n)\tau_{ijk} \le 1; u_{ijk}(t_n) \ge 0. \end{aligned} \qquad (5)$$

The cumulative production of part type *j* at the end of the *k*th operation on machine *i* is defined as:

$$y_{ijk}(t_{n+1}) = y_{ijk}(t_n) + u_{ijk}(t_n)\delta t_n \qquad (6)$$

where $\delta t_n = t_{n+1} - t_n$.

Buffers are located between two consecutive operations. Buffer B_{ijk} holds only parts of type *j* just gone through the *k*th operation on m_{ijk}. Then, the buffer levels are given by:

$$x_{ijk}(t_{n+1}) = x_{ijk}(t_n) + [u_{ijk}(t_n) - u_{ljm}(t_n)]\delta t_n \qquad (7)$$

where u_{ljm} define the production rate of the next operation of part type *j* taken from buffer B_{ijk}, with the possibility of $i=l$ and $k \neq m$ since re-entrant flow is allowed. The buffer level must satisfy:

$$0 \le x_{ijk}(t_k) \le x_{ijk}^{\max} \qquad (8)$$

where $x_{ijk}^{\max}$ the maximum capacity of buffer B_{ijk}.

Let us define the *system surplus* as:

$$s_{ijk}^{K_j}(t_{n+1}) = s_{ijk}^{K_j}(t_n) + [u_{ijk}^{K_j}(t_n) - d_j]\delta t_n \,; \forall j \qquad (9)$$

It defines the set of cumulative difference between the production and demand at the last machine (or last operation) of the production process of part type *j*. If it is positive, it represents the inventory surplus of part type j and if it is negative that means a backlog has occurred.

The concept of the system surplus may be generalised by defining the *local surplus*:

$$s_{ijk}(t_n) = y_{ijk}(t_n) - d_j t_n \quad (10)$$

It defines the difference between the total number of parts of type j that have had their kth operation at time step t_n on machine i and the cumulative demand for part type j up to that time. The local surplus summarizes the downstream information in a way that helps decentralized decision making [8].

3.2 Capacity allocation

Given the dynamic of a single machine M_i as described in the previous section, the local control objectives reside in:

- Reducing the difference between the cumulative production and demand (the surplus),
- Avoiding overloading and eliminating machine starvation or blocking,
- Maintaining high machine utilisation.

These objectives may have achieved when regulating the production flow through each machine by allocating the appropriate capacity for each part type. Since we split (virtually) the original machine M_i in as many sub-machines as the number of the operations to be performed on it, thus the sub-machine m_{ijk} represents the controller of the original machine M_i regulating the execution of the operation k on parts of type j.

When considering this problem, there are quit few papers dealing with analytical methods to obtain optimal or approximately optimal solution. Most of them address only single machine models requiring no state constraint [6]. Flow shop and re-entrant system with multiple-part-type are much harder to solve. Approximate computational methods have been developed [3,10], but only two and three machine systems are studied. Thus, it appears that obtaining optimal solutions for large, complex systems is not a realistic goal. Since neither analytical nor computational solutions are achievable, simulation techniques are suggested to control job flow within production systems [11,14].

However, in this case, when considering discrete event simulation, it may require extremely long computation times. The situation becomes worse in the optimization problems where many runs are needed to evaluate various candidate designs. In [2], an efficient continuous-flow simulator is proposed for discrete-part production lines. Its efficiency over piece-by-piece methods results from the reduced number of events that are observed by the simulator, namely a machine fails or recovers and a buffer fills or empties. The system evolution between successive events is tracked analytically. In [14] production networks are analyzed by considering two approximations. The first one convert the random processing times into piecewise deterministic variables, while the second one approximates the discrete traffic by continuous-flow. An immediate implication of the second approximation is that when the buffer becomes full or empty, its inflow or outflow rate is reduced instantly whereas the flow rate increases instantly when the buffer becomes not-full or not-empty. Compared to the conventional simulator, the model combines accuracy with speed for a wide range of network topologies and parameters.

Another drawback in the current use of the simulation is the lack of an automatic integration of the knowledge system based on the human experts, particularly in the presence of multiple and possibly conflicting objectives. Indeed, in order to achieve satisfactory surplus, one has to reduce the production throughput if an important finished inventory (positive surplus) is occurred. However, this may lead to low machine utilization. In the other hand, if a backlog is occurred (negative surplus), one has to increase the production throughput; as a result there is an immediate improvement of the machine utilisation. Unfortunately, this is achieved at the cost of increased inventory which may lead to some throughput bottleneck.

In this field, fuzzy control theory offers a good tool allowing the integration of human expert knowledge in the control procedure which makes a system work with practical rules. Indeed, it has been shown the usefulness of the fuzzy control approach in comparison to the simple crisp control particularly in handling conflicting objectives. Indeed, in the control flow problem’s, the expert is sensitive to the events that may affect the buffer levels, surplus and the production rate. In this case, the buffers may be “Empty”, “Almost Empty”, “Normal”, “Almost Full” or “Full”, while the surplus may be “Negative” (backlog), “Zero” (demand satisfied), and “Positive” (important finished inventory). According to his perception, the expert determines the action to be adopted concerning the value of production rate: “Maximum production”, “Large production”, “Average production”, “Small production” or “Stop production”. These terms can be interpreted with regard to the maximum production capacity of the machine by constant values giving the fraction of the machine devoted to the processing: 1 (100%), 0.75 (75%), 0.5 (50%), 0.25 (25%) and 0 (0 %).
The general tendency of the previous statements can be summarized as follows:

- If the surplus level is satisfying, then try to prevent starving or blocking by increasing or decreasing the production rate of the machine.

- If the surplus is either too low or too high, then produce respectively with maximum or zero rate.

This heuristic strategy is applied to the continuous-flow simulation process which is more convenient to cope with the combinatorial nature of the problem. In addition, fuzzy control theory is more easily used in the continuous processes rather than discrete one.

3.3 Local fuzzy controller synthesis

Observing the numerical nature of the control output, the above knowledge can be formalized as a Takagi-Sugeno fuzzy system [17] for each sub-machine. The rules base for the case of transformation process in Fig. 1 is given by a collection of rules in the following form:

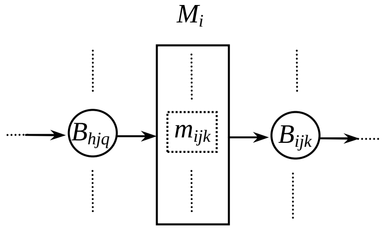


Fig. 1 Virtual sub-machine.

$$R_{ijk}^{(i_1,i_2,i_3)}: \quad IF\ x_{hjq}\ is\ X_1^{i_1}\ and\ x_{ijk}\ is\ X_2^{i_2}\ and\ s_{ijk}\ is\ X_3^{i_3}$$
$$THEN\ r_{ijk} = \phi_{ijk}^{(i_1,i_2,i_3)}.$$

where:

- x_{hjq}, x_{ijk} are respectively the levels of the upstream and downstream buffers of sub-machine m_{ijk} and s_{ijk} the local surplus given according to (10).
- $X_l^{i_l}\ (l = 1,2,3)$ is the i_lth linguistic term of the lth input variable $x_l\ (x_1 = x_{hjq}, x_2 = x_{ijk}$ and $x_3 = s_{ijk})$ of the fuzzy controller of m_{ijk}.
- $\phi_{ijk}^{(i_1,i_2,i_3)}$ is the real value involved in the rule conclusion.

The output generated by the fuzzy controller $0 \le r_{ijk}(t_n) \le 1$ is given by:

$$r_{ijk}(t_n) = \begin{cases} 0 & \text{if } \alpha_i(t_n) = 0 \\ FC_{ijk}(x_{hjq}, x_{ijk}, s_{ijk}) & \text{if } \alpha_i(t_n) = 1 \end{cases} \tag{11}$$

where

$$FC_{ijk}(x_{hjq}, x_{ijk}, s_{ijk}) = \sum_{(i_1,i_2,i_3)\in I} \xi^{(i_1,i_2,i_3)}(x_{hjq}, x_{ijk}, s_{ijk}) \cdot \phi_{ijk}^{(i_1,i_2,i_3)}$$

and:

- $\xi_{ijk}^{(i_1,i_2,i_3)}(x_{hjq}, x_{ijk}, s_{ijk}) = \prod_{l=1}^{3} \mu_{X_l^{i_l}}(x_l)$ represents the truth value of the set of rules $R_{ijk}^{(i_1,i_2,i_3)}$.
- $\mu_{X_l^{i_l}}(x_l)$ is the degree of the membership function of x_l in $X_l^{i_l}$.
- $I = I_1 \times I_2 \times I_3$ indicates the set of labels representing the rule base.

The partitioning of different universe of discourse associated with the input variables of the controller is illustrated in Fig. 2.

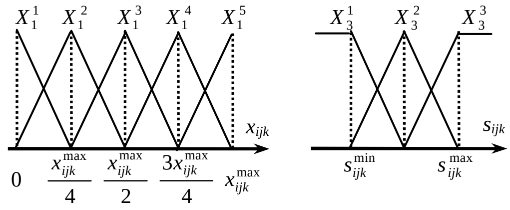


Fig. 2 Fuzzy partitioning of universe of discourse.

3.4 Distributed fuzzy control architecture

Let us consider a general manufacturing system composed of N subsystems, i.e. N machines interconnected by M intermediate buffers. The concept of fuzzy controller detailed in the previous section is exploited to control each sub-system in order to ensure local control performances. In this case, the control structure can be viewed as a distributed fuzzy control system.

The major advantage of this control architecture resides in its modularity and distributivity able to enhance the flexibility of the system and to make easy the implementation phases in complex manufacturing systems.

Although the distributed structure can give sometimes good control performance results [19], it does not guarantee optimal control performances since the information concerning the overall production system is not integrated in the local control. Indeed, the production objectives to be satisfied are measured in term of global performance indicators which are often contrasting (e.g., reduce the tardiness, increase the system's throughput, decrease the WIP). So, a fusion methodology between different objectives has to be sought. In order to achieve this control goal, higher supervisor based on global performance indicators aggregation is proposed.

4 Supervised control architecture

In order to achieve the global control performances, the distributed control structure is augmented by a supervisor, placed on the top of the local controllers. In previous work, fuzzy supervisory controller has been proposed in [12]. The authors attempt to minimize the WIP, cycle time and backlog by controlling the material released into the system. In this case, the overall production control system is viewed as a surplus-based control [8], i.e., the decisions are made on the basis of how far the cumulative production is ahead or behind the cumulative demand. The approach that we use in this paper is based on the last idea.

Although the method shows significant improvements in the control performances, it does not take into account the evolution of internal system dynamics (related to local information) which can generate problems such as throughput bottleneck. Then, in this paper, a new supervisor synthesis based on a capacity allocation principle is proposed [18]. In this case, the supervisor evaluates the system state and recomputes the production rate for each local controller according to the remaining capacity of the corresponding machine (local information). In contrast to the approaches developed in [12], our supervision decision mechanism combines both local and global information for producing an additive control component to reinforce the local control action, and thus modifies the evolution of the internal dynamics system.

4.1 Supervisor performance indicators

The global control objectives consist in:

- Tracking the surplus between the final cumulative production and demand to avoid backlog (negative surplus) or large finished inventory (positive surplus).
- Reducing the total average WIP as much as possible.
- Keeping the cycle time as low as possible.

These control objectives are related to each part type. Then we need as many supervisory controllers as the number of part types to be produced. In order to satisfy the control objectives, let us define the global performance indicators P_{jl} related to the part type j , as follow:

- P_{j1} : The mean system surplus of the end product of part type j, $\bar{s}_{ijk}^{K_j}$.
- P_{j2} : The instantaneous system surplus of the end product of part type j, $s_{ijk}^{K_j}$.
- P_{j3} : The relative error e_{wip_j} of WIP of part type j, given by:

$$e_{wip_j} = \frac{\mathrm{WIP}_j(t_n) - \overline{\mathrm{WIP}}_j(t_n)}{\overline{\mathrm{WIP}}_j(t_n)} \tag{12}$$

where:

$$\mathrm{WIP}_j(t_n) = \sum_{\{i | i \in I_j\}} \sum_{k=1}^{K_{ij}} x_{ijk}(t_n) \tag{13}$$

and:

$$\overline{\mathrm{WIP}}_j(t_n) = \frac{1}{t_n} \sum_{l=0}^{n} \mathrm{WIP}_j(t_l)\delta t_l \tag{14}$$

where I_j is the index set of the machines of the process sequence of part type j.

- P_{j4} : The relative error of lead time T_j for part type j, computed as in (12), where the average lead time is given by [3]:

$$\overline{T}_j(t_n) = \sum_{\{i | i \in I_j\}} \sum_{k=1}^{K_{ij}} \left[\frac{\overline{x}_{ijk}(t_n)}{d_j} + \tau_{ijk} \right] \tag{15}$$

The average lead time (sometimes called cycle time) for a part to go through the production system includes waiting times in buffers and processing times on machines. Since the part type j travels through the system with an average rate d_j (the local control objective), then $\overline{x}_{ijk}(t_n) / d_j$ denotes the average waiting time for part type j in buffer B_{ijk}.

Both indicators P_{j1} and P_{j2} are used to keep the production close to the demand (reduce the backlog or the inventory surplus), while P_{j3} and P_{j4} are used to maintain the WIP and the lead time of parts of type j as low as possible.

As stated in [12], the WIP relative error is used as a measure of WIP performance, since an analytical measurement of the optimal WIP cannot be estimated. This is based on the assumption that WIP needed to smooth operations is approximately equal to its mean value, and large deviation from it should be avoided. The same observation holds for the lead time since it is proportional to the WIP.

The mean value of the final surplus P_{j1} is computed as in (14).

Each global performance indicator P_{j1}, P_{j2}, P_{j3} and P_{j4} is defined with its tolerance interval:

$$P_{jl}^{obj} = [P_{jl}^{\min}, P_{jl}^{\max}], \ (l = 1, 2, 3, 4)$$

Given an interval P_{jl}^{obj} , its *Midpoint* $Mid[P_{jl}^{obj}]$ and its *Radius* $Rad[P_{jl}^{obj}]$ are defined by:

$$Mid\left[P_{jl}^{obj}\right] = \frac{\left(P_{jl}^{\min} + P_{jl}^{\max}\right)}{2}, Rad\left[P_{jl}^{obj}\right] = \frac{\left(P_{jl}^{\max} - P_{jl}^{\min}\right)}{2}$$

where:

$$P_{jl}^{obj} = [P_{jl}^{\min}, P_{jl}^{\max}] = \left[Mid[P_{jl}^{obj}] - Rad[P_{jl}^{obj}], Mid[P_{jl}^{obj}] + Rad[P_{jl}^{obj}]\right]$$

The supervisor aims at maintaining each performance indicator P_{jl} inside its tolerance interval.

4.2 Supervision based mechanism

The supervisor is built according to the global performance indicators and their admissible zones defining the operating modes of the manufacturing system (normal or abnormal mode). As illustrated in Fig. 3, the supervision mechanism is based on:

- An operating mode selector (normal or abnormal) which determines the supervisor action associated to each performance indicator.

- An aggregation mechanism which computes the different actions to produce the global supervisor action.

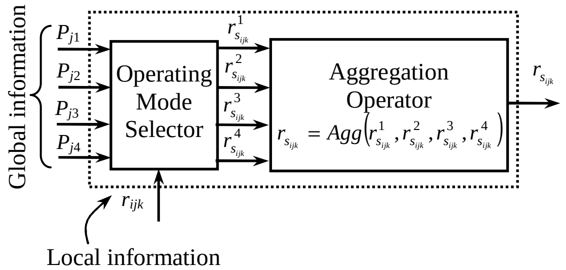


Fig. 3 The synopsis of the supervisory control mechanism.

When considering a local controller FC_{ijk} producing a local action r_{ijk}, the supervisor objective is to augment this local action by an additive component $r_{s_{ijk}}$ according to the aggregated performance indicators, and the system's state behaviour.

The controlled manufacturing system behaviour is in normal mode if:

$$\forall l, P_{jl} \in P_{jl}^{obj} = [P_{jl}^{\min}, P_{jl}^{\max}], \forall l \tag{16}$$

This condition can be rewritten as:

$$P_{jl} \in P_{jl}^{obj} \Leftrightarrow \begin{cases} P_{jl} \geq P_{jl}^{\min} \Rightarrow P_{jl} \geq Mid\left[P_{jl}^{obj}\right] - Rad\left[P_{jl}^{obj}\right] \\ P_{jl} \leq P_{jl}^{\max} \Rightarrow P_{jl} \leq Mid\left[P_{jl}^{obj}\right] + Rad\left[P_{jl}^{obj}\right] \end{cases}$$

which can be reformulated as:

$$\forall l, \left| P_{jl} - Mid[P_{jl}^{obj}] \right| \leq Rad[P_{jl}^{obj}], \forall l \tag{17}$$

In this case, the supervisor does not modify the local control law (no supervisor action).

On the other hand, if the condition (17) is violated for at least one performance indicator, i.e.

$$\exists l, P_{jl} \notin P_{jl}^{obj} = [P_{jl}^{\min}, P_{jl}^{\max}], \forall l \tag{18}$$

That is:

$$\exists l, P_{jl} \notin P_{jl}^{obj} \Leftrightarrow \begin{cases} P_{jl} < P_{jl}^{\min} \Rightarrow P_{jl} < Mid\left[P_{jl}^{obj}\right] - Rad\left[P_{jl}^{obj}\right] \\ P_{jl} > P_{jl}^{\max} \Rightarrow P_{jl} > Mid\left[P_{jl}^{obj}\right] + Rad\left[P_{jl}^{obj}\right] \end{cases}$$

which can be given by:

$$\exists l, \left| P_{jl} - Mid[P_{jl}^{obj}] \right| > Rad[P_{jl}^{obj}], \forall l \tag{19}$$

The system switches on the abnormal mode. In this context, according to the performance indicator P_{jl}, the supervisor provides the appropriate action $r^l_{s_{ijk}}$ related to the operating zones as shown in Fig. 4.

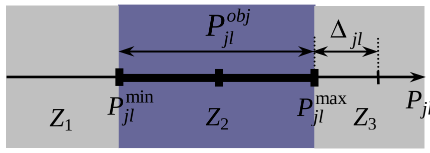


Fig. 4 The different operating modes.

The supervisory control action decides to either increase the production capacity or reduce the production throughput according to the remaining capacity at each machine.
If P_{jl} evolves in its admissible zone (Z_2), the system is then in normal mode and no supervision action is given:

$$P_{jl} \in Z_2 \Leftrightarrow r^l_{s_{ijk}}(t_n) = 0 \tag{20}$$

However, if P_{jl} value is located in the non admissible zones Z_1 or Z_3, the supervisor gives an action to the local controller according to whether there is an advance (Z_3) or a delay (Z_1) with regard to the admissible zone (Z_2). When the indicator is located in Z_3:

$$P_{jl} \in Z_3 \Leftrightarrow P_{jl} > P_{jl}^{\max} \Rightarrow P_{jl} - Mid[P_{jl}^{obj}] > Rad[P_{jl}^{obj}]$$

In this case, the supervisor action is given by:

$$r^l_{s_{ijk}}(t_n) = -r_{ijk}(t_n) \tag{21}$$

that provides an inverse action of the local controller in order to stop the processing of the operation k of part type j on machine i.
In the other case, when the indicator is located in Z_1:

$$P_{jl} \in Z_1 \Leftrightarrow P_{jl} < P_{jl}^{\min} \Rightarrow P_{jl} - Mid[P_{jl}^{obj}] < -Rad[P_{jl}^{obj}]$$

this means that there is a delay in production, then the supervisor action is given as:

$$r^l_{s_{ijk}}(t_n) = 1 - r_{ijk}(t_n) \tag{22}$$

This action attempts to allocate the remaining capacity, if any, in order to maximise the production of the submachine m_{ijk}.

A supervisor control action is defined according to the behaviour of each performance indicator P_{jl}. Our objective is to propose a way combining the different behaviours related to each performance indicator in order to determine the adapted supervisor action.

For this goal, an aggregation mechanism based on the weighted sum operator is proposed. The key issue is to determine the weights associated for each performance indicator. In our case, each weight is evaluated according to the dissatisfaction degree of each control objective given by the metric between the measured performance indicator and its tolerance interval bounds (Fig. 4).

ISBN 978-3-901608-32-2

Formally, let us consider $r_{s_{ijk}}$ as the aggregated result of the different actions $r^{l}_{s_{ijk}}$ associated to the performance indicator P_{jl}, and let us define:

$$\Delta_{jl} = \begin{cases} \left|P_{jl} - P_{jl}^{\min}\right| & \text{if } P_{jl} - Mid\left[P_{jl}^{obj}\right] < -Rad\left[P_{jl}^{obj}\right] \\ \left|P_{jl} - P_{jl}^{\max}\right| & \text{if } P_{jl} - Mid\left[P_{jl}^{obj}\right] > Rad\left[P_{jl}^{obj}\right] \\ 0 & \text{if } \left|P_{jl} - Mid\left[P_{jl}^{obj}\right]\right| \le Rad\left[P_{jl}^{obj}\right] \end{cases}$$

which evaluates the metric between the measure of the performance indicator P_{jl} and its tolerance interval bounds (Fig. 4). In this case, the larger the value of Δ_{jl} is, the worse the performance of P_{jl} is and thus more important is the assigned weight with $r^{l}_{s_{ijk}}$.

Then, the weight associated for the action related to the performance indicator P_{jl} is given as follow:

$$w_{jl} = \frac{\Delta_{jl}}{\sum_{l=1}^{4} \Delta_{jl}}$$

where the two following properties are satisfied:

- $\sum_{l=1}^{4} w_{jl} = 1$;
- $w_{jl} \in [0,1],\ l = 1,\ldots,4$

Then the supervisory action is given as the weighted sum of the different actions related for each objective as follow:

$$r_{s_{ijk}}(t_n) = \sum_{l=1}^{4} w_{jl} \cdot r^{l}_{s_{ijk}}(t_n) \qquad (23)$$

Note that the supervisory action is restricted to the operating machines that have not reached their maximum capacity limit, i.e., for that where $0 < r_{ijk} < 1$.

4.3 Supervisory allocation capacity

In order to allocate capacity for each part type, the supervisory controller first evaluates the remaining capacity at the original machine M_i. This is done as follow:

$$RT_i(t_n) = \left(1 - \sum_{\{j \mid r_{ijk}(t_n)=1\}} \sum_{k=1}^{K_{ij}} u_{ijk}^{\max} \tau_{ijk} r_{ijk}(t_n) - \sum_{\{j \mid 0<r_{ijk}(t_n)<1\}} \sum_{k=1}^{K_{ij}} [r_{ijk}(t_n) + r_{s_{ijk}}(t_n)] u_{ijk}^{\max} \tau_{ijk}\right) \qquad (24)$$

The first sum in (24) represents the working time of the sub-machines on the original machine M_i which produce at their maximum capacity, while the second sum denotes the working time of the sub-machines for which there is some remaining capacity allocated by the supervisor.

The remaining machine capacity is distributed to the operations with the highest priority. The highest priority is given to the operations having r_{ijk} equal to one. This means that the specific operations should be processed in the maximum feasible rate [12]. Then, the extra machine capacity devoted to the operation with the highest priority is:

$$E_{ijk}(t_n) = RT_i(t_n) \cdot \left(\frac{d_j \tau_{ijk} \beta_{ijk}(t_n)}{\sum_{j=1}^{P_i} \sum_{k=1}^{K_{ij}} d_j \tau_{ijk} \beta_{ijk}(t_n)} \right)$$

where:

$$\beta_{ijk}(t_n) = \begin{cases} 0 & \text{if } r_{ijk}(t_n) < 1 \\ 1 & \text{if } r_{ijk}(t_n) = 1 \end{cases}$$

identifies the operations with the highest priority as it has been stated in [12]. In this case, the production rate of the submachine m_{ijk} is:

$$u_{ijk}(t_n) = r_{ijk}(t_n) u_{ijk}^{\max} + E_{ijk}(t_n) \cdot \left(\frac{1}{\tau_{ijk}} \right) \qquad (25)$$

The production rate given by (25) is a continuous time one. Therefore, some dispatching strategy has to be employed to determine the loading times of the parts for each machine. For this purpose, the heuristic known as the *staircase strategy* is used [7]. That is, whenever the actual cumulative production is less than the expected production (the integral of the production rate), load a part type j into the machine M_i. If more than one part type is eligible for loading, choose the one which is farthest behind [3].

Finally, the steps of the supervisory based allocation capacity are summarized as follow:

Step 1. Evaluate the performance measures of the global indicators P_{jl} with regard to their tolerance interval P_{jl}^{obj} (j=1,…,M ; l=1,…,L).

Step 2. If (17) is satisfied, no supervisory action is needed. Then, calculate the production rate according to (1) and go to step 4.

Step 3. If (19) is satisfied, then the supervisor evaluates the remaining capacity at the original machine according to (24), and distributes it, if any, to the operations with the highest priority according to (25).

Step 4. Calculate the loading times according to the *staircase strategy*.

5 Simulation results and comparison

To illustrate the feasibility of the proposed method, intensive simulations have been carried out through continuous-flow simulator. A multiple-part-type production network studied in [12] is considered (Fig 5).

The production system under consideration consists of seven machines with 24 intermediate buffers and produces three part types. All intermediate buffer capacities $x_{ijk}^{\max}$ are equal to 10. The input and output buffers have infinite storage capacities. The failure and repair rates of all the machines are $f_i = 0.1$ and $g_i = 0.5$ $(i = 1,\ldots,7)$. The different product type routes are shown in Table 1, in which a part type may visit a machine more than once. Assembly and disassembly operations are involved.

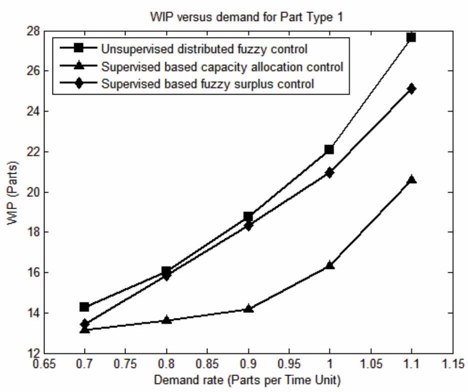
The production networks described above along with the hierarchical control strategy is simulated by means of Simulink and Floulib toolbox [7] (available at http://www.listic.univ-savoie.fr). The obtained results are compared with the unsupervised distributed approach and the supervised surplus-based approach proposed in [12].

Ten simulation runs of 10000 time units for each scenario have been performed (5 different scenarios). All the performance indicators are given by the average over the total ten simulation runs.

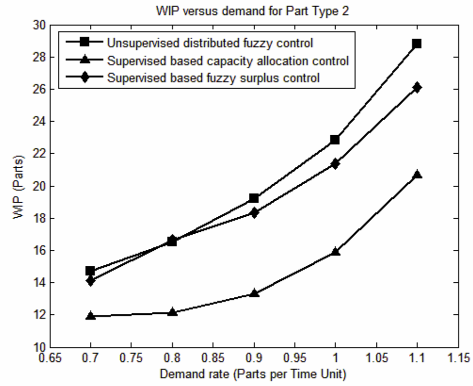
Comparative results for the WIP, backlog and lead time in relation to demand are shown in Fig. 6, Fig. 7 and Fig. 8 respectively.

Tab. 1 Product processing routes

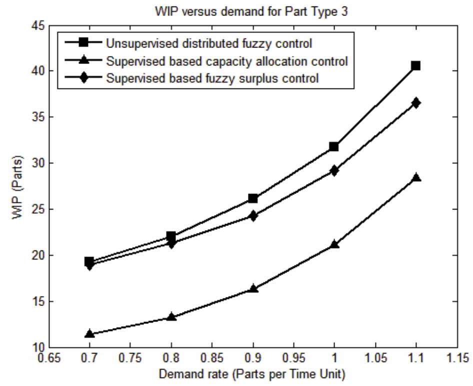
Part type (j)	Operation (k)	Machine (i)						
		1	2	3	4	5	6	7
1	1	0.2	0.2	0.2	-	0.15	-	0.2
	2	-	0.15	0.15	-	-	-	-
2	1	0.2	-	-	0.15	0.15	0.15	0.2
	2	-	-	-	0.15	-	0.15	-
3	1	0.2	0.2	0.2	0.15	0.15	0.15	0.2
	2	-	-	-	0.15	0.15	0.15	-



(a)

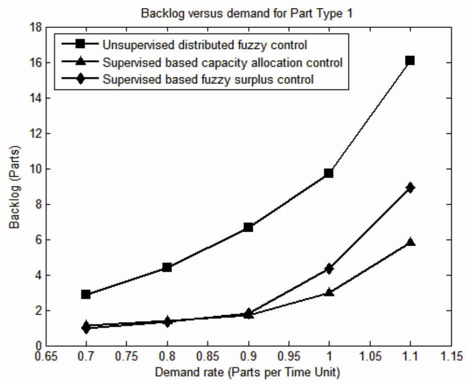


(b)

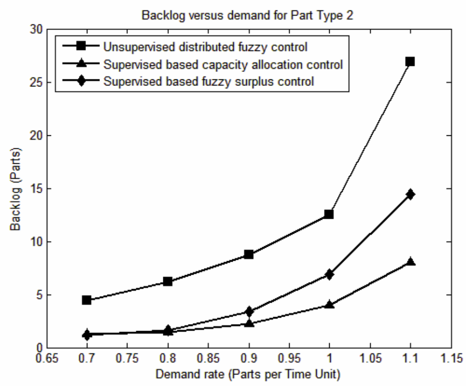


(c)

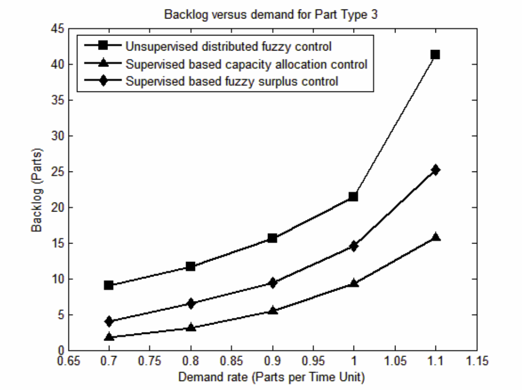
Fig. 6 WIP of each part type for various demand rates.



(a)

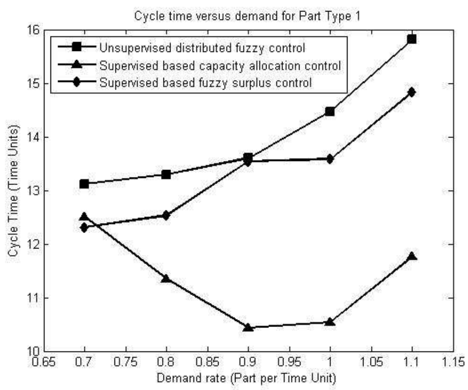


(b)

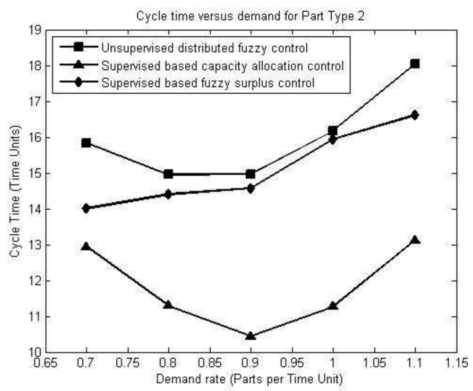


(c)

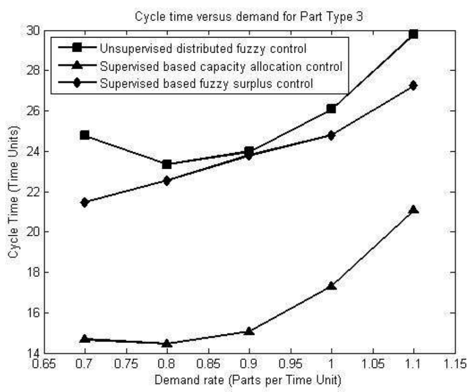
Fig. 7 Backlog of each part type for various demand rates.



(a)



(b)



(c)

Fig. 8 Lead time of each part type versus demand.

From the obtained results, it can be stated that:

- When the demand is low, the distributed control structure can easily satisfy the objectives. In this case, the supervisory control strategy modifies the local control law according to the level of the WIP and the finished inventory. This action aims at reducing the surplus and WIP by decreasing the production throughput.
- When the demand increases and reaches the maximum system productivity, the supervisory based capacity allocation control achieves a substantial reduction of WIP, backlog and Lead times compared with the distributed and the surplus-based control (Fig. 6, Fig. 7 and Fig. 8). Indeed, as the demand is high the parts enter the system faster causing some throughput bottleneck that leads to an important WIP, and thus a backlog of demand. Consequently, the supervisory action attempts to allocate the remaining capacity on each machine by some fraction in order to increase the production throughput.

We can conclude that the proposed supervisory control strategy exhibits better performance in almost every case. These results are very promising, since the decision method is very flexible and combining several control objectives with the local control one. This may help to cope with the multiple and conflicting control objectives and make the simulation process more efficient.

6 Conclusion

In this paper, an approach based on hierarchical control architecture for continuous-flow simulation of manufacturing systems has been presented. The proposed control system is hierarchical, consisting of a basic level of fuzzy logic controller supervised by a higher level of decision-maker. The lower level is responsible for the sequencing and routing decisions since it attempts to control the production flow by adjusting the machine's processing rates. It uses the fuzzy decision method based on the expert's knowledge. The higher level of supervision consists of monitoring the system by the use of global performance indicators. The supervisor allocates the production capacity or reduces the production throughput drived by the dissatisfaction's degree of the different and conflicting objectives.

For the studied case, the obtained results show a promise improvement for control performances compared with the unsupervised distributed control and the technique developed in [12]. To summarize, the main advantages of the proposed control approach are:

1. It combines global and local information to synthesize the local control law,

2. It takes into account the different degrees of importance for each control objective,
3. It facilitates the implementation phase due to the modularity and the distributivity of the control architecture.

In the future, it would be interesting to consider uncertain demand and to investigate the way of reducing the chattering phenomena that may occur around the upper and lower bounds of the tolerance interval. In this case, using a fuzzy interval with the α-cut representation of the tolerance interval of each performance indicator may avoid an abrupt switching and provide tools for characterising result reliability in the manufacturing contexts.

7 References

[1] A. Angsana and K.M. Passino. Distributed Fuzzy Control of Flexible Manufacturing Systems. *IEEE Transactions on Control Systems Technology*, 2(4):423–435, 1994.

[2] H. D'Angelo, M. Caramanis, S. Finger, A. Mavretic, Y.A. Phillis and E. Ramsden. Event driven model of unreliable production lines with storage. International Journal of Production Research, 26(7):1173-1182, 1988.

[3] X. Bai and S.B. Gershwin. Scheduling Manufacturing Systems with Work-In-Process Inventory Control: Multiple-Part-Type Systems. *International Journal of Production Research*, 32(2):365–385, 1994.

[4] L.M.M. Custodio, J.J.S. Sentieiro, and C.F.G. Bispo. Production planning and scheduling using a fuzzy decision system. *IEEE Transactions on Robotics and Automation*, 10(2):160–168, 1994.

[5] N.A. Duffie and V.V. Prabhu. Heterarchical control of highly distributed manufacturing systems', *International Journal of Computer Integrated Manufacturing*, 9(4):270–281, 1996.

[6] Y. Feng and B. Xiao. Optimal threshold control in discrete failure prone manufacturing system. *IEEE Transactions on automatic and control*, 7(47):1167–1174, 2002.

[7] L. Foulloy, R. Boukezzoula, and S. Galichet. An educational tools for fuzzy control. *IEEE Transactions on Fuzzy Systems*, 14(2):217–221, 2006.

[8] S.B. Gershwin. Design and Operation of Manufacturing Systems - The Control Point Policy. *IIE Transactions*, 32(2):891–906, 2000.

[9] S.B. Gershwin and M.H. Burman. A Decomposition Method for Analyzing Inhomogeneous Assembly/Disassembly Systems. *Annals of Operations Research*, 93:91–116, 2000.

[10] S.B. Gershwin. Manufacturing Systems Engineering. Englewood Cliffs, NJ: Prentice-Hall. 1994.

[11] G. Habchi, and C. Berchet. A model for manufacturing systems simulation with control dimension. *Simulation Modelling Practice and Theory*, 11:21–44, 2003.

[12] S. Ioannidis, N.C. Tsourveloudis, and K. Valavanis. Fuzzy supervisory control of manufacturing systems. *IEEE Transactions on Robotics and Automation*, 20(3):379–389, June 2004.

[13] V.S. Kouikoglou. An efficient discrete event model of assembly/disassembly production networks. *International Journal of Production Research*, 40:4485–4503, 2002.

[14] V.S. Kouikoglou and Y.A. Phillis. A continuous-flow model for production networks with finite buffers, unreliable machines, and multiple products. *International Journal of Production Research*, 35:381–397, 1997.

[15] C. Sohyung. A distributed time-driven simulation method for enabling real-time manufacturing shop floor control. *Computers & Industrial Engineering*, (49)4:572–590, 2005.

[16] A. Suhail and Z. A. Khan. Fuzzy control with limited control opportunities and response delay—a production-inventory control scenario', *International Journal of Approximate Reasoning*, 38(1):113–131, 2005.

[17] T. Takagi and M. Sugeno. Fuzzy identification of systems and its applications to modelling and control. *IEEE Transactions on Systems, Man and Cybernetics*, 15:116–132, 1985.

[18] K. Tamani, R. Boukezzoula and G. Habchi. Fuzzy supervisory based capacity allocation control for manufacturing systems. *FUZZ-IEEE, International Conference on Fuzzy Systems*, London, UK, July 2007.

[19] N.C. Tsourveloudis, E. Dretoulakis, and S. Ioannidis. Fuzzy work-in-process inventory control of unreliable manufacturing systems. *Information Sciences*, 127:69–83, 2000.

[20] D.J. van der Zee. Modeling decision making and control in manufacturing simulation. *International Journal of Production Economics*, 100(1):155-167, 2006.

[21] M.N. Yuniarto and A.W. Labib. Optimal control of an unreliable machine using fuzzy-logic control: from design to implementation. *International Journal of Production Research*, 43(21):4509–4537, 2005.

[22] A.R. Pritchell, S. Lee, D. Huang, and D. Goldsman. Hybrid-system simulation for national airspace system safety analysis. *In Proceeding of the 2000 Winter Simulation Conference*, 1132-1142, 2000.

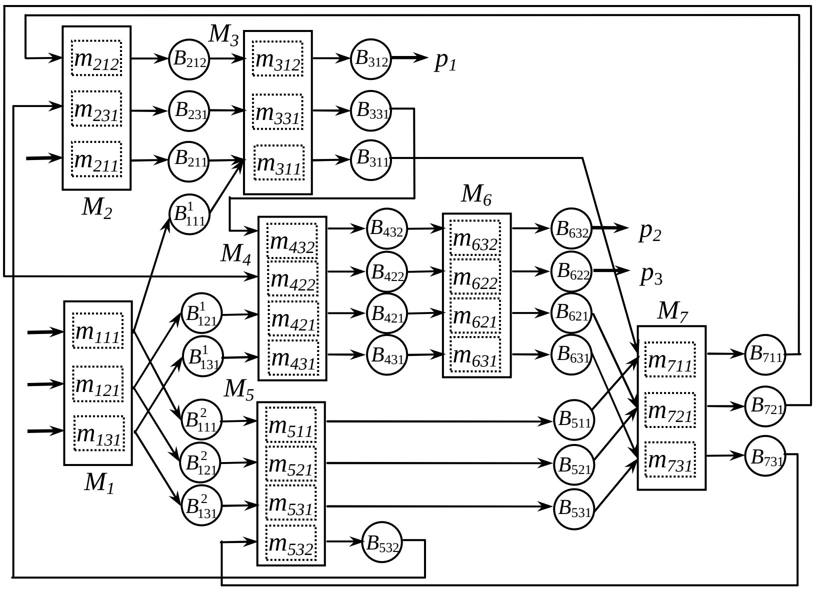


Fig. 5 Production network: multiple-part-type and reentrant flow.