Production Control

-

Keep it Simple

Rui S. Gonçalves
Fachhochschule München 96

Production Control

-

Keep it Simple

Rui S. Gonçalves
Fachhochschule München 96

Contents

Acknowledgements

I would like to thank Professor Helfrich for his help, attention and essentially for making me ask questions.

Thanks also to the city of München for receiving me so gently.

1. Introduction

KISS is a well known buzzword: Keep It Simple and Stupid. This should of course be read with a good portion of humour. It is no serious science. But like most jokes, this one has a slight touch of truth. Maybe the joke is "stupid" and the truth is "simple".

Production planning and control today is far from simple. Most planning managers would agree that it is actually quite complex. But the principles are very simple. It is somewhat a paradox that the systems grow so complex. There are of course several explanations.

One reason is the lack of consistent theory. We don't have the same situation as mathematicians which can apply operators on data where the operators are unambiguously defined and yield the same result every time. In production planning and control the operators are not clearly defined. We do not even know precisely what the operators are. And we cannot agree on what the basic data and the basic operations are.

To compensate for the lacking mathematical approach, tailor-made systems are supplied. The vendors have their perception of a production environment and make systems fitted to this environment. But the data and routines in the company in question are different from the scenario of the vendor. So customer adaptation becomes necessary. The result is that complexity grows.

And it grows even complexer when the manager starts maintaining the system. There is no end to what type of refinement that can be done. A good planning manager knows the company well, and he is usually able to modify the planning system so that it matches exactly the situation of the company. The result is a system that is totally dependent on the planning manager who in turn never will have time for further education.

Another reason is of course that production planning in reality is a complex task, not because of complex principles but because of large volumes of data. It is a problem to focus on and even detect the parts or operations that need new decisions. They usually obtain attention for replanning when a delay has occurred.

Computers also have their share in growing systems unnecessarily complex. The development has been fast. To have better control naturally has been obtained by controlling more items and updating more frequently. Online systems are of course a must. The price we have paid is complex systems.

It is about time this development is reversed. We need to simplify our planning and control systems. They simply require too many resources. We need to focus on fewer data and less tight control. Goldratt's theory of constraining resources is one contribution in this direction. Delegating production planning to a larger degree to work teams is another approach that will eliminate some of the laborious planning tasks.

1.1 Work Structure

The question that triggers this work - Production Control - Keep it Simple - is a very simple one:

Do we need to keep a precise control on every resource and material within the manufacturing system at central level?

Some systems have tried to do that and have failed. This leads us to play a little with words. And if instead of asking - "Do we need..." - we changed the question to - "Can we keep a precise control..." or "Is it possible to keep a precise control...". For instance, MRPII, one of the most widespread production control systems, tries to keep precise control at a central level of every resource and material and is now dealing with many acceptance and functional problems.

This work is devided in two main parts - Figure 1.1.1. The first one tries to answer the question - "Why do we need new production control methods?" - and is, in turn, devided into two chapters: The Failure of Computers in Manufacturing Operations Management and New Competitive Pressures.

The second one tries to provide answers to the question - "How do we build new methods for production control?" - and is devided in the chapters: OPT Philosophy, Priority Management, Fuzzy Logic, New Methods - Basic Concepts, and Fuzzy Logic in Production Control.

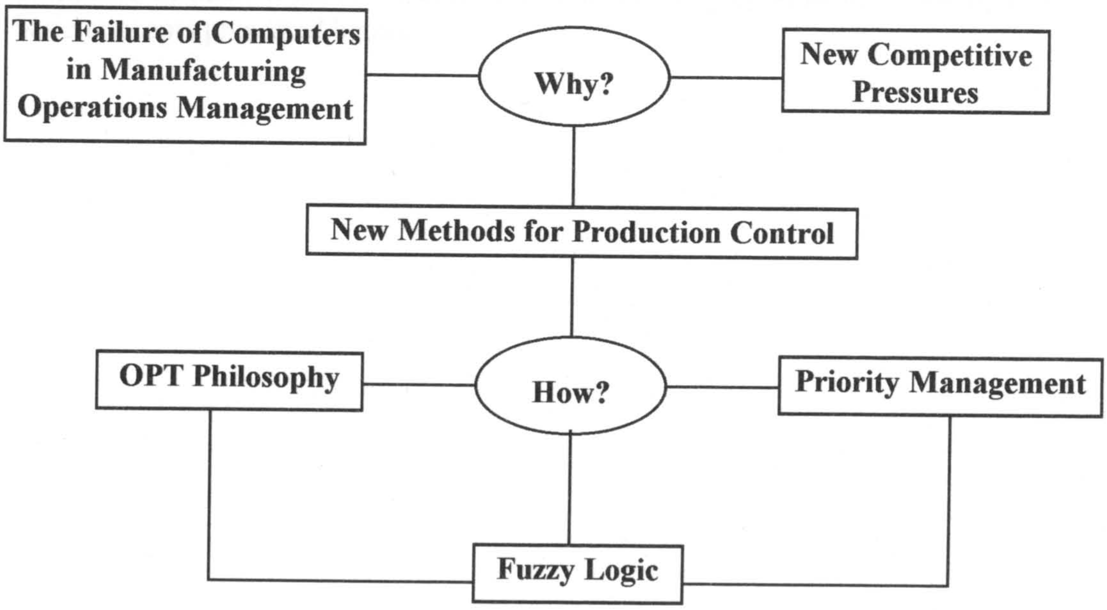


Figure 1.1.1 - Work Structure

The Failure of Computers in Manufacturing Operations Management basically exposes the inadequacies of the "standard" software packages to the real practices in many companies and shows, also, some less correct practices in manufacturing companies, probably due to the lack of computer support in production control.

New Competitive Pressures is dedicated to explore the relation between Order Winning Criteria - today time is the cutting edge of competitive advantage - in today's global markets and production control systems.

OPT Philosophy, derived by Dr. Goldratt, sustains that controlling the constraints within the system is enough to guarantee good performance for that same system. Here we review this philosophy and present some supporting studies from other researchers.

Priority Management departs from the observation that managers try to compromise more than to optimize. This is the cause of the existing gap between academics and practitioners. Aspects like customer importance, which are by nature subjective, are never included in models built by academics and are always taken into account by managers.

Fuzzy Logic is here presented as a possible solution to many sort of problems in decision making systems. As an extension of the Boolean Logic, fuzzy logic provides a way of computing with subjective and uncertain factors existing in a manufacturing system.

New Methods - Basic Concepts derives the rules that in the future, in order to reduce the complexity of production control and adapt to market globalization new requirements, should be used in manufacturing companies.

Fuzzy Logic in Production Control, through some examples tries to show the usefulness of fuzzy logic in the control of a manufacturing system, on the light of the basic concepts presented before.

2. The Failure of Computers in Manufacturing Operations Management

One continuing problem in manufacturing operations management is the comparative lack of success in computer-based planning and control systems. Computer systems are typically founded on formal planning and control methods, mainly MRP, but such models are either absent from, or resisted in, every company. Clearly, there is a gap between the ideas of the designers and the installers of these systems and those of their client organizations.

This chapter lists several reasons that contribute for this situation and were found in operations management literature. The three main groups of reasons are the environment where these systems operate, here we focus on batch manufacturing companies, concepts themselves and people. All of them have a share in this situation.

2.1 Instability of the Manufacturing Environment

Despite of using several well-known operations management techniques, just as they are found in almost every OM publication, many **batch engineering companies** have, as a normal situation in their shop-floor, late orders. If we are dealing with a stable internal and external environment, that is, if we can define perfectly every factor that influences our system they are effective methods. But batch manufacturing systems are not that stable and its instability diminishes the ability of existing techniques to effectively control them.

We can agree that batch manufacturing systems are complex in design, but that is not what makes them unstable. The sources of instability must also be sought in the nature of the task which the system is expected to perform, and of the resources which it uses.

Sources of instability - Variety, Variation, Volume
We may categorize a manufacturing system as consisting of physical materials which undergo a process according to procedures, which in turn are controlled by data aggregated into information. We assume here, that people are part of the process. In the batch engineering companies, each of these three broad elements - materials, process and data - are subject to three causes of instability - variety, variation and volume.

Variety of product types and process routes denies companies simplicity in their search for process economies, making economic product grouping and task sequencing complicated and elusive. Variation of product demand and supplies lead time, of product quality and material availability, as well as process and labour performance, makes planning a complex act and plans too rapidly invalid. Volume of live orders, of order items in progress, and the consequential control paperwork, makes the task of physical organization - which at the lowest level means merely knowing where everything is - very hard. The effects of instability on the three system elements in companies are summarized in Table 2.1.1.

The three elements of instability - variety, variation and volume - combine, especially in batch manufacturing, to defeat the efforts of managers to optimize a process which is so volatile and so lacking in constants. One could say, "Manufacturing systems are so complex in general, that is not possible for a control theory approach to be used for continuous, real-time overall optimization". Where it is difficult to optimize, it becomes necessary to compromise, and what cannot be modeled, must be managed.

Elements of Manufacturing Systems	Causes of Instability		
	Variety	Variation	Volume
Materials/Products	Offering wide range denies companies the stability of dedicated processes. Changes (setups) to general process will be necessary.	Material shortages and supplies delays create lead time instability. Variation to specification can cause quality problems. Mix changes affect lead time unpredictably.	Large volumes require more space and larger effort in storage and control areas. May also be difficult to obtain sufficient material, or other order may be "robbed", affecting their progress. But predictable: larger volumes will permit more stability.
Process/Procedure	Routing variety creates unpredictable queues and moving bottlenecks. This increases WIP and the control task.	Physical processes vary in stability and tolerance tightness. Machines can break down, operator skills varies (as does operator time keeping: absenthism).	Very large orders tie-up capacity for long periods and restrict flexibility.
Data/Information	Different functions have varied data needs in relation to the same entity. This increase complexity, inhibits aggregation and frustrates functional integration.	Data can be incomplete, inaccurate or late. The more changes are made to product and process, the greater the chance of information not being correct.	A single order creates numerous types of documents for differing functions. So, increasing number of orders (not order size), have a disproportionate effect on administration, complexity and cost.

Table 2.1.1 - Causes of Instability in a Manufacturing System. [20]

2.2 Reasons for Failure

Many researchers agree on the relative failure of many Computer Aided Production Management (CAPM) and MRP projects. This section provides a eight-heading table with reasons pointed out by these researchers between the years of 1977 and 1986 - Table 2.2.1.

Authors	MRP philosophy hardware or software	Data problems	Fail to involve users in design	Fail to involve functions outside production	Neglect of the informal system	Resistance to change	Degree of organizational commitment	Education
De Maio *et al.* [2]			√		√			
Kochhar [3]	√	√		√		√	√	√
Safizadeh and Raafat [4]	√				√	√	√	√
Lee [10]				√	√	√	√	√
Plossl [5]	√	√		√			√	
Fischer [7]		√		√				√
Lathum [9]		√		√	√		√	√
Etienne [6]	√	√		√			√	
Wacher and Hills [8]	√	√		√		√	√	√
Wight [11]							√	√

Table 2.2.1 - Reasons for Failure of Computers in Operations Management of Ten Articles. [20]

The principal reasons given for system failure are:

MRP Philosophy, Hardware and Software

Kochhar [3] queries whether the MRP philosophy itself may be to blame, especially its dependence on a forecast, its tolerance of disturbances, and its simplistic view of capacity. Safizadeh and Raafat [4] are more concerned with technical issues, the underestimation of computing power needs, reliability of hardware, and a lack of integration between software modules. Plossl [5] regretted the complexity and cost of MRP software. Etienne [6] argues “that there are features endemic to MRP that make it unsuitable for some companies”, principally the degree of discipline it requires.

Data Problems

MRP is very sensitive to data inaccuracy, which can lead to wrong information and wrong decisions. Fischer [7] places particular emphasis on this: “The first task of phase 2 will be bringing the inventory accuracy to 95 percent…this must be accomplished before the pilot program can begin, and is absolutely vital to the success of MRP”.

Failure to Involve Users in the Design of the System

Failure to involve users leads to misunderstanding by analysts of the current procedure and lack of a feeling of ownership by the eventual users of the system. Wacker and Hills [8] advised: “At time of system design get the input of employees. The purpose is to capture the idiosyncrasies of your production system and gain commitment of employees”.

Failure to Involve Other Functions

Since it links end-product demand to demand for bought-out components, MRP systems also affect sales, marketing, purchasing and, through inventory control parameters, cost accounting [9]. Systems cannot be introduced solely with reference to production management.

Neglect of Informal System

In response to inadequacies of formal systems, informal systems of control are developed over time, and these systems underpin relationships and confer roles and even status on some individuals in a company. In formalizing many of the activities of the informal system, MRP may threaten these roles and relationships.

Resistance to Change

Many of the above issues, and especially failure to involve users in system design and neglect of the informal system, lead to resistance from supervisors and from middle management. Wacker and Hills [8] and Lee [10] take a behavioral view of the implementation, proposing that MRP requires organizational changes which need sensitive handling if natural resistance is to be overcome.

Degree of Organizational Commitment

This is not so much a criticism of MRP, but top management. As in the case of many programs, MRP needs to be endorsed and continually supported by top management. An MRP system is doomed to failure if management believes its responsibilities end with authorizing purchase of the program and turning over responsibility for running the computer system to the MIS group. Continuous reinforcement and encouragement are needed: everyone must be convinced that the system is worth its time and expense. It also means spending money on training, and perhaps changing the internal measurement and reward system. If this is not done, shop-floor personnel ignore MRP schedules and use their own priorities in job selection and in process batch sizes.

Education

Resistance to change and lack of commitment require an MRP education program alongside the technical development. Wight [11] claims that class D users of MRP may have spent 80 percent of the resource spent by class A users, and that the 20 percent shortfall will typically be in education.

Even those companies which have successfully installed a formal Manufacturing Planning and Control (MPC) have not always eradicated the problems which the system was installed to tackle. A 1988 survey of 42 US companies found that none of them met the MPC performance goals which they had set themselves, and most had not improved since a 1983 survey [12].

2.3 MRP - Critics to the Concept

In section 2.2 we pointed out some reasons for the failure of computers in manufacturing operations management. The reasons were either organizational or related to the concept itself. Now, we will focus on the critics to the MRP concept.

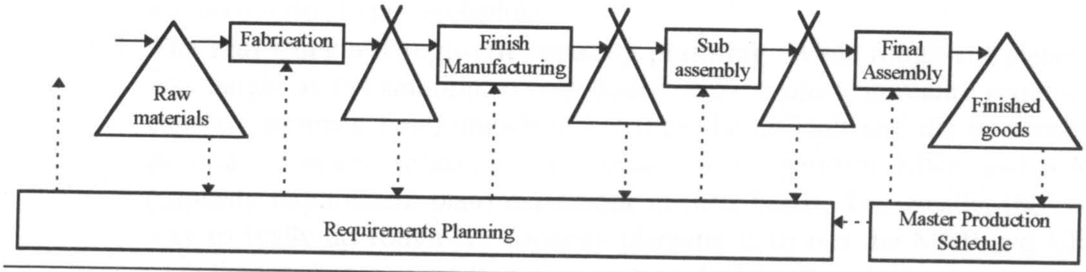


Figure 2.3.1- MRP System

The question asked at the beginning of this work - Do we need to keep a precise control over every resource and material within the manufacturing system?-, from the MRP point of view, has a positive answer. In fact, as we can perceive from Figure 2.3.1, the MRP system tries to keep, at every point in the shop-floor and over every resource, perfect and precise information.

Problems encountered in using MRP include [1]:

1. **The fallacy of static lead time**. MRP software programs treat lead time as a fixed number, while in reality lead time changes for a variety of reasons, such as normal variations in processing time, waiting for parts, delays in processing due to expedited jobs, breakdown or normal maintenance of machines, and so on.
2. **The misdefinition of lead time**. Manufacturing lead time consists of:
    - Make-ready time - to write order, enter, prepare job packet, release order, and issue material.
    - Queue time - time at the operation center waiting for operations to begin.
    - Setup time - time to prepare equipment for operations.
    - Run time - time to perform operations (produce product).
    - Wait time - time waiting after operations ends.
    - Move time - time to physically move between operations.
3. **Lead time versus fabrication/production quantity**. MRP software, because it considers lead time fixed, does not account for the fact that run time (part of the lead time) varies depending on the quantity of units to be produced.
4. **Bill-of-materials**. MRP software programs use the bill-of-materials as bottom-up product structure representing the way firms produce products. For many firms, however, especially those producing on an assembly line, products may be produced in a very different sequence than the engineering bill-of-materials.
5. **Material revision control**. Many MRP systems do not easily allow changes to be made in part numbers or in the way the product is produced.
6. **Lead time versus routing**. Since many MRP programs use the bill-of-materials structure to schedule the shop-floor, poor schedules may result. For example, there may be several routing steps at the same bill-of-materials level, which require more time than allowed.
7. **Fallacy of infinite capacity scheduling**. Few MRP programs can recognize a shop overload and reschedule.
8. **The real story of rough-cut capacity planning**. While rough-cut planning was taught as the solution to overloaded work centers, in reality rough-cut capacity planning lies somewhere between the difficult and the impossible. In effect, master schedulers are expected to perform MRP and CRP (capacity requirement plan) explosions in their heads. Technically, the only way to really do rough-cut capacity planning is to run the MRP and CRP each time a change is made in the master schedule. Required computer time makes this impossible.

9. **Capacity planning versus MRP logic**. Because MRP and CRP are not run together when changes are made in the master production schedule, such changes often create "floating bottlenecks". Bottlenecks appear and disappear, depending on the master production schedule.
10. **MRP logic - a user confuser**. MRP logic differs from system to system. The user should test how the system reacts to, say, accelerating, decelerating, or canceling an order.

Non-feasible Schedules

In addition to the problems previously mentioned in installing and using an MRP system, there are other criticisms as well. Many critics state that MRP schedules are either impossible or are only true the day that they were created. Too many changes take place in the system for MRP to be able to adjust to all of them.

Accuracy requirements

Because MRP uses detailed files to schedule, MRP cannot tolerate inaccuracies. In fact, for many years since MRP was introduced, companies have been rated into classes based on the accuracy of their records. Class A companies, for example, have 99 percent accuracy. MRP's failures in its scheduling performance had been blamed on inaccurate records. Now we recognize that inaccuracy was not completely to blame; the MRP scheduling technique was also a fault.

Excess of data needs

When computers try to control every resource and material in the system at a central level, the data needs grow fast. The growing of data needs means increasing of the system complexity, thus increasing also the work of controlling it, leading to the inefficiencies discussed before.

Safety stock

Ordinarily, adding a safety stock to required quantities is not advised in an MRP system that is based on derived demand. There is some feeling, however, that when the availability of parts could suffer from a long and inflexible lead time or is subject to strikes or cancellation, a safety stock offers protection against production delays. A safety stock is sometimes intentionally created by planning for excess. One of the main arguments against using safety stock is that the MRP system considers it a fixed quantity, and the safety stock is never actually used.

Lot sizing in MRP systems

The determination of lot sizes in an MRP system is a complicated and difficult problem. Lot sizes are the part quantities issued in the planned order receipt and planned order release sections of an MRP schedule. For parts produced in house, lot sizes are the production quantities or batch sizes. For purchased parts, these are the quantities ordered from the supplier. Lot sizes generally meet part requirements for one or more periods.

Most lot-sizing techniques deal with how to balance the setup or order costs and holding costs associated with meeting the net requirements generated by the MRP planning process. Many MRP systems have options for computing lot sizes based on some of the more commonly used techniques. It should be obvious, though, that the use of lot-sizing techniques increases the complexity in generating MRP schedules. When fully exploded, the numbers of parts scheduled can be enormous.

2.4 A Study of US Machine Shops

A survey was developed to gain a better knowledge on the production control practices in US machine shops [13]. The questionnaire addressed four broad topics of manufacturing operations:

1. basic facility operating characteristics
2. machine operator characteristics
3. shop-floor control policies
4. shop performance measures

Here I will focus on some points in shop-floor control policies and performance measures, to show how wrong practices in production control can cause some problems.

The respondent firms were classified as either JIT oriented or non-JIT oriented machine shops, based on their ratio of JIT jobs to total jobs. JIT jobs were defined as jobs that were required to be completed and delivered within a specific time window - nor before and not after. Of the respondent firms, 40.8% were considered as non-JIT against 46.9% as JIT oriented.

Computer systems
Table 2.4.1 shows that computers are not widely used in production control, with larger evidence in those companies that were classified as non-JIT oriented. As seen in the previous chapters, an informal system was developed to overcome the inadequacies of the existing manufacturing planning and control systems.

	Non-JIT companies (%)	JIT companies (%)
Computer system is used to:		
Track job progress in shop	15.1	47.5
Determine job release date	9.4	32.8
Prioritize jobs at each machine	7.5	39.3

Table 2.4.1 - Computer Systems in US Machine Shops. [13]

Job release and sequencing policies
The job processing policies used by respondents are shown in Table 2.4.2. Both sets of respondents tended to rely heavily on immediate job release to the shop-floor. Fewer respondents used some form of workload oriented or backward loading technique. Relatively few respondents stated use of some form of forward loading technique.

This may tend to explain the higher levels of finished goods inventory among the JIT-oriented machine shops. These firms may have been starting jobs immediately to reduce the risk of late deliveries, resulting in many jobs that were finished early and then placed in inventory - Table 2.4.3. As for the non-JIT oriented firms, finished jobs with less than one day in finished goods inventory are almost a rule. This may point that these jobs are finished with delay and have to be shipped the same day.

	Non-JIT companies (%)	JIT companies (%)
Job release policies		
Immediate release	63.1	52.5
Release when workload permits	18.7	16.6
Backloaded from due date	7.5	13.8
Back loaded from due date, then modified based on workload	10.3	15.5

Table 2.4.2 - Job Release Policies in US Machine Shops. [13]

	Non-JIT companies (%)	JIT companies (%)
Percentage of jobs in finished goods inventory		
(1) Less than 1	44.5	19.5
(2) 1-5	28.9	33.2
(3) 6-10	7.2	12.6
(4) 11-30	7.0	19.3
(5) 31-60	6.8	8.2
(6) Over 60	3.2	6.2
Sample average(category)	(2.1)	(2.8)

Table 2.4.3 - Number of Days of Jobs in Finished Goods Inventory. [13]

Focusing on the non-JIT companies, those we can assume that are batch manufacturing companies, we can take several conclusions pointing in the same direction as the authors mentioned above have.

So, it is clear that an informal system is ruling the operations in these companies, just has seen before. Three techniques mentioned in every OM book are less used than **immediate release** - the mirror of complete lack of control.

Consequences? We might be led to think that inventory of finished goods would accumulate, but that is not so. On the contrary, almost **half** of the finished goods spend less than one day in inventory - what could be seen as a very good performance, but what these numbers really mean is that half of the orders are late - **terrible customer service**. Even investment in inventory is high, since **work-in-progress accumulates** everywhere because of the immediate release policy.

And why? The environment itself is unstable and the way the system is designed and managed brings out its instability, instead of diminishing it.

Even though, these companies trust more this policy than they trust computer systems that implement techniques inadequate to their reality.

3. New Competitive Pressures

Today, time is the cutting edge. In fact, as a strategic weapon, time is the equivalent to money, productivity, quality, even innovation. The ways leading companies manage time - in production, in new product development, and in sales and distribution - represent the most powerful new sources of competitive advantage.

Reducing elapsed time can make the critical difference between success and failure. Give customers what they want when they want it, or the competition will. Time-based companies are offering greater varieties of products and services, at lower costs, and with quicker delivery times than their most pedestrian competitors. Moreover, by refocusing their organization on responsiveness, companies are discovering that long held assumptions about the behaviour of costs and customers are not true: costs do not increase when lead times are reduced, they decline; costs do not increase with greater investment in quality, they decrease; costs do not go up when product variety is increased and response time is decreased, they go down. And contrary to a commonly held belief that customer demand would be only marginally improved by expanded product choice and better responsiveness, the actual results have been an explosion in the demand for the product or service of a time-sensitive competitor, in most cases catapulting it into the most profitable segments of its markets [16].

Like cost, time consumption is quantifiable and therefore manageable. Today's new generation companies recognize time as the fourth dimension of competitiveness and, as a result, operate with flexible manufacturing and rapid-response systems, and place extraordinary emphasis on R&D and innovation. Factories are close to the customers they serve. **Organizations are structured to produce fast responses rather than low costs and control**. Companies concentrate on reducing, if not eliminating, delays and using their response advantage to attract the most profitable customers.

3.1 The Importance of Time-related Factors

In a 1988 survey by Boston University [1], senior management executives were asked to rate the importance of 11 designated manufacturing capabilities. The relative change of eight of these capabilities since 1984 is displayed in Figure 3.1.1. The ranking of these strategic capabilities was:

1. Conformance quality
2. On-time delivery
3. Performance quality
4. Delivery speed
5. Product flexibility
6. After-sale service
7. Price
8. Broad line
9. Broad distribution
10. Volume flexibility
11. Promotion

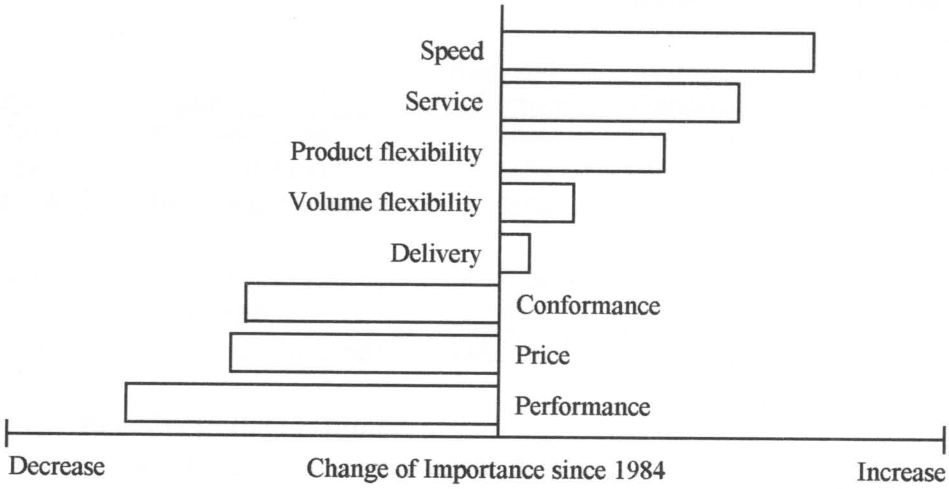


Figure 3.1.1 - Time-based Capabilities Are Becoming More Important. [1]

In the last few years we have assisted to an increased emphasis on time in manufacturing organizations. In a 1993 survey of 1300 manufacturing executives from 11 different countries [14], 93 percent of US and European managers identified reliable delivery times as having above average importance over the next three years. Overall delivery speed was cited by 88 percent and 89 percent, respectively. At the same time, 82 percent of Japanese managers rated the ability to introduce new products rapidly as having above average importance - Table 3.1.2.

Country	Time-related factor	Percentage of respondents indicating importance
USA	Reliability of delivery time	93
	Overall delivery speed	88
	Ability to reduce manufacturing lead times	86
Europe	Reliability of delivery time	93
	Overall delivery speed	89
Japan	Reliability of delivery time	87
	Ability to introduce new products rapidly	82

Table 3.1.2 - Importance of time-related factors. [14]

3.2 Benefits of Time-based Competition

Companies are obtaining remarkable results by focusing their organization on responsiveness. Each of the companies in Table 3.2.1 uses its response advantage to grow at least three times faster in the US than other companies in the industry and with profitabilities that are more than twice the US industry average. The five examples in Table 3.2.1 illustrate the competitive force of timely responsiveness to customer needs.

Other type of benefits, those not directly connected to financial measurements, are divided into benefits to the customer and benefits to the organization [24].

Company	Business	Response Difference	Growth Advantage vs. Average Competitor	Profit Advantage vs. Average Competitor
Wal-Mart	Discount Stores	80%	3X	2X
Atlas Doors	Industrial doors	66%	3X	5X
Ralph Wilson Plastics	Decorative laminates	75%	3X	4X
Thomasville	Furniture	70%	4X	2X
Citicorp	Mortgages	85%	33X	N/A

Table 3.2.1 - Time-based Competitors (estimated performance). [16]

1. **Benefits to the customer**:
    - shorter time from order to delivery
    - greater product range
    - fresher product designs
    - improved customer support
    - higher-quality goods and services
2. **Benefits to the organization**:
    - simplified, more flexible organizations
    - lower development costs
    - **simplified production control**
    - improved manufacturability
    - higher efficiencies

3.3 How to measure performance

Time-based companies go back to basics when they decide how they are going to keep track of their performance - Table 3.3.1. Time is already widely used to measure performance in business. Managers use terms like lead-time, on-time delivery and response time almost instinctively in describing how well a company is serving its customers. But time-based companies go a step further. They use time-based metrics as diagnostic tools throughout the company and set basic goals of the operation around them. In effect, they use time to help them design how the organization should work. They will often compare their performance in time with that of their best competitors or best practices anywhere. The core view that they have on performance tracking is this: Time is the best diagnostic measure and design parameter available. If we can provide the product that customers want and still compress time, we are also going to be solving cost and quality problems in our value-delivery process.

Traditional companies	Time-based companies
Cost is the metric	Time is the metric
Look to financial results	Look first to physical results
Utilization-oriented measures	Throughput-oriented measures
Individualized or department	Team measures

Table 3.3.1 - Differences between Time-based and Traditional Companies in Performance Measurement. [16]

How do time-based companies measure time? They follow two rules: Keep the measure physical, and measure as close to customer as possible. Table 3.3.2 summarizes the mainstream indicators of time compression in the four performance areas - developing new products, making decisions, processing work along the main sequence and servicing customers.

New-product development	Decision making
Time from idea to market	Decision cycle time
Rate of new-product introduction	Time lost waiting for decisions
Percent first competitor to market	
Processing and production	**Customer service**
Value added as percent of total elapsed time	Response time
Uptime × yield	Quoted lead time
Inventory turnover	Percent deliveries on time
Cycle time (per major phase of main sequence)	Time from customer's recognition of need to delivery

Table 3.3.2 - Time-based Performance Measures. [16]

Time's major advantage as a management tool is that it forces analysis down to a physical level. Putting together a time line of activity - a chart that says what happened every hour or every day to an order, or to a project, or to whatever you want to track - tells you what actually goes on in your company. Once physical activity is laid bare, the right question can be asked: Why do we do this step twice? Why are these tasks done serially and not in parallel? Why does this process work only half the time? Why do we invest to speed this up and then let its output sit and wait for the next process? Answers to these questions **lead managers to where the cost and quality problems of the company actually are.**

Senior management must shift its focus from cost to time, and its objectives from control and functional optimization to providing resources to compress time throughout the organization.

3.4 The Flexible Factory

Time-based competitors are flexible manufacturers. Their factories are organized and have a different functional philosophy from traditional factories. Flexible manufacturers have different policies and practices from traditional manufacturers. The main differences are related to:

- the length of a typical production run
- the organization of the process components
- the complexity of the scheduling procedures

Lot Sizing

Traditional factories attempt to maximize the length of their production runs in an effort to amortize lengthy and costly setups over the maximum number of pieces.

Flexible manufacturers, on the other hand, try to shorten their production runs as much as possible - Table 3.4.1. To prevent the costs of short runs from getting out of control, management focuses on reducing the complexity and hence the length and costs of setups and changeovers. The logic behind this approach is as simple as it is fundamental to competitive success: reduced run lengths mean more frequent production of the complete mix of products and faster response to customers' demands.

	Traditional	**Flexible**
Lot size	Large batches	Small batches
Flow pattern	Move through process technology centers	Organized by product
Scheduling	Centrally	Locally
Lead Times	100	10
Productivity	100	150-300

Table 3.4.1 - Differences between Traditional and Flexible Factories. [16]

Layout

Factory layout can contribute to reduce production complexity and, thus, time consumption. Traditional factories are organized by process technology centers. For example, metal goods manufacturers often organize their factories into shearing, punching, wave-soldering, wire harnesses, assembly, testing and packaging departments. Parts are moved from one process technology center to the next. Each step consumes valuable time: parts sit, wait to be moved, are then moved, then wait to be used by the next step. Amazingly, in the traditional manufacturing systems, **products usually receive value for only 0.05 percent to 5 percent of the time they are in the system**. The rest of the time products are waiting to receive value.

A survey to 148 UK companies competing for the Management Today - Cranfield School of Management Best Factory Awards [15] on throughput efficiency, produced the results showed in Figure 3.4.2. Throughput efficiency was defined as:

$$\text{Throughput efficiency} = \frac{\text{Total value adding time to produce a product or a component}}{\text{Total manufacturing throughput time}} \times 100$$

To increase value-adding time, flexible manufacturers organize their factories by product or set of products - those with similar routings, creating a **cellular layout**. To minimize the handling and movement of parts, the manufacturing processes for a component or a product are located as close together as is possible. Parts move from one activity to the next with little or no delay. Because the production process reduces or eliminates the need to pile and re-pile parts, they flow quickly and efficiently through the factory.

Scheduling

In traditional factories, **scheduling is another source of delay and waste of time. Most traditional factories use central scheduling that requires sophisticated**

materials resource planning and shop-floor control systems. Even though these systems can be quite sensitive, they are inaccurate and they waste time: work orders usually flow to the factory on a monthly or weekly basis; parts of the process can be scheduled for overtime while other parts are idle; new parts that are not needed can be produced because the same parts are being expedited by humans through the system. All the while, parts sit idle - not receiving value.

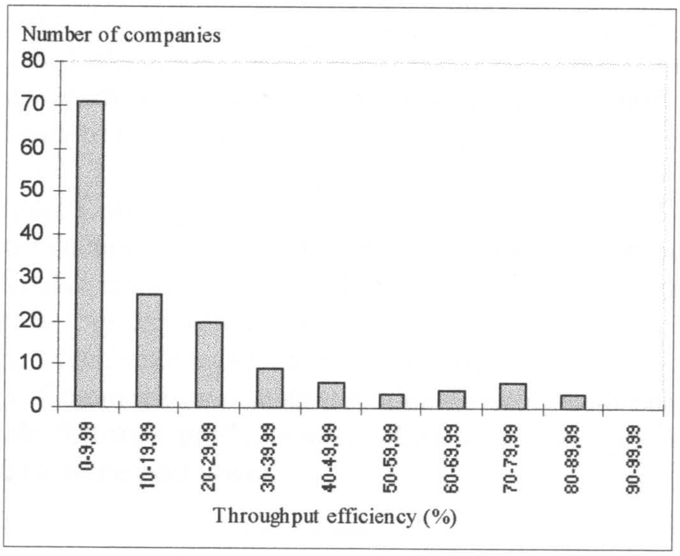


Figure 3.4.2 - Throughput Efficiency in UK Companies. [15]

In flexible factories, a combination of small run lengths and simplified plant layout makes production scheduling and management operate more smoothly. Since process layouts are product-oriented, they can be directed locally. **Local scheduling enables employees to make more production control decisions on the factory floor without time-consuming loop-backs to management for decisions and approvals**. Moreover, once a part starts into production, much of its movements between manufacturing steps are purely automatic and require no intermediate scheduling.

3.5 Costs of Variety

If a company produces only one product to only one customer, production control is very simple and costs can be reduced to maximum. If a product is produced day-in day-out, the needs for a tight production control does not exist. Economies of scale can be easily achieved.

This paradise can be destroyed by adding additional products to satisfy additional customers. Production schedules for each product now must be created and managed; there will be changeovers that require both scheduling and people to manage them; time will be lost to setup; quality will become more expensive, since with each changeover, the process must be brought back into tolerance; more steps are likely to be required and it is very unlikely that these steps can be operated in unison; inventories will increase to accommodate uncertainty; customer priorities must be weighted against the priorities for smooth operation of the factory; the process rarely will be in balance. In this factory, almost nothing is predictable and management costs will be higher.

In choosing to reduce product variety, however, executives attack the symptoms rather than the root causes of the cost of complexity. Costs go down because complexity has been removed, not because the drivers of the cost of complexity have been eliminated. These drivers are hard to identify because they are management rather than activity cost and are therefore less visible. They include the costs of various decision-making process involving the tasks at hand, as well as the costs of remaking early decisions. To support variety and market responsiveness, managers must identify these costs and reduce the complexity of the process that generate them.

Variety and Volume Costs

We can devide production costs into variety and volume driven costs. Time-based companies, has flexible manufacturers, perform better than traditional factories. Very simply, a flexible factory can accommodate more variety at lower costs than can a traditional factory, which must make a trade-off between variety and scale at an earlier point Figure 3.5.1. Compared to a similar traditional factory with only half its product variety, the flexible factory's productivity is often 50 to 150 percent higher and its costs are often 20 to 30 percent lower.

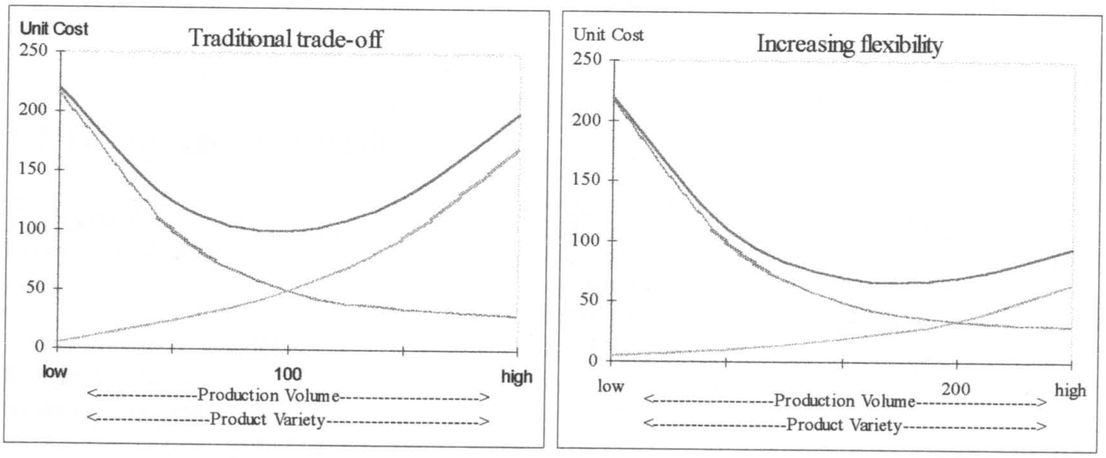


Figure 3.5.1 - Braking the Variety Barrier. [16]

Cost reduction Opportunities

Time-based companies are more productive. When time consumption in a value-delivery system is reduced by half, or even more, costs decrease substantially. If an order for a semi-custom-made product can be processed in 35 instead of 70 days with the same number of people, the cost per order processed is reduced by half. In actuality, though, the number of man-hours needed in the faster processing systems is probably less than half the number needed in the slower ones. In many manufacturing-based business, the additional reduction in the costs of goods can be expected to be at least 20 to 25% as time-consumption is reduced by 50 to 75%. And for people intensive business the additional cost reduction can be even higher - between 35 to 50%.

3.6 The Damaging Planning Loop

A basic test of management understanding of the systemic nature of its business is whether or not it is caught in the Planning Loop. All business must do some sort of planning for the future to be sure they are ready to make the sale. Manufacturers are challenged by the need to order raw material, schedule facilities, add labour, and so on. Traditional manufacturing requires long lead times to resolve conflicts between various jobs or activities that require the same resources. The long lead times, in turn, require sales forecast to guide planning. But sales forecast are inevitably wrong; by definition they are guesses, however informed. Naturally, as lead times lengthen, the accuracy of sales forecast declines. With more forecasting errors, the need for safety stocks at all levels and excess capacity at all levels increases, and inventories balloon. Errors in forecasting also mean more unscheduled jobs that have to be expedited, thereby crowding out scheduled jobs. The need for even longer lead times grows greater and the planning loop expands - driving up costs, increasing delays and creating system inefficiencies.

Time-based companies, by compressing the time to perform every task within their organization, are reducing risks. They reduce the need of forecasting and, as a consequence, the errors produced by those same forecasts. They don't have, contrary to traditional companies, to sell excess finished-goods inventory at marked-down prices and don't incur in high opportunity cost for not having the right products that customers need. The planning task is also simplified because they only schedule firm orders.

3.7 Two distinct perspectives

TBC perspective	JIT-based	NPD-based
Performance objectives	Speed Efficiency Reliability	Speed Flexibility Responsiveness
Implementation strategies		
System simplification	√	√
System integration		
• *Ad-hoc* tactics, including temporary cross-functional teams, multiskilled workers, query based DSS's		√
• Streamlining tactics, including product-based layouts, kanban and/or automated MPC systems	√	
Standardization	√	
Parallel activities	√	√
Variance control	√	
Automation		
• For flexibility		√
• For efficiency/reliability	√	
Excess capacity		√

Table 3.7.1 - TBC - Two Different Perspectives. [24]

One way to get a handle on the TBC paradigm is to consider two distinct perspectives. Table 3.7.1 summarizes the two perspectives, and matches them up with seven TBC implementation strategies. Although there is some overlap in implementation strategies between the two, the specific tactics may differ. Consider system integration, for instance. On the one hand, the JIT-based perspective would use system integration to rationalize ongoing workflows across functional boundaries. The online order entry system at Toyota, which was linked directly to production planning, is on example. In contrast, the new product development (NPD)-based perspective would use system integration to help reduce cycle times, increase flexibility and improve responsiveness in situations where the activity is not well understood or highly structured. Tactics such as *ad-hoc* cross-functional teams, multiskilled workers (who can be reassigned quickly), and query-based information systems help provide speed and flexibility in these environments.

3.8 Strategic implications

The refocusing attention from cost to time is enabling the early innovators to become time-based competitors who can, literally, run circles around their slower competition. Time-based competitors are offering greater variety of products at lower costs and in less time than their more pedestrian competitors. A set of empirical rules is emerging as more and more companies become time-based competitors. These are the **Rules of Response** [16]:

- the 0.05 to 5 rule
- the 3/3 rule
- the 1/4-2-20 rule
- the 3 X 2 rule

The 0.05 to 5 Rule

Across a spectrum of business, the amount of time required to execute a service or an order, manufacture and deliver a product is far less than the actual time the service or product spends in the value-delivery system. The 0.05 to 5 rule highlights the poor "time productivity" of most organizations since most products and many services are actually receiving value for only 0.05 to 5 percent of the time they are in the value delivery systems of their companies.

The 3/3 rule

During the 95 to 99.95 percent of the time a product or service is not receiving value while in the value-delivery system, the product or service is waiting.

The waiting time has three components, which are the amounts of time lost while waiting for

- completion of the batch a particular product or service is a part of as well as the completion of the batch ahead of the batch a particular product or service is a part of;
- physical and intellectual rework to be completed;
- management to get around to making and executing the decision to send the batch on to the next step of the value-adding process.

Generally the 95 to 99.95 percent of the time lost divides almost equally among these three categories.

The amount of time lost is affected very little by working harder. But working smarter has a tremendous impact. Companies that reduce the size of the batches they process - whether the batches are physical goods or packets of information - streamline the work flows and significantly reduce the time lost in their value-delivery systems. For example, when a manufacturer of hospital equipment reduced standard production lot sizes by half, the time required to manufacture the product declined 65 percent. After the production flow was streamlined to reduce material handling and the number of intermediate events requiring scheduling was reduced, the total time was reduced by another 65 percent for a total reduction of 58 percent.

While these improvements were dramatic, this company barely escaped the 0.05 to 5 rule. Its time productivity increased over 200 percent from 3 to 7 percent.

The 1/4-2-20 rule
Companies that attack the consumption of time in their value-delivery system experience remarkable performance improvements. For every quartering of the time interval required to provide a service or product, the productivity of labour and of working capital can often double. These productivity gains result in as much as a 20 percent reduction in costs.

A North American manufacturer of a consumer durable has reduced its time interval from five weeks to slightly more than one week. Labour and asset have more than doubled, and profits are approaching extraordinary levels.

The 3 X 2 rule
Companies that cut their time consumption on their value-delivery systems turn the basis of competitive advantage to their favour. Growth rates of three time the average of their industry with two times the industry profit margins are exciting and achievable targets.

A manufacturer of prefinished building materials reduced the time required to meet any and all customer orders to less than ten days. Most orders can be on the customer's site one to three days from when their order was placed. Some other competitors required 30 to 45 days to fill any and all orders.

This time-based competitor has grown over 10 percent a year for the last ten years, to become the market leader. The U.S. industry average growth rate has been less than 3 percent per year over the same period. The pretax return on net assets of this time-based competitor is 80 percent - more than double average of the industry in the U.S.

Western companies are obtaining remarkable results by focusing their organizations on responsiveness. Consistent with the 3X2 discussed above, each of the companies in Table 3.2.1 uses its response advantage to grow at least three times faster than the growth of its industry and with profitabilities that are more than twice the average of the others in its industries.

Time-based companies have, as we conclude from this analysis, successfully overcome the traditional problems in production control, found in traditional companies, by completely reshaping their manufacturing systems, shifting its focus from cost to time. While traditional companies track cost and size, the new competitor derives advantage from time, staying on the cutting edge and leaving its rivals behind.

Contrary to what used to happen, the role played by computers in the task of controlling production systems has been reduced in time-based organizations. Flexible companies, drastically changing the ways they manage their systems (lot sizing, product-oriented layouts, local scheduling, dedicated work-teams), have reduced the needs for complex computer systems and, also, the functions those systems perform.

After this revolution, no longer tasks such as central scheduling of every machine will be the objective of a computer system in production control. The implication of this is that old systems no longer fit in new organizations and have, therefore, to be replaced by newer ones that are able to shaped accordingly to the needs of every different and particular situation existing in manufacturing companies.

4. OPT Philosophy

It is nowadays obvious that we need to simplify our planning and control systems. They simply require too many resources. We need to focus on fewer data and less tight control. Has we saw in the previous chapter, leading companies are already changing their control techniques in this direction.

Opposite to traditional MRP systems, Dr. Goldratt developed a philosophy that changes the view of manufacturing organizations - Optimized Production Technology (OPT). Instead of controlling "everything", he argues that what is really important is concentrating on system limitations imposed by capacity-constrained resources.

The pillars of this theory are the new definitions of some concepts, such as throughput, inventory and operating expenses, that are against the traditional vision.

Here we will review the important particularities of OPT, so we can understand how this new philosophy will be able to cope, better than others, with new competitive pressures and the requirements of new time-based organizations.

The underlying philosophy presented here - the vital importance of - has led Goldratt to boarden his view of the importance of system limitations and to develop his five-step "general theory of constraints".

This chapter is built out of "The Goal - Excellence in manufacturing" [16] and "Production & Operations Management: A Life Cycle Approach" [1].

To begin with, the observation that led to formulate this philosophy, the hockey-stick phenomenon.

4.1 Hockey-Stick Phenomenon

Just about every company faces a problem called the hockey-stick phenomenon - rushing to meet quotas at the end of the time period. If the time period is a month, then this is an end-of-the-month-syndrome; if the period is a quarter, it is an end-of-the-quarter-syndrome - Figure 4.1.1. The reason that this is a problem is primarily because of the chaos that occurs at the end of the month. The system never runs smoothly; everyone works under pressure during the early flat part of the cycle as well as during the end of the cycle. The cause of the problem is that two sets of measurements are being employed: at the beginning of the period, cost-accounting efficiency measurements are used for utilizations and variations from standards. These are local measurements. This encourages minimizing setups through large batches and minimizing variances. As the end of the month approaches, however, pressure mounts to meet a different set of measurements, ones that relate to financial performance. These global measurements vary at different points within a plant. They are stated in terms such as dollars of output shipped. On the financial statements these measurements are expressed as net profit, return on investment, and cash-flow. As soon as the end of the month passes, pressure decreases and everyone again looks at the cost-accounting measurements of variances and utilization. And so the cycle repeats.

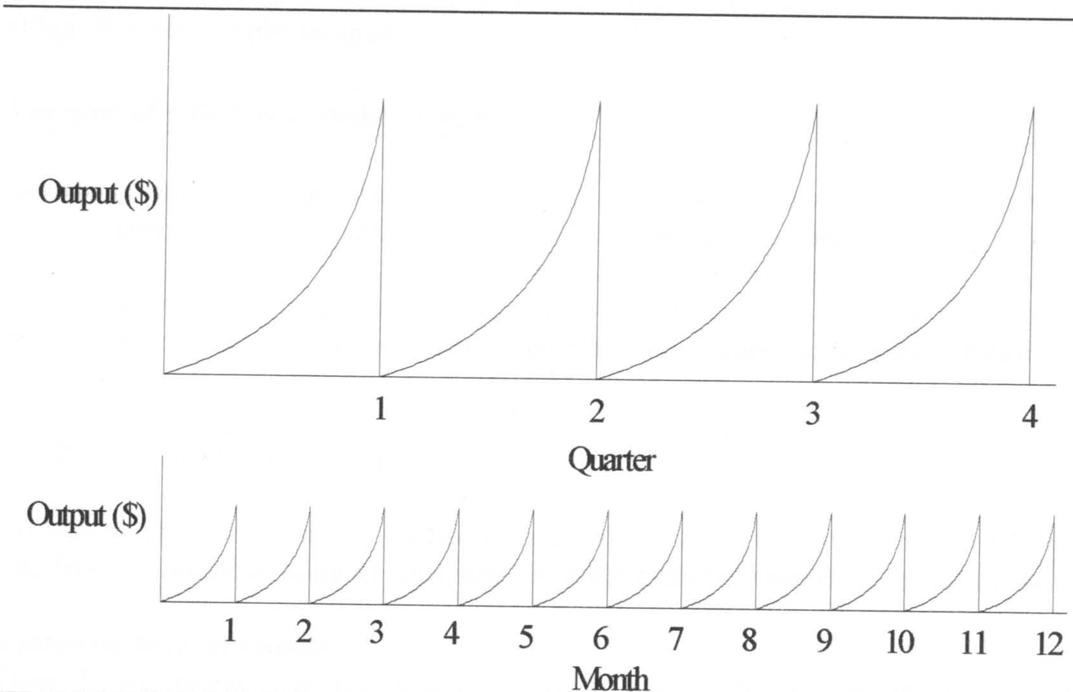


Figure 4.1.1 - Hockey-Stick Phenomenon: The End-of-the-Period Rush.[1]

To emphasize the value of techniques such as OPT, Goldratt tells two stories of what can happen. One is about a firm that did not need the third shift because of a slump in the market. The workers were highly skilled and, because the company expected the market to rebound in three months, it decided to keep the workers. With a weekly payroll of $25,000, the company allocated $300,000 to carry the workers for three months, within which time they expected renewed market demand. To the surprise of management, however, these workers consumed $400,000 during the second week, so the workers were laid off. There were so much pressure to keep the workers busy that they were not allowed to just be idle (nor would the worker feel right just sitting around waiting). Two months later market demand increased, but not for the products the workers had been working on.

The second is about a plant whose demand went down because the overall market had deteriorated. Because the company was having some cash-flow problems, it decided to cut expenses. Noticing that the setup workers were the highest paid in the plant, management decided to double all the batch sizes (which meant only the half setups were needed) and lay off half the setup workers. The result put the plant out of business! With doubled batch sizes, the increase of work in process drained all the company´s cash so that it could no longer operate.

4.2 Goal of the Firm

The goal of a firm, in OPT, is very clear and leaves no doubts about its pragmatism, although many people disagree:

The goal of a firm is to make money

Despite having many purposes - providing jobs, consuming raw materials, increasing sales, increasing market share, developing technology, or producing high-quality products - these do not guarantee long-term survival of the firm. They are means to achive the goal, not the goal itself. If the firm makes money - and only then - it will prosper. When a firm has money, then it can place more emphasis on other objectives.

4.3 Performance Measurements

To adequately measure a firm´s performance, two sets of measurements must be used: one from the financial point of view, and the other from the operational point of view.

Financial Measurements
How do we measure the firm´s ability to make money? In financial terms, we keep track of:

1. **Net profit**, which is an absolute measurement;
2. **Return on investment**, which is a relative measurement;
3. **Cash-flow**, which is a survival measurement.

To accurately evaluate a firm´s performance in financial terms, all three measurements must be used together. For example, net profit of $10 million is important as one measurement, but it has no real meaning until we know how much investment it took to generate that $10 million. If the investment was $100 million, then this is a 10% return on investment. Cash-flow is important since cash is important to pay bills for day-to-day operations; without cash, a firm can go bankrupt even though it is very sound in normal accounting terms. A firm can have a high net profit and a high return on investment, but can still be short on cash if, for example, the profit is invested in new equipment or tied up in inventory.

Operational Measurements
While financial measurements work well at the higher level, they cannot be used at the operational level. Here we need another set of measurements that will give us guidance. There are three:

1. **Throughput** - the rate at which money is generated by the system through sales
2. **Inventory** - all the money that the system has invested in purchasing things which it intends to sell
3. **Operating Expense** - all the money that the system spends to turn inventory into throughput

Throughput is specifically defined as goods sold. An inventory of finished goods is not throughput, but inventory. Actual sales must occur. It specifically defined this way to prevent the system from continuing to produce under the illusion that the goods might be sold. Such action simply increases costs, builds inventory, and consumes cash. Inventory that is carried - regardless of the stage (work in process or finished goods) - is valued only at the cost of the materials it contains. Money spent to convert raw materials into throughput is ignored. (In traditional accounting terms, money spent is called value added).

While this is often an arguable point, using only the raw material cost is a conservative view. When using the value-added method (which includes costs of production), inventory is inflated and presents some serious income and balance sheet problems. Consider, for example, work-in-progress or finished-goods inventory that has become obsolete, or for which a contract was canceled. It is a difficult management decision to declare large amounts of inventory as scrap since it is often carried on the books as assets, even though it may really have no value. Using raw materials also avoids the problem of determining which costs are direct and which are indirect.

Operating expenses include production costs, such as direct labour, indirect labour, inventory carrying costs, equipment depreciation, materials and supplies used in production, and admimistrative costs. The key difference here is that there is no need to separate direct and indirect labour.

As shown in Figure 4.3.1, the objective of a firm is to treat all three measurements simultaneously, and continually; this achieves the goal of making money.

In these operational measurements, the goal of the firm is to:

Increase throughput while simultaneously reducing inventory and reducing operating expense.

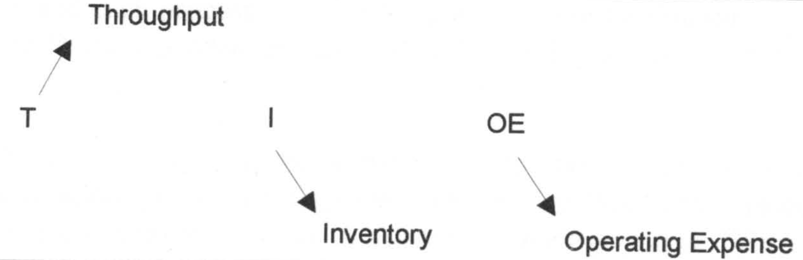


Figure 4.3.1 - Operational Goal.[1]

Productivity
Typically, productivity is measured in terms of output per labour hour. However, this measurement does not ensure that the firm will make money (for example, when extra output is not sold but accumulates as inventory). To test whether or not productivity has increased, we should ask these questions: Has the action taken increased throughput? Has it decreased inventory? Has it decreased operational expense? This lead to a new definition:

Productivity is all the actions that bring a company closer to its goal .

4.4 Unbalanced capacity

Historically (and still typically in most firms), manufacturers have tried to balance capacity across a sequence of processes, in an attempt to match capacity with market demand. Using systems manufacturing logic, however, this is the wrong thing to do - unbalanced capacity is better. Consider a simple process line with several stations, for example. Once the cycle time (or average output rate) of the line has been established, production people try to make the capacities of all stations the same. This is done by adjusting machines or equipment used, work loads, skill and type of labour assigned, tools used, overtime budgeted, and so on.

In the OPT Philosophy thinking, however, making all capacities the same is viewed as a bad decision. Such a balance would be possible only if the output times of all stations were constant or had a very narrow distribution. A normal variation of output times causes downstream stations to have idle time when upstream stations take longer to process. Conversely, when upstream stations process in a shorter time, inventory builds up between the stations. The effect of statistical variation is cumulative. The only way that this variation can be smoothed is by increasing work in process to absorb the variation (a bad choice, since we should be trying to reduce work in process), or increasing capacities downstream to be able to make up for the longer upstream times. The rule here is that capacities within the process sequence should not be balanced to the same levels. Rather, attempts should be made to balance the flow of product through the system. When flow is balanced, capacities are unbalanced. This idea is further explained in the next section.

Dependent events and statistical fluctuations

The term dependent events refers to a process sequence. If a process flows from A to B to C to D, and each process must be completed before passing on to the next step, then B, C and D are dependent events. The ability to do the next process is dependent on the preceding one.

Statistical fluctuations refers to the normal variation about a mean or average. When statistical fluctuations occur in a dependent sequence without any inventory between workstations, there is no opportunity to achieve the average output. When one process takes longer than average, the next process cannot make up the time. We follow through an example to show what could hapen. This situation exists in assembly lines, although most of them - electronic products, cars, - have shorter processing times than our example. The other extreme for processing times also exists: airplanes, prefab houses, and even ships are built on assembly lines.

Suppose that we wanted to process five items that could come from the two distributions. The processing sequence is from A to B with no space for inventory in between. Process A has a mean of 10 hours and a standard deviation of 2 hours. This means that we could expect 95.5 percent of the processing time to be between 6 and

14 hours (plus or minus 2 sigma). Process B has a constant processing time of 10 hours - Figure 4.4.1.

We see that the last item was completed in 64 hours, for an average of 12.8 hours per item, although the expected time of completion was 60, for an average of 12 hours per item (taking into account the waiting time fot the first unit by process B).

We were not able to achieve the expected output average rate. Why? Because the time lost when the second process is idle cannot be made up.

	Mean (hours)	Standard Deviation (hours)
Process A	10	2
Process B	10	0

Item Number	Start Time (hours)	Processing Time (hours)	Finish Time (hours)	Item Number	Start Time (hours)	Processing Time (hours)	Finish Time (hours)
1	0	12	12	1	12	10	22
2	12	12	22	2	22	10	32
3	24	10	34	3	34	10	44
4	34	8	42	4	44	10	54
5	42	8	50	5	54	10	64
		Average =	10 hours		Average =	10 hours	

Figure 4.4.1 - Processing and Completion Times, Process A to Process B.[1]

This example is intended to challange the theory that capacities should be balanced to an average time. Rather than balancing capacities, the flow of product through the system should be balanced. This is the first rule of the OPT Philosophy:

1st rule
Do not balance capacity - balance flow.

4.5 Bottlenecks, Nonbottlenecks, and Capacity-constraind Resources

A **bottleneck** is defined as any resource whose capacity is less than the demand placed upon it. A bottleneck, in other words, is a process that limits throughput. It is that point in the manufacuring process were flow thins to a narrow stream. A bottleneck may be a machine, scarce or highly skilled labour, or a specialized tool. Observations in the industry have shown that most plants have very few bottleneck operations, usually just several.

Capacity is defined as the available time for production. This excludes maintenance and other downtime. A **nonbottleneck** is a resource whose capacity is greater than the demand placed on it. A nonbottleneck, therefore, should not be working constantly since it can produce more than is needed. A nonbottleneck contains idle time.

Capacity-constrained resource (CCR) is one whose utilization is close to capacity and could be a bottleneck if it is not scheduled carefully. For example, a CCR may be receiving work in a job-shop environment from several sources. If these sources schedule their flow in a way that causes occasional idle time for the CCR in excess of its unused capacity time, the CCR becomes a bottleneck. This can happen if batch sizes are changed or if one of the upstream operations is not working for some reasons and does not feed enough work to the CCR.

4.6 Basic Manufacturing Building Blocks

All manufacturing processes and flows can be simplified to four basic configurations, as shown in Figure 4.6.1. In Exhibit 4.6.1A, product that flows through process X feeds into process Y. In section B, Y is feeding X. In section C, processes X and Y are creating subassemblies, which are then combined, say to feed the market demand. In section D, process X and process Y are independent of each other and are supllying their own markets. The last column in the Figure shows possible sequences of nonbottleneck resources, which can be grouped and represented as Y, in order to simplify the representation.

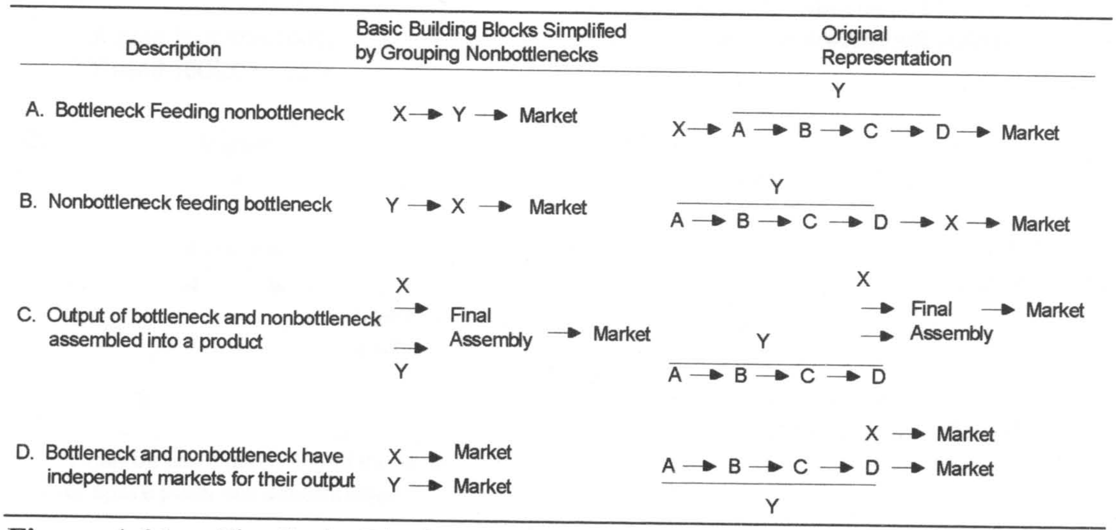


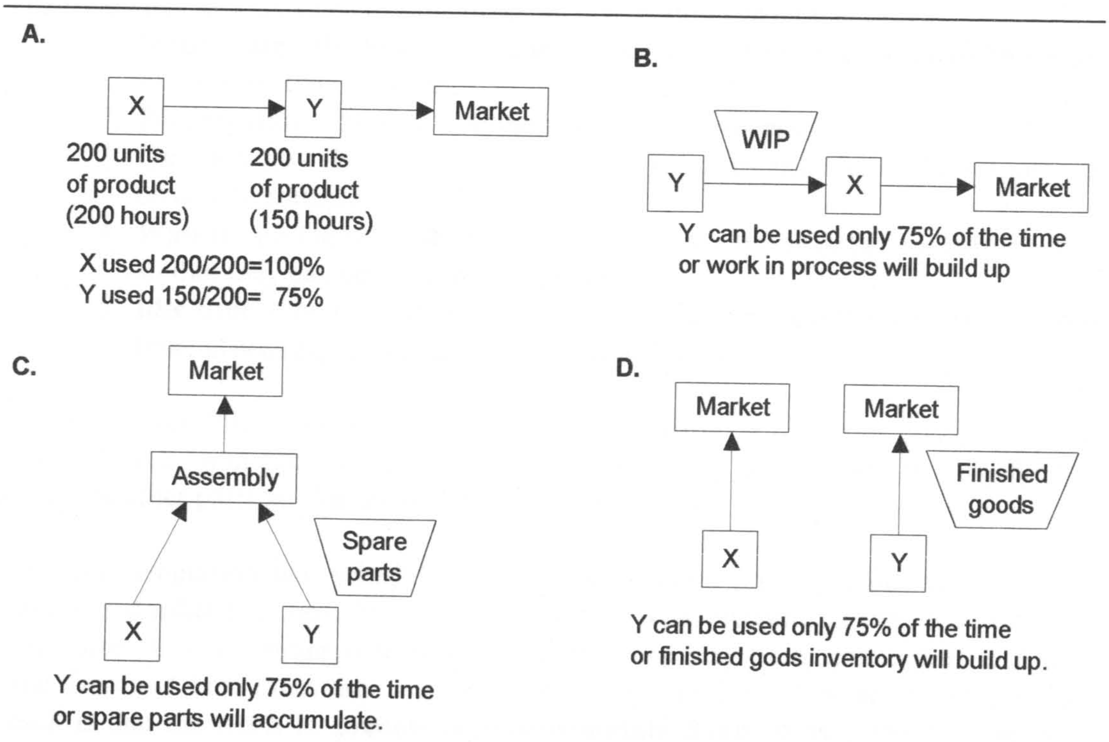
Figure 4.6.1 - The Basic Blocks of Manufacturing Derived by Grouping Process Flows.[1]

The value in using these simple basic building blocks is that a production process can be greatly simplified for analisys and control. Rather than track and schedule all the steps in a production sequence through nonbottleneck operations, for example, attention can be placed at the beginning and endpoints of the building block groupings.

4.7 Methods for control with OPT

In Figure 4.7.1, we illustrate the way bottlenecks and nonbottlenecks resources should be managed, using the four basic building blocks in manufacturing described in Figure 4.6.1. Assume that resource X has a market demand of 200 units per month and resource Y has a market demand of 150 units per month, regardeless of the building block configuration. Assume also that both X and Y have 200 hours per month available. To help understand our example, assume that each unit or product passing through X takes one hour of processing time and each unit passing through Y takes 45 minutes. With the available times, 200 hours of resource X can produce 200 units of product and 200 hours of resource Y can produce 267 units of product. Let´s consider each of the following building blocks.

In Figure 4.7.1A, the flow of product is in a dependent sequence, that is, it must pass through both resource X (machine X) and resource Y (machine Y). Resource X is the bottleneck since it has a capacity of 200 units whereas Y has a capacity of 267 units. Resource Y can be used only 75 percent of the time, since it is starved for work (X is not feeding it enough to allow it to work longer). In this case, no extra product is produced.



Y can be used only 75% of the time or work in process will build up

Y can be used only 75% of the time or spare parts will accumulate.

Y can be used only 75% of the time or finished gods inventory will build up.

X is a bottleneck; Y is a nonbottleneck. Both X and Y have 200 hours available.

Figure 4.7.1 - Product Flow Through the Four Basic Building Blocks.[1]

Section B of the Figure is the reverse of section A: a nonbottleneck is feeding a bottleneck. Since X can put through only 200 units, we must be careful not to produce more than 200 on Y, or inventory will build up as work is process.

Section C of the Figure shows that the outputs from X (a bottleneck) and Y (a nonbottleneck) are assembled into a product. As a nonbottleck, Y has more capacity than X, so it can be used only 75 percent of the time, otherwise spare parts accumulate.

In section D, market demands for X and Y are 200 units each. Since 200 units from Y corresponds to only 75 percent of its capacity, Y can be used only 75 percent of the time or finished goods inventory will build up. The situations we have just discussed are important because current industry practice considers percentage of resource utilization as one of the measures of performance. Such practice encourages the overuse of nonbottleneck resources, resulting in excess inventories. This leads us to derive the 2nd and 3rd rules of OPT:

2nd rule
The level of utilization of a nonbottleneck is not determined by its own potential but by some other constraint in the system.

3rd rule
Utilization and activation of a resource are not the same.

Time components
The following various kinds of time make up production cycle time:

1. **Setup time** - the time that a part spends waiting for a resource to be set up to work on this same part
2. **Process time** - the time that the part is being processed
3. **Queue time** - the time that a part waits for a resource while the resource is busy with something else
4. **Wait time** - the time that a part waits not for a resource but for another part so that they can be assembled together
5. **Idle time** - the unused time, that is, the cycle time less the sum of the setup time, processing time, queue time and wait time.

For a part waiting to go through a bottleneck, queue time is the greatest. For a nonbottleneck, wait time is the greatest. The part is just sitting there waiting for the arrival of other parts so that an assembly can take place.

Note the temptation that schedulers have to save setup times. Supposing the **batch sizes are doubled** to save half the setup times. What happens is that with doubled batch size all of the other time would double: processing time, queue time and wait time. Because these times are doubled while saving only half of the setup time, the net result is that the **work in process is approximataly doubled** as is the investment in inventory.

Saving time on bottleneck and nonbottleneck resources
Recall that a bottleneck is a resource whose capacity is less than the demand placed on it. Since we focus on bottlenecks as restricting throughput (defined as sales), a bottleneck´s capacity is less than the market demand. There are number of ways we can save time on a bottleneck (better tooling, higher quality labour, larger batch sizes, reducing setup times, and so forth), but how valuable is this extra time? Very, very valuable!

4th rule
An hour lost at a bottleneck is an hour lost for the entire system.

How about an hour saved at a nonbottleneck?

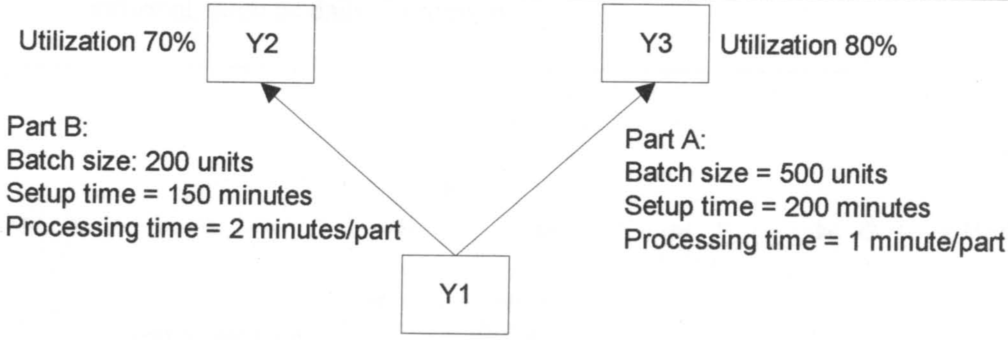
5th rule
An hour saved at a nonbottleneck is a mirage.

Because a nonbottleneck has more capacity than the system needes for ist current throughput, it already has idle time. Implementing any measures to save more time does not increase throughput but only serves to decrease its idle time.

Avoid changing a nonbottleneck into a bottleneck
When nonbottleneck resources are scheduled with larger batch sizes, this action could create a bottleneck which we certainly would want to avoid. Consider the case in Figure 4.7.2, where Y1, Y2, Y3, are nonbottleneck resources. Y1 currently produces part A, which is routed to Y3, and part B is routed to Y2. To produce part A, Y1 has a 200 minute setup time and a processing time of 1 minute per part. Part A is currently produced in batches of 500 units; utilization is 70 percent. To produce part B, Y1 has a setup time of 150 minutes as 2 minutes processing time per part. Utilization is 80 percent, and part B is currently produced in batches of 200 units.

Since setup time is 200 minutes for part A, both worker and supervisor mistakenly believe that more production can be gained if fewer setups are made. Let´s assume that batch size is increased to 1,500 units. What happens? The illusion is that we have saved 400 minutes of setup. (Instead of three setups taking 600 minutes to produce three batches of 500 units each, there is just one setup with a 1,500 unit batch.)

The problem is that the 400 minutes saved serves no purpose, but this delay did interfere with the production of part B. Y1 also produces part B for Y2. The sequence before any changes were made was part A (700 minutes), part B (550 minutes), part A (700 minutes), part B (550 minutes), and so on. Now, however, when the part A batch is increased to 1,500 units (1,700 minutes), Y2 and Y3 could well be starved for work and have to wait more time than they have available (30 percent idle time for Y2 and 20 percent for Y3). The new sequence would be: part A (1700 minutes), part B (1350 minutes), etc. Such an extended wait for Y2 and Y3 could be disruptive. Y2 and Y3 could become temporary bottlenecks and lose throughput for the system.



Resource Y1 produces Part A for resource Y3 and Part B for resource Y2.

Figure 4.7.2 - Nonbottleneck Resources.[1]

Drum, Buffer, Rope
Every production system needs some control point or points to control the flow of product through the system. If the system contains a bottleneck, the bottlenck is the best palce for control. This control point is called the **drum**, for it strikes the beat that the rest of the system (or those parts which it influences) uses to function. Recall that a

bottleneck is defined as a resource that does not have capacity to meet demand. Therefore, a bottleneck is working all the time and one reason for using it as a control point is to make sure that the operations upstream do not overproduce and build up excess WIP inventory that the bottleneck cannot handle.

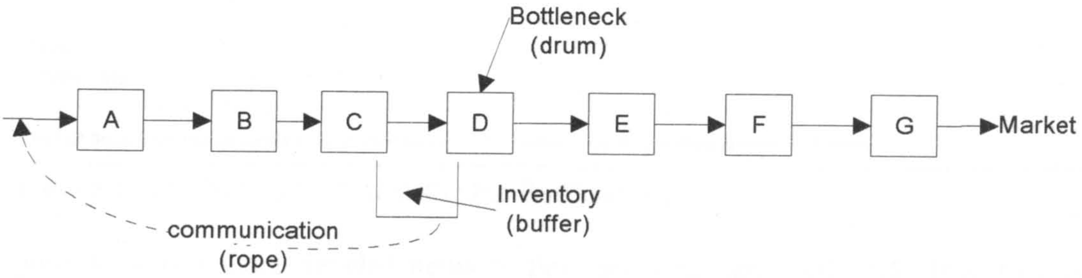
If there is no bottleneck, the next best place to set the drum would be a capacity-constrained resource (CCR). A CCR, remember, is one that is operating near capacity, but on the average has adequate capability as long as it is not incorrectly scheduled (for example, with too many setups, causing it to run short of capacity, or producing too large a lot size, thereby starving downstream operations).

If neither a bottleneck nor a CCR is present, the control point can be designated anywhere. The best position would generally be at some divergent point where the output of the resource is used in several downstream operations.

Dealing with the bottleneck is most critical, and our discussion focuses on assuring that the bottleneck always has work to do. Figure 4.7.3 shows a simple linear flow. Suppose that resource D, which is a machine center, is a bottleneck. This means that the capacities are greater both upstream and downstream from it. If the sequence is not controlled, we would expect to see a large amount of inventory in front of machine center D and very little anywhere else. There would be little finished goods because (by the definition of the term bottleneck) all the product produced would be taken by the market.

There are two things that we must do with this bottleneck:

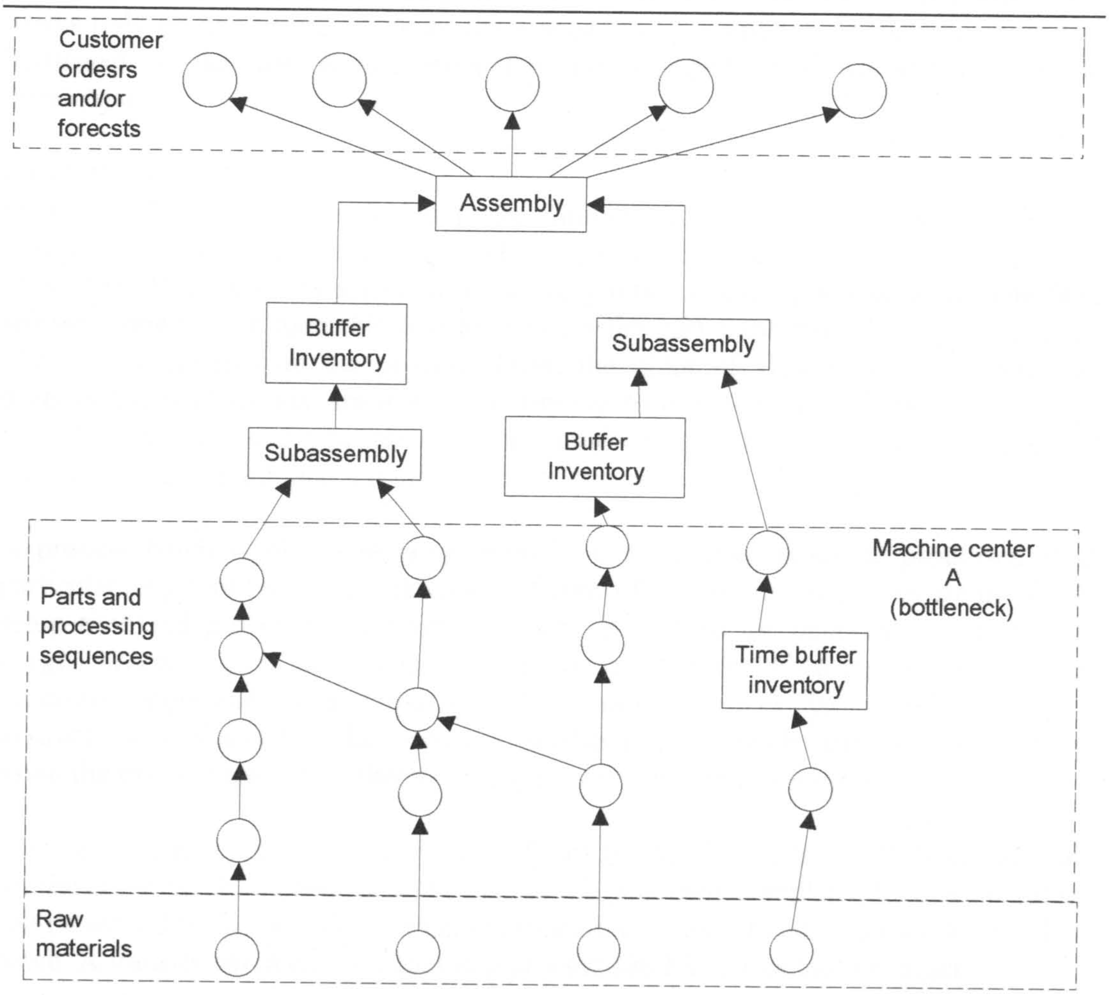
1. Keep a **buffer** inventory in front of it to make sure that it always has something to work on. Because it is a bottleneck, its output determines the throughput of the system.
2. Communicate back upstream to A what D has produced so that A provides only that amount. This keeps inventory from building up. This communication is called the **rope**. It can be formal, such as a schedule, or informal, such as daily discussion.



Product flows from work centers a through G. Work center D is a bottleneck.

Figure 4.7.3 - Linear Flow of Product with a Bottleneck.[1]

The buffer inventory in front of a bottleneck operation is a time buffer. We want to make sure that machine center D always has work to do, and it doesn´t matter which of the scheduled products are worked on. We might ask: How large should the time buffer be? The answer: As large as it needs to be to ensure that the bottleneck continues to work. By examining the variation of each operation, we can make a guess. Theoratically the size of the buffer can be computed statistically by examining data, or the sequence can be simulated. In any event, precision is not critical. We could start with an estimate of the time buffer as one fourth of the total lead time of the system. If during the next few days or weeks the buffer runs out, we need to increase the buffer size. We do this by releasing extra material to the first operation. On the other hand, if we find that our buffer never drops below three days, we might want to hold back releases and reduce the time buffer to three days. Experience is the best determination of the buffer size.



Product flows from raw materials through processing to the market. inventory buffers protect throughput.

Figure 4.7.4 - Network Flow with One Bottleneck.[1]

Figure 4.7.4 is a more detailed network flow showing one bottleneck. Inventory is provided not only in front of that bottleneck but also after the nonbottleneck assembly to which it is assembled. This ensures that the flow of the product after it leaves the bottleneck is not slowed down by having to wait. With this, we are now able to derive the 6th OPT rule:

6th rule
Bottlenecks govern both throughput and inventory.

Importance of Quality

An MRP system allows for rejects by building a larger batch than actually needed. A JIT system cannot tolerate poor quality since JIT success is based on a balanced system. A defective part or component can cause a JIT system to shut down, thereby loosing throughput of the total system. Optimized production technology, however, has excess capacity throughout the system, except for the bottleneck. If a bad part is produced upstream of the bottleneck, the result is that there is a loss of material only. Because of the exces capacity, there is still time to do another operation to replace the one just scrapped. For the bottleneck, however, extra time does not exist, so that there should be a quality inspection just prior to the bottleneck to ensure that the bottleneck works only on good product. Also, there needs to be assurance downstream from the bottleneck so that the passing product is not scrapped - that would mean lost of throughput.

Batch Sizes

In an assembly line, what is the batch size? Some would say "one", because one unit is moved at a time; others would say "infinity", since the line continues to produce the same item. Both answers are correct, but they differ in their point of view. The first answer, "one", in an assembly line focuses on the part tranferred one unit at a time. The second focuses on the process. From the point of view of the resource, the process batch is infinity since it is continuing to run the same units. Thus, in an assembly line, we have a process batch of infinity (or all the units until we change to another process setup) and a tranfer batch of one unit.

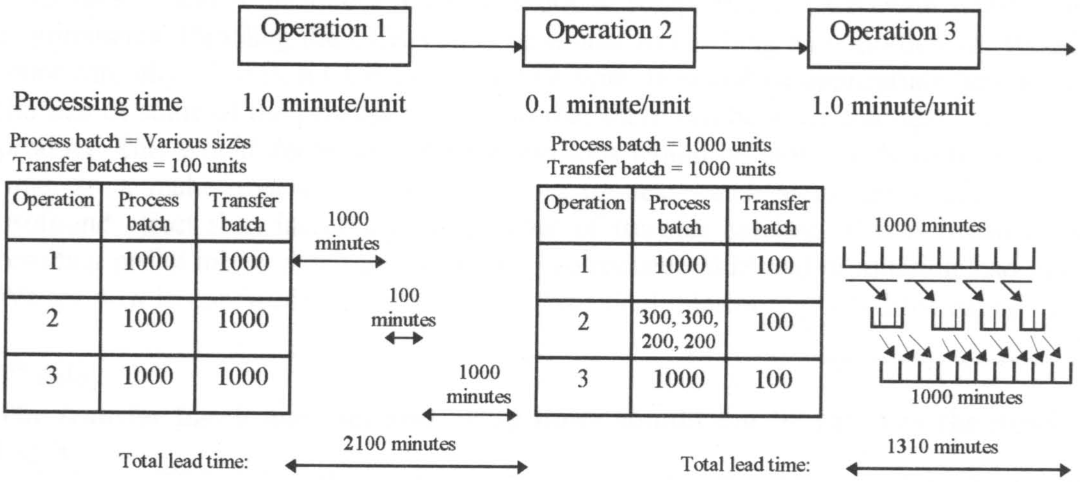
A process batch is of a size large enough or small enough to be processed in a particular length of time. From the point of view of a resource, two times are involved: setup time and processing run time (ignoring downtime for maintenance or repair). Larger process batch sizes require fewer setups and therefore can generate more processing time and more output. For bottleneck resources, larger batch sizes are desirable. For nonbottlenecks resources, smaller process batch sizes are desirable (by using the existing idle time), thereby reducing work-in-process inventory.

Transfer batches refer to the movement of part of the process batch. Rather than wait for the entire batch to finished, work that has been already completed by that operation can be moved to the next downstream workstation so that it can begin working in that batch. A transfer batch can be equal to a process batch but it cannot be larger.

The advantage of using transfer batches that are smaller than the process batch quantity is that the total production time is shorter and therefore the amount of work in process is smaller. Figure 4.7.5 shows a situation where the total production lead time was reduced from 2,100 to 1,310 minutes by using a transfer batch of 100 rather than 1000, and reducing the process batch sizes for operation 2.

How to determine process batch and transfer batch sizes

Logic would suggest that the master production schedule (however it was developed) be analyzed as to its effect on various work centers. In an MRP system, this means that the master production schedule should be run through the MRP and CRP (capacity requirements planning program) to generate a detailed load on each work center. Some managers state from their experience that there are too many errors in the manufacturing database to do this.



Operation	Process batch	Transfer batch
1	1000	1000
2	1000	1000
3	1000	1000

Operation	Process batch	Transfer batch
1	1000	100
2	300, 300, 200, 200	100
3	1000	100

Figure 4.7.5 - Effect of Changing the Process Batch Sizes on Production Lead Time for a Job Order of 1000 units.[1]

Rather than try to adjust the master production schedule to change resource loads, it is more practical to control the flow at each bottleneck or CCR to bring the capacities in line. The process batch sizes and transfer batch sizes are changed after comparing past performances in meeting due dates.

Smaller transfer batches give lower work-in-process inventory but faster product flow (and consequently shorter lead time). More material handling is required, however. Larger transfer batches give longer lead times and higher inventories, but there is less material handling. Therefore, the transfer batch size is determined by a trade-off of production lead times, inventory reduction benefits, and costs of material movement.

When trying to control the flow at CCR´s and bottlenecks, there are four possible situations:

1. A bottleneck (no idle time) with no setup required when changing from one product to another
2. A bottleneck with setup times required from one product to another
3. A CCR (with small amount of idle time) with no setup required to change from one product to another
4. A CCR with setup time required when changing from one product to another.

In the first case, a bottleneck with no setup time to change products, jobs should be processed in the order of the schedule so that delivery is on time. Without setups, only the sequence is important. In the second case, when setups are required, larger batch sizes combine separate similar jobs in the sequence. This means reaching ahead into future time periods. Some jobs will therefore be done early. Since this is a bottleneck resource, larger batches save setups and thereby increase throughput (the setup time saved is used for processing). The larger process batches may cause the early scheduled jobs to be late. Therefore, frequent small-sized tranfer batches are necessary to try to shorten lead time.

Situations 3 and 4 include a CCR without a setup and a CCR with setup time requirements. Handling the CCR would be similar to handling a nonbottleneck, though more carefully. That is, a CCR has some idle time. It would be appropriate here to cut the size of some of the process batches so that there can be more frequent changes of product. This would decrease lead time and jobs would be more likely to be done on time. In a make-to-stock situation, cutting process batch sizes has a much more profound effect than increasing the number of transfer batches. This is because the resulting prdict mix is much greater leading to reduced WIP and production lead time. We are now in conditions of enunciate the three final rules of the OPT Philosophy:

7th rule
The transfer batch may not and many times should not be equal to the process batch.

8th rule
The process batch should be variable, not fixed.

Caskey and Storch [19], by enumerating a series of possibilities in a shop with three machines and various routings, arrive to the same conclusion as Goldratt. They found that the way a nonbottleneck is planned as a minimal impact on global performance. As long as the bottleneck is planned in a way that allows good performance, the choice of rules on the other machines have little influence on that same performance for the whole shop. The implication of this, as Goldratt sustains, is that management need only to be concerned with the bottlenecks, the operators of nonbottleneck machines can be given plenty of freedom to operate as they see fit. Only a very poor choice in sequencing rules can impact the rest of the shop.

9th rule
Schedules should be established by looking at all of the constraints simultaneously. Lead times are the result of a schedule and cannot de predetermined.

How to treat inventory and where to charge inventory costs

The traditional view of inventory is that its only negative impact on a firm´s performance is its carrying costs. We realize now inventory´s negative impact also comes from lengthening lead times and creating problems with engineering changes. (When an engineering change on a product comes through, which is frequent, product still within the production system often must be modified to include the changes. Therefore, less work-in-process reduces the number of engineering changes to be made).

Fox and Goldratt propose to treat inventory as a loan given to the manufacturing unit. The value of the loan is based only on the purchased items that are part of the inventory. As we stated earlier, inventory is treated as material cost only, and without any accounting type value added from production. If inventory is carried as a loan to manufacturing, we need a way to measure how long the loan is carried. One measurement is dollar days.

Dollar days

A useful performance measurement is the concept of dollar days, which is a measurement of the value of inventory and the time it stays within an area. To use this measure, we could simply multiply the total value of inventory by the number of days.

Supposing Department X carries an average inventory of $40,000, and on the average, the inventory stays within the department five days. In dollar days then, Department X is charged with $40,000 times five days or $200,000 dollar days of inventory. At this point one cannot say the $200,000 is high or low, but it does show where the inventory is located. Management can then see where it should focus attention and determine acceptable levels. Techniques can be instituted to try to reduce the number of dollar days while being careful that such a measure does not become a local objective (i.e., minimizing dollar days) and hurt the global objectives (such as increasing ROI, cash-flow, and net profit).

Dollar days could be beneficial in a variety of ways. Consider the current practice of using efficiencies of equipment utilization as a performance measurement. To get high utilizations, large amounts of inventory are held to keep everything working. However, high inventories would result in a high number of dollar days, which would discourage high levels of work in process. Dollar days measurements could also be used in other areas:

- Purchasing, to discourage placing large purchase orders that on the surface appear to take advantage of quantity discounts. This encourages JIT-purchasing.
- Manufacturing, to discourage large WIP and producing earlier than needed. This would promote rapid flow of material within the plant.
- Marketing, to discourage holding large amounts of finished-goods inventory. The net result would be to encourage sale of finished products.

4.8 Comparing OPT to MRP and JIT

MRP uses backward scheduling after having been fed a master production schedule. MRP schedules production through a bill-of-materials explosions in a backward manner - working backward in the time from the desired completion date. As a secondary procedure, MRP, through its capacity resource planning module, develops capacity utilization profiles of work centers. When work centers are overloaded, either the master production schedule must be adjusted or enough slack capacity must be left unscheduled in the system so that work can be smoothed at the local level (by work-center supervisors or the workers themselves). Trying to smooth capacity using MRP is so difficult and would require so many computer runs that capacity overloads and underloads are best left to local decisions, such as at machine centers. An MRP schedule becomes invalid just days after it was created.

The OPT manufacturing approach uses forward scheduling because it focuses on the critical resources. These are scheduled forward in time, ensuring that loads placed on them are within capacity. The noncritical (or nonbottleneck) resources are then scheduled to support the critical resources. (This can be done backward to minimize the length of time inventories are held). This procedure enables a feasible schedule. To help reduce lead time and WIP, in OPT the process batch size and transfer batch size are varied - a procedure that MRP is not able to do.

Comparing JIT to OPT , JIT does an excellent job in reducing lead times and work-in-process, but has several drawbacks:

1. JIT is limited to repetitive manufacturing;
2. JIT requires a stable production level (usually about a month long);
3. JIT does not allow very much flexibility in the products produced (products must be similar with a limited number of options);
4. JIT still requires work-in-process when used with the Kanban so that there is "something to pull". This means that completed work must be stored on the downstream side of each workstation to be pulled by the next workstation.
5. Vendors need to be located nearby because the system depends on smaller and more frequent deliveries.

This philosophy comprehends many of the techniques used by time-based companies. So, we may consider it as an overall solution for many of the problems encountered in batch manufacturing systems. Its strength lies in its capacity to focus on what is really important in systems controlling, allowing other resources to be managed in a more free way, being able adapt to the necessities of each particular situation.

To be transformed into computer language, we will need some type of language that allows us to cope with some factors that are not mathematically expressed in its definitions but they can rather be seen as human logic. Later, Fuzzy Logic will be presented as a tool in a very good position to become, for its own characteristics, the solution to this question

5. Priority Management

Managers, rather than seeking to optimize, must usually compromise. Formal systems don't recognize this fact, and so, they are always overlapped with an informal system. Priority management realizes this fact and tries to develop a theory that enhances better performance for manufacturing planning and control systems in the future.

5.1 Definition and Theory

In batch manufacturing companies, priority management manifested itself mainly in terms of orders: pursuing certain kind of orders to fill capacity, progressing supplies orders and choosing which customer orders are to have priority. In several situations, a pressure on cash-flow is typically responded to, by delaying (some) supplies orders and hastening the dispatch of higher-value customer orders - and budget pressures may lead to the same action (the "end-of-the-month" syndrome, as seen in chapter 4). Priority management addresses the question of which supply orders can be delayed, which customer orders can be hastened and, above all, what is the consequence, in terms of likely future priorities, of these choices. Of course, priorities may appear which are unrelated to specific orders, such as a moving bottleneck leading to low capacity utilization, or a specific capacity constraint leading to lengthening queues and lead times. But even these situations arise because of a volatile order mix or a more permanent change in the order-book pattern since process capacity was last adjusted, and, in both instances, the short- to medium-term response by management will be priority choices for specific orders or groups of orders. Thus, priority management might be defined as:

Priority Management
The allocation of resources, or the expression of preference, to specific order or order groupings (whether supplies, production, or customer orders), in response to current pressures on operational productivity and/or customer service, with the aim of relieving those pressures while at the same time promoting, or minimizing the deleterious impact upon, the wider economic and strategic goals of the company.

The ubiquity of this phenomenon, across all industries, has prompted the development of an emerging theory of Priority Management. The four main elements of this theory can be baldly stated:

1. Instability is a norm in many manufacturing operations, not a disruption of normal stability. It is endemic, and derives principally from variation, variety and volume effects on the physical system and the administrative system. (Section 2.1).
2. In these conditions, it is order administration which is the key activity for operations managers, and the focus for the continual rebalancing of supply (people, processes, material) and the demand (the order-book), and the trade-off between service level and resource utilization.
3. The allocation of order priority is typically a matter of judgment rather than technical analysis, although some analysis may precede the making of

choices. The lack of constants in a situation will normally preclude optimal "solutions".

4. Problem finding will prevail over problem solving. The questions asked, are not of the class "how can I resolve this conflict" but rather "what are the probable consequences of this choice".

5.2 Priorities at the Operational Level

Within the operations level, there are always two broad priorities: customer service and productivity. But these two goals are themselves frequently in conflict. Good service may require sequencing certain tasks to meet customers' due dates, while productivity may require a quite different to minimize machine change-over. So, even without the disturbance caused by the instability already discussed, there is an inherent conflict in the task of the operations manager. It is that extra instability, on top of a system which is itself required to pursue conflicting goals, which produces the preconditions for priority management.

Priority management does not mean merely "fire fighting" or moving from one production crisis to the next. Priority management implies a recognition that plans often cannot be completely executed, that forecasts are rarely completely accurate. For these and other reasons, running tight prescriptive planning and control systems such as MRP can be as much of a challenge, or even more, than coping with the uncertainty which they are meant to displace.

Most companies tend to make use of priority techniques based on due dates, critical ratios, or management assigned priority values. The major problem with most of the production scheduling techniques is that even when used for scheduling the work for the next day or shift, the schedule generated is out of date very soon due to random events such as machine/tool breakdowns, or the non-availability of materials. The foreman

5.3 Priority Management and Related Disciplines

The academic world, or indeed the manufacturing companies themselves, have not generally recognized the management of priorities as a discipline or general approach, and it does not, apparently, fit well with other approaches which are so recognized. Among other recognized approaches are operational research (OR) as an example of a rigorous academic subject, industrial engineering (IE) as an example of a manufacturing function, and manufacturing policy as an example of a recent approach relating the operation to market needs. In this section, priority management is further defined in terms of its relationship to these disciplines.

Priority management operates at a level between the strategic processes of policy decision making and the detailed focus, away from customer pressures, on a specific problem solving of OR and IE. Manufacturing policy makers are looking at the medium to long term, because the data which they use (past sales patterns, order-winning criteria, competitor analysis) take time to come into focus and the decisions which they make (e.g. on product range or capacity investment) take time to implement. Also the strategists are concerned less with customers and orders than with markets and sales levels. OR and IE are often both concerned with process rather than the whole operation, and their "customer" is management, not end customers. Their focus is problem solving, whereas priority management is more likely to be faced with "problem finding", i.e. trying to foresee the consequences of priority choices or comprehend the dynamics underlying the recurrent pressure from shifting priorities.

Thus, priority management occupies the middle ground between policy determination and the technical approaches for enhancing shop-floor efficiency. Its aim is to trade off customer service and productivity goals to ensure the continuing effectiveness of company operations, and since the skill with which this is done is a major factor in competitiveness, there is a link here with the policy approach, which is mainly concerned with competitiveness. Priority management is usually practiced at middle and even junior management levels, down to the first-line supervisors who are sometimes left with daily shop-floor initiatives which affect priority choices. The information needed to manage in this way will be primary concerned with status, with individual orders and their progress, with materials and capacity availability and with indications of relative priority (due date, value, grouping convenience). In all these respects, priority management contrasts with, and occupies a middle position between, the level of policy and the level of techniques - Table 5.3.1.

	Action timespan	Managerial Level	Focus	Information needs	Output	Aims
Strategy (manufacturing policy)	Long term (one to five years)	Top management	Markets, competitors	Environmental, sales patterns and projections, competitors results	Policy documents on products, locations, growth plans	Reposition company to enhance competitiveness
Priority management	Short to medium term (one week to six months)	Middle and junior management	Customers, suppliers, orders	Current status, availability, relative priority	Priority choices, order placements or delays, capacity allocations	Trade off customer service and productivity goals
Efficiency techniques (OR, IE, OM)	Determined by the need for result and complexity of problem	Professional management services	Problem solving	Synthetic or historic data on system	Optimized solution	Process efficiency (IE), establish model viability (OR)

Table 5.3.1 - Priority Management in Relation to the Strategic and Technical Approaches.[20]

The problems discussed in chapter two, with special relevance to the existence of an informal system that is always preferred to the formal one because the last one is not appropriate to the company reality, are faced by Priority Management which tries to provide a way to deal with them. A very important point in this new vision is that it recognizes that human thinking and judgement have to be brought in discussion and no longer "academic" models of a "false" reality will have, and not denying their utility, the monopoly of methods to control production environments that suffer from instability. Once again, it seems that Fuzzy Logic, as we states in the former chapter, will be able to provide useful contributions to erase the problems once found in those systems.

6. Fuzzy Logic

This chapter consists of a presentation of the fuzzy logic basic concepts and is the compilation of references [21], [22] and [23]. The precise mathematical model will not be seen. Through an example of creditworthiness evaluation, the power of fuzzy logic in decision supporting and making will be demonstrated.

6.1 Fuzzy Logic Primer

What is fuzzy logic?
Fuzzy Logic is an innovative technology that enhances conventional system design with engineering expertise. Using fuzzy logic, you will circumvent the need for rigorous mathematical modelling. Fudge factors in your control system will be replaced by self-explanatory linguistic description of the control strategy.

Fuzzy Logic provides an easy and transparent way to incorporate common-sense type reasoning in informatic applications. By introducing a means of coping with "soft facts", "soft criteria", and "fuzzy data", you can implement humanlike decision making in your applications.

Why must software make decisions?
Enabling software to make humanlike decisions yelds many benefits:

- In decisions that need to be taken in large quantities, such as buy/sell decisions in a stock trading system, automation of decision making greatly expands capacity at low cost.
- Automation of decision making leads to a completely reproducible decision making process.
- Complex decision processes become transparent and can, thus, explicitily be evaluated and optimized.
- The experience of more than one person can be agglomerated into a single system.

Why fuzzy Logic?
However, to develop systems that shall take humanlike decisions, mathematical models come short. This is due to the fact that human judgement and evaluation simply does not follow Boolean Logic nor any other conventional mathematical discipline.

Most of our traditional tools for formal modeling, reasoning, and computing are crisp, deterministic, and precise in character. By crisp we mean dichotomous, that is, yes-or-no type rather than mor-or-less type. In conventional dual logic, for instance, a statement can be true or false - and nothing in between. In set theory, an element can either belong to a set or not; and in optimization, a solution is either feasible or not.

Fuzzy Logics Underlying Priciple
As humans we often rely on imprecise expressions like "usually", "expensive", or "far". But the comprehension of a computer is limitd to a black-white, everything-or-nothing,

or true-false mode of thinking. In this context, Lofti Zadeh, who published in 1965 the first work on Fuzzy Logic, emphasizes the fact that we easily let ourselves be drug along by a desire to attain the highest possible precision without paying attention to the imprecise character of reality.

As succintely as possible, what fuzzy logic offers in the realm of systems analysis and design is a **methodology for computing with words**.

6.2 Basic Concepts of Fuzzy Logic

Mathematical Principles of Uncertainty
Many mathematical disciplines deal with the description of uncertainties, such as probability theory, information theory and fuzzy set theory. It is most convenient to classify these by the type of uncertainty they treat. In this section, I will confine to only two types of uncertainty, stochastic and lexical uncertainty.

Stochastic Uncertainty
Stochastic uncertainty deals with the uncertainty toward the occurrence of a certain event. Consider the statement:

Statement 1
The probability of hitting the target is 0.8.

The event itself - hitting the target - is well defined. The uncertainty in this statement is whether the target is hit or not. This uncertainty is quantified by a degree of probability. In the case of the statement, the probability is 0.8. Statements like this can be processed and combined with other statements using stochastic methods, such as the Bayesian calculus of conditional probability.

Lexical Uncertainty
A different type of uncertainty lies in human languages, the so-called lexical uncertainty. This type of uncertainty deals with the imprecision that is inherent to most words humans use to evaluate concepts and derive conclusions. Consider words such as "tall men", "hot days", or "stable currencies", where no exact definitions underlie. Whether a man is considered "tall" hinges on many factors. A child has a different concept of a "tall" man than an adult. Also the context and the background of a person making an evaluation plays a role. Even for one single person, an exact definition on whether a man is considered "tall" does not exist. No law exists that determines the threshold above which a man is conceived "tall". This would not make sense anyhow, since a law that defines all men taller than 6' 4" to be "tall" would imply that a man with 6' 3" is not tall at all. The science that deals with the way humans evaluate concepts and derive decisions is psycho linguistics. It has been proven that humans use words as "subjective category" to classify figures such as "height", "temperature", and "inflation". Using these subjective categories, things in real world are evaluated to which degree they satisfy the criteria.

Even though most concepts used are not precisely defined, humans can use them for quite complex evaluations and decisions that are based on many different factors. By using abstraction and by thinking in analogies, a few sentences can describe complex contexts that would be very hard to model with mathematical precision. Consider the statement:

Statement 2
We will probably have a successful financial year.

On a first glance, statement 2 is very similar to statement 1. However, there are significant differences. First, the event itself is not clearly defined. For some companies, a successful financial year means that they deferred bankruptcy, for others it means to have surpassed last years profit. Even for one company, no fixed threshold exists to define whether a fiscal year is considered to be successful or not. Hence, the concept of a "successful fiscal year" is a subjective category.

Another difference lies in the definition of expressing probability. While in statement 1, the probability is expressed in a mathematical sense, statement 2 does not quantify a probability. If someone expresses that a certain type of airplane probably has problems, the actual probability can well be lower than 10%, still justifying this judgement. If someone expresses that the food in a certain expensive restaurant is probably good, the actual probability can well be higher than 90%. Hence, the expression of probability in statement 2 is a perceived probability rather than a mathematically defined probability as in statement 1. In statement 2, the expression of probability is also a subjective category just as "tall men".

Modeling Linguistic Uncertainty

Statements using subjective categories such as statement 2 play a major role in the decision making process of humans. Even though these statements do not have quantitative contents, humans can use them successfully for complex evaluations. In many cases the uncertainty that lies in the definition of the words we use adds a certain flexibility. Consider for illustration the annual wage increase negotiations between unions and industry. Both want to achieve the same goal: an appropriate wage increase. The problem only starts when they have to express in percentage, what they mean with "appropriate".

The flexibility that lies in words and statements we employ is made use of widely in our society. In most western societies, the legal system consists of a certain number of laws, that each describe a different situation. For example, one law could express that theft of a car should be punished with 2 years of prison. Another law could define diminished responsibility. In one case in court, the judge now has to decide the exact number of days in prison for a thief that stole a car under the influence of 0.1% blood alcohol, that had a bad childhood and was left by his spouse the day before. As not for each "real" case a specific law exists, the judge has to combine all applying laws to derive a fair decision. This is only possible due to the flexibility in the definition of the words and statements used in each law.

Fuzzy Logic as Human Logic

The basic idea is simple: in reality, you cannot define a rule for each possible case. Exact rules (or laws) that cover the respective case perfectly, can only be defined for a few distinct cases. These rules are discrete points in the continuum of possible cases and humans approximate between them. Hence, for a given situation, humans combine the rules that describe similar situations. This approximation is possible due to the flexibility in the definition of the words that constitute the rules. Likewise, abstraction and thinking in analogies is only rendered possible by the flexibility of "human logic".

To implement this human logic in engineering solutions, a mathematical model is required. Fuzzy logic has been developed as such a mathematical model. It allows to represent human decision and evaluation processes in algorithmic form. There are limits to what fuzzy logic can do. The full scope of human thinking, fantasy and creativity can not be mimiced with fuzzy logic. However, fuzzy logic can derive a solution for a given case out of rules that have been defined for similar cases. So, if you can describe the desired performance of a technical system for certain distinct cases by rules, fuzzy logic will effectively put this knowledge to a solution.

Fuzzy Logic vs. Probability Theory

Especially people working extensively with probability theory have denied the usefulness of fuzzy logic in applications. They claim, that all kinds of uncertainty can be expressed with probability theory. Rather than embarking on a discussion of whether this is true or just another case of the previously cited "hammer principle", consider statement 3. If you find such a statement in a medical textbook, and you want to implement it in a system, it looks very easy on the first glance. If you got a patient that suffers from strong fever, has no yellowish colored skin but nausea, you can compute the probability for a hepatitis infection using the Bayesian calculus.

Statement 3

Patients suffering from hepatitis show in 60% of all cases strong fever, in 45% of all cases a yellowish colored skin, and in 30% of all cases nausea.

Although this looks very easy, the problem starts when you have to define what a "strong fever" is. If you read medical books or ask doctors, you will not get an equivocal answer. Even if most doctors will agree that the threshold is at about 39°C (102°F), this does not mean that a patient with 101.9°F does not at all have a strong fever while another patient with 102°F fully has a strong fever.

If a threshold for "strong fever" would exist, the reverse must also exist. That is, that a very precisely measured body temperature will result in a very precise diagnosis. If this would be true, you can measure your body temperature up to the fifth significant figure and expect a doctor to tell you just out of this very precise information, what disease you suffer from. In contrast to this, a doctor will get a competent diagnosis not from the precision of a single parameter but much rather from evaluating many symptoms' parameters. Here, the precision of each parameter does, for the most part, not imply the quality of the result. If the doctor asks you whether you sweat at night, he is most likely not interested in the precise amount in ounces but rather a tendency.

As the example illustrates, stochastic uncertainty and linguistic uncertainty are of different nature. Stochastic uncertainty deals with the uncertainty of whether a certain event will take place and probability theory lets you model this. In contrast, Lexical uncertainty deals with the uncertainty of the definition if the event itself. Probability theory cannot be used to model this as the combination of subjective categories in human decision processes does not follow its axioms.

6.3 A "fuzzy" Set

Now how can you model linguistic uncertainty adequately? If a doctor does not have a precise threshold in mind when evaluating whether a patient suffers from "strong fever", how does it work? Psycho-linguistic research has shown that a doctor would compare the patient with two "prototypes". On one side the "perfect" strong fever patient, pale, sweating, with the shivers. On the other side the "perfect" well-tempered patient that does not show any signs of fever at all. Comparing with these two extremes, a doctor evaluates where in-between the two his patient ranks.

How can this be mathematically modeled? Consider set theory, where you would first define the set of all patients with strong fever. Then you define a mathematical function that indicates for each patient whether he is a member of this set or not. In conventional math, this indicator function has to uniquely identify each patient as member or non-member of the set. As pointed out before, a doctor rather evaluates the degree to which his patient matches the prototype of a strong fever patient. Figure 6.3.1 gives an example of a set where certain elements can also be "more-or-less" members. The "shade of gray" indicates the degree to which the body temperature belongs to the set of strong fever. This "shade of gray" that looks as if the black area in the figure is "fuzzy" gave the name "fuzzy logic".

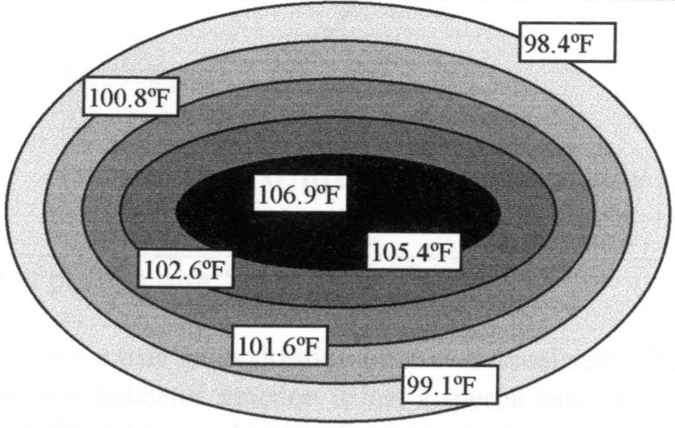


Figure 6.3.1 - Fuzzy Set of patients with “strong fever”

In Figure 6.3.1, each body temperature is associated with a certain degree to which it matches the prototype for "strong fever". This degree is called the "degree of membership" of the element $x \in X$ to the set "strong fever" SF. The body temperature is called a "base variable" x with the universe X. The range of μ is from 0 to 1, representing absolutely no membership to the set and complete membership, respectively. As a temperature of 94°F - Example 1 - would have no membership at all, a temperature of 110°F would have complete membership. Temperatures between are member to the set only to a certain degree.

Example 1

$\mu_{SF}(94^\circ F) = 0$
$\mu_{SF}(96^\circ F) = 0$
$\mu_{SF}(98^\circ F) = 0$
$\mu_{SF}(100^\circ F) = 0.1$
$\mu_{SF}(102^\circ F) = 0.35$
$\mu_{SF}(104^\circ F) = 0.65$
$\mu_{SF}(106^\circ F) = 0.9$
$\mu_{SF}(108^\circ F) = 1$
$\mu_{SF}(110^\circ F) = 1$

Membership Functions

The degree to which the value of a technical figure satisfies the linguistic concept of the term of a linguistic variable is called degree of membership. For a continous variable, this degree is expressed by a function called membership function (MBF). The membership functions map each value of the technical figure to the membership degree to the linguistic terms. The technical quantity is called the base variable. Usually, one draws the membership functions for all terms in the same diagramm.

The degree of membership - in the figure 6.3.2, μSF(x) of the temperature x - can be represented by a continuous function. Note, that a temperature of 102°F and a temperature of 101.9°F are evaluated differently, but just as a slight bit and not as a threshold.

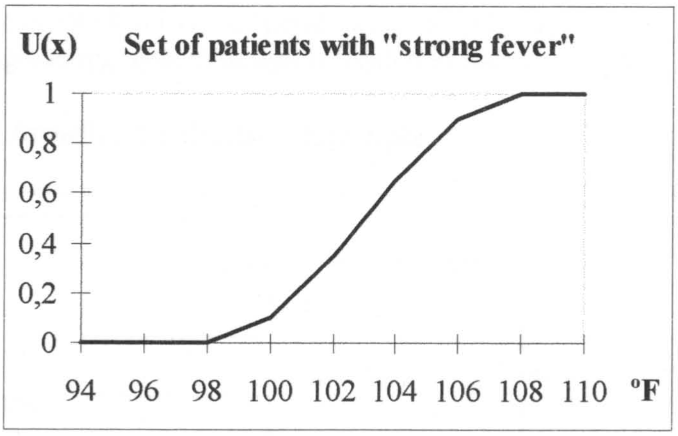


Figure 6.3.2 - Set of patients with “strong fever”

Note, that fuzzy sets are a true generalization of conventional sets. The cases $\mu=0$ and $\mu=1$ of the conventional indicator function is just a special case of the fuzzy set. The use of fuzzy sets defined by membership functions in logical expressions is called "fuzzy logic". Here, the degree of membership in a set becomes the degree of truth of a statement. For example, the expression "the patient has strong fever" would be true to the degree of 0.65 for a temperature of 104°F.

Linguistic Variables

The primary building block of any fuzzy logic system is the so-called "linguistic variable". Here, multiple subjective categories describing the same context are combined. In the case of fever, not only strong fever but also raised temperature, normal temperature, and low temperature exist. These are called "linguistic terms" and represent the possible values of a linguistic variable. Figure 6.3.3 plots the membership functions of all terms of the linguistic variable fever into the same graph.

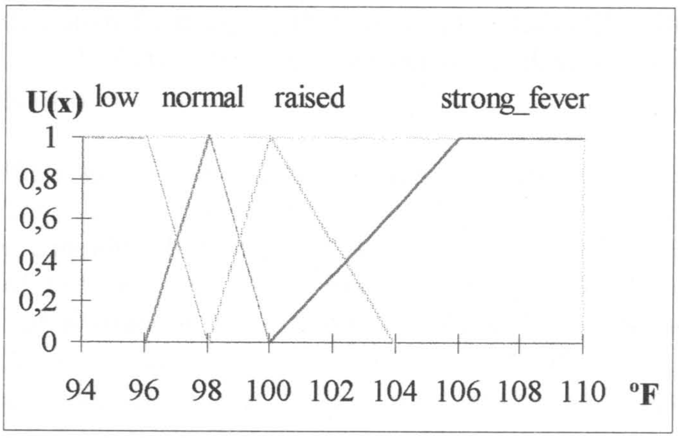


Figure 6.3.3 - Linguistic Terms of the Linguistic Variable “Body Temperature”

A linguistic variable translates real values into linguistic values. This linguistic variable now allows for the translation of a measured body temperature, given in Fahrenheit, into its linguistic description. For example, a body temperature of 102°F would be evaluated as "pretty much raised temperature, jut slightly high fever". How to use this technology in engineering system design is treated in the next section.

6.4 Financial Liquidity Evaluation Example

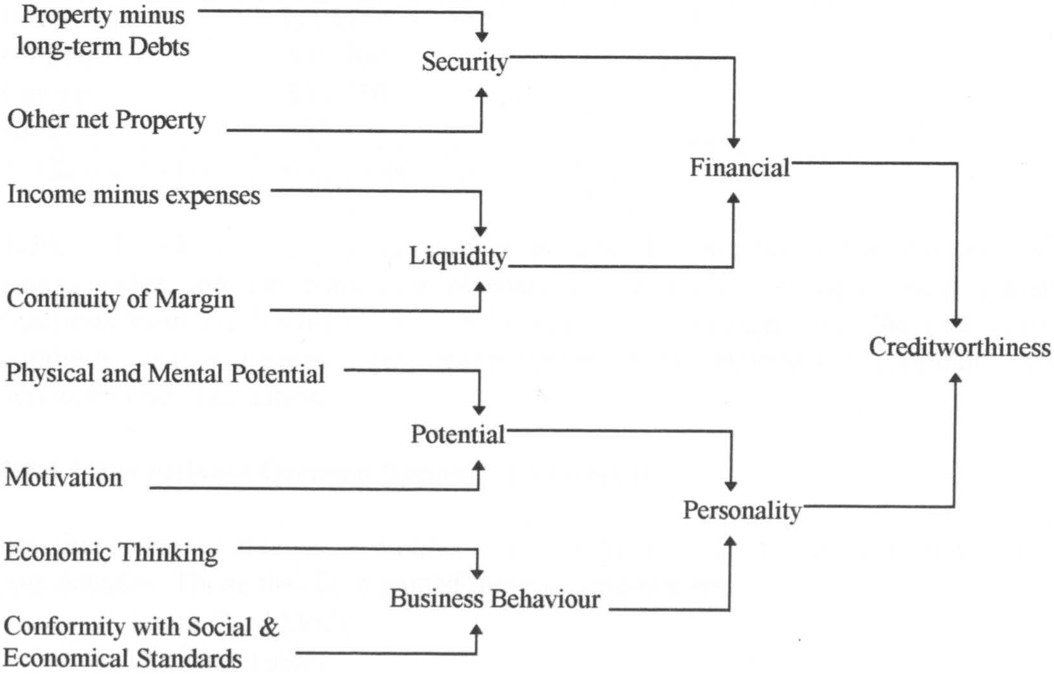


Figure 6.4.1 - Aggregation Hierarchy of a Fuzzy Logic Creditworthiness Estimation

This case study is taken from a real-world fuzzy logic system implemented in 1986 in Europe. To evaluate the creditworthiness of bank customers for consumer loans, multiple financial and personal factors are used to decide on whether a loan is granted or not. Figure 6.4.1 shows a simplified layout of the system's structure. A total of 8 input variables is used for the evaluation. Note that the upper 4 variables denote "hard facts", numbers that stem from inputs that are given numerically. The lower 4 input variables denote "soft facts", that are subjective evaluations of the applicants personality and lifestyle.

The structure of the system is hierarchical, at each node, two elements are aggregated to a new one. This makes three layers of abstraction. The first layer contains the elements Security, Liquidity, Potential and Business Behaviour, which each comprise the information of two input variables. Because information is condensed in each nodes, we speak of abstraction here. Similar to a human, which takes many input variables into account to come up with one abstract judgment, the aggregation hierarchy proceeds until the output node, Creditworthiness is reached, the most abstract information in the hierarchy.

In a way, such hierarchical decision model stepwise "squeezes" the desired information on creditworthiness out of the raw date (the 8 input variables).

While the design of such a decision hierarchy is relatively easy, it is difficult to decide on how information shall be combined at each node. For example, the node defining Liquidity, uses the two inputs Income minus Expenses and Continuity of Margin. As stated before, such an evaluation exposes no problem for a human, but is not trivial to put into a mathematical model.

Customer	Income minus Expenses	Continuity of Margin
Boris	$118,000	0.12
Clinton	$45,700	0.71
Cavaco	$94,250	0.89
Kohl	$37,400	0.22

Table 6.4.2 - Four Sample Customers

Table 6.4.2 shows four examples of customers. The income minus expenses is the annual value and the continuity of margin is a value computed using statistical functions from the fluctuation of the income minus expenses over the past years. A continuity near 1 denotes a very stable income minus expenses, a continuity near 0 denotes a high fluctuation.

6.4.1 Conventional Decision Support Techniques

For these type of models, a number of different techniques was developed over the past decades. Those that have gained more acceptance are:

- Score Card Models
- Decision Tables
- Expert Systems

Expert Systems

Before it is demonstrated how fuzzy logic can better handle such evaluation process modeling, let me introduce expert systems. Expert systems contain rules that specifically address cases such as that of a customer with no or low cash flow is not considered liquid, no matter how continuous his financial situation is. Often, these rules are formulated in the "If-Then" format, such as:

(1) IF Income < \$20,000 THEN Liquidity = 0.0
(2) IF \$20,000 ≤ Income < \$50,000 AND Continuity < 0.2 THEN Liquidity = 0
(3) IF \$20,000 ≤ Income < \$50,000 AND 0.2 ≤ Continuity < 0.4 THEN Liquidity = 0.2
(4) IF \$20,000 ≤ Income < \$50,000 AND 0.42 ≤ Continuity THEN Liquidity = 0.3
(5) IF \$50,000 ≤ Income < \$80,000 AND Continuity < 0.2 THEN Liquidity = 0.3
(6) ...

The total number of rules necessary to design a system of the complexity as shown in Figure 6.4.1 will still be in the multiple hundreds. This is because each rule of an expert system can only describe one specific situation that is identified by the condition intervals of the rule's premise. Hence, to have a smooth decision response, the discretisation must be chosen very fine, keeping the number of rules rather high. Humans, in contrast, do not use that many rules. They use rules that rather represent "tendencies" and "decisions in general" than each and every situation. In these rules, no "hard" criteria such as "\$20,000 ≤ Income < \$50,000 AND 0.42 ≤ Continuity" exists. Humans rather use soft criteria, such as "medium income" or "high continuity" that can be represented by fuzzy logic.

In expert systems, the criteria has to be hard as the rules are evaluated by a so-called inference engine that bases on Boolean logic. In Boolean logic, a statement can only be 100% true or 100% false, no "somewhat", "slightly", or "pretty much".

6.4.2 Linguistic Decision Making

The last section showed that while "If-Then" rules are a transparent way to express human decision making, the hard facts stated by Boolean expressions in these rules make them inappropriate. Hence, most fuzzy logic applications use "If-Then" rules, but the way they are computed is entirely different from expert systems. Using fuzzy logic for this computation allows the use of "soft facts". For example, a fuzzy logic rule could be:

(A) IF Income IS high AND Continuity Is Medium THEN Liquidity IS High

At first glance, this rule (A) looks similar to rules (1)...(5) shown in Section 6.4.1. At a second glance, they are very different. Rule (A) uses words such as "high" or "medium" that are not as clearly defined as "\$20.000 ≤ Income < \$50.000". As discussed earlier, such concepts can be represented well as linguistic variables.

A linguistic variable translates a numerical value into a linguistic value. Just as the numerical description of a body temperature of 100°F (37.8°C) is translated to the linguistic description "raised temperature, almost no high fever" by the linguistic variable shown in Figure 6.3.3; in the Liquidity assessment example, the numerical input variables Income and Continuity need to be translated to linguistic values. This step is called "fuzzification" as it uses fuzzy sets for this translation.

Figure 6.4.3 shows the complete structure of a fuzzy logic system. Once all input variable values are translated into respective linguistic variable values, the so-called "fuzzy inference" step evaluates the set of If-Then rules that define the evaluation. The result of this is again a linguistic value for the linguistic variable Liquidity. For example, the linguistic result could be "a little less than medium". The so-called "deffuzification" step translates this linguistic result into a numerical value that represents the Liquidity as a number.

The benefit of this is sketched by Figure 6.4.3. The decision process can better be described on a linguistic level using rules with soft facts as rule (A) exemplifies. However, input information is given numerically, and the output is required to be numerical too. Thus, to compute the fuzzy logic rules, two translation steps, fuzzification and defuzzification, are required. These two translation steps are the cost involved for getting the benefit of having rules that are computed in humanlike way. For both transition steps, a link between the numerical representation and the linguistic representation of a variable is needed. This link is the set of membership function contained in each linguistic variable definition. Both fuzzification and defuzzification use the membership function definition to compute the translation step.

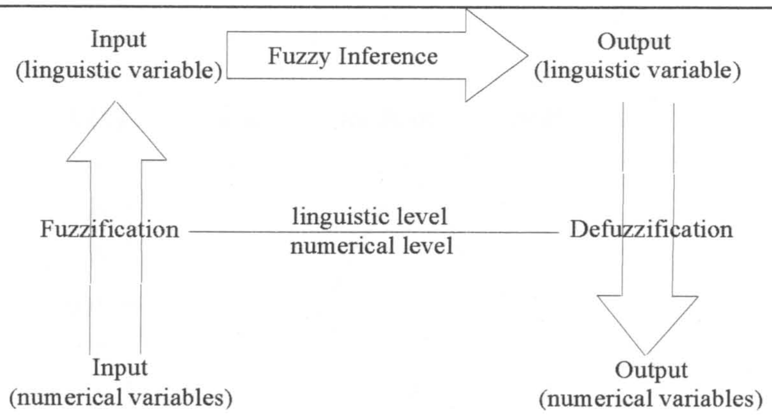


Figure 6.4.3 - Structure of a Fuzzy Logic System. The Fuzzy Logic Computation Consists of Three Steps: Fuzzification, Fuzzy Inference and Defuzzification.

6.4.3 The Fuzzy Logic Algorithm

In a nutshell, fuzzy logic is a technology that translates natural language descriptions of decision policies into an algorithm by a mathematical model. This mathematical model consists of three major sections: fuzzification, inference and defuzzification.

6.4.3.1 Fuzzification using Linguistic Variables

Linguistic variables have to be defined for all variables used in the if-then rules. Possible values of a linguistic variable are called terms. For the liquidity assessment example, the terms are:

Linguistic Variable	Possible Values - Terms
1. Income	∈ {low, medium, high}
2. Continuity	∈ {low, medium, high}
3. Liquidity	∈ {very_low, low, medium, high, very_high}

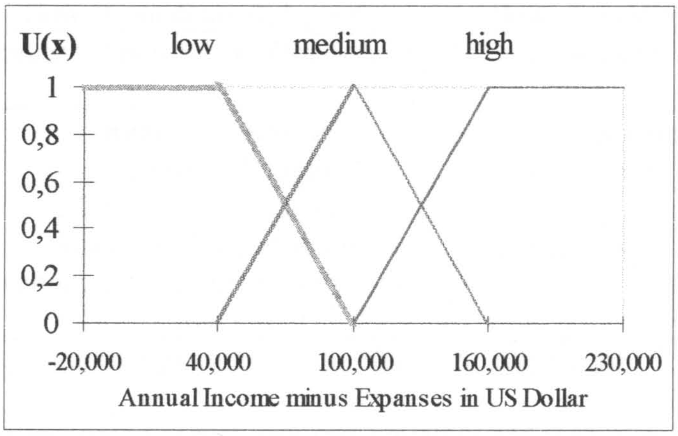


Figure 6.4.4 - Linguistic Variable "Income"

For every linguistic variable, each term is defined by its membership function. Figures 6.4.4 and 6.4.5 show possible definitions for the two input variables.

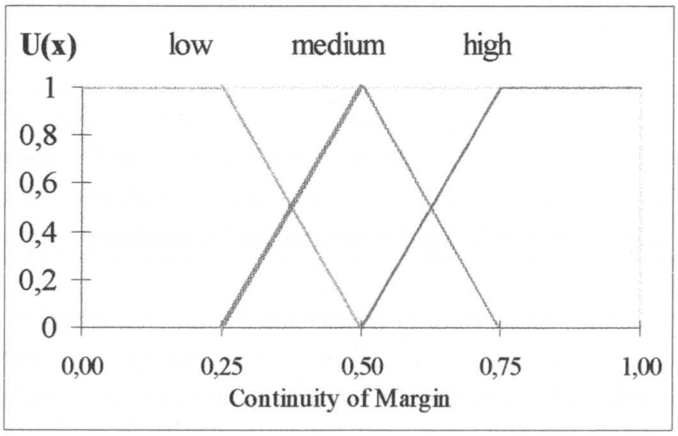


Figure 6.4.5 - Linguistic Variable "Continuity"

Consider the sample of customers of Table 6.4.2. Table 6.4.6 gives the translation of these numerical values into linguistic ones.

Note, that the value of a linguistic variable can be expressed in two ways. One is a vector of degree of membership to the fuzzy sets that define the terms of a linguistic variable. This is displayed in Table 6.4.6. Also, one can express this linguistic value only by words. The degrees of membership are then expressed by words such as "slightly", "almost", "little", "fairly", "rather", "very much" or "completely". Table 6.4.7 shows a possible interpretation of the values.

Customer	Linguistic Value of Income minus Expenses	Linguistic Value of Continuity of Margin
Boris	{low=0, medium=0.85, high=0.15}	{low=1, medium=0, high=0}
Clinton	{low=0.8, medium=0.2, high=0}	{low=0, medium=0.1, high=0.9}
Cavaco	{low=0.1, medium=0.9, high=0}	{low=0, medium=0, high=1}
Kohl	{low=1, medium=0, high=0}	{low=1, medium=0, high=0}

Table 6.4.6 - Linguistic Translation of the Numerical Customer Data

Customer	Linguistic Value of Income minus Expenses	Linguistic Value of Continuity of Margin
Boris	rather medium, little high	low
Clinton	rather low, little medium	very much high, slightly medium
Cavaco	very much medium, slightly low	high
Kohl	low	low

Table 6.4.7 - Linguistic Interpretation of the Values of Table 6.4.6

Section 6.1 discusses how to design linguistic variables and their membership functions for a given application.

6.4.3.2 Fuzzy Logic Inference Using If-Then Rules

Now that all numeric input values have been converted to linguistic values, the fuzzy inference step can identify the rules that apply to the current case and can compute the values of the output linguistic variables. Here we show a subset of four possible rules for illustration:

Rule α: IF Income = low THEN Liquidity = very_low
Rule β: IF Continuity = low AND Income = medium THEN Liquidity = low
Rule χ: IF Continuity = high AND Income = medium THEN Liquidity = high
Rule δ: IF Continuity = medium AND Income = high THEN Liquidity = medium

The computation of the fuzzy inference consists of two components:

- **Aggregation**: computation of the IF part of the rules
- **Composition**: computation of the THEN part of the rules

Aggregation

The IF part of rule χ combines the two conditions "Continuity = high" and "Income = medium". The IF part defines whether the rule is valid for the current case or not. In conventional logic, the combination of the two conditions can be computed by the Boolean AND, that is shown in the following table:

A	B	A∧B
0	0	0
0	1	0
1	0	0
1	1	1

In the case of fuzzy logic, the Boolean AND cannot be used as it cannot cope with conditions that are "more-or-less" true. Hence, new operators had to be defined for fuzzy logic to represent logical connectives such as AND, OR and NOT. The first set of operators that has been proposed are give below. These three operators are used in the majority of today's fuzzy logic applications:

AND: $\mu_{A \wedge B} = min \{ \mu_A; \mu_B \}$
OR: $\mu_{A \vee B} = max \{ \mu_A; \mu_B \}$
NOT: $\mu_{\neg A} = 1 - \mu_A$

If you use the *min* operator to represent the logical AND, the IF parts of the rules of the previous example would compute for customer Boris as shown here:

Rule α: *min* $\{0.8\} = 0.8$
Rule β: *min* $\{0.0, 0.2\} = 0.0$
Rule χ: *min* $\{0.9, 0.2\} = 0.2$
Rule δ: *min* $\{0.1, 0.0\} = 0.0$

These results ate the degrees of truth of the IF parts and hence indicate how adequate for the current case each rule is.

Composition
Each rule defines the evaluation result for certain prototypical case in the THEN part. The degree to which the evaluation result is valid is given by the adequateness of the rule to the current case. This adequateness is computed by the aggregation as the degree of truth of the IF part. Hence, rule α has the evaluation result "Liquidity = very_low" to the degree 0.8, and the rule χ the result "Liquidity = high" to the degree 0.2.

Hence, the value of the linguistic variable "Liquidity" for customer Boris would be:

Liquidity = {very_low = 0.8, low = 0.0, medium = 0.0, high = 0.2, very_high = 0.0}

A linguistic interpretation of this result could be "very much low, slightly high liquidity".

In some applications, a linguistic interpretation of the result is sufficient. For example, if the result is used to give a language-type answer. In other applications, a numerical value as output is required, for example to rank cases or - as in the creditworthiness example - decide on acceptance or rejectal of a credit application. In case a numerical output is required, a defuzzification step has to follow the fuzzy logic inference.

Fuzzy Associative Memories
Of this simple fuzzy logic inference principle, a number of extensions exist. Of these, only few have gained practical relevance. The most common is the association of rules with a weight factor. Such a weight represents the importance of the rule in relevance to the other rules in the system. The use of weights is the most simple and transparent implementation of more general concepts such as Fuzzy Associative Memories or the

Compositional Rule of Inference. In the fuzzy logic inference, the weight in the interval [0,1] is multiplied with the aggregation result in the composition step.

Fuzzy Rules

A possible interpretation of the weight factor is to consider rules themselves as fuzzy. That is, in the simple fuzzy logic inference principle, each rule either is member to the set of valid rules or nor. That is Boolean logic again, only truth degrees of 0 and 1 allowed. If you extend this truth degrees to the continuous interval [0,1], the set of valid rules becomes a fuzzy set and, hence, allows for the definition of "more-or-less" valid rules.

6.4.3.3 Deffuzication using Linguistic Variables

At the end of the fuzzy logic inference, the result for Liquidity is given as the value of a linguistic variable. To use this value for comparisons or ranking, it has to be translated into a numerical value. This step is called deffuzification. The relation between linguistic values and corresponding real values is always given by the membership function definition. Figure 6.4.8 plots the membership functions for the linguistic variable "Liquidity".

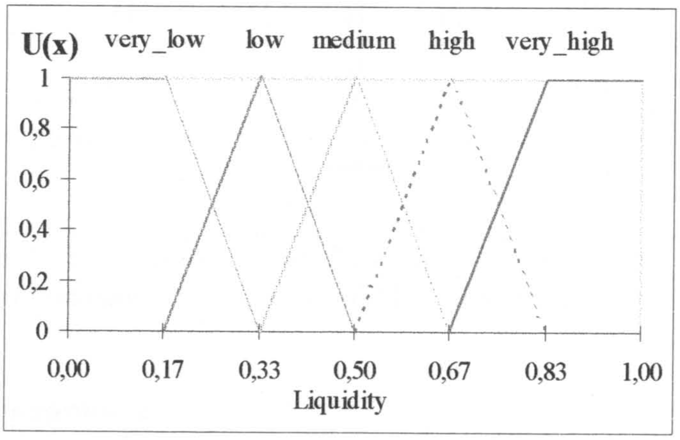


Figure 6.4.8 - Linguistic Variable "Liquidity"

The result of the fuzzy inference given in preceding section is both fuzzy and ambiguous as two different actions have non-zero truth degrees. How can two conflicting actions that are defined as fuzzy sets be combined to a numerical output for Liquidity? Consider how humans solve the problem of combining two fuzzy and conflicting actions in the following example:

Consider yourself in an apartment house at 11pm. You would like to listen to some music, such as Wagner. Music that requires some volume to be fun. On the other hand, your neighbours have already suffered quite a bit from your recent music sessions. Now, when you set the volume on your stereo, you have to combine these two conflicting and fuzzy goals into a crisp value, as only such a value can be set at the volume knob of your stereo. To find a volume that compromises the two goals, you could turn on the music and tune the volume until you balanced out the two goals .

As fuzzy logic mimics the human decision and evaluating process, a good defuzzification method should approximate this approach. Most deffuzification methods use a two step approach for this. In the first step, a "typical" value is computed for each term in the linguistic variable. In the second step, the "best compromise" is determined by "balancing" out the results.

Compute the "Typical" Values

The most common approach to compute the typical values of each term is to find the maximum of the respective membership function. If the membership function has a maximizing interval, the median of the maximizing set is chooses. For the linguistic variable Liquidity as shown in Figure 6.4.8, the computation of the typical values is illustrated in Figure 6.4.9.

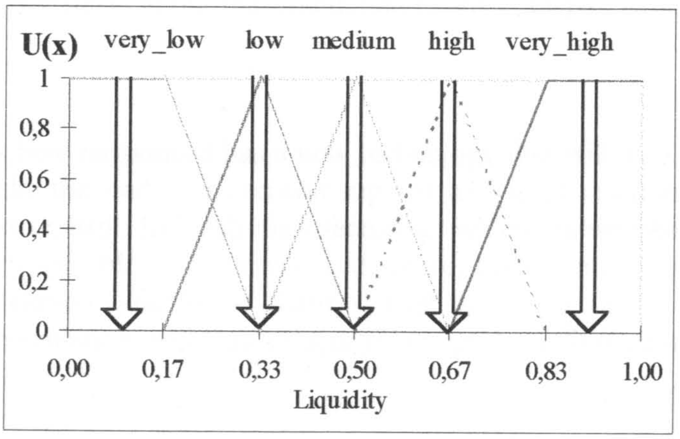


Figure 6.4.9 - In the First Step of Defuzzification, the Typical Value for each Term is Computed as the Maximum of the Respective Membership Function

Find the Best Compromise

In the second step, the best compromising crisp value for the linguistic result is computed. Figure 6.4.10 illustrates this step. At the horizontal position of the typical values, a weight of a size proportional to the degree to which the action is true is put. The weights are shown as the heights of the black arrow. The compromising crisp value is then determined by balancing the weights. In the example, the position that balances the fuzzy inference result is at the position of 0.27. This value is considered the best compromise for the Liquidity assessment.

This method of deffuzification is called "Center-of-Maximum" and is identical to the Center-of-Gravity" method using singleton membership functions. These defuzzification methods are used in most fuzzy logic implementations.

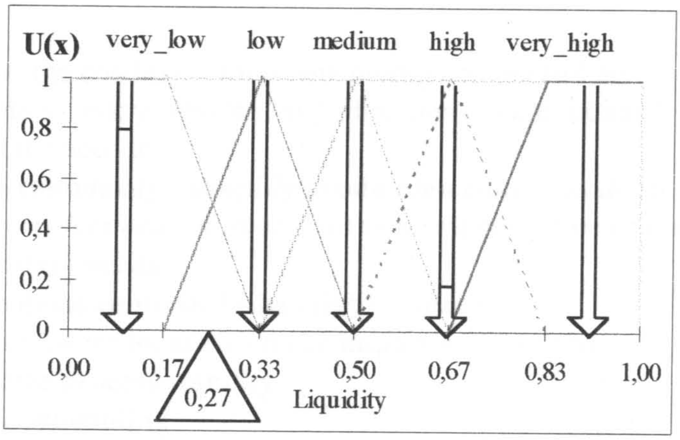


Figure 6.4.10 - By Balancing out the Conflicting Results, a Crisp Result is Found

Fuzzy Logic was here introduced has a new technology that will, in a near future, and in my opinion, take the lead as a decision support tool in production control, as it is already in process control. In batch manufacturing, which, as we saw in the previous chapters, is a very unstable environment, and where, the despite of being known, the variables have a relation which is very hard to establish, Fuzzy Logic, allied with other tools, especially Simulation, will play a major role in problem finding and solving.

7. Production Control - Basic Concepts

As seen until now, the conditions for a new production control age exist. Through the 6 previous chapters, many reasons and directions were pointed out. The main conclusions show the need to:

- **reduce complexity - simplify production control systems**
- **shift focus from cost and control to produce fast responses**
- **reduce data needs**
- **reduce items controlled - less tight control**
- **improve buffer location - more important than size**
- **guarantee process stability**
- **feasible scheduling**
- **shape control system to company instead of imposing formal system**

7.1 Concepts

The new vision on production control has some points that are essential to its understanding. This points are:

1. What to control centrally and locally
2. Layout
3. Batch sizes
4. Where to place buffer barriers to accommodate uncertainty
5. Time fences for flexibility
6. Scheduling - rough but feasible
7. Process stability
8. Enterprise objectives combination

7.1.1 What to Control Centrally and Locally

Machines

Has discussed in chapter 4 - OPT Philosophy, new production control methods should focus centrally only on bottleneck resources. Nonbottlenecks, for all the reasons discussed, are better controlled locally. Bottlenecks govern both throughput and inventory. So, as bottleneck performance is vital to system success, production control must first be placed on bottleneck controlling.

- **Bottlenecks** - controlled at a central level by a production planning and control system. This control will be a precise one, for the waste of time in a bottleneck represents a loss for the entire system.
- **Nonbottlenecks** - controlled at local shop-floor level by workers. Shop-floor workers are more aware of the constant changes in the situation - "they know it better" -, improving this way the control task performance and reducing the waste of time in information loop-backs to management.

Materials

As for materials, we can take a look to a standard Pareto Analysis - Figure 7.1.1. The graphic displays the typical configuration of the ABC curve. We can separate the materials into three classes:

- **Class A** - with about 20% of materials, corresponding to 80% of inventory value.
- **Class B** - with about 30% of materials, corresponding to 15% of inventory value.
- **Class C** - with about 50% of materials, corresponding to 5% of inventory value.

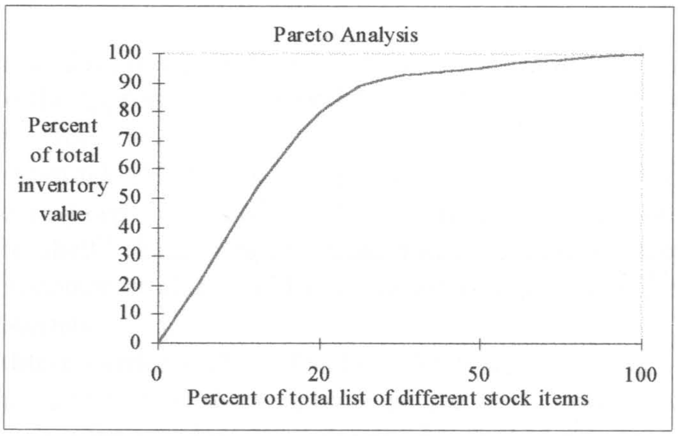


Figure 7.1.1 - Pareto Analysis Graphic

Trying to keep a less tight control and simplifying the production control task, **class C** materials should be controlled by workers on the shop-floor, once again, - "they know it better". Their high usage and low value do not justify a tight control. Workers on the shop-floor should be able to take decisions on this matter, they have a more realistic view on the needs and availability of this type of materials. A bad decision hardly would imply serious financial problems.

As for materials from **class A**, the picture is a little different. From the financial point of view, they can cause serious troubles. They have a very high value, and a bad management could lead to cash-flow problems. Their usage is low, is special for some orders, and only at the production planning and control (PPC) level exists the information on the needs of this class A of materials.

Class B material lie in between those other two classes. Perhaps local control would be better since they have high usage.

7.1.2 Layout

It was proved through the analysis of time-based competitors, who are flexible manufacturers, that a layout organized by product, even with high product variety, yields better results that a standard job-shop layout and simplifies the scheduling task.

7.1.3 Batch Sizes

OPT Philosophy, lean production and time-based competition point in the same direction, leaving no margin for possible discussion. Reduced batch sizes is the best solution for reducing WIP, streamlining the workflow, increase productivity, etc.

7.1.4 A Buffer-barrier View of Manufacturing Organizations

To accommodate uncertainty, companies are forced to use some sort of buffer barrier - inventory or capacity. Rather than size, location is a more important decision. A framework for its location is here discussed as well as its implications on the managerial needs up- and downstream of that barrier [24].

The nature of the product to be produced and its effects on the organization processes are determinant to the appropriate managerial focus. Consider that most manufacturing firms fall into one of four categories:

1. **Make-to-stock** (MTS). Products have a standard design and large enough volume to keep a finished goods inventory which customers buy directly "off the shelf". Examples are basic tools (hammers, screw drivers, etc.), basic customer products sold in retail establishments, and many commodity raw materials.
2. **Assemble-to-order** (ATO). Products have standard major options, but the existence and mix of these options on the final product are determined by specific customer orders. In general, the final configuration is determined only at the final assembly stage. Most domestic automobiles fall into this category. Other operations have standard products which may be finished or even packed to customer specifications. This is fairly common in some commodity goods such as chemicals.
3. **Make-to-order** (MTO). Products tend to use standard components, but the final configuration of those components is customer-specific. It is likely, for instance, that an MTO product has never been made in a particular final form before. An example is Bailey Engineered Structures, which builds an endless variety of customized walk-in-coolers or refrigerators from a standard set of panels.
4. **Engineer-to-order** (ETO). Products are specially designed from engineering specifications. While the products might use some standard components, at least some of the components or arrangements of components have been specifically designed by the customer or the customer working with the producer. A classic example of this type of operation is a tool and die shop.

As one moves on the continuum from MTS to an ETO environment, the customer has increasing influence and involvement in the final design and production of the product. As a result, the impact on the production process will shift considerably. **Unless or until the cumulative throughput time in the delivery chain is less than the required delivery time to the customer, the company will have to use some sort of buffer barrier at some point in the process to accommodate uncertainty in the customer demand**. Such a barrier can be extremely important in absorbing demand fluctuations. In manufacturing environments this buffer can be capacity (including flexible workers and processes), inventory or a combination of the two. Inventory theory has recognized this need for many years, with most of the safety stock calculations for both independent and dependent demand being based on the classic trade-off between inventory holding costs and the level of customer service desired by the process manager.

The critical issue is not the size of this safety stock or **buffer barrier**. Rather, **it is the location** of this barrier and the managerial objectives that emerge on either side of it. Figure 7.1.2 illustrates the location of the buffer barrier in MTS, ATO, MTO and ETO environments.

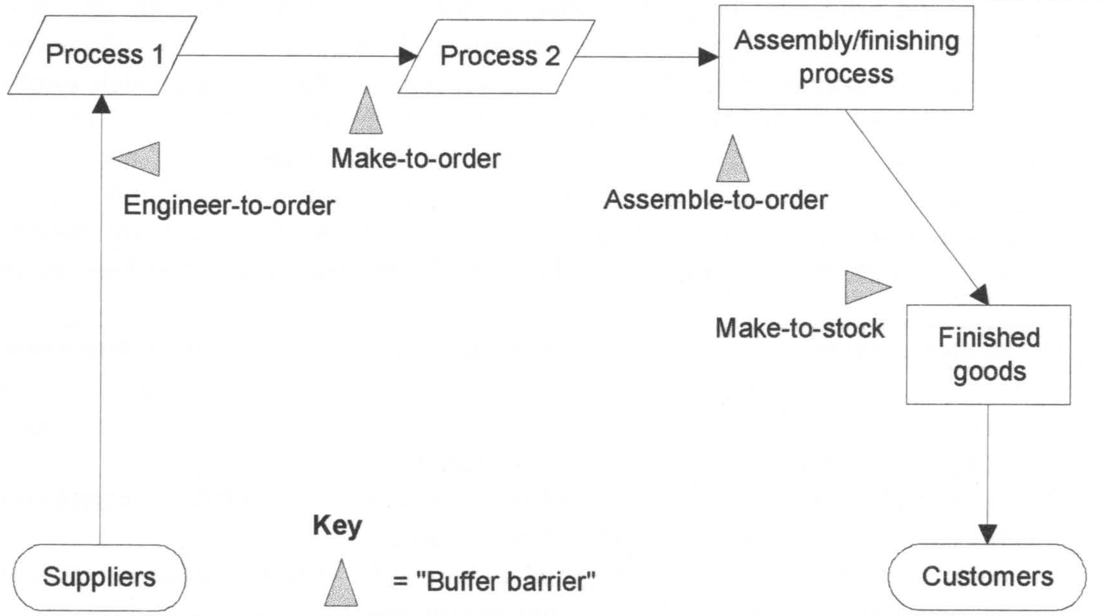


Figure 7.1.2 - Buffer Barriers in MTS, ATO, MTO and ETO Environments.[24]

Make-to-stock

In most cases, the buffer barrier appears in the form of finished goods inventory, as the typical customer for an MTS product generally aspects rapid delivery. The manufacturing task in such environment is fairly straightforward, in that the primary goal is to replenish the stock according to some established rules.

Assemble-to-order

The buffer barrier in an ATO generally occurs at the final assembly (or finishing/packaging) area. It will usually occur as a combination of assembly or finishing capacity and some inventory on the option modules and components.

Make-to-order

The buffer in MTO will usually occur at some intermediate processing area, again consisting of some capacity and some inventory of standard components. Exactly how far upstream in the production process the buffer appears depends on the extent of product uniqueness from one customer to the next, and on the stage of production where this uniqueness becomes evident. The general rule for most MTO organizations is to place the buffer barrier as far downstream (late) in the production process as possible to maximize the number of processes protected against normal variation in demand.

Engineer-to-order
In this type of environment, the barrier is forced even further upstream as compared to MTO, often occurring at the design stage, and in some cases even within the supplier base.

The placement of the buffer barrier as far downstream in the process as possible will occasionally make traditional managers and cost accountants a bit uneasy, as the maintenance of higher value-added inventory is typically discouraged for two reasons. Not only does it tend to be a more expensive form of inventory, but since more processing steps are completed, it also tends to be less flexible. The buffer barrier allows stability and therefore efficiency and high productivity in virtually every process prior to the barrier. As a result, many firms successfully utilizing time-based tactics find that, while some isolated costs do rise, the overall system costs will significantly decline. With regard to flexibility, it must be remembered that the buffer is strategically located given the nature of the customer demand. This means that flexibility is maintained, but only at the level which it makes sense for the production environment.

Managerial focus	Upstream from barrier	Downstream from barrier
Demand fluctuation management	Generally smooth demand, as most fluctuations absorbed	Fluctuations accommodated with flexibility and capacity
Processing flexibility	Needs not as high, focused on barrier replenishment	Much flexibility required dependent on customer
Information systems focus	Focused on barrier needs and processing requirements	Focused on customer orders and timing
Key management task	Barrier maintenance and processing efficiencies	Customer orders and processing effectiveness

Table 7.1.3 - Managerial Focus Differences between Up- and Downstream Activities.[24]

A major implication of this view is to recognize the differences in managerial focus which occur on each side of the buffer barrier. These differences are significant, and affect virtually every aspect of management for these operations. On the downstream (customer) side of the barrier, the critical operational task is to respond to the specific customer order. Here, flexibility and responsiveness are paramount. Information systems should be linked to the customer order, as should all production planning and control activities. Small capacity buffers and maximum flexibility should exist in virtually all downstream processes to absorb the fluctuations in customer demand. This implies higher cost, and these costs grow as the barrier moves further upstream. This is typically not a major concern, however, since as products move from MTS to ETO, price usually declines in importance relative to capability, delivery and quality issues - those very attributes aided by the customer-focused flexibility downstream from the buffer barrier.

Upstream from the barrier represents significantly different set of managerial tasks. Here, efficiency and reliability become more important. The information systems and production planning and control systems are not focused on specific customer orders, but on replenishment of the buffer barrier that is being used to absorb the variations in demand. The replenishment focus usually allows for more effective scheduling, meaning that greater efficiencies, higher utilization's and fewer disruptions are possible. Many of these differences can be seen easily, as shown in Table 7.1.3.

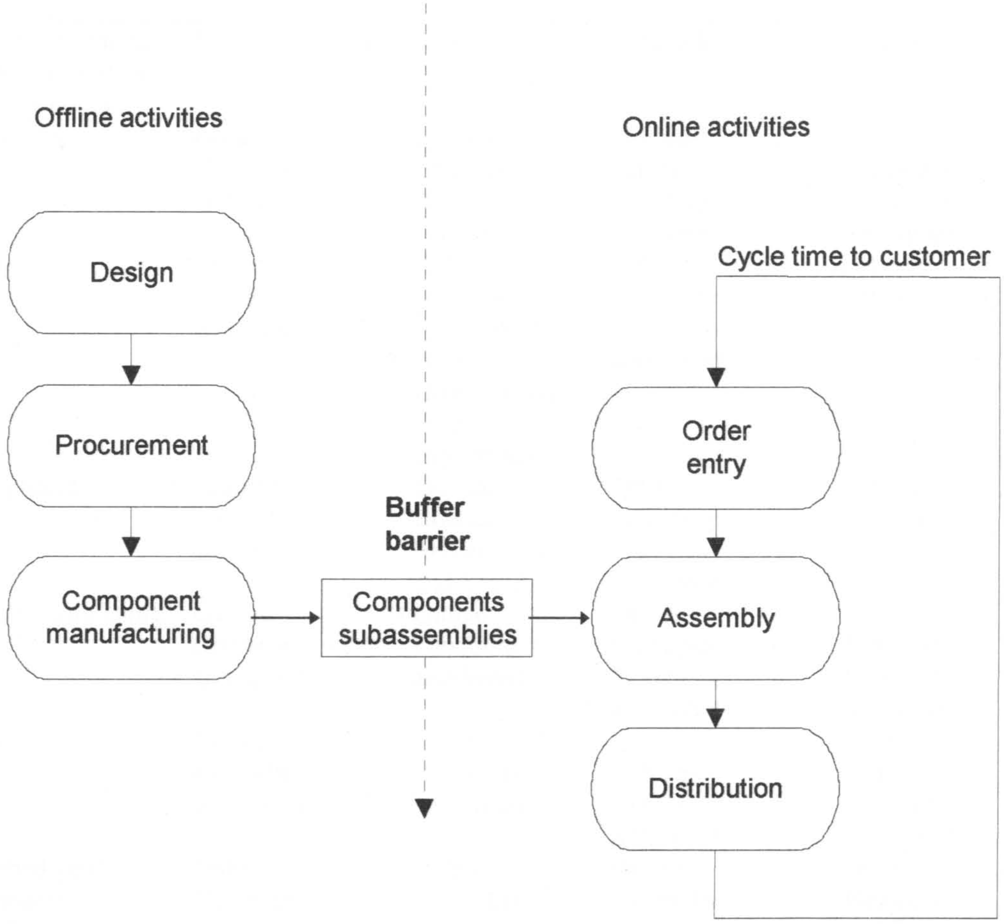


Figure 7.1.4 - Break between Down- and Upstream Processes for ATO manufacturer.[24]

The break between downstream and upstream processes by the buffer-barrier can be seen by considering the value-added delivery system for a typical ATO manufacturer - Figure 7.1.4. The buffer barrier occurs at the final assembly/finishing stage. All processes which feed this stage can be completed prior to the arrival of a specific customer order. These processes should be moved "off-line" so they do not add to the customers' cycle time. Flexibility is nor critical since these processes do not have to respond in real time to individual customer orders. As a result, we should focus on tactics that allow us to improve the efficiency and reliability, as well as cycle times, for these processes.

In contrast, the downstream processes are "online" - they can only be started once a customer order arrives. In order to shrink cycle time to the customer, these processes must be flexible enough to respond quickly to changes in demand level or customer requirements. The desired performance results demand a different set of tactics. The exception to this rule is the order entry process, which is, by definition, always online. The relative flexibility needed here will depend on the nature of the product. As one moves from ETO to an MTS environment, flexibility becomes less of an issue and the order entry process can be streamlined.

Objectives of time-based competition implementation	Engineer-to-order	Make-to-order	Assemble-to-order	Make-to-stock
Procurement	Online Flexibility and Speed	Off-line Efficiency Reliability and Speed	Off-line Efficiency Reliability and Speed	Off-line Efficiency Reliability and Speed
Design	Online Flexibility and Speed	Off-line for components; online for final product Differs based on above requirements	Off-line Efficiency Reliability and Speed	Off-line Efficiency Reliability and Speed
Component manufacturing	Online Flexibility and Speed	Off-line Efficiency Reliability and Speed	Off-line Efficiency Reliability and Speed	Off-line Efficiency Reliability and Speed
Assembly/finishing	Online Flexibility and Speed	Online Flexibility and Speed	Online Efficiency Reliability and Speed	Online Efficiency Reliability and Speed
Order entry	Online Flexibility and Speed	Online Flexibility and Speed	Online Efficiency Reliability and Speed	Online Efficiency Reliability and Speed
Finished good inventory/ distribution	Online Flexibility and Speed	Online Flexibility and Speed	Online Flexibility and Speed	Online Flexibility and Speed

Table 7.1.5 - Differences between Manufacturing Environments.[24]

Table 7.1.5 summarizes the important differences between ETO, MTO, ATO and MTS manufacturers across six major processes, and identifies the objectives associated with each step in the delivery system. Consider the requirements associated with component manufacturing in the ETO versus MTS environment. In the MTS, component manufacturing is an off-line, upstream activity. Efforts should: shrink the setup times, standardize tasks, control variances in time requirements and use smaller batch sizes to speed up the flow of items throughout the manufacturing process, thereby reducing WIP. In short, use time-based tactics to improve the efficiency within component manufacturing. In the ETO, component manufacturing is an online, downstream process. The more appropriate perspective is one which focuses on improving the responsiveness of manufacturing to specific customer requirements. Highly flexible processes and workers, and excess resources are more appropriate tactics here.

Flexibility

There has been a large discussion on flexibility measures. Questions like - how much flexibility do we need - and - how much flexibility can we afford - have been the drivers of the discussion. The term flexibility used in this work has no emphasis on a particular set of measures, it is more a global definition. Flexibility just happens to be a convenient word which helps describe the fact that manufacturing facilities must be capable of dealing with change and uncertainty [29].

7.1.5 Time Fences for Flexibility

The objective of time fences is to compromise the required planning flexibility with a certain degree of reliability. If we have too much freedom, we will risk constant expediting, because recent orders with higher priority may be released to the shop-floor. If we have a too rigid, long time fence, we will not be able to cope with uncertainty, i.e., some demanding customer may have to wait too much for the delivery. It is known that demanding customers don't like to wait and, before this situation, they would look in the competition for better service.

It is up to the designer to set the fence to a month or a week (or any other time interval). The first choice will decrease the flexibility to react to changes in this period. A time fence of weeks leads to a more nervousness in the last three weeks when compared to a time fence of a month.

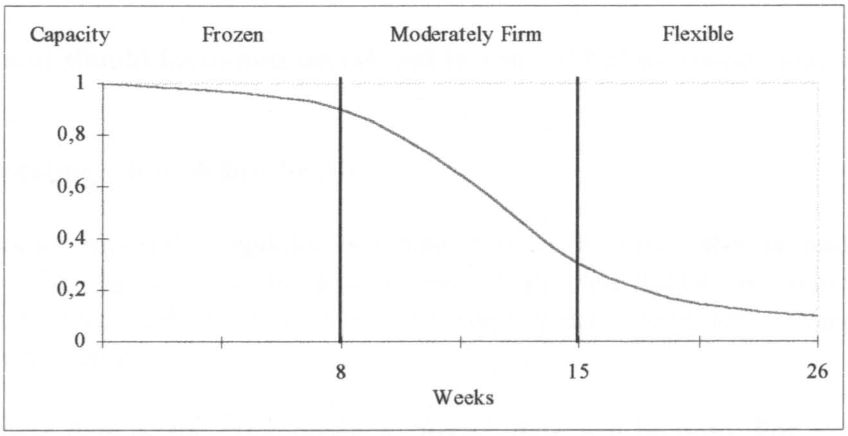


Figure 7.1.6 - MPS Time Fences.[18]

The question of flexibility within a master production schedule (MPS) depends on several factors: production lead time, commitment of parts and components to a specific item, relationship between the customer and vendor, amount of excess capacity, and the reluctance or willingness of management to make changes.

Figure 7.1.6 shows an example of a MPS time fence. Management defines time fences as periods of time having some specified level of opportunity for the customer to make changes. Note in the Figure that for the next eight weeks this particular master schedule is frozen. Each firm has its own time fences and operating rules. Under these rules, frozen could be defined as anything from absolutely no changes in one firm to only the most minor of changes in another. Moderately firm may allow changes in specific products within a product group, so has long as parts are available. Flexible may allow almost any variations in products, with the provision that capacity remains about the same and there are no long lead time involved.

The purpose of time fences is to maintain a reasonably controlled flow through the production system. Unless some operating rules are established and adhered to, the system could be chaotic and filled with overdue orders and constant expediting.

7.1.6 Process Stability

Essential to feasibility of schedules is process stability. If we don't have a stable process, i.e., if a batch one week takes 2 days to be processed and the other week it takes 4, for whatever reason that happens, there's no possible feasibility in any schedule.

As discussed in section 7.1.4, reliability is one major issue in downstream activities. If there is reliability processes that feed bottlenecks, bottlenecks can be planned in a precise way without risking not being fed by previous activities. So, throughput will not be affected in a negative way.

Total Quality Management and Total Productive Maintenance/Zero Defects will play here a major role.

Management should focus and invest, not in control but in process stability.

7.1.7 Scheduling - Rough but Feasible

Bottlenecks are precisely scheduled. For that, it must be certain that components will be ready on time for when the bottleneck needs it. Process stability is a mean of doing that. Usually, bottleneck will have a buffer inventory that allows for continuous work (stops only for setups).

As for release date of the same order, a "fuzzy" date will be used. For example, we know that it takes "about 3 days" for a batch to reach the bottleneck. If we know that it will be processed at the bottleneck today at 3pm, would it have to be released exactly 3 days ago? The answer is no. We will work with a time-window for order release. The moment when it actually starts being processed depends on shop-floor control decisions by workers. They know it must be started within that same time window and must be at the bottleneck until a certain time, according to the central schedule.

Of course, the schedule will be dependent on material availability (especially class A materials which have lower usage and, therefore, should be bought only when needed, creating the possibility of delays in its delivery). Working JIT with suppliers is essential to success of this type of control.

7.1.8 Enterprise Objectives Assessment

The definition of objectives which a company intends to fulfill is always a "fuzzy" question. The importance of all factors is not well defined. For the production manager, flexibility is a good objective, as for cost accounting cost reduction is the best objective.

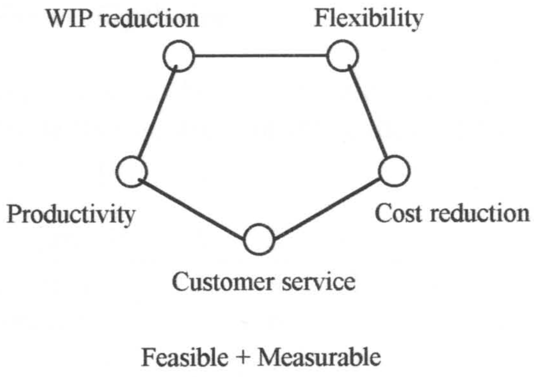


Figure 7.1.7 - Example of some Goal. [18]

As discussed earlier, time-based performance measures are, today, the most important to achieve good results in. With good performance in time-based factors, such as delivery speed and delivery reliability (customer service) it is known that WIP will be affected in a positive way, for instance.

New production control systems should take this into account and find a way to better assess the future situation of the company in terms of capacity required, materials, and financial parameters.

An estimate points to data needs reduction to only 5% of the actual needs in a standard software package for production control.

8. Fuzzy Logic in Production Control

Fuzzy Logic was presented has a technology able to cope with "soft data" and uncertain factors. "Soft data" and uncertain factors exist in a manufacturing environment, so fuzzy logic may contribute with something new to the area of production control.

8.1 Compromise vs. Optimize

With a simple example it will be demonstrated the potential of fuzzy logic in priorities assessment. Suppose that we are controlling the bottleneck resource in a company. The demand placed upon it is:

Order (in order of arrival)	A	B	C	D	E
Processing time (days)	3	4	2	6	1
Due date (days hence)	5	6	7	9	2

Total time needed for production = 16 days
Latest due date = 9 days

Clearly we are in the presence of a bottleneck. Usually this situation is solved by academics, by trying to optimize various performance parameters like: minimize number of tardy jobs, minimize maximum tardiness, minimize total tardiness, etc.

8.1.1 Optimize

It is known that, for one machine, the sequencing rule shortest processing time (SPT) minimizes the total tardiness. Other techniques can be seen in Gupta and Ramnaryanan [30], again trying to optimize. For the SPT rule, orders would be processed in the sequence:

Order	E	C	A	B	D
Start	0	1	3	6	10
Processing time	1	2	3	4	6
Finish	1	3	6	10	16
Due date	2	7	5	6	9
Lateness	0	0	1	4	7

As we saw in priority management, this is not the normal procedure in most companies. What about customer importance? Suppose that the customer that placed order B is a JIT client, thus an "important" customer. It is clear that we cannot deliver to a JIT customer with 4 days of delay. Now, how can we model this situation? Fuzzy logic is able to cope with fuzzy factors such as "customer importance", providing this way a more realistic measure of customer service.

The lateness itself is fuzzy. Is a two days delay a big delay or a small delay? Is it better to have no delay for a JIT customer and 10 days for a "not very important customer" or just three days of delay for each one?

8.1.2 Compromise

Trying to solve this problem with fuzzy logic, we can start by identifying three linguistic variables:

- customer importance
- lateness
- order score

Customer Importance

For example, we can create three terms for this linguistic variable. Their membership functions are displayed in Figure 8.1.2. The terms are:

- high
- normal
- small

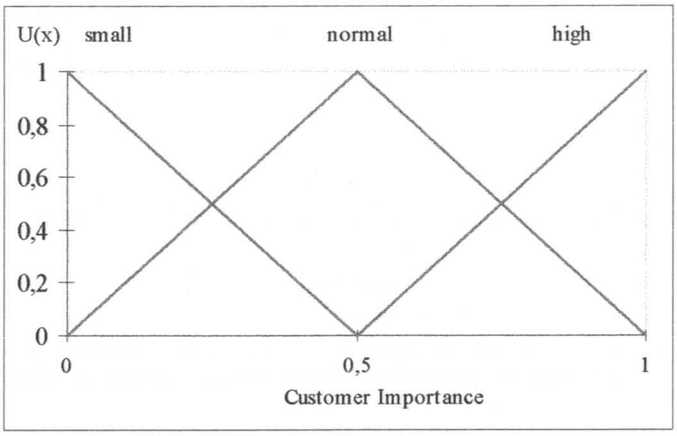


Figure 8.1.2 - Linguistic Terms of the Linguistic Variable "Customer Importance"

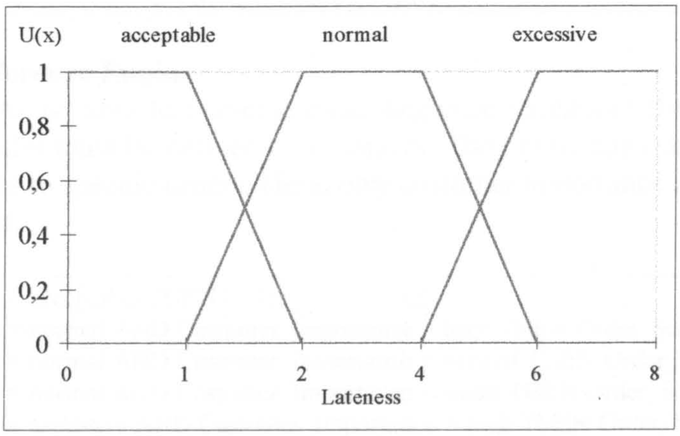


Figure 8.1.3 - Membership Functions of the Terms of "Lateness"

Lateness

As for the linguistic variable "lateness", Figure 8.1.3 shows its terms and membership functions (these are possible definitions, they would be different from company to company). The terms are:

- acceptable
- normal
- excessive

Order Score

"Customer Importance" and "lateness" will be combined in way that enables the correct evaluation of customer service. For that, the linguistic variable "order score" was created. Figure 8.1.4 presents its membership functions. Its terms are:

- high
- medium
- low

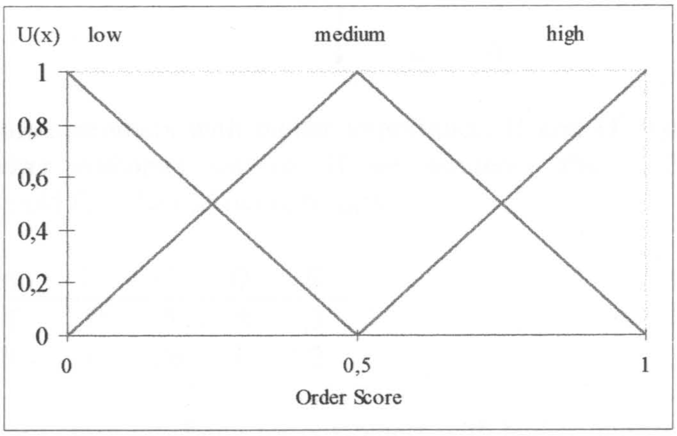


Figure 8.1.4 - "Order Score's" Terms and Membership Functions

Linguistic Variable	Possible Terms
1. Customer Importance	∈ {small, normal, high}
2. Lateness	∈ {acceptable, normal, excessive}
3. Order Score	∈ {low, medium, high}

Fuzzy Logic Inference Engine

The question now is: how to combine these linguistic variables? If-then rules is the answer. These rules must be derived by managers. They must copy the way managers attribute priorities to specific orders. Here, only customer importance is included, other reasons may exist.

Rule 1: IF Lateness = acceptable THEN Order_Score = high
Rule 2: IF Lateness = normal AND Customer_Importance = high THEN Order_Score = low
Rule 3: IF Lateness = normal AND Customer_Importance = normal THEN Order_Score = medium
Rule 4: IF Lateness = normal AND Customer_Importance = small THEN Order_Score = high
Rule 5: IF Lateness = excessive AND Customer_Importance = high THEN Order_Score = low
Rule 6: IF Lateness = excessive AND Customer_Importance = normal THEN Order_Score = low
Rule 7: IF Lateness = excessive AND Customer_Importance = small THEN Order_Score = medium

Now, lets evaluate Customers importance. To each order a classification between 0 and 1 will be given.

Order (in order of arrival)	A	B	C	D	E
Processing time (days)	3	4	2	6	1
Due date (days hence)	5	6	7	9	2
Importance [0; 1]	.5	1	.7	.8	.3

Suppose also that the sequence is the same as the one found with the SPT rule. What happens to customer service?

Order	E	C	A	B	D
Importance [0; 1]	**.3**	**.7**	**.5**	**1**	**.8**
Start	0	1	3	6	10
Processing time	1	2	3	4	6
Finish	1	3	6	10	16
Due date	2	7	5	6	9
Lateness	0	0	1	4	7
Score	**1**	**1**	**1**	**0**	**0**

As we can see, the customers with higher importance, B and D, were the ones who achieved the lower customer service. If we sequence the orders by customer importance we would find the following results:

Order	A	B	C	D	E
Importance	.5	1	.7	.8	.3
Score	0	1	.26	1	.2

Now, the results are more satisfying for customers with higher importance. Those who have less importance were left with worse performances. This doesn't mean that this is the better option, other parameters are missing for a more complete evaluation, it just tries to demonstrate how common rules can lead to deceiving results when they don't take into account the priorities that usually managers consider.

Obviously, evaluating the score relatively to the customer importance requires another fuzzy inference. How important is to achieve, for a customer with importance .3, the score 1? How bad is to achieve, for a customer with importance 1, the score 0? Also, a global measure could be derived by weighting individual orders importance and scores, all together.

I think that the usefulness of fuzzy logic was demonstrated. Each case is a different one. Each company has to design the system and find other places where it can apply fuzzy logic. The knowledge of the operation rules within each company is essential to the successful design of fuzzy logic applications. Where human evaluation exists, a possible application also exists.

9. Conclusions

In the past, production planning and control systems have tried to control precisely what is not precise. This state of affairs led to an increase in control complexity. Many examples were provided to demonstrate this. Concepts like MRP require to many data because they try to control too many resources.

Today, many managers with large experience in the sector state that production control systems simplification is a must. Leading this revolution are time-based competitors that are shifting their focus from control to providing resources that enable better global performance.

Goldratt's "general theory of constraints" is gaining every day more supporters. He advocates that, for good manufacturing system performance, controlling the constraints in the system is enough. Trying to control precisely everything else just wastes time and money.

Priority management is built on the knowledge of inadequacies of formal systems to companies real practices. Most systems are designed in a way that optimization is one of the main objectives. In reality, managers don't try to optimize, they must compromise. Simply, there is too many instability and customer pressures.

The basic concepts described for new production control will have some problems in finding a widespread acceptation. They touch too many "axioms" that managers of traditional companies have accepted without questioning for long time. For instance, it will be almost impossible to convince a traditional manager that he does not need a system like SAP and, instead, he must focus its efforts in process stability, today much more important.

Fuzzy logic is taking its first steps in decision support systems design for production control. The example here presented is just the "tip of the iceberg". As a technology able to cope with "soft data" and uncertainty, characteristics of batch manufacturing environment, fuzzy logic certainly will have a strong contribution for production control systems simplification, user friendliness and performance improvement.

In the future, production control with "soft data" will become a reality to which every company will have to adapt to stay competitive in a global market.

10. Bibliography

[1] Chase, R.B., and Aquilano, N.J., "Production & Operations Management: A Life Cycle Approach", Sixth Edition, Irwin, Homewood, Boston, ISBN-0-256-10039-X.

[2] De Maio, A., Bartezzaghi, E. and Zanarini, G., "A New System Analysis Method Based on the STS (Socio-technical System) Approach", in Schneider, H.J. (Ed.), *Formal Models and Practical Tools for IS Design*, Proceedings of the WG8.1 Working Conferences, Oxford, 17-20 April 1979, pp. 167-93.

[3] Kochhar, A., "Why Computerized Production Control Systems Continue to Disappoint", *The Production Engineer*, December 1982, pp. 37-8.

[4] Safizadeh, M.H. and Raafat, F., "Formal/informal Systems and MRP Implementation", *Production and Inventory Management*, First Quarter, 1986, pp. 115-20.

[5] Plossl, G.W., "MRP Yesterday, Today, and Tomorrow", *Production and Inventory Management,* Third Quarter, 1986, pp.1-10.

[6] Etienne, E.C., "MRP May Not Be Right for You: At Least Not Yet", *Production and Inventory Management,* Third Quarter, 1983, pp. 33-46.

[7] Fischer, K.R., "How to Implement MRP - Successfully", *Production and Inventory Management*, Fourth Quarter, 1981, pp. 36-54.

[8] Wacker, J., and Hills, F.S., "The Key to Success or Failure of MRP: Overcoming Human Resistance", *Production and Inventory Management*, Fourth Quarter, 1977, pp. 7-16.

[9] Latham, D.A, "Are You among MRP's Walking Wounded?", *Production and Inventory Management*, Third Quarter, 1981, pp. 33-41.

[10] Lee, R.S., "Material Requirements Planning and the Organization", *Proceedings of the British Production and Inventory Control Society Annual Conference*, 1981.

[11] Wight, O., "Why People Come First in Computer based Systems", *MRPII: Manufacturing Resource Planning*, "Modern Materials Handling", 6 May 1981.

[12] Wharton, T.J., and Reid, R.D., "Manufacturing Planning and Control: How well Are We Doing?", *Production and Inventory Management*, Third Quarter, 1990, pp.51-55.

[13] Wisner, J.D., "A study if US machine shops with just-in-time customers", *International Journal of Operations Production and Management*, Vol.16, No.7, 1996, pp.62-76.

[14] Roth, A., Giffi, C., Shinsato, D., and Fradette, M., "Vision in manufacturing planning for the future", funded research for Deloitte Touche Tohmatsu International, 1993.

[15] Sweeney, M.T., and Szwejczewski, M., "Manufacturing strategy and performance - A study of the UK engineering industry", *International Journal of Operations Production and Management*, Vol.16, No.5, 1996, pp.25-40.

[16] Stalk, G. and Hout, T., "Competing Against Time: How Time-based Competition is Reshaping Global Markets", The Free Press, New York, NY, 1990.

[17] Goldratt, E.M., and Cox, J., "The Goal: Excellence in Manufacturing", Croton-on-Hudson, N.Y.: North River Press, 1984.

[18] Helfrich, C., "PPS-Praxis: Fertigungsteuerung und Logistik im CIM-Verbund", 1989, Resch-Media Mail Verlag GmbH, ISBN 3-87806-118-8.

[19] Caskey, K., Storch, R.L., "Heterogeneous dispatching rules in job and flow shops", *Production Planning and Control*, 1996, Vol.7, No.4, pp.351-361.

[20] Westbrook, R., "Priority Management: New Theory for Operations Management", *International Journal of Operations Production and Management*, Vol.14, No.6, 1994, pp.4-24.

[21] Zimmermann, H.-J., "Fuzzy Set Theory - and its applications", Second Revised Edition (1991), Boston, Dordrecht, London, ISBN 0-7923-9075-X.

[22] von Altrock, C., "Fuzzy Logic and Neuro Fuzzy Applications Explained", Prentice Hall (1995), ISBN 0-13-368465-2.

[23] "FuzzyTech for Business - Users Manual", Inform GmbH.

[24] Bozarth, C., and Chapman, S., "A Contingency View of Time-based Competition for Manufacturers", *International Journal of Operations Production and Management*, Vol.16, No.6, 1996, pp.56-67.

[25] Courtois, A., Pillet, M., et Martin, C., "Gestion de Production", Les Éditions d'Organisation, ISBN 2-7081-1116-7.

[26] Porter, J.K., Jarvis, P., Little, D., Laakman, J., Hannen, C., and Schotten, M., "Production Planning and Control Systems Developments in Germany", *International Journal of Operations Production and Management*, Vol.16, No.1, 1996, pp.27-39.

[27] Wortman, J.C., Euwe, M.J., Taal, M., and Wiers, V.C.S., "A review of capacity planning techniques within standard software packages", *Production Planning & Control*, 1996, Vol.7, No.2, pp.117-128.

[28] Karlsson, C., and Ahlström, P., "Assessing changes towards lean production", *International Journal of Operations Production and Management*, Vol.16, No.1, 1996, pp.27-39.

[29] Primrose, P.L., Verter, V., "Do companies need to measure their production flexibility?", *International Journal of Operations Production and Management*, Vol.16, No.6, 1996, pp.4-12.

[30] Gupta, J.N.D., and Ramnarayanan, R., "Single facility scheduling with dual criteria: minimizing maximum tardiness subject to minimum number of tardy jobs", *Production Planning & Control*, 1996, Vol.7, No.2, pp.190-196.

[31] Rolstadas, A., "Editorial: Keep it Simple", *Production Planning & Control*, 1996, Vol.7, No.1, pp.1.