Chapter 1

Introduction

Motivation: This chapter is intended to place Production Systems Engineering into the general framework of Manufacturing at large and outline the main problems addressed in this textbook.

Overview: Five main areas of Manufacturing are described and three main problems of Production Systems Engineering are characterized.

1.1 Main Areas of Manufacturing

The informal definitions and classifications given below are subjective and based solely on the authors' experience and understanding.

Manufacturing – the process of transforming raw materials into a useful product. Everything, which is done at or for the factory floor operations, we view as manufacturing.

Manufacturing matters. Indeed, the wealth of a nation can be either taken from the ground (natural resources and agriculture) or manufactured (value added by processing materials). Thus, being one of just two ways of generating national wealth, manufacturing is of fundamental importance.

Manufacturing can be classified into two groups: *continuous* and *discrete*. Examples of industries with continuous manufacturing are chemical, materials, and power, where core processes evolve continuously in time. Discrete part manufacturing is practically everything else: automotive, electronics, appliances, aerospace and other industries. This textbook is concerned with discrete part manufacturing.

Quite informally, manufacturing can be classified into the following five areas:

- *Machine tools and material handling devices.* The main problem here is: Given a desired material transformation and/or relocation, design, implement, and maintain a machine or a material handling device, which carries out its function in an efficient manner. This is a mature engineering field with numerous achievements to its credit.

- *Production systems.* Main problem: Given machines and material handling devices, structure a production system so that it operates as efficiently as the machines in isolation. This can be achieved by maintaining smooth flow of parts throughout the system, so that mutual interference of the machines does not cause losses of production. The term "structure" is used here to include both design of new and improvement of existing production systems. To-date, this field lacks in rigorous quantitative methods and fundamental engineering knowledge.
- *Production planning and scheduling.* Main problem: Given a production system and customer demand, calculate a production plan and schedule delivery of materials so that the demand is satisfied in an economically efficient manner. Numerous quantitative methods, often based on optimization, are available in this relatively mature field of manufacturing.
- *Quality assurance.* Main problem: Structure and operate the production system so that parts produced are of the desired quality. To-date, statistical quality control is a major quantitative tool for maintaining product quality.
- *Work systems.* Main problem: Organize personnel training and operation so that the production process is carried out safely and efficiently. This includes, in particular, designing wage and incentive systems so that the maximum of the utility function of an individual worker coincides with that of the manufacturing enterprise as a whole and, thus, self-interest of the worker leads to high efficiency of the manufacturing enterprise. At present, this field is still in its infancy.

In addition to the above classification, discrete part manufacturing can be subdivided into two groups: *job-shop* and *large volume* manufacturing. Job-shop is concerned with manufacturing "one-of-a-kind" products: unique instruments, highly specialized equipment, some aerospace systems, etc. Large volume manufacturing is intented to produce parts and products in multiple copies: cars, computers, refrigerators, and other items of wide use.

This textbook is devoted to one area of manufacturing - *production systems*, although some structural issues of *quality assurance* are also addressed. While some methods included here may be useful for job-shops as well, the emphasis is on production systems in *large volume manufacturing*.

1.2 Main Problems of Production Systems Engineering

1.2.1 Complicating phenomena

Examples of typical production systems are shown in Figures 1.1 and 1.2, where the circles represent the machines and rectangles are material handling devices in their function as buffers. The system in Figure 1.1 is referred to as a *serial production line* while that of Figure 1.2 as an *assembly system.*

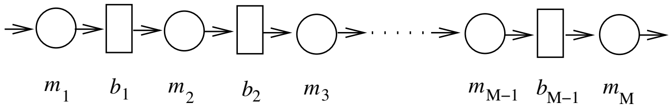

Figure 1.1: Serial production line

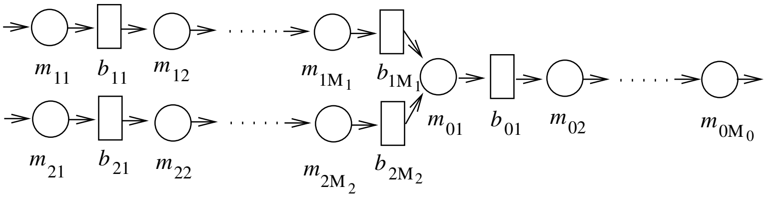

Figure 1.2: Assembly system

If machines never broke down, the situation would be simple: to ensure smooth part flow, balance the capacity of the machines and speeds of the material handling devices so that the desired throughput is achieved. (The notions of machine capacity, throughput, etc., are defined precisely in Chapter 3; at this point, an intuitive understanding of these terms is sufficient.)

In reality, however, the machines always experience random breakdowns. This leads to a complex phenomenon of *perturbation propagation.* Indeed, when, say, machine m_2 in the serial line fails, machine m_3 may eventually become *starved* for parts and machine m_1 *blocked* by full buffer b_1. If m_2 continues staying down, the starvations will propagate downstream reaching eventually the last machine, m_M, and the throughput is lost. Similarly, the blockage will propagate upstream, causing all machines to stop, wasting their production capacity. Thus, failure of one machine may affect all other machines in the system – up- and downstream. This is what makes the investigation of production systems difficult (and necessary too – in order to develop methods for alleviating this perturbation propagation).

Similar phenomena take place in assembly systems as well: failures of m_{12} (see Figure 1.2) may lead to the starvation of m_{13}, …, m_{1M_1}, m_{01}, …, m_{0M_0} and to blockage of m_{11} and m_{21}, …, m_{2M_2}. (Note that it is typically assumed that the first machines m_1 and (m_{11}, m_{21}) are never starved and the last machines m_M and m_{0M_0} are never blocked.)

To alleviate this mutual interference, material handling devices are used not only as means for moving parts but also as buffers – to protect against blockages and starvations. If the buffers are infinite, only starvations are possible, and the parts flow is improved (since the disturbances propagate only downstream). However, infinite buffers are impractical, economically detrimental, and, as it turns out, unnecessary as well.

This textbook presents rigorous engineering methods for reducing mutual machine interference by selecting buffer capacities, which are "*just right*" (rather than *just-in-time*) for guaranteeing a sufficiently smooth flow of parts and, thus, resulting in an acceptable system behavior. To accomplish this, three main problems of production systems engineering are addressed. These problems are described next.

1.2.2 Analysis, continuous improvement, and design problems

Analysis: The problem of *analysis* is addressed in this book in two formulations:

Problem A1: Given a production system (i.e., the machine and buffer characteristics), calculate its performance measures, such as its throughput, work-in-process, and the probabilities of blockages and starvations.

The second formulation is concerned with systems having a Finished Goods Buffer (FGB), as shown in Figure 1.3 for the case of a serial line; similarly, assembly systems may have FGBs. The purpose of FGBs is to filter out production and demand randomness and, thus, improve the level of customer demand satisfaction.

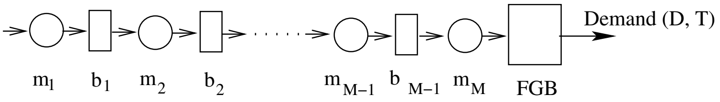

Figure 1.3: Serial production line with a finished goods buffer

Problem A2: Given a production system and customer demand specifications (i.e., the shipment size, D, and the shipping period, T), calculate the probability of demand satisfaction.

Continuous improvement: The problem of *continuous improvement* is also considered in two formulations – constrained and unconstrained improvability.

Problem CI1 (Constrained improvability): Given a production system, determine if its performance can be improved by re-allocating its limited resources (such as buffer capacity and/or workforce) and suggest an improved or even optimal allocation.

Problem CI2 (Unconstrained improvability): Given a production system, determine the machine and the buffer, which impede system performance in the strongest manner; such a machine and a buffer are referred to as a *bottleneck machine* (BN-m) and a *bottleneck buffer* (BN-b), respectively.

Design: *Design* issues are formalized as three problems:

Problem D1: Given the machine characteristics and the desired system throughput, determine the smallest (i.e., lean or "just-right") buffers capacity, which ensures this throughput.

Problem D2: Given the machine and buffer parameters, determine the locations of quality inspection devices so that no defective parts are shipped to the customer and the throughput of non-defective parts is maximized.

Problem D3: Given a production system and customer demand specification, calculate the smallest (i.e., lean or "just-right") capacity of FGB so that the demand is satisfied with the desired probability.

Solutions to all problems listed above are provided in Part II of this textbook for the case of serial lines with the Bernoulli model of machine reliability. A generalization for exponential and arbitrary machines is given in Part III. For the case of assembly systems, these problems are addressed in Part IV.

1.2.3 Fundamental laws of Production Systems Engineering

Each engineering discipline is based on a fundamental law that is the foundation for all its analysis and design methods. For instance, in Mechanical Engineering this is Newton's law:

$$F = ma.$$

Given system parameters, i.e., the force F and mass m, this law allows us to calculate the acceleration a and, subsequently, the behavior, i.e., trajectory, of a moving body. Similarly, in Electrical Engineering, Kirchoff's law,

$$E = \sum_{i=1}^{n} V_i,$$

allows us to investigate the behavior of an electric circuit, since the voltage drops V_i on each of its elements (i.e., resistors, capacitors, etc.) can be calculated, given the applied voltage E.

In Industrial Engineering, Little's law is often mentioned as a tool for production systems analysis. This law states that, in the steady state of system operation,

$$TP = \frac{WIP}{RT}, \tag{1.1}$$

where TP is the average throughput, WIP is the average work-in-process, and RT is the average residence time, i.e., the average time that the part spends in the system – from the moment it enters the first machine to the moment it leaves the last machine. Although this is an important relationship, it does not offer a direct way to investigate production systems. The reason is that all three quantities involved in (1.1) are unknown: neither TP, nor WIP, nor RT are given in advance. Given are only parameters of the machines and buffers and, unless at least two of these quantities are calculated, (1.1) does not offer any quantitative information concerning performance of a production system.

In this textbook, methods for calculating TP, WIP and other performance measures, based on machine and buffer parameters, are developed and various relationships among them are established. Some of these relationships, which can be viewed as fundamental facts of production systems, are summarized in Chapter 18.

1.2.4 Techniques used in this textbook

Often, production systems are studied using the methods of Queuing Theory. The classical model of Queuing Theory is illustrated in Figure 1.4. Here, customers arrive into a queue according to some randomly distributed inter-arrival times. There are M servers defined by their processing capacity. After reaching the head of the queue, the customer is processed by the first available server during a randomly distributed service time. Given this model, Queuing Theory develops methods for analysis of its performance characteristics: the average queue length, throughput, residence time, etc.

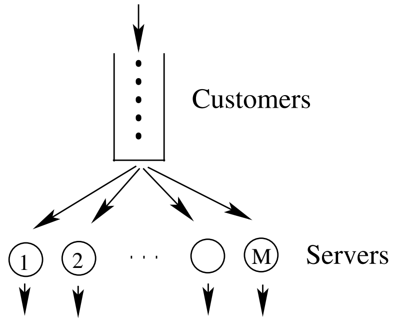

Figure 1.4: Queuing model

Unfortunately, production systems typically have a qualitatively different structure. Indeed, instead of servers operating in parallel, they have a serial connection as shown in Figures 1.1 and 1.2. In some cases, however, methods of Queuing Theory may be modified to be applicable to production systems as well.

This approach is not pursued here. Instead, methods, which are directly applicable to serial lines and assembly systems, are developed. In most cases, they are based on a three-step approach: First, the simplest systems, consisting of two machines, are investigated using the methods of Markov processes. Second, systems with more than two machines are analyzed using the results of the first step and recursive aggregation procedures; the convergence of these procedures and their accuracy are investigated. Third, these results are extended to non-Markovian cases by discrete event simulations and subsequent analytical approximations. In this framework, the recursive procedures mentioned above are viewed as governing equations of production systems. Their analyses, described throughout this book, lead to formulation of the fundamental facts of Production Systems Engineering.

1.3 Summary

- This textbook addresses one of the five main areas of Manufacturing described in Section 1.1: Production Systems with unreliable machines and finite buffers.
- The fundamental difficulty in studying these systems is due to the phenomena of blockages and starvations caused primarily by machine breakdowns.
- The topics covered in this textbook are
  - analysis,
  - continuous improvement, and
  - design

  of production systems in large volume manufacturing environment.
- A system-theoretic approach is used throughout this volume.

1.4 Problems

In all problems listed below, use your common sense to provide and justify your answers. Later in this textbook, you will learn rigorous techniques to solve these and many other problems of the same nature.

Problem 1.1 Consider a serial production line with ten identical machines and no buffers. Where do you think a single buffer should be placed so that the throughput of the system is maximized?

Problem 1.2 Consider a serial production line with eleven identical machines and ten identical buffers. Assume that one of the machines can be replaced by a more efficient one. Which machine should be replaced so that the throughput is maximized?

Problem 1.3 Assume that two types of machines are available: those with short average up- and downtime and those with long average up- and downtime but with the same efficiency

$$e = \frac{T_{up}}{T_{up} + T_{down}} = \frac{1}{1 + T_{down}/T_{up}}, \tag{1.2}$$

where T_{up} and T_{down} are the average up and downtime, respectively. Which types of machines will you buy for your production line so that the throughput is optimized?

Problem 1.4 If the efficiency of a machine in a production line can be improved, is it better to increase its uptime by a factor of α or decrease its downtime by the same factor, so that the system throughput is maximized?

Problem 1.5 In a five-machine production line with identical machines and identical buffers, which machine is the bottleneck?

Problem 1.6 In a production line where parts are transported on carriers, will the throughput necessarily increase if the number of carriers is increased?

Problem 1.7 Is the machine with the smallest throughput in isolation necessarily the bottleneck machine of a production line? Is the smallest buffer necessarily the bottleneck buffer?

Problem 1.8 Is an in-process buffer of capacity $N = 1000$ lean or not? Is a finished goods buffer of capacity $N = 1000$ lean or not?

Problem 1.9 If the only machine that produces defective parts in a serial line is the first one, where should a single quality inspection machine be placed so that the throughput of good parts is maximized?

Problem 1.10 Let the term "transients" describe the process of reaching steady state values of either throughput (TP) or work-in-process (WIP), or buffer occupancy. Which transients are faster, those of TP, WIP, or buffer occupancy?

Problem 1.11 Will the transients of buffer occupancy become faster or slower when machine efficiency is increased?

Problem 1.12 What should be the smallest buffer occupancies at the beginning of the shift so that no production losses due to transients take place?

1.5 Annotated Bibliography

There are literally millions of publications in all five areas of Manufacturing. In fact, a quick Google Scholar search returns over five million entries. However, when we started writing this book (2003), the title "Production Systems Engineering" did not appear even once! The references given below are the closest to the subject matter of this textbook.

Perhaps, the first publications on quantitative analysis of production lines appeared in Russia:

[**1.1**] A.P. Vladzievskii, "The Theory of Internal Stocks and their Influence on the Output of Automatic Lines," *Stanki i Instrumenty*, vol. 21, no. 12, pp. 4-7, 1950 and vol. 22, no. 1, pp. 16-17, 1951 (in Russian).

[**1.2**] A.P. Vladzievskii, "The Probability Law of Operation of Automatic Lines and Internal Storage in Them," *Automatika i Telemekhanika*, vol. 13, pp. 227-281, 1952 (in Russian).

[**1.3**] B.A. Sevast'yanov, "Influence of Storage Bin Capacity on the Average Standstill Time of a Production Line," *Theory of Probability Applications*, pp. 429-438, 1962.

It should be pointed out that in 1957, A.N. Kolmogorov, the founder of Probability Theory, gave a lecture at a meeting of the Moscow Mathematical Society devoted to production systems. Unfortunately, no record of this presentation could be found. (Since Sevastianov was a student of Kolmogorov, it is reasonable to assume that Kolmogorov's lecture contained ideas close to those of [1.3].)

One of the first papers on quantitative analysis of production systems published outside Russia is

[**1.4**] J.A. Buzacott, "Automatic Transfer Lines with Buffer Stocks," *International Journal of Production Research*, vol. 5, pp. 183-200, 1967.

The decomposition approach to production systems analysis appeared in

[**1.5**] S.B. Gershwin, "An Efficient Decomposition Method for the Approximate Evaluation of Tandem Queues with Finite Storage Space and Blocking," *Operations Research*, vol. 35, pp. 291-305, 1987.

Among the textbooks and monographs devoted largely, but not exclusively, to Production Systems Engineering, the following are well known:

[**1.6**] J.A. Buzacott and J.G. Shantikumar, *Stochastic Models of Manufacturing Systems*, Prentice Hall, Englewood Cliffs, NJ, 1993.

This is an encyclopedic book covering many topics of manufacturing, including production systems, scheduling, and others. Mostly the queueing theory approach is used.

[**1.7**] H.T. Papadopoulos, C. Heavey and J. Browne, *Queueing Theory in Manufacturing Systems Analysis and Design*, Chapman & Hall, London, 1993.

One of the first monographs on the queueing theory approach to production systems. It includes results on serial lines, assembly systems, and flexible manufacturing operations.

[**1.8**] S.B. Gershwin, *Manufacturing Systems Engineering*, Prentice Hall, Englewood Cliffs, NJ, 1994.

A pioneering book as far as the system-theoretic approach to manufacturing is concerned. It addresses the issues of production systems performance analysis and scheduling. The performance analysis is based on a decomposition technique and scheduling uses methods of optimal control.

[**1.9**] D.D. Yao (ed.), *Stochastic Modeling and Analysis of Manufacturing Systems*, Springer-Verlag, Series in Operations Research, New York, 1994.

[1.10] H.G. Perros, *Queueing Networks with Blocking*, Oxford University Press, Oxford, 1994.

[1.11] T. Altiok, *Performance Analysis of Manufacturing Systems*, Springer, New York, 1997.

Also is based on queuing theory. A characteristic feature is the study of non-exponential reliability models, e.g., co-axial distributions of machine up- and downtimes.

Additional results on production systems engineering can be found in

[1.12] N. Viswanadham and Y. Narahari, *Performance Modeling of Automated Manufacturing Systems*, Prentice Hall, Englewood Cliffs, NJ, 1992.

[1.13] R.G. Askin and C.R. Standridge, *Modeling and Analysis of Manufacturing Systems*, Wiley, 1993.

[1.14] W.J. Hopp and M.L. Spearman, *Factory Physics*, Irwin/McGraw-Hill, Boston, 1996.

[1.15] C.L. Curry, *Manufacturing Systems Modeling and Analysis*, Duxbury Press, to appear.

and in the proceedings of the following conferences:

[1.16] *Proceedings of the International Workshop on Performance Evaluation and Optimization of Production Lines*, Samos Island, Greece, 1997.

[1.17] *Proceedings of the Second Aegean International Conference on Analysis and Modeling of Manufacturing Systems*, Tinos Island, Greece, 1999.

[1.18] *Proceedings of the Third Aegean International Conference on Design and Analysis of Manufacturing Systems*, Tinos Island, Greece, 2001.

[1.19] *Proceedings of the Fourth Aegean International Conference on Analysis of Manufacturing Systems*, Samos Island, Greece, 2003.

[1.20] *Proceedings of the Fifth International Conference on Analysis of Manufacturing Systems - Production Management*, Zakynthos Island, Greece, 2005.

[1.21] *Proceedings of the Sixth International Conference on Analysis of Manufacturing Systems*, Lunteren, The Netherlands, 2007.

The aphorism, "*Manufacturing matters*," cited in Section 1.1, was advanced by Gary Cowger of the General Motors Corporation. He stated it during a seminar at the University of Michigan in 1988.