CHAPTER 3

ASSEMBLY LINE: HISTORY AND FORMULATION

Keywords: history of manufacturing and assembly line, line balancing, line design approaches, line layout.

Assembly work has a very long history. Ancient people already knew how to create useful objects composed of multiple parts. However, the objective of modern assembly processes is to produce high quality and low cost products. Manufacturing and assembly systems evolved from single hunter-gatherer to present day architectures. This chapter starts with the history of the evolution of manufacturing. Next, the assembly line balancing problem is introduced. Thereafter, a sort of classification of the ALD problems is given. The question why is the balancing problem hard to solve is discussed, we finally draw some conclusions in section 7.

1. Evolution of today's manufacturing issues

As long as man has been a reasoning being, he has been a manufacturer. One of the things that distinguishes man from the higher primates is his ability to fashion tools

.which make it easier for him to find and process food supplies and make the world a better place to live in is. The evolution of manufacture has passed through distinct stages at various periods of history in different parts of the world (Kirton, 1994) (Shimokawa, 1997). This section examines some stages of the evolutionary cycle of manufacturing and assembly (see Figure 3.1). It shows how the important manufacturing issues which face us today have developed, and looks ahead to what the future might bring us.

1.1. First metals

From the earliest days, organised manufacturing started to appear. Hunter-gatherers originally formed a simple society that exploited the resources found around them, but they did not attempt to process them. Manufacturing as a specialisation developed first to provide simple hand-tools to hunter-gatherers. Development of manufacture in the classical world was also based on a local nobility, demanding goods and services and having enough surplus food to support the workers who produced them. With the establishment of a nobility, a two-class society developed. There was no middle class, only nobles and peasants. The demand increased for the manufacture of luxury items, jewellery, weapons, chariots, all the accoutrements which support the life-style of nobility. Religion was also a powerful force and the production of religious artefacts played significant role in the development of manufacturing.

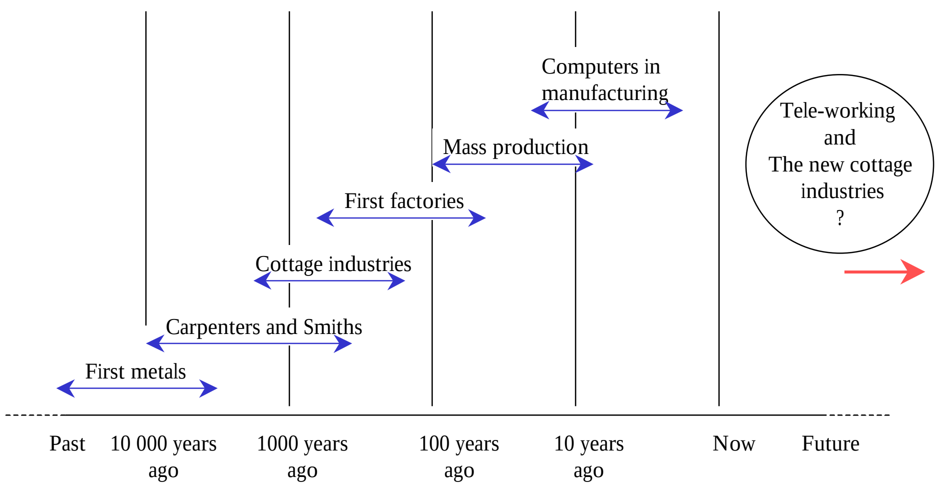

Figure 3.1. Evolution of manufacturing (Kirton, 1994).

1.2. Carpenters and Smiths

As people concentrated around the places where surpluses were brought to be traded, towns developed, and this led to the beginning of a middle class: the independent burgesses who traded these surpluses, transporting them from where they were produced to where they were wanted.

In towns, the guild system emerged to regulate entry to crafts, setting out qualifications for achieving each grade: apprentice, journeyman, master, etc. Manufacturing was *concentrated* in a number of specialised areas but it was totally based on skills of individual artisans; *the person who started to make something, finished it*.

Basic processes for extracting metals brought about the earliest concentration of manufacturing industry around certain areas. Smelting had to take place where both raw materials and a source of energy needed to convert it were available. The earliest iron smelting were concentrated in forests growing on top of iron ore deposits.

The evolution of manufacture was very slow (Figure 3.1). It took thousands of years to move from farms to towns, to craftsmen with specialised skills, to focusing on a specific location. There were no factories. The only concentration of labour took place in the extractive industries and these conditions prevailed for centuries.

1.3. Cottage industries

The next stage in the evolution of manufacturing was the development of cottage industries (artisan or home industry). They began to develop when peasants who worked on the land began to practice a craft in order to earn a little more. This was an early stage of specialisation. Then, the cottage industry entrepreneurs saw the opportunities of the factory system. However, it was not the skilled cottage industry workers, but the land-less peasants who eventually went into the factories. Manufacture could have remained with artisans in towns for much longer, but several things happened which precipitated the emergence of the factory system.

1.4. Factory system

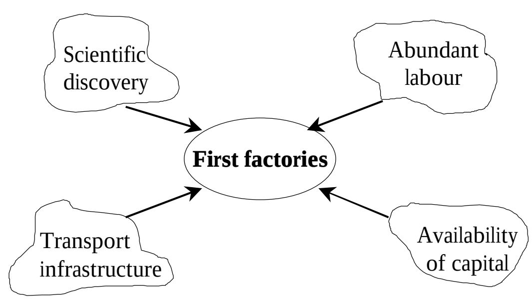

Figure 3.2. Four prerequisites of the development of the factory system.

A factory is a *concentration* of people and resources in one locality which specialises in the manufacture of one group of products, for competitive advantage and superior quality. The rise of factories was dictated by the need for specialisation and concentration. The four prerequisites which were necessary for factories to develop came together in England during the late eighteenth and early nineteenth centuries (Figure 3.2).

1.5. Mass production

The end of the nineteenth century saw a coming together of good and bad events– the era of mass production and interchangeable parts had begun. This came about as a consequence of the economic and technological developments that resulted from rising demand. However, mass production puts focus onto managing tiny slices of the job. Managing the whole was not an issue; all went quite well at the beginning. The emergence of TV advertising increased the demand for mass-produced goods. Customers were introduced to a whole new world variety that came into their homes via the tiny screen. Just having a kind of machine was no longer enough. People wanted to choose between *36* different models, with a variety of features, in different sizes and colours. Manufacturers were confronted with the need to offer a variety or lose market shares. This meant that companies had to find a way to react quickly to market trends.

The thing which people most appreciate about mass production in the USA is Henry Ford's model T. He founded the 'ford motor company', where he built on the concept of interchangeable parts by breaking down the process of making each part into standard pieces and by creating a production line. With the beginning of mass production the social organisation of labour was forced to change. Division of labour meant breaking down the job into specialised, repetitive tasks and setting up a track which brought the job to the man, instead of letting the man go to the job. Mass production and division of labour made it necessary to invent the concept of *work measurement*. Management determined the standard number of hours it ought to take to do a job, then measured productivity in man-hours, or by piece-work.

Taylor, who originated the principles of '*scientific management*', was one of the first to experiment with ways of increasing productivity in factories organised around mass production. Taylor and his colleagues observed that one employee ran to and from work every day. Further investigation revealed that he was also building his own house during evenings and weekends. They reasoned that if he could still run home in the evening, and then work some more, he must have some *energy left* so they designed an experiment to find out how to get the maximum output from him. They asked the employee to participate to an experiment. After observing the results over several weeks, they redesigned tasks and employee productivity went up. However, he probably no longer had enough energy left to run home, let alone build his house in the evenings. The results of time and motion study and careful job measurement reached industrial centres world-wide.

1.6. Computer in manufacturing

In the late of the *1960s* the unthinkable happened. The commodities prices fell, this was the end of the period where the problem was not selling, but making enough. Manufacturers were confronted with the need to offer variety or lose market share. The need to superimpose flexibility and variety on a system based on mass production and economies of scale began to cause problems. Confronted with these circumstances, manufactures first started to look around to see how they could improve their flexibility and responsiveness. Approaches like flexible manufacturing, just-in-time, MRP (I and II), numerical control: 'let the computer control your factory for you!', total quality, group technology, etc. raised at this moment. The need for manufacturing to be more agile prompted the development of production control systems. This hurried up the introduction of computers in manufacturing (Kirton, 1994). The evolution of manufacturing for the time seems to be more complex, and needs more and more reflection. Only few know how this will develop or what will happen in the new millennium, the changes due to the GATT[1], and so on.

2. Introduction to assembly line balancing

Most people think that Henry Ford invented the automobile. In fact he did not, the automobile is the result of a combination of technologies developed by many people over the time. However, he invented the assembly line (AL), which revolutionised the way cars are made, and how much they cost. Ford was the first who introduced a moving belt in the factory. Employees were able to build cars one piece at a time, instead of one car at a time. This principle, called 'division of labour', allowed workers to focus on doing one thing very well, rather than being responsible for a number of tasks[2]. The actor Charlie Chaplin immortalised in his film 'Modern Times' the way the workers try to repeat themselves on an assembly line in a factory.

One of the innovations in the history of assembly manufacturing is the division of the assembly job. If an assembly task has a long process time or involves too many parts, the work may be broken into a number of smaller tasks. Since each task has a relatively limited content, skill can be developed in a short time. Thus, assembly speed may be increased and quality improved.

An assembly line is a sequence of stations connected together by a material handling system. Stations are used to assemble components into a final product. The assembly process consists of a sequence of tasks or work-elements. A task consists of some elementary operations which are tied together because of the use of a common tool, jig or fixture. Accordingly, a task cannot be subdivided and must be completed at its assigned station. The tasks in an assembly process are typically ordered. The *assembly line balancing problem* (ALBP) aims to assign tasks to stations. The term 'balancing' comes from the fact that the aim is to balance the workload of the stations. Because tasks may require widely different times, the assignment of task-times to stations is rarely equal. This leads to idle time at stations. One of the classical objectives of the ALBP is to minimise this idle time.

---

[1] GATT: the general agreement on tariffs and trade.

[2] It takes several hours to assemble an automobile; how can a car assembly plant produce a car every few minutes? They use an assembly line with many stations: each station adds one part or many to the frame in a few minutes and then sends the frame to the next station. After a warm up period, it may take several hours for one given frame to go through the whole assembly line but every few minutes the line produces another car. Few minutes rather than several hours of work: kind of magic... !

For manual assembly line the global cost of the line is directly influenced by the number of stations. Thus, the main objective of the classical *line balancing* is to minimise the number of stations. A second kind of line is the line where the tasks can be executed either manually, by robots or by hard automated equipments. These lines are called hybrid assembly line (HAL). In general, the operating time and cost depend on the resource used. Given a list of candidate equipment available to complete the operations, the problem is to decide which resources to use and which tasks to assign to each of them. The problem is called in the literature the *resource planner* (RP) (Delchambre, 1996).

3. Assembly line systems

An assembly line is a familiar example from the realm of manufacturing; flow lines are found in all types of industries, wherever 'products' may be imagined to move along, from station to station. Products are progressively assembled as they move down the line toward completion. Assembly is a process by which subassemblies, manufactured parts and components are put together to make the finished products. The assembled product takes shape gradually, starting with one part (usually called base part), the remaining parts being attached at various stations the product visits.

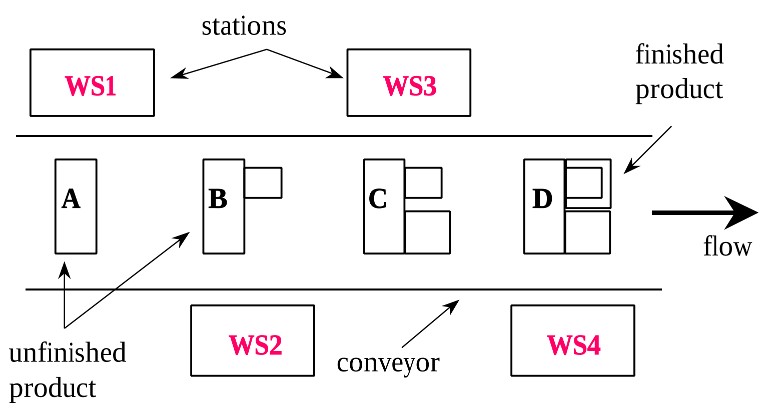

Figure 3.3. Assembly line concept.

The concept of a paced line is quite simple: a number of stations (four, in the figure: WS1 through WS4) are connected by a conveyor, each station performs one or more tasks (addition of components, inspection, etc.) on the partially finished product in front of the station (see Figure 3.3). Tasks are accomplished by a group of trained workers, machines or robots. After a lapse of time called the cycle time (C), the conveyor moves, thus positioning each product in front of the next station in the line. The product previously at the last station is finished and leaves the line. Operations are subject to precedence constraints certain tasks have predecessors, i.e. they can only be performed after one or more other ones have been completed. This means that tasks must be assigned to stations in such a way that precedence constraints always follow the sense of the conveyor.

In *paced* assembly line if a station finishes in less than *C* time units it will be idle for the remaining period. On the other hand, if the station is unable to complete the assigned tasks in *C* time units the remaining elementary-tasks will be finished forward. To ensure that this does not happen it is common to provide an additive time at each station so that the chances of incomplete work are minimised. Typically, if a part does not get completed it is taken to a repair and rework area for completion. If there is too much variability in the task process time, it is preferable to have *unpaced* or asynchronous line. In such line each station works at its own pace and advances the part to the next station whenever it completes its assigned tasks.

4. Notations and definitions

In this section we give some definitions, terms and notations used to describe the single/multi-product assembly line balancing problem. Some terms have already been intuitively introduced in previous section or chapters.

Assembly

An assembly is the process fitting together various parts in order to create a finished product. Parts may be subdivided into *components* and *sub-assemblies*. The unfinished units of the product are called *work in progress*.

Assembly line

An assembly line is a flow-line production system composed by a *number of stations* (n) arranged along a conveyor system (material handling system). The pieces are consecutively launched down the conveyor system and are moved from one station to another. A certain part of a total work needed to assemble a product is performed at each station (except control station, etc.).

Task

A task is a portion of the total work content (including control task, cleaning, etc.) in an assembly process. The necessary time to perform task is called *task process time*. Tasks (*operations*) are considered indivisible, they cannot be split into smaller work elements without unnecessary additional work.

Precedence constraints

Is the ordering in which tasks must be performed (technological restrictions). The partial ordering of tasks can be illustrated by means of a *precedence graph* (Sacerdoti, 1977). The nodes represent tasks and the directed arcs *(i, j)* constitute precedence relationships. In Figure 3.4 for instance, task 4 is preceded by tasks 1 and 2.

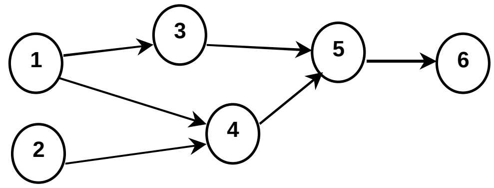

Figure 3.4. Precedence graph.

Cycle time (C)
Is defined as the time between the exit of two consecutive products from the line. In the case of paced assembly line, it also represents the maximal amount of work processed by each station, while in unpaced flow-lines, the cycle time is the maximal station time. Except in the case of parallel stations, the cycle time cannot be smaller than the largest process time. A positive difference between the cycle time and the station process time is called *idle time*. The sum of idle times of all stations of the line is called the *delay time*. The *desired cycle time* (C) is what the planning department asks for, while the *effective cycle time* (EC) is the real cycle time by which the line will operate, due to failures, setup-times, etc.

Capacity supply (CS)
The capacity supply is defined as $CS = n * C$. It is the total time available to assemble each product. Obviously, the CS is greater or equal to the sum of process time of all tasks work content. The CS, depends also on the transfer system of the assembly line (free or linked transfer).

Imbalance
The multi-product assembly line are balanced on average, thus, the process time of each station depends on the variant-product. The imbalance (IB) is measured by the difference between C and the total duration of tasks concerning a given variant on each station (Figure 3.5).

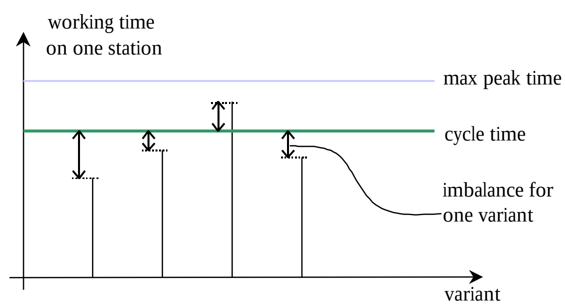

Figure 3.5. Task duration is variable according to the variant.

Time interval
The time interval is the interval TI = $[t_{min}/C, t_{max}/C] \in [0, 1]$, it measures the interrelation between the cycle time and the task times. Problems are expected to be relatively complex if TI is close to 1.

Time variability ratio
The time variability is defined by TVR = t_{max}/t_{min}. Small values of TVR indicate that the operation times vary only in a small range.

Times
A (work) station is a part of an assembly line where a certain amount of work (a set of assigned tasks) is performed. The work content of a station is referred to as *station*

load, the total process time as *station time* (ST). The sum of station times of the whole assembly line is the *total assembly time*. Each station is characterised by its dimension, a set of tasks, a set of equipment as well as its kind. Work is manually done by human operators using simple tools or by semi-automated machines. The process time of tasks belongs to the interval [t_{min}, t_{max}], (where t_{min} is the minimum process time and t_{max} is the maximum process time).

Work content

The work content is the sum of process time of all tasks: $WC = \sum_{i=1..n} T_i$

Line efficiency (E)

The line efficiency is defined by: $E = {}^{WC}\!/_{CS}$ measures the capacity utilisation of the line. The unused (idle) capacity is reflected by the *balance delay time* which is defined by $BD = CS - WC$.

Makespan

The maximum completion time required to process all operations for a given set of products. It is also the total time until the last job is done.

Maximum peak time

The maximum peak time, fixed by the user, is introduced to deal with multi-product assembly line. It allows some variants process time to exceed C. It lays in the interval [C, 2C]. This maximum peak time may not be exceeded by any variant process time on a given station, while the cycle time must not be overstepped by the average working time on a station.

Smoothness index (SX)

The Smoothness index is defined through $SX = \sqrt{\sum_{i=1}^{n} (C - ST(i))^2}$ which, measures the standard deviation of the distribution of work among the stations.

Station idle time

The station idle time is the positive difference between the cycle time and the station time. The sum of stations idle time is given by: $I = CS - WC$.

Throughput time

Denotes the *average* time interval between launching a work-piece down the line and removing the finished product from the line.

5. Assembly line balancing problems

Many classifications of assembly line problems are given in the literature (Baybars, 1986) (Scholl, 1999). Ghosh and Gagnon (Ghosh , 1989) classify the ALB studies into four categories: single model deterministic (SMD), single model stochastic (SMS), multi/mixed model deterministic (MMD) and multi/mixed model stochastic (MMS). In this section we relate some classification schemes of AL problems.

5.1. Assembly line models

The planning in production management begins with demand forecasting, the process of estimating future demand of a single product or variants of family of products. The reason doing this at all is to get an idea of what will happen in the future. It is more-or-less like driving a car, your actions depend on what you are expecting.
The demand forecast in assembly and in manufacturing in general may depend among other things on the number of product units to be made at each period. Meeting this demand generates requirements for people, equipment, and other resources. A demand forecast therefore helps to develop appropriate plans for production. The more products to be made, the greater the resource requirements. Thus, the number of products to be assembled and their proportion has an important influence on the line architecture. Three types of productions exist, namely single product line, family of products line and multiple products line. The main factors that influence the choice among the three approaches are summarised below.

5.1.1. Single product line

Used to produce only one product. The workload of all stations remains constant over time if dynamic phenomena are neglected. The stations perform the same tasks on the product. Each time the following conditions are true, it is better to use a single assembly line to produce only one product in large quantities rather than a multiple products line.

- The demand of the product is constant and allows absorption of line cost.
- The product must be delivered in a very short time (no delivery delays).
- The product has a structure which is different from other products.
- The cost of product interdict mistakes during production (parts made of gold).
- The assembly of the product needs heavy and bulky machines (resources).
- The setup-time of the dedicated assembly line needs a lot of time (several shifts).

The demand for a single product is often not sufficient to justify an independent assembly line. In such cases it is common to design mixed product or multiple assembly line.

5.1.2. Mixed-production line

A family of products is a set of distinguished products, called variants, whose main functions are preferably similar. In the automotive industry, a typical example is a family of cars with different options: some of the cars will have a sunroof, some will have the ABS brake system and so on. For such lines, the workload on each station depends on the variant at hand. Such line can be used if:

- The cycle time is greater than a minute.
- The demand of products changes quickly, and the line price cannot be amortised by a single product.
- The product must not be delivered in a short time.
- Each product is quite similar to others (the whole constitute products family).
- The same resources are needed to assemble all the products.
- The setup-time of the assembly line needs a short time.

Typically, the products being made on the assembly line tend to have very similar tasks and precedence diagrams. While designing such assembly lines, the assembly line operation stage must be taken into account. More, since all the products follow the same direction (flow line), special attention must be paid when assigning tasks to stations (precedence between tasks must be respected). In case of mixed production, scheduling methods have to be used to ensure workflow smoothness along the line.

5.1.3. Batch production line

Used in the case of multiple different products, or family of products which presents significant differences in the production processes. In this case, the batch production will be applied most of the time. Using batch production leads to scheduling and lot sizing problems. The main factors that make the difference compared to the family of products line are:

- The demand of products changes slowly;
- The product must be delivered in a short time;
- The assembly of some products needs heavy and bulky machines (resources);
- The setup-time of the assembly line is short.

5.2. Batch or mixed production

With the increased diversified demand in the production of white goods and cars, more and more manufacturers are using mixed-model assembly line. The ALD has to take this evolution into account. Two approaches are used in modern industry; the choice is made (or guided) by the marketing and the resource departments.

- **Mixed production (MP)** Several products (variants of the product family) or all variants are produced all over the time,
- **Batch production (BP)** A given amount of products of the same product are produced over a certain period.

A mixed-production assembly line (MPAL) is an assembly line capable of producing a variety of different product models (called variants) simultaneously and continuously. Stations are sufficiently flexible to perform their respective tasks on the different variants. All these variants are represented by a generic product model which contains the necessary information. In MPAL, the right part must be at the right place at the right time and any program to control this process must run in real time. One of the big deals is to manage the storage space, the resources, etc.

Other industries use batch production, and produce different products on separate assembly lines. The problem seems to be solved: each product may have its assembly line. Therefore, if the company produces 100 products (which is an average), it may have 100 assembly lines. The cost will be too high, and none has used such philosophy due to plant space constraints. Thus, designers tend to re-design and reuse existing assembly line to produce the different products (batches).

The most important factors which influence the use of one of these approaches are:

- The MP must be discarded in the case of high setup-time (several shifts);
- The family of products presents great irregularities between its variants, BP will be preferred;
- The demand forecast of the variants greatly oscillates, BP will be preferred;
- Small cycle time (less than a minute) reinforce the use of BP;
- For small number of stations, BP helps to avoid a high imbalance among variants.

5.3. Variability of tasks process times

The task process time is an essential parameter in the assembly line balancing. Indeed, depending on the nature of tasks and the skills of operators or the reliability of the dedicated machine or robot, the operation time may vary. Simple tasks may have a small process time variance, while complex and unreliable tasks may have high varying execution times. In the case of human workers, factors such as skills, motivation, communication among the group, etc. have a great influence on the whole assembly line. The following situations may arise.

5.3.1. Deterministic time

In the case of manual assembly line, the task time will stay constant only in the case of highly qualified and motivated workers. Modern machines and robots are able to work permanently at a constant speed. The use of deterministic times is always justified if the expected process time variability is sufficiently small (this happens in the case of simple and well defined tasks). One can reduce the task time variation by increasing the line's automation degree.

5.3.2. Stochastic time

In automated flow line varying production rates may result from machine breakdowns. In the case of humans, significant variation may result from non qualified workers, motivations of the employees, lack of training, etc. Thus, such variations may considerably influence the performance of the system. In the case of hybrid assembly line, humans may be hung up by the machine, since a dedicated machine or a robot has the same throughput over time, while human effort varies.

To incorporate process time variability, operation time may be modified by adding the stochastic component reflecting the MTBF (mean time between failures) and the MTTR (mean time to repair). The varying process time of AL, lead to buffer sizing problem, and resources duplication, etc. The fuzzy logic (Zadeh, 1965) concept can be used to tackle the stochastic nature and the variability of process time.

5.3.3. Hidden times

In the case of automated stations, it is often not an easy task to determine the operating time of a complex task (grouping of two or more tasks). Indeed, the global duration of the operation is not always the sum of the operating times of each equipment in the group because of so-called hidden times (see Chapter 7).

5.3.4. Dynamic time

In the case of human workers, systematic reductions are possible due to learning effects or successive improvements of the production process. Indeed, for new tasks, operators take a longer time to execute the operation than after becoming familiar with the work. The process time reduction depends on the task complexity degree, skills and motivations.

5.4. Line configuration

In the plant layout problem, emphasis is often placed on material flow between departments or stations. The material flow requires material handling which is a costly and non value adding operation. Ideally, the preliminary analysis of the product and the plant shape leads to a first global layout of the assembly line. This helps the designer to reduce the problem complexity. Thereafter, the whole assembly system is divided into small connected workcenters. Each workcenter is a small assembly line associated to one or more sub-assemblies. The space, the cost, the balance and the material handling problem, must be taken into account while deciding about workcenters. The plant is then constituted by a set of workcenters. Several line configurations (flow patterns) are possible.

5.4.1. Serial lines

Single stations are arranged in a straight line along a conveying system (Figure 3.6). Each station perform one or more tasks on the partially finished product (Baybars, 1986). It can be a simple unit of a complex system.

Figure 3.6. Serial line configuration.

5.4.2. U-shaped lines

As a consequence of introducing the just-in-time production principle, it has been recognised that arranging the stations in a U-line has several advantages over the traditional configuration. Indeed, serial line shape may render a worker's job quite boring.

Both tops of the line are close to each other forming a 'U' shape (see Figure 3.7). Thus, visibility of the whole production process and communication between operators are improved. Thus, workers acquire multiple skills leading to higher motivation, improved quality of products and increased flexibility. This flexibility is required in JIT production systems, because demand rates of products change dynamically from time to time. In this case, the workers are placed in the centre of the U and can monitor each other's progress and collaborate easily whenever required. The close proximity also helps in team-building (Miltenburg, 1994).

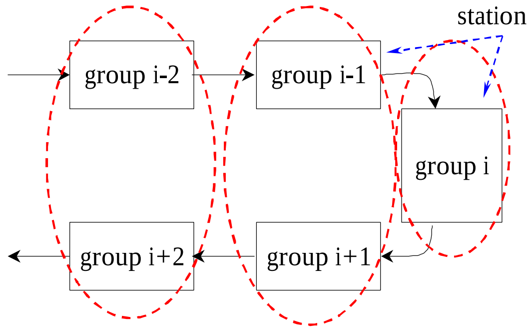

Figure 3.7. U-shaped line configuration.

The U-line pattern is often used when it is necessary to keep the receiving and shipping ends of the line at the same end of the plant. This may be due to material handling considerations (the same forklift can be used for both purposes) or external access considerations (road access, truck docks).

5.4.3. Parallel stations

With high production rates, the longest task time sometimes exceeds the specified cycle time. A common remedy is to create stations with parallel or serial posts, where two or more workers perform an identical set of tasks. This procedure reduces the average value of the task duration proportionally to the number of workers on the station. Figure 3.8 gives a schematic of a simple serial system and one with a parallel station.
- parallel: the same job is done on two parallel stations,
- serial: the same job is done on two serial stations.

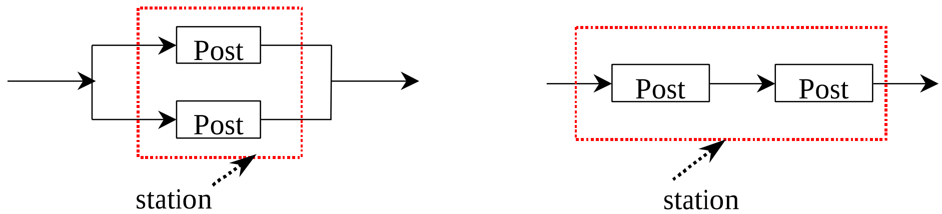

Figure 3.8. Parallel and serial stations.

Parallelism can be used if a task process time exceeds cycle time or at the will of the designer. There are cases where having multiple workers (posts) in station will result in a higher utilisation as compared to having a single worker in station. Paralleling also permits one to deal with cycle time which is shorter than the shortest task time. The main factor which helps to decide on the utilisation of parallel or serial stations is the space dedicated to the resources. The serial configuration is more adapted for the automotive industry, while the parallel one is more preferred in the case of small products.

5.4.4. Parallel lines

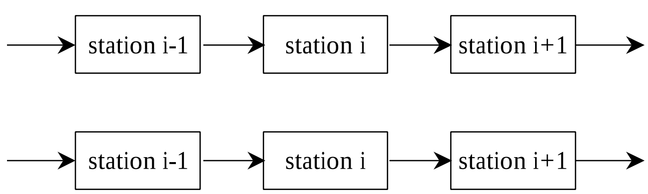

Figure 3.9. Parallel assembly lines.

If we decide to duplicate the assembly line, the cycle time will be multiplied by two (Figure 3.9). When the demand is high enough it is not uncommon to duplicate the entire assembly line. This has the advantage of shortening the assembly line, but may require more equipment and tooling. Parallel assembly lines have other advantages as well. The use of parallel lines allows the enlargement of the cycle time. Thus, the extent of labour division is relatively small, and the same number of operators is employed in each parallel line. Since each station has a larger amount of time to complete its tasks, more tasks can be assigned to the station, thereby enriching the work content. Also, if failure occurs at a given station, other lines can continue to run. A single serial line would have to be shutdown whenever there is failure at any station. These units are organised as autonomous work teams. Paralleling assembly lines also permits to increase the flexibility of the system.

5.4.5. Workcenters

The plant layout problem is concerned with the location and arrangement of departments, cells or machines, etc. on a plant. For complex products, the assembly system, is most of the time, decomposed into sub-systems (workcenters) which are easier to manage than the entire one (see Figure 3.10). The routing of a product between workcenters is fixed, as it works according to a flow topology. The main topology of the line is not necessarily a linear one (see Chapter 9 for more details).

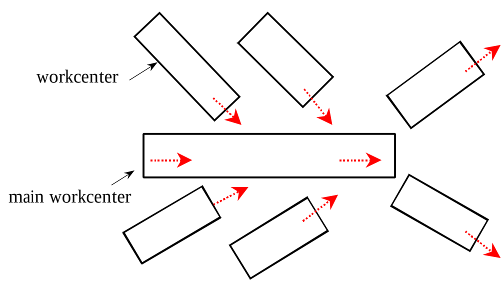

Figure 3.10. Example of plant topology.

5.5. Additional constraints

The classical line balancing problem can be improved to include other factors in order to deal with real-world problems. Dealing with constraints and user's preferences reduces the number of task-station combination and at the same time the search space of the problem. Such constraints have the same influence on the complexity of the problem as the precedence constraints of the product.

In some cases, it is necessary that a particular pair of tasks should not be done at the same station (*dissociative* designer's preferences). Such restrictions are required because of dirty tasks, fire hazards, vibration problems, etc. In such cases we must add additional constraints to the model that allow us to deal with incompatible tasks on stations. Similarly, we may wish two or more tasks to be assigned to the same station (*associative* designer's preferences). This is often done because of the use of expensive tooling, hazardous chemicals, etc.

One of the industrial preoccupations is the recovery of existing stations (special machines or robots) during layout changes. Two types of operations are introduced to deal with this kind of user's preferences.

5.5.1. Fixed operations on stations

Some operations have to be fixed on a given station (control station, paint station, etc.) and no additional operation can be added to this station. Figure 3.11 represents two different layout where the content as well as the position of the fixed station remains constant. It occupies the second position.

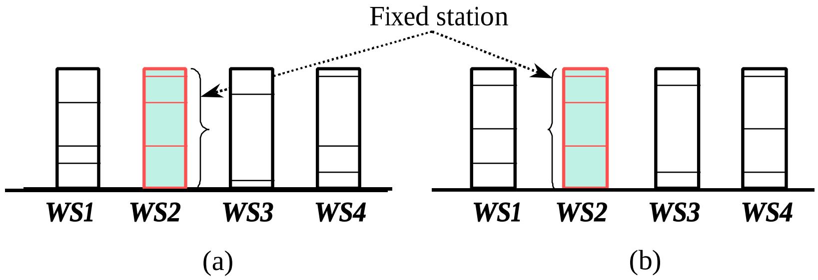

Figure 3.11. Two positions of the fixed station.

5.5.2. Linked operations

A set of operations must be grouped on the same station, but additional operations can be added. In Figure 3.12 a set of linked tasks occupy the fourth position in the first layout, while they occupy the third position along the line in the second layout. The two solutions have to respect all precedence constraints of the product.

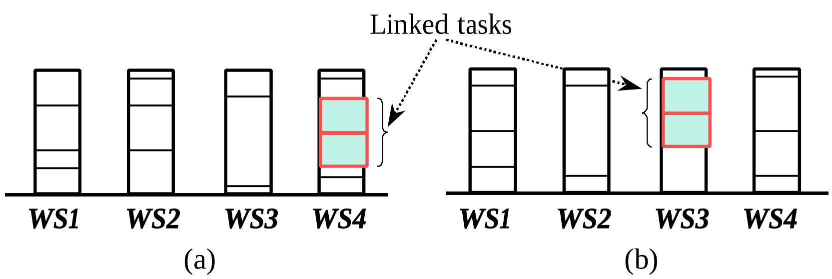

Figure 3.12. Two positions of a set of linked tasks.

5.5.3. Must directly precede

In order to install safe, non expensive and stable assembly line, designers have to think about the stability of the product being assembled. This lead us to include a '*must directly precede*' relation, which means presence of a direct precedence between tasks rather than an indirect precedence constraint. Figure 3.13 illustrates the idea. In Figure 3.13(a) task *1* must only precedes task *5*. Thus, the solution where the first station contains task 1, the second tasks {3, 4} and the third station task 5 is accepted since the precedence is not restricted in this case. Figure 3.13(c) represents a valid solution where task *1* must directly precede task *5*. Indeed, task 1 is directly followed by task 5. Figure 3.13(b) represents a non valid solution since stations containing task 1 and 5 are separated by a station which contains tasks {3, 4}.

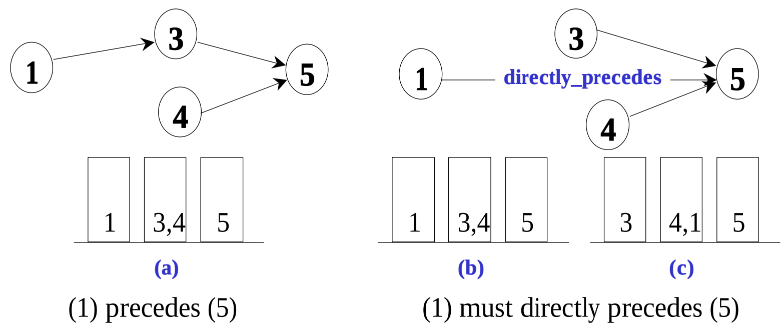

Figure 3.13. Precedence and must-directly precedence relations.

5.5.4. Associative and dissociative constraints

In a more general setting and especially when we have to design hybrid assembly line (line with manual, robotic and automated operations), sparing a station does not necessarily mean making economies. Indeed, it is usually possible to save time by investing in faster equipments or by increasing workforce assigned to a task. Consequently, under a constant cycle time, the number of stations can be reduced by increasing the price of the line.

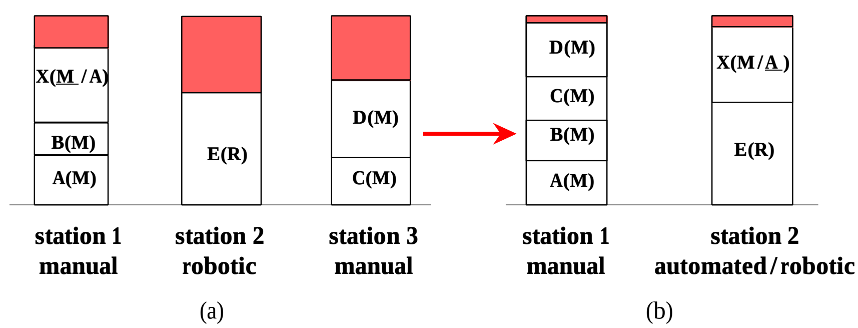

Figure 3.14. Result when X is manual (a) and the result when X is Automated (b).

Three possible methods for each operation (manual, robotic and automated) are considered. This yields a set of *associative* preferences (manual operations have to be grouped together, and the robotic or automated operations have to be grouped together) and *dissociative* preferences (manual operations cannot be grouped with robotic or automated operations). Those constraints are hard, and by the way they cannot be violated.

The operating mode of tasks will be fixed by the line balancing, to obtain the best logical layout. Let us suppose that we have six operations to realize:

- four operations (A, B, C, D) have to be executed manually.
- one operation (E) has to be executed with a robot.
- one other operation (X) has to be executed manually or automatically, with a preference for the manual method.

Suppose that when we consider the case where X is manual, the best configuration we can obtain for the global line is the one represented in (a), whereas if we take the case where X is automated, the best configuration we can obtain is the one showed in (b) (see Figure 3.14). In this example, it is clear that the second solution is the best one, if we do not take the equipment cost into account.

5.5.5. Positions of operations

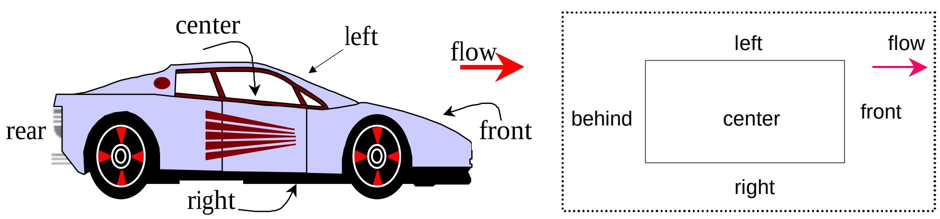

Figure 3.15. Different position of operations.

In automotive industry, designers have to be aware of the position of operations inside the car (influence the time needed for operators to move among positions). In order to solve this problem, five positions are introduced (Figure 3.15): (*C*: operation on the centre of the car, *R*: operation on the right side, *F*: operation on the front, *L*: operation on the left side, *B*: operation on the rear). Thus, station duration will be composed of the sum of the operations duration and the additional time for (1) *operator moves*, and (2) *reading code* of different variants (time needed to make difference between two successive different variants).

5.6. Assembly line design problems

As mentioned in Chapter 2, the design of assembly line comprises the logical and the physical layout problems. The physical layout fixes the space requirements taking into account station dimensions and material storage, etc. The logical layout is referred to in the literature as the assembly line balancing (ALB) or the resource planning (RP). In order to deal with mixed production, the ordering variants (OV) allows us to take into account the operation phase of the line during the design phase. In the next sections, the three modules will be introduced.

5.6.1. Line balancing

Assembly line balancing is a traditional approach to the elaboration of a first rough layout of a production line. This approach is appropriate for manual and mono-product assembly environment where the primary concern is the balancing of tasks among assembly workers, while keeping labour costs to a minimum (see Chapter 6).

The classical definition of the simple ALB specifies the following assumptions (Baybars, 1986):

1. All the parameters relating to the line must be known with certainty.
2. An operation cannot be divided between two or several stations.
3. Tasks cannot be treated in an arbitrary order due to the precedence constraints.
4. All the tasks of an assembly line must be carried out.
5. All the stations are equipped with various resources, and can carry out any task.
6. The operation duration is independent of the station on which it will be carried out.
7. Any operation can be made on any station.
8. The assembly line is serial, and does not contain either feeding system, nor lines in parallel.
9. The assembly system is to be designed for a unique model of a single product.
10. The cycle time is fixed, and the goal is to minimise the number of stations. Or, the number of stations is fixed, and the goal is to minimise the cycle time.

This problem formulation has a natural application in environments where the production means can be considered universal, because the definition of ALBP implies that (1) any task can be performed at any station and (2) a task always takes the same time, regardless of the station it is performed at. These conditions are usually met by the fully manual lines or the fully robotic ones (homogeneous line). In these environments, a reduction in the number of stations corresponds directly to savings in economic terms, because workforce or robots are saved. Each product features precedence constraints (3), i.e. some tasks can only be performed after one or more other ones have been done. This means that tasks must be assigned to stations in such a way that precedence constraints always follow the conveyor sense.

Installing a production system requires large capital investments and thus needs a long-term decision. Depending on the line layout, many objectives may be sought after. Six types of assembly line balancing problems are proposed.

The mathematical formulation of these problems is given below. Let us have a directed acyclic graph G=(T, P), where the nodes T represent the tasks and the arrows P represent the precedence constraints, a constant L_i (task length) assigned to each node T_i.

SALBP-1 (simple assembly line balancing problem-1)
Consists in assigning tasks to stations so that the number of stations is minimised for a given production rate. It is dedicated to a single product with deterministic process times and serial line. Given a constant C (the cycle time) and a constant N, can the nodes T be partitioned into N or less subsets S_j (the j-th station's tasks) in such a way that (1) for each of the subsets, the sum of L_i associated with the nodes in the subset does not exceed C, and (2) there exists an ordering of the subsets such that whenever two nodes in distinct subsets are jointed by an arrow in G, the arrow goes from a high ordered (earlier) to a lower ordered (later subset)? The aim is to minimise the number of stations.

SALBP-2
Aims to maximise the production rate, or equivalently, to minimise the sum of the idle times for a given number of stations. Used in the case of single product, deterministic process times and serial line. Given a constant N (number of stations), can the nodes of T be partitioned into the N subsets S_j. Since the number of stations is fixed, the aim is to minimise the cycle time of the assembly line.

SALBP-E
Aims to minimise the sum of the idle times with a fixed production rate and a fixed number of stations. It is the generalisation of the SALBP-1 and SALBP-2. Given a constant N (number of stations) and a constant C (cycle time), can the nodes of T be partitioned into the N subsets S_j. The aim is to find a solution where the sum of the L_i associated with the nodes of the N subset does not exceed C.

EPALP (equal piles for assembly line problem)
The method warrants the obtaining of the desired number of stations, and tries to equalise the station loads. Given a constant N (number of stations), can the nodes of T be partitioned into the N subsets S_j. In contrast to the SALBP-2, the aim is to equalise the station loads and not to minimise the cycle time.

MWkCALB (multiple workcenters assembly line balancing problem)
Given a set of W directed non cyclic graphs $G_i = (T_i, P_i)$ and a set of N_i (number of stations); given a set of links between these graphs, (graphs can be connected by their ends or their beginning), can the nodes T_i of each graph be partitioned into the N_i subsets S_{mi} (the m-th station tasks). The aim is the balancing of a set of workcenters using the different links between them.

MOALBP (multi-objective assembly line balancing problem)
Assembly line designers deal with objectives like: line efficiency, smoothness index, imbalance (Chow, 1990). It represents most assembly line design problems. The aim is to optimise a set of objectives like: the line efficiency, the balance, cost, etc.

For a fixed cycle time, a valid assembly line is the one where the sum of duration of all tasks assigned to any station does not exceed the cycle time, otherwise the line could not keep up the desired pace. The ALBP with a fixed number of stations may have two objectives: minimise the cycle time (SALBP-2) or equalise the station loads (EPALP). Since these objectives are not the same, the results are generally different. The first case may lead to unbalanced lines (by minimising the maximal idle time), whereas the second one leads to balanced ones.

5.6.2. Resource planning

The main task of the resource planner is to design a flexible assembly system that will be able to assemble a product at a least cost. The tasks can be executed either manually, by robots or by automated equipments. In general, the operating time and the cost depend on the resource used. Given a list of candidate equipments available to complete the operations, the design problem is to decide which resources to use and which tasks to assign to each resource in order to meet the production requirements at minimum cost (see Chapter 7).

In formal terms we defined the RP problem as the following decision problem. For sake of simplicity, we will only take the process time and the cost of tasks to define the problem. The same reasoning can be made if other design objectives are added. Let G = (T, P) be a directed acyclic graph, the set of nodes T representing the tasks and the arrows P the precedence constraints between the tasks. Each node T_i is characterised by a set of couples $\{L_{i,j}, C_{i,j}\}$($L_{i,j}$ is a possible duration of the task and $C_{i,j}$ the cost of the corresponding used resource). Let N (number of stations), CT (cycle time), and C (cost) be three constants. We define the cost of a subset of T as the sum of $C_{i,j}$ of the nodes belonging to this subset. Is it possible to find a partition into N subsets of the set of operations and for each of them select a couple ($L_{i,j}$, $C_{i,j}$) so that the sum of $L_{i,j}$s in a partition is less than or equal to CT, and the sum of all subsets costs less than or equal to C? The subsets must be ordered in such a way that whenever two nodes in distinct subsets are joined by an arrow in G, the arrow goes from a higher ordered (earlier one) to a lower ordered (later one).

It is easy to show that all these problems are similar and are generalisation of the (SALBP-1). The difference consists in the way of treating the problem, the aim is to group tasks in stations in order to minimise a certain function. Since the objectives are not the same the results are generally different.

5.6.3. Ordering variants

Multi-product assembly line (MPAL) have become an important issue in industry because of market pressures for diversity, shorter life and competition. However, diversity is in contradiction with other current goals as low costs, high productivity and standardisation. Thus a company's ultimate success depends on its ability to deal with the complexity of a product and process designs. It has become clear that assembly and manufacturing questions must be taken into account from the product design stage. Methods as design for manufacturing (DFM), design for assembly (DFA) have reported significant benefits including product simplification, lower assembly and manufacturing costs, improved quality and reduced time to market (Delchambre, 1996).

The aim of multi-product line balancing is to provide unique assembly line balancing, valid for all the variants of the product. Traditionally, because of the variability of the duration of work at each station due to the fact that some operations are missing for some variants, the MPAL is balanced on average. The aim is to minimise the imbalance measured by the difference between the average cycle time and the total duration of operations concerning a variant on each station. A large imbalance of the workload among different variants has to be avoided. Indeed, even when the station load is balanced on average, the production can easily be faced with work overflow

or starvation on individual stations. Thus, work imbalance due to the variants should be distributed among the stations.
MPAL are more difficult to balance than single product assembly line. There are two ways to design such line.

- Balance the line for each product and use a batch production. This solution can create a high volume of inventory.
- Balance the line for all products and use a mixed production. Since it is difficult to balance a line for all products, scheduling (ordering variants) problem must be considered during the design phase of the line.

The ordering variants problem occurs when a number of products (variants of family of products) are produced on an assembly line at the same time. The objective is to determine the sequence of variants which maximise the utilisation of the stations. Line balancing and ordering variants are interrelated because the balancing solution affects the determination of the launching model (see Chapter 8).

6. Why is the balancing problem hard to solve?

While it is known that some solving-methods are difficult to juggle with, it is not enough to formulate a problem as such solving-methods to claim that the problem is difficult. Some special property may exist (problem's constraints) which may allow to solve the problem easily.

To show how the assembly line balancing problem is computationally difficult we consider a special case of the problem. In this special case, we assume that there are no precedence, no grouping preferences, etc. The resulting problem is reduced to packing the tasks into the fewest number of stations. This is the well-studied 'bin-packing problem' (BPP). The BPP belongs to a class of problems which are known to be computationally difficult (NP-Hard problems) (Wee, 1982). Although, we cannot claim that an easy (more formally, polynomial) method to solve these problems does not exist, it is widely believed that this is indeed true. The reason for this belief is that the BPP belongs to a class of problems which are all related to each other in such a way that if any one of these problems is shown to be easy then all these problems can be shown to be easy. Since these problems have been the subject of studies for several decades without yielding an easy method, it is widely believed that no such method exists.

Now we know that the problem is computationally hard, we can feel comfortable using any method: enumerative methods, evolutionary approaches, heuristics, etc. Since classical methods are quite time consuming for larger problems, our emphasis was on genetic algorithms (see Chapters 4, 5). As seen above, the simple line balancing problem can be augmented by additional constraints to deal with real-world problems. Thereafter, it is an important tasks to find solution that compromises between the balance and all these additional constraints.

7. Conclusions

In this chapter, we started with the history of the evolution of manufacturing, we showed how manufacturing and assembly systems evolved from single hunter-gatherer to present day architectures. Then, we introduced the assembly line balancing problem. Thereafter, a sort of classification of the assembly line design problems was given. This classification is based on line models, production philosophy, process time models, line configuration, design objectives, etc. Finally, a discussion on the *complexity* of the balancing problem was presented.

The logical line layout is composed of two modules namely the ALB for the manual assembly line and the RP for hybrid assembly line. The solving methods for the two problems will be presented in Chapter 6 and 7. The methods are based on the grouping genetic algorithm (GGA). Thus, in the next chapter, the basics of genetic algorithms are presented.

8. References

(Baybars, 1986) Baybars I., ‘A survey of exact algorithms for the simple assembly line balancing’, Management Science, Vol. 32, pp. 909-932, 1986.

(Chow, 1990) Chow W.-M., ‘Assembly line design - methodology and applications’, Marcel Dekker, New York, 1990.

(Delchambre, 1996) Delchambre A., (Ed.), CAD method for industrial assembly concurrent design of products equipment and control systems’, John Wiley & Sons Inc., Chichester, England, 1996.

(Ghosh, 1989) Ghosh S. and Gagnon R.J., ‘A comprehensive literature review and analysis of the design, balancing and scheduling of assembly systems’, Int. J. of Prod. Res., Vol. 27, pp. 637-670, 1989.

(Kirton, 1994) Kirton J. and Brooks E., Cells in industry: managing teams for profit’, McGraw-Hill, 1994.

(McMullen, 1998) McMullen P.R. and Frazier G.V., Using simulated annealing to solve a multiobjective assembly line balancing problem with parallel stations’, International Journal of Production Research, Vol. 36, pp. 2717-2741, 1998.

(Miltenburg, 1994) Miltenburg G.J. and Wijngaard J., ‘The U-line balancing problem’, Management Science, Vol. 40, pp. 1378-1388, 1994.

(Pellichero, 1999) Pellichero F., Computer-aided choice of assembly methods and selection of equipment in production line design’, Ph.D. Thesis, Université Libre de Bruxelles, 1999.

(Sacerdoti, 1977) Sacerdoti E.D., ‘A structure for plans and behavior’, Elsevier, Amsterdam, Netherlands, 1977.

(Scholl, 1999) Scholl A., ‘Balancing and sequencing of assembly lines’, Heidelberg Physica, 1999.

(Shimokawa, 1997) Shimokawa K., Jürgens U. and Fujimoto T., (Eds.), ‘Transforming automobile assembly: experience in automation and work organization’, Springer-Verlag, Berlin Heidelberg, Germany, 1997.

(Zadeh, 1965) Zadeh L., 'Fuzzy sets', Information and Control, Vol. 8, pp. 338-353, 1965.

(Wee, 1982) Wee T.S. and Magazine M.J., 'Assembly line balancing as generalized bin packing', Operations Research Letters, Vol. 1, pp. 56-58, 1982.