

# An MGTS 352 introduction to *Constraint Management* using the *Theory of Constraints*<sup>1</sup>

## Grain Export Bottleneck

Crops grown in Western Canada and sold overseas (for example, Asia) pass through a number of steps in the physical export system. First, the crops are harvested at farms and temporarily stored on the farm or at local storage facilities. The crops are then transported by truck to regional grain elevators for short-term storage while waiting for rail transport to a sea port (such as Prince Rupert, BC). Rail cars are positioned at the grain elevators by railways (CN Rail and/or CP Rail), filled, and then transported to the port. At the port, the crops may be stored once again while waiting for ocean transport to a foreign port, eventually to be loaded into dry bulk ocean vessels for transport across the Pacific Ocean.

During the winter of 2014, a bottleneck emerged in this system when the railways were unable to move the crops fast enough and the crops became backlogged at storage facilities and on farms. Put simply, the capacity of the rail lines, due in part to the need to move other more profitable cargo such as oil and containerized goods, could not keep up to the demand for agriculture exports (CBC News, 2014).

The following document presents a methodology for managing constraints (bottlenecks) in systems by way of a set of management principals known as The Theory of Constraints.

## Learning Goals (after this chapter, the student should be able to:)

- Describe the *Theory of Constraints*, in general, and be proficient with the terminology.
- Discuss the concept of *system capacity* and the implications of activities working in series.
- Explain the “*The Five Steps of Focusing*” and apply them to examples and cases.
- Describe and solve *product-mix* problems.

## Outline

1. Context, capacity, and terminology
  - 1.1 Context (chair factory)
  - 1.2 Capacity
2. The Theory of Constraints (TOC)
  - 2.1 Origin
  - 2.2 TOC as “The Five Steps of Focusing” (with application to the chair factory)
  - 2.3 Other contributions and representations of TOC
3. Application to “Product Mix” problems
  - 3.1 Case Study
  - 3.2 Aggregate workload
  - 3.3 Trade-off: profit versus consumption of constrained resource
  - 3.4 Finding the optimal solution
  - 3.5 Example two
  - 3.6 Multiple bottlenecks

Appendix One – Process chart for single-bottleneck product mix problems

References

---

<sup>1</sup> Document authored by Ryan Orchard, June, 2015. Thank you to Mark Arnison for editing.

# 1. Context, capacity, and terminology

## 1.1 Context (chair factory)

Consider the following hypothetical system (not meant to be perfectly factual) for producing wooden chairs:

- *Raw material* (wood) is cut into pieces of a specific size and shape.
- The cut pieces are passed to a drilling station, where holes will be drilled as per specifications.
- After drilling, pieces are put through a sanding machine to make them smooth and to remove rough edges from the cutting and drilling steps.
- After sanding, a finishing step is performed in which the wooden pieces are stained and lacquered to seal them.
- Finally, finished pieces are assembled into *finished goods*, ready for sale to customers.

A diagram of this production system is provided below (figure 1). Note that additional *raw materials* may be required by different activities in the process - for example, wood stain and lacquer will be used for finishing, and glue, brackets, and fasteners will be needed for the assembly activity. For convention, we will say that work-in-progress (WIP) moves from “upstream” activities in the production system to “downstream” activities, and that WIP accumulates “in front of” an activity if the WIP has yet to be processed by that activity but has been processed by all upstream activities. WIP generally accumulates at a given activity when all upstream activities are working at a faster rate than the activity, although WIP can also accumulate due to the tendency to produce in *batches* at each activity. This will all be discussed further later in this document.

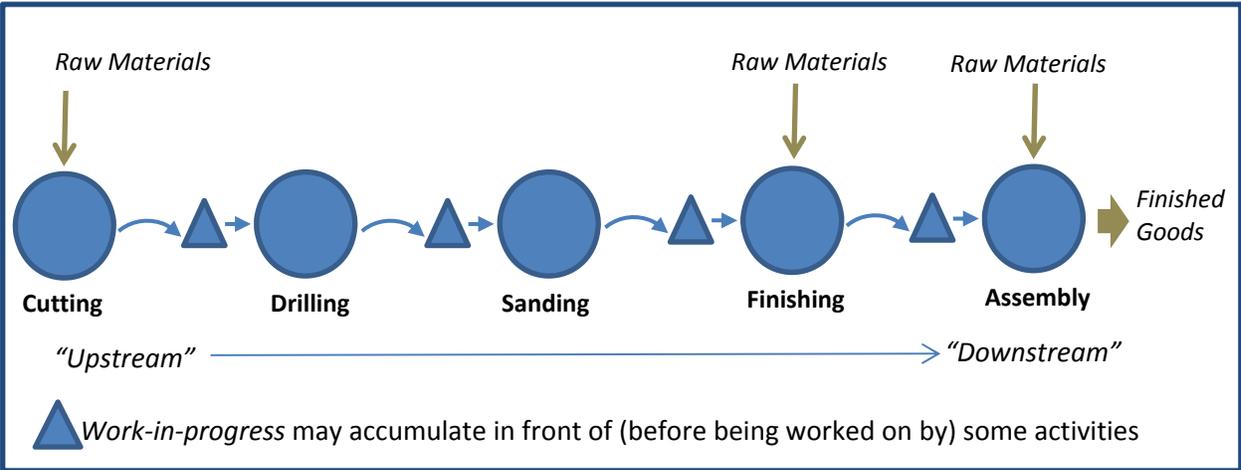


Figure 1 – Production system for wooden chairs

## 1.2 Capacity

We will use the term **capacity** to refer to the *output capacity* of an activity or system, which could also be described as the **maximum output per unit of time**. For example, a city's light rail transit (LRT) system may be capable of moving 4,500 passengers per hour into the downtown, or a workstation in a clothing factory (where the term *workstation* simply refers to a distinct activity in a process/system) may have a capacity of 1,200 pieces per day. We will use the term **utilization** to refer to the **actual output divided by the capacity** for a given activity or workstation. So, if the clothing factory workstation from the above example processed 800 pieces/day, then the utilization of the workstation would be 66.7%.

Returning to our chair factory example, suppose that we express the capacity of each of the five activities in terms of the number of units per day that an individual worker can process. If some of the workstations have more than one worker, then the total capacity of the workstation would simply be the number of workers multiplied by the capacity per worker. (Note that there are some obvious simplifying assumptions here - this is not meant to be a scientific calculation by any means, it is simply meant for providing context.) This is illustrated in the table below.

Activity (Workstation)	Capacity per worker	Number of workers	Workstation Capacity
Cutting	60 units/day	2	120 units/day
Drilling	45 units/day	3	135 units/day
Sanding	22 units/day	5	110 units/day
Finishing	200 units/day	1	200 units/day
Assembly	50 units/day	3	150 units/day

Table 1 – workstation capacities for the chair factory

It is important at this point to discuss the capacity of the *system* (the entire chair factory). Given that the sanding workstation is able to process no more than 110 units/day, it will not be possible for the factory to produce more than 110 chairs/day, since all chairs must go through the sanding workstation at some point. Thus, we see that the **system capacity** will be determined by the **system constraint (bottleneck)**, similar to how the volume that can flow through a pipe will be determined by the narrowest point (see figure two). We will discuss system capacity and bottlenecks in much more detail in section 2.3.

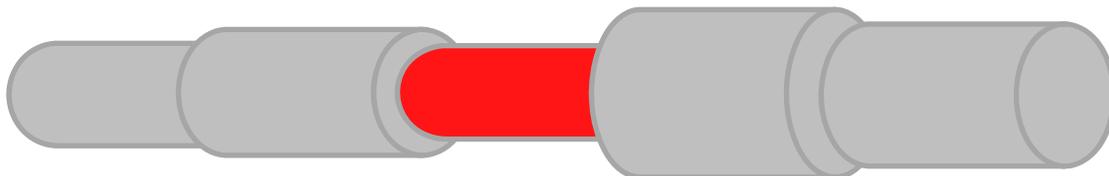


Figure 2 – The output capacity of the pipe will be determined by the narrowest point

## 2. The Theory of Constraints (TOC)

### 2.1 Origin

The Theory of Constraints (TOC) was created by Eli Goldratt in the mid-1980s and became most-recognized by a novel titled *The Goal* which uses an effective analogy (a children’s hike), a fictitious factory and management team (led by Plant Manager Alex Rogo) and a professor/management guru (Jonah) to illustrate key points of TOC (Goldratt & Cox, 1986). Subsequent books by Goldratt as well as a movie version of the book followed. In simple terms, TOC purports that *every system must have at least one constraint*, and further that *constraints provide an opportunity for improvement* (Rahman, 1998). Goldratt himself describes TOC as “an overall theory for running an organisation” (Rahman, 1998). Dettmer (1997) characterizes TOC as “a collection of...principles and tools or methods for improving overall system performance.” Indeed the interpretations, representations, and applications of TOC are vast; for example, in a later book Goldratt weaves many business and organizational dimensions into TOC including organizational psychology and encouraging change (Goldratt, 1990). For the purpose of understanding TOC within the scope of an introductory operations management course, we will focus on only a few, but key representations and applications, beginning with what Goldratt refers to as the “five steps of focusing” (1990). These steps provide an effective and somewhat concrete framework for focusing system improvement efforts, and are a useful articulation tool.

### 2.2 TOC as “The Five Steps of Focusing” (with application to the chair factory)

Recall our chair factory example from section 1, summarized in an abbreviated version, below:

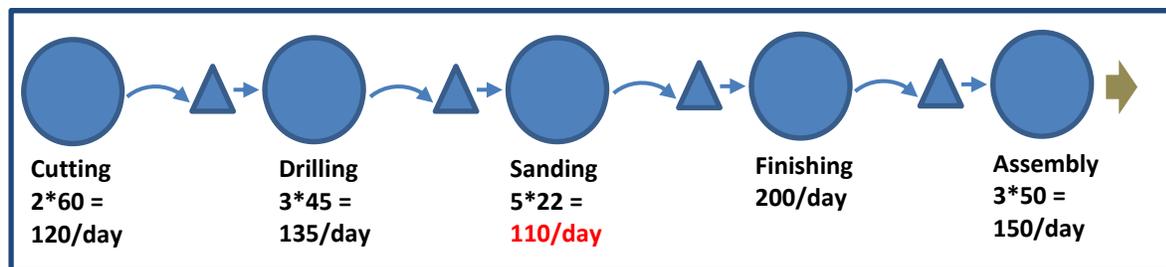


Figure 3 – summary of chair factory capacities.

As previously mentioned, the maximum output (capacity) of the *system* (factory) will be 110 units/day. This is because the production process requires that all five activities be performed, in sequence, and thus the factory will not be able to produce more than the workstation with the lowest capacity – sanding, in this case – which we will call the **bottleneck** or the system constraint. This brings us to the first focusing step...

**TOC Step One: Identify the Bottleneck** – “the bottleneck resource is that for which the demand upon it exceeds its’ capacity; the bottleneck determines the output of the entire system” (Goldratt & Jackson, 2007). In this case, where we have a simple linear process, we found the bottleneck by identifying the workstation with the lowest capacity. (We may not always have a simple linear process like this one, where all materials must pass through all activities, but for now the point is to understand that system capacity is determined by bottleneck capacity).

**TOC Step Two: Exploit the Bottleneck** – now imagine that the sanding activity is located on the opposite side of the factory as the drilling activity, and thus materials are batched and moved in a cart from the drilling activity to the sanding activity. Depending on how often batches are moved, it is possible that at times the sanding activity will be without material to work on (i.e., while waiting for a cart of materials from drilling). Since we know that the output of the system depends on the output of the bottleneck (sanding), we don't ever want to "starve" the bottleneck by not having WIP available to it when needed. Any idle time on the bottleneck is lost *system* capacity. *Exploiting* the bottleneck refers to ensuring that *all* of the existing capacity is realized. One example of how to prevent idle time is to ensure that a "buffer" of WIP inventory is always available to the bottleneck (this is often easier said than done given inherent output variability in most process activities). There are numerous other means of exploiting the bottleneck, some of which are illustrated in the book and movie versions of *The Goal*, such as staggering worker breaks so that the bottleneck does not sit idle.

**TOC Step Three: Subordinate to the Bottleneck** – consider what would happen if all activities attempted to function at their full capacity. Clearly the output of the activities that are "upstream" of the bottleneck (i.e., cutting and drilling) would pass on more work than the bottleneck could handle, which would result in a constantly growing pile of work-in-progress (WIP) inventory immediately in front of (yet to be worked on by) sanding. In general, *excessive* and unnecessary inventory is not a good thing, as it requires working capital to fund it, it must be stored, and there are risks such as damage, etc. Thus, it would be ill-advised for the activities upstream of the bottleneck to function continuously at their full capacity; in fact, they should try to run at the same rate of output as (i.e. *subordinate* to the pace of) the bottleneck. Note, in this case, that this would result in utilization of less than 100% for some activities - all except for the bottleneck, in fact. This might be counter-intuitive for some managers - why not keep producing if there is capacity? Note also that the two activities that are upstream of the bottleneck (cutting and drilling) will need to *deliberately* work at a utilization of less than 100%. The two activities that are downstream of the bottleneck won't have a choice, however, since they receive WIP at the pace in which it is passed to them by the bottleneck. (Understanding system flows and how activities affect each other is an important to applying TOC.)

There is a subtle contradiction between steps two and three – exploiting requires that there is always some WIP in front of the bottleneck, while subordinating to the bottleneck requires that upstream activities slow their pace to the bottleneck's pace. Indeed the presence of variability will require that WIP in front of the bottleneck be monitored – too little risks starving the bottleneck, and thus upstream activities should be sure to continue to produce, while too much WIP can be expensive, and upstream activities should potentially slow down.

**TOC Step Four: Elevate the Bottleneck** – if we have ensured that every bit of bottleneck capacity is being used (i.e. we have *exploited* the bottleneck), then the only way that the capacity of the *system* can be increased is by *elevating* (increasing) the capacity of the *bottleneck*. In our example, this could be done, for example, by fundamentally changing the sanding process (e.g. by automating it in some way) or by adding another machine/worker. We are essentially removing a system constraint when we do this. For example, if a worker is added to the sanding activity, the capacity of that activity will become 132 units/day, and it will no longer be a system bottleneck. Note, though, that although the capacity of

sanding has increased to 132 units/day, the capacity of the *system* won't increase to 132 units per day. That is because there will be a new bottleneck – cutting – and the capacity of the system will depend on the capacity of this activity, which is 120 units/day (see figure 4).

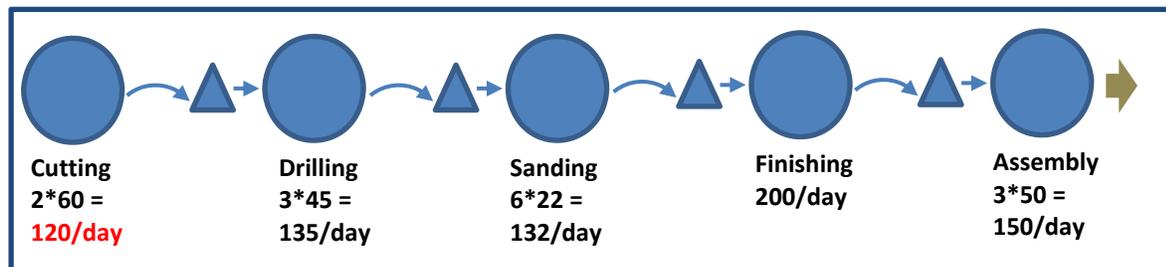


Figure 4 – Updated capacities

**TOC Step Five: Repeat** – now we are back where we started – we have identified the new system bottleneck (cutting) and we should repeat all of the previously described “steps” of the Theory of Constraints to subordinate to, exploit, and elevate the new bottleneck. Our final step is thus to go back to step one! The Theory of Constraints is an *on-going process* (or perhaps we could describe it as a philosophy, or a set of tools for managing flows and capacity, or a way of thinking); as capacity is increased and new bottlenecks (system constraints) emerge, the Theory of Constraints steps need to be repeatedly applied.

(Note on sequencing of steps – although bottleneck identification would generally need to be performed first, and elevation should not occur until the bottleneck has been thoroughly exploited, subordination to and exploiting of the bottleneck technically don't have to be completed in a specific sequence – examples would exist where either of them are done before the other).

**The Movie** - it is advised that students watch the entire movie, for context; however, the application of *the five focusing steps* is illustrated by a hiking trip (16:00 – 22:00) and at the factory (22:00 – 34:00).

## 2.3 Other contributions and representations of TOC

Please note that only a few are presented here; those interested in the full contributions and representations of TOC are encouraged to consult the reference list at the end of this document.

### 2.3.1 Throughput, Inventory and Operating Expenses

Consider the chair factory presented previously, where cutting was the bottleneck. If there was an opportunity to improve the drilling activity through automation (e.g. installing robotics), would it be worthwhile to do so? In terms of *system capacity*, no it wouldn't, since drilling is not the system constraint and therefore increasing the capacity of the drilling activity would not increase the capacity (throughput, in terms of total sales less total cost of raw materials) of the system. However, this doesn't mean that it wouldn't still make sense to automate the drilling function, if doing so resulted in enough of a decrease in operating expenses (e.g. through reduced labour requirements) or inventory (e.g. by being able to produce on demand, thus reducing the need for buffers of WIP). In essence, this is a basic framework for “determining the effect that any local action has on progress toward the system's goal” (Dettmer, 1997).

**The Movie** – the first 16:00 minutes of the movie discusses the use of robots, but where the robots did not in fact increase system capacity and were actually being used in a way that was detrimental to the factory because they ran the robots at capacity (in order to reduce the cost per part by creating economies of scale), but since the bottleneck could not keep up it resulted in excessive (and expensive) WIP.

### 2.3.1 Drum, Buffer, Rope (DBR)

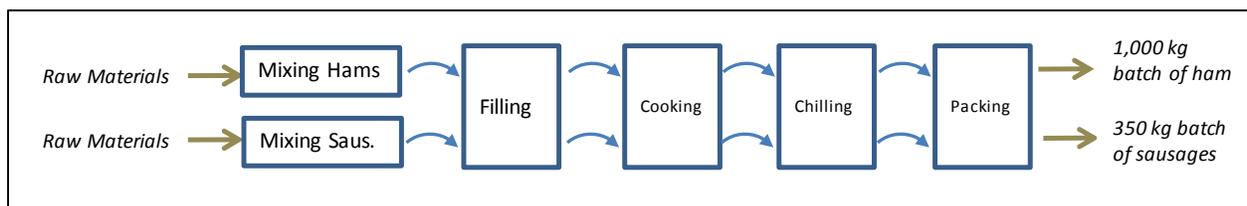
DBR is an operations scheduling methodology that is used to create structure and discipline in implementing TOC in order to maximize the effectiveness of the system (Goldratt UK). It can be summarized as follows:

The bottleneck sets a **drumbeat** that the factory marches to (e.g. through the production schedule). This provides a way to ensure that the bottleneck always has work to do (exploited), by ensuring that there is a **buffer** (i.e. some WIP inventory immediately in front of the bottleneck) to prevent idle time in the case of variability and/or unexpected events (machine breakdowns, delays in supplies arriving, etc.). At the same time, if all materials are “tied” to the drumbeat (hence the **rope** analogy), then they will only be released through the system at appropriate times so as not to create excessive WIP in the system.

## 3. Application to “product mix” problems

### 3.1 Case Study

Mabin and Gibson (2013) present a case study based on a manufacturing company that produces a variety of convenience food products; in particular, “the case focuses on the production of manufacturing hams (used in commercial products and sandwiches) and pre-cooked sausages” (Mabin & Gibson, 2013). Production of each of these two product types requires processing time on five workstations, as per the diagram below. Note that each product has their own mixing workstation, but otherwise they compete for time on the other workstations.



**Figure 5 – production process for batches of hams and sausages**

The amount of time that each batch requires on each workstation is provided in the following table, as well as the capacity (total time available per week) on each workstation.

Workstation	Process Times (hrs/batch)		Total Process Time Available (hrs/week)
	Hams	Sausage	
Mixing (Ham)	5	0	40
Mixing (Sausage)	0	0.5	40
Filling	8	1	40
Cooking	8	1	160
Chilling	10	2	200
Packing	1	7	160

**Table 2 – process times (hrs per batch) for each product type on each workstation**

Consider now that weekly demand is forecast to be 8 batches/wk for hams and 20 batches/wk for sausages. We can calculate the total amount of processing time that would be required at each workstation in order to meet all of demand for both product types; we will call this the **aggregate workload**.

### 3.2 Aggregate workload

The following table shows the aggregate workload for each workstation, given processing times and demand. For example, we could say that “in order to meet weekly demand of 8 batches of hams and 20 batches of sausage, given that each batch of hams needs 8 hours at the cooking workstation and each batch of sausages needs 1 hour at the cooking workstation, we would need 84 hours of processing time at the cooking station.”

Workstation	Process Times (hrs/batch)		Total Process Time Required (hrs/wk)	Total Process Time Available (hrs/week)
	Hams	Sausage		
Mixing (Ham)	5	0	= $8*5 + 20*0 = 40$	40
Mixing (Sausage)	0	0.5	= $8*0 + 20*0.5 = 10$	40
Filling	8	1	= $8*8 + 20*1 = 84$	40
Cooking	8	1	= $8*8 + 20*1 = 84$	160
Chilling	10	2	= $8*10 + 20*2 = 120$	200
Packing	1	7	= $8*1 + 20*7 = 148$	160
Demand (batches/wk)	8	20		

**Table 3 – aggregate workload (total processing time required to meet all of demand) for each workstation**

Comparing the aggregate workload (total processing time *required*) to the capacity (total processing time *available*) for each workstation may reveal a **bottleneck/constraint**; in the above example, there is not enough capacity (40 hours/week) at the filling workstation to meet total required processing time (84 hours/week, based on demand). This constraint motivates the fundamental business problem of the product mix problem – we need to determine how many batches of each product type to produce in light of our system constraint (i.e. since we can’t meet all of demand of both products). Assuming that the objective is to maximize total profit, and further assuming that only one bottleneck exists, there is a relatively simple methodology for determining this optimal “product mix”.

### 3.3 Trade-off: profit versus consumption of constrained resource

Assume that we have been told by the ham and sausage manufacturer that the profit margin on a batch of hams is six times that of a batch of sausages. This might lead us to jump to the conclusion that the

optimal product mix would include as many batches of hams as can be made (up to the demand amount). However, closer inspection reveals that a batch of hams requires eight times as much processing time on the filling workstation (our constraint) as a batch of sausages requires. Thus, even though ham batches are six times more profitable than sausage batches, we would be able to make eight times more sausage batches using the same amount of the constrained resource (time at the filling workstation). Thus, and perhaps contrary to initial intuition, the product mix that maximizes profit would include as many batches of sausages as can be made (up to the demand amount). In other words, we should prioritize for inclusion in the product mix according to the **profit per unit of consumption of the constrained resource** (see footnote<sup>2</sup>). This can be computed as follows:

Hams= (\$6 profit/batch)/(8 hours on the filling workstation/batch) = **\$0.75** profit per hour of filling workstation time.

Sausages= (\$1 profit/batch)/(1 hour on the filling workstation/batch) = **\$1.00** profit per hour of filling workstation time.

Note that processing time on other (non-bottleneck) workstations are not factors in the product mix decision at hand, since they are not system constraints.

### 3.4 Finding the optimal solution

To find the optimal product mix, start with the product that has the highest profit per unit of consumption on the constrained resource, make as many as possible (i.e., until demand is met), and if capacity (on the constrained resource) remains, continue on to the next most profitable product until capacity is depleted. In the example provided above,

*Sausages give the highest profit per hour on the filling workstation. Since each batch of sausages requires 1 hour of time on the filling workstation, and weekly demand is 20 batches, we have enough capacity (40 hours) to meet all 20 batches of demand for sausages. After allocating (20 batches)\*(1 hours/batch) = 20 hours of time on the filling workstation, we have 40 – 20 = 20 hours of capacity remaining.*

*Each batch of the next product (hams) requires 8 hours of time on the filling workstation, and weekly demand for hams is 8 batches, thus 8\*8 = 64 hours would be required to meet all of demand. Since we only have 20 hours remaining, we will only be able to produce 20/8 = 2.5 batches of hams (assuming that partial batches are possible). Thus:*

*The **optimal product mix is to make 20 batches of sausages and 2.5 batches of hams** for total profit of  $20*\$1 + 2.5*\$6 = \$35$ .*

*(Note that had we considered only the profit per unit, without regard for the amount of the constrained resource needed, we would have prioritized ham first and therefore made 40/8 = 5 batches of hams and zero batches of sausages for total profit of 5\*\$6 = \$30. This is referred to as the “traditional approach”, since it considers only profit per unit and not the consumption of the constrained resource, thus possibly providing a sub-optimal solution. The methodology that we are using is referred to as the “TOC approach” or the “bottleneck approach”. Example two in the next section will explain this further).*

---

<sup>2</sup> Profit may be expressed as **contribution margin**, or, in the case of Blackstone (2001), **throughput**

For larger size problems (more products and/or resources), the procedure for determining the optimal product mix is the same; please see the following section for a problem and the Appendix for a flow chart to help understand the methodology.

### 3.5 Example two

The following example is based on the work of Blackstone (2001).

*Consider a hypothetical company that produces three products – X, Y, and Z. X sells for \$90 and has weekly demand of 50 units; Y sells for \$100 and has weekly demand of 75 units; Z sells for \$70 and has weekly demand of 100. Each unit of each product may require processing time on one of five workstations (A, B, C, D and E) and will require raw materials. Each workstation has 40 hours (2,400 minutes) of capacity available per week. This information is summarized in tables three and four.*

		PRODUCT		
		X	Y	Z
WORKSTATION	A	20	10	5
	B	0	10	10
	C	15	5	5
	D	15	25	10
	E	5	10	5

**Table 3 - Capacity (minutes) required per unit produced**

	X	Y	Z
Demand	50	75	100

**Table 4 - Weekly demand per unit**

As a first step (see Appendix One), we will determine the aggregate workload of each workstation and compare the results to capacity to see if it is possible to meet all of demand, and if not then to determine the workstation(s) that is the bottleneck(s). For example, to find the aggregate workload on workstation A, we would take the demand for product X (50/week) and multiply it by the number of units that each unit of product X requires on workstation A (20 units) to get the total amount of capacity per week that product X requires of workstation A. Doing the same for all products on all workstations and summing across the three products for each workstation will give the results in table five.

		X	Y	Z	Total
WORKSTATION	A	1,000	750	500	2,250
	B	-	750	1,000	1,750
	C	750	375	500	1,625
	D	750	1,875	1,000	3,625
	E	250	750	500	1,500

**Table 5 - Capacity (minutes) required per week**

Given that each workstation has 2,400 minutes of capacity, we see that workstation D is a bottleneck – there is not enough capacity on workstation D to meet all of demand for all three products. Thus, we need to determine the optimal (profit-maximizing) product mix – how much of each of products X, Y and Z to produce.

Table six provides profit information for each product – expressed by Blackstone for the purpose of the product mix problem as “throughput/unit”, which is simply the selling price less raw materials costs; (we will continue to use the term “profit” in the following). Please note that this implies that labour and overhead costs are treated as fixed; we will leave a more detailed discussion of this point for cost accounting classes, but for those interested please see Blackstone and/or Perkins et al. for more.

	<u>X</u>	<u>Y</u>	<u>Z</u>
<b>Selling Price</b>	\$ 90	\$ 100	\$ 70
<b>Raw Materials</b>	\$ 40	\$ 30	\$ 25
<b>Profit/unit</b>	\$ 50	\$ 70	\$ 45

**Table 6 - Product price and raw materials costs**

As mentioned in the previous section, a “traditional approach” would simply use the profit/unit to prioritize inclusion in the product mix. If the profit/unit were used for this, product Y would be considered most profitable. However, referring back to table three, we see that product Y is also the biggest user of time on workstation D, our constrained resource. In order to accurately evaluate the profitability of each product, we need to calculate the *profit per constraint minute*. This is done in table 7. Once again note that we are only concerned with the time required on workstation D, since the other four workstations have enough processing time to meet all of demand and therefore are not a constraint in our product mix decision (they are moot.)

	<u>X</u>	<u>Y</u>	<u>Z</u>
<b>Profit/unit</b>	\$ 50	\$ 70	\$ 45
<b>Minutes of D per unit</b>	15	25	10
<b>Profit/minute on D</b>	\$ 3.33	\$ 2.80	\$ 4.50

**Table 7 - Profit per minute on workstation D**

The results show that product Z is the most preferred product for production, followed by product X. We can now use the processing requirements and demand information in tables three and four to determine the optimal product mix. (Please see appendix one if you require further explanation of the methodology.) Blackstone provides the following description of the optimal solution:

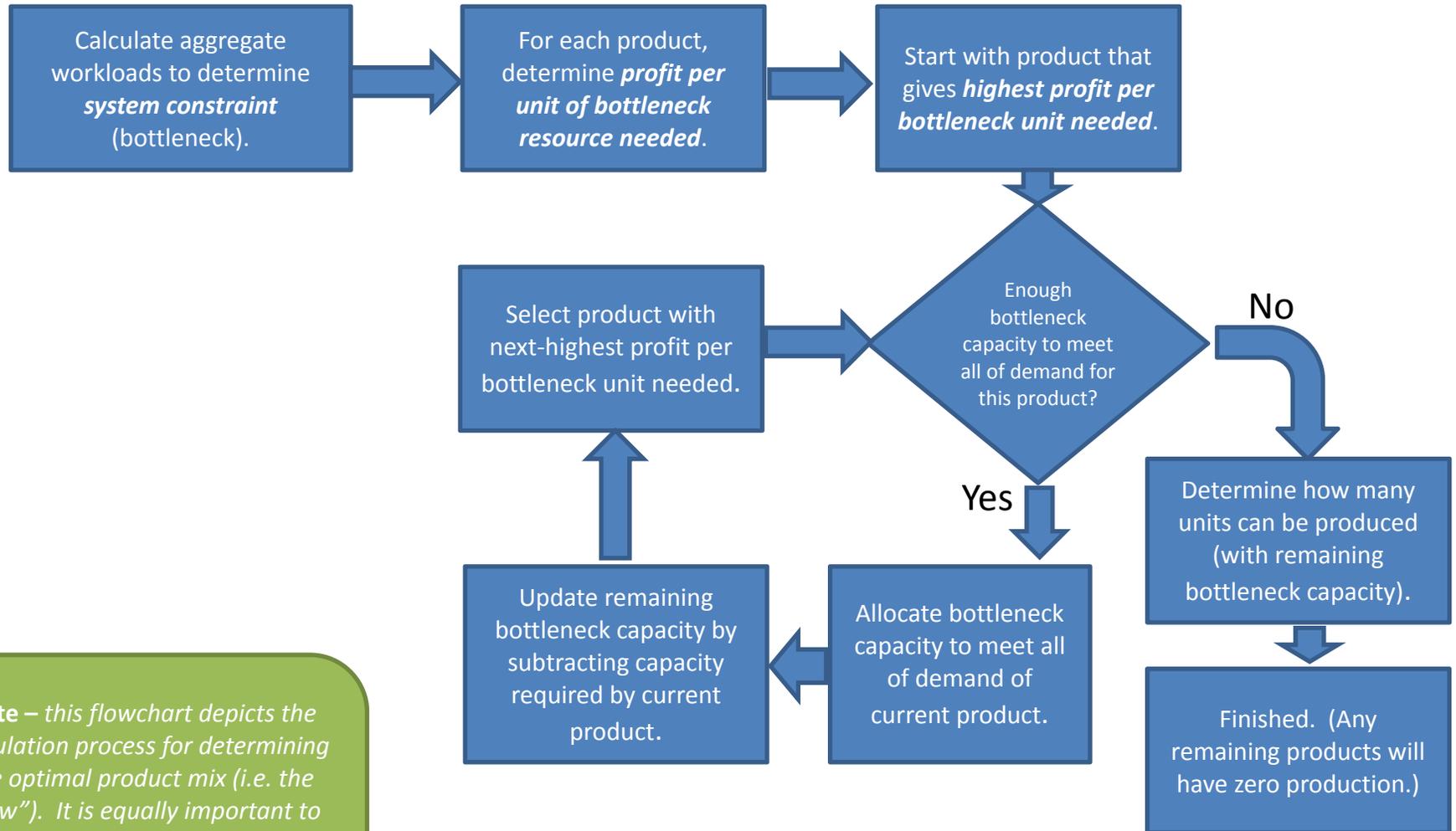
*Making 100 Zs requires 1,000 minutes [100\*10] at Station D. Making 50 Xs requires 750 more minutes [50\*15]. There are 650 minutes remaining [2,400 – 1,750] of the 2,400 minutes available to Station D; these 650 minutes can be used to make 26 Ys at 25 minutes each.*

Thus, the optimal (profit-maximizing) product mix is to make 100 of Z, 50 of X, and 26 of Y.

### **3.6 Multiple bottlenecks**

If more than one resource is a bottleneck (i.e., total required is less than total available), then the methodology described above will not be usable. In the case of multiple bottlenecks, a Linear Programming model is required, and can be implemented using an Excel Spreadsheet and the 'Solver' tool. The specific methodology is beyond the scope of this document, but in general terms the objective is to maximize profit by determining how much of each product type to produce, subject to the constraints that production amounts don't exceed demand (for each individual product) and the total amount required of each bottleneck resource does not exceed the capacity. Perkins, Stewart and Stovall (2002) provide a paper on this topic for those interested.

## Appendix One – Process chart for single-bottleneck product mix problems



**Note** – this flowchart depicts the calculation process for determining the optimal product mix (i.e. the “how”). It is equally important to understand the business problem and the general context around product mix problems (i.e., the “what” and the “why”.)

## References

- Blackstone, J.H., (2001). Theory of constraints: a status report. *International Journal of Production Research*, 39 (6), pp. 1053-1080.
- CBC News (January 21, 2014). *Farmers frustrated as grain stuck in railway bottleneck*. Retrieved from <http://www.cbc.ca/news/canada/calgary/farmers-frustrated-as-grain-stuck-in-railway-bottleneck-1.2504820>.
- Dettmer, H.W. (1997). *Goldratt's Theory of Constraints: A systems approach to continuous improvement*. Milwaukee, WI: Quality Press, American Society for Quality.
- Goldratt, E. M. and Cox, J. (1986). *The Goal: A Process of Ongoing Improvement*. New York: North River Press.
- Goldratt, E.M. (1990). *What is this thing called Theory of Constraints and how should it be implemented?* Great Barrington, MA: North River Press.
- Goldratt, E.M., Jackson, D., Axotis, J. (2007; originally produced in 1995). *The goal [videorecording]: the how-to version*. Virginia Beach, VA : Coastal Training Technologies Corp. [distributor].
- Goldratt UK (n.d.). *The Theory of Constraints and Drum-Buffer-Rope*. Retrieved May 14, 2015 from [http://www.goldratt.co.uk/resources/drum\\_buffer\\_rope/](http://www.goldratt.co.uk/resources/drum_buffer_rope/)
- Mabin, V.J., Gibson, J., (1998). Synergies from Spreadsheet LP used with the Theory of Constraints - A Case Study. *The Journal of the Operational Research Society*, 49 (9), pp. 918 – 927.
- Perkins, D., Steward, J., & Stovall, S., (2002). Using Excel, TOC, and ABC to Solve Product Mix Decisions with More Than One Constraint. *Management Accounting Quarterly*, 3 (3), pp. 1-10.
- Rahman, Shams-ur (1998). "Theory of constraints: A review of the philosophy and its applications". *International Journal of Operations & Production Management*, 18 (4), pp. 336 – 355.