
Anantaram Balakrishnan
and
Stuart Brown

WP# 3432-92-MSA May, 1992
Process Planning for Metal Forming Operations:
An Integrated Engineering-Operations Perspective

Anantaram Balakrishnan
and
Stuart Brown

WP# 3432-92-MSA May, 1992
Process Planning for Metal Forming Operations: An Integrated Engineering–Operations Perspective †

Anantaram Balakrishnan
Sloan School of Management
M. I. T.

Stuart Brown
Department of Materials Science and Engineering
M. I. T.

May 1992

† Supported by research grants from MIT’s Leaders for Manufacturing program and Aluminum Company of America
Abstract

Metal forming operations such as rolling, extrusion, and drawing offer many opportunities for operations improvement through better process understanding and improved planning practices. This paper addresses short and medium term planning issues in sheet, plate, and tube manufacturing operations. First, we identify certain distinctive characteristics—the inherent process flexibility, close interdependence between successive stages, and economies of scale—of metal forming operations, and identify the planning and performance tradeoffs. We argue that, just as product design plays a critical role in manufacturing of discrete parts, process planning has strategic importance in the metal forming context. To be successful, process planning must be closely coupled with process engineering efforts, and must simultaneously consider the facility's entire product mix. In contrast, current process engineering efforts are mainly reactive, focusing on fixing problems at individual operations, ignoring the interactions between successive stages. Similarly, planning activities are incremental, considering only individual products or orders one at a time rather than the entire range of product sizes to be manufactured. By working together, planners and engineers can develop principled practices and process plans, analyze the sensitivity of production performance to current process constraints, and adopt a proactive process improvement strategy that focuses on critical constraints. We illustrate these concepts using examples from aluminum tube and sheet manufacturing. We present an integrative process engineering–planning framework, and identify several interdisciplinary research opportunities spanning management science, materials science, and mechanical engineering.

Keywords: Manufacturing, process planning, process improvement, metal forming operations, process modeling.
1. Introduction

The metal forming industry, including rolling facilities, extrusion plants, and tube and wire drawing facilities, plays an important and critical role in global manufacturing competitiveness. This industry supplies plates and sheets, extrusions, and tubes to most major manufacturing enterprises including the automobile, aircraft, housing, and food service and beverage industries. These user industries often perform additional operations such as stamping, drilling, machining, welding and other finishing operations before assembling the metal components in finished products. In 1987, shipments of aluminum alone to United States' markets exceeded 15 billion pounds, valued at over $12 billion. (Aluminum Statistical Review, 1987 and Minerals Year Book, 1988). The containers and packaging market segment accounted for approximately 30% of the total usage, while building, construction and transportation consumed 20% each. Over 50% of the shipments were sheet, plate and foil products, with extrusions and tube constituting another 17%.

This paper addresses planning and process improvement issues in metal forming operations. Although metal forming has a broad interpretation, for the purposes of our discussions we exclude "discrete" operations such as stamping, forging, cutting, drilling, welding, and machining operations. Thus, we are concerned mainly with operations such as hot and cold rolling, extrusion, and drawing.

The metal forming industry is characterized by large investments in plant and equipment, a wide range of product offerings, the strategic importance of process technology, and universal standards for specifying, measuring, and testing product quality. Because of the significant economies of scale, installing machines with large capacities, maintaining high levels of utilization for prime equipment, and improving process yield or recovery are important strategic objectives for the industry. In turn, these objectives of improving recovery and utilization have led metal forming companies to continually upgrade their equipment to handle larger ingot and lot sizes. In contrast, customers are placing smaller, more frequent orders as they move towards just-in-time procurement and production. The industry also faces increasing pressures to improve quality, reduce cost and lead times, and meet more stringent specifications. These trends make the planning and process engineering functions critical for competitive survival.
Despite the metal forming industry's distinctive characteristics and considerable economic importance, the manufacturing and management science literature does not adequately emphasize planning and process improvement models that are tailored to this industry; in contrast, the literature on planning, scheduling, process improvement, and product design for discrete parts manufacturing and assembly operations is quite extensive. Likewise, the materials science and mechanical engineering literature focuses on studying material properties and understanding individual processing steps from a process development rather than a manufacturability or planning perspective. This paper attempts to highlight and explore some of these issues at the interface between engineering and management science using two specific examples—aluminum tube manufacturing, and sheet and plate rolling—derived from our experience in the aluminum industry. The discussions are aimed at a broad audience including management science researchers concerned with modeling manufacturing operations and developing process planning systems, engineering researchers dealing with deformation processes, and practicing managers in the metal forming industry. Specifically, the paper offers the following contributions and insights:

• We argue that, just as product design has considerable impact over the manufacturability of discrete parts (see, for example, Nevins and Whitney [1990], Clark and Fujimoto [1991]), process planning has strategic importance for metal forming operations. This importance stems from the wide flexibility, but close coupling, between successive metal processing operations, combined with the significant impact that the choice of processing paths has over manufacturing performance. The current prevalent practice of treating process planning mainly as an operational function does not exploit its strategic potential.

• Exploiting the strategic potential of process planning requires the simultaneous consideration of multiple products rather than the product-by-product (or order-by-order) incremental planning procedure that both dominates the process planning literature (see, for example, Alting and Zhang [1989], and Chang [1990]) and is found commonly in practice. A systems view, incorporating multiple stages as well as multiple products, is essential to make appropriate strategic and tactical decisions regarding product grouping, the level of commonality, and capacity requirements.

• We also emphasize the close interrelationship between engineering models and planning models to respectively characterize the process constraints and
capabilities and choose effective processing paths. We propose an iterative framework in which the planning model uses as input empirical and model-based process constraints; in turn, the sensitivity of manufacturing performance to these constraints prioritizes engineering process improvement efforts. This framework represents a departure from conventional process modeling efforts that focus on troubleshooting and expanding process capability without explicit linkages to the planning function. The engineering–planning linkage that we describe has implications both for collaborative, interdisciplinary research within universities and with industry, and for the design of organization structures, incentives, and multi-function coordination processes between engineering, planning, and manufacturing.

- Finally, we identify several new problems and issues for further research. These research issues span topics such as tactical operations modeling and production planning, inventory management, and deformation process modeling.

To summarize, this paper provides an overview of two metal forming operations, proposes an integrative process engineering and planning framework, and identifies new research issues based on our joint experience with several plants in the aluminum industry. We provide only a broad description of modeling alternatives, and do not discuss any specific analytical results, solution algorithms, or computational results.

Sections 4 and 5 develop and illustrate the main themes of this paper. Since this discussion requires familiarity with metal forming processes and operations, Sections 2 and 3 provide the necessary background. Section 2 introduces a generic two-stage process description that applies to both of our subsequent examples–tube manufacturing and rolling. We identify the common characteristics and tradeoffs underlying these two examples, and motivate the importance of the process planning function in the metal forming context. Sections 3 and 4 focus on tube manufacturing. Section 3 describes the engineering principles underlying tube manufacturing, and discusses typical current practices in process engineering and planning. Section 4 characterizes process flexibility in tube manufacturing, identifies the factors affecting extrusion and drawing workload, and develops an integrative framework that links engineering and planning. Section 5 outlines process planning issues related to sheet manufacturing (rolling), and Section 6 offers concluding remarks. The paper places much greater emphasis on tube manufacturing since this example
captures most of the principles that we wish to highlight. We discuss rolling operations only briefly in order to emphasize that the same underlying principles apply more broadly to other metal forming operations. Also, these principles impact both strategic and operational decisions. We therefore consider a medium-term planning problem for tube manufacturing, and a short-term decision for rolling operations.

2. Generic Description of Metal Forming Operations

This section presents a generic description of metal forming operations encompassing the two examples—aluminum tube manufacturing and rolling operations—that we discuss later. We will limit the discussion to the two important successive stages in the metal forming process, hot forming and cold forming, which we refer to as upstream and downstream operations.

A flat (sheet or plate) or tubular product is identified by its alloy, its temper and other mechanical or microstructural specifications, the physical dimensions of each piece and their variance, and geometric tolerances (e.g., flatness, eccentricity). For (hollow) tubular products, the physical dimensions are outer diameter, wall thickness, and tube length. Sheet or plate products are specified by their width, gauge (i.e., thickness), and length. Both of our subsequent examples are based on facilities that produce primarily to order, manufacturing several thousand different product specifications each year.

Continuous metal forming consists of deforming the shape of the raw material, e.g., rectangular or cylindrical ingots, into the desired final shape and dimensions. This transformation typically entails successively reducing the cross-sectional dimensions (for instance, tube diameter or sheet thickness) while elongating the workpiece. The transformation from ingot to the final product dimensions is achieved through a combination of hot forming and cold forming processes, possibly with intermediate annealing operations. The process might also include some preliminary steps such as cutting, drilling, or scalping the ingot and preheating it, and some finishing operations such as stretching, coating, slitting, and cutting. In tube manufacturing, the process of extrusion corresponds to hot forming, while tube drawing represents cold forming (see Figure 1). For sheet and plate manufacturing the two corresponding stages are hot and cold rolling.
All of our subsequent discussions focus on the interactions between the hot and cold forming stages, although the concepts that we discuss also extend to more than two stages (including prior stages such as ingot casting, and subsequent finishing operations). An intermediate inventory stocking point might decouple these two stages. Each stage can have multiple steps and parallel workcenters, with possible reentrant flows (multiple passes at the same workcenter) in the downstream operation. We will refer to the intermediate product produced by the upstream (hot forming) stage as semi-finished stock.

Hot and cold forming processes share some similarities, but also have important differences. Hot forming, as the name implies, consists of deforming the metal at an elevated temperature, while cold forming processes the metal at or near room temperature. Consequently, hot forming is more "efficient" than cold working, i.e., it permits greater amount of deformation per unit input of energy. However, since controlling the process at higher temperatures is more difficult, hot forming cannot achieve very tight dimensional tolerances, and is also limited in terms of the smallest size (gauge, outer diameter, wall thickness) that it can produce. Furthermore, cold working can introduce some desirable material properties such as strength and uniformity.

The process plan for a product is the "recipe" specifying its entire processing path. This specification includes:

(i) the alloy and size of the ingot to be used,
(ii) the sequence of processing steps,
(iii) the type of equipment required and the processing parameters (machine setups, processing speeds, special operating instructions) at each step, and
(iv) the intermediate and final product dimensions and metallurgical specifications.

Equipment limitations, workpiece characteristics and the underlying physics of the deformation process together impose upper limits on the amount of deformation that can be achieved in each hot or cold forming step. The limits differ substantially for hot and cold forming, and depend both on equipment and product specifications. In general, the number of processing steps increases as the differential between the geometries of the initial workpiece and final product increases. For our examples, we will express this differential as the difference between the cross-sectional areas of the workpiece and the finished product.

Let us now describe three important common features that characterize our two metal forming examples.
• Metal forming operations permit a wide range of flexibility, i.e., each finished product has numerous alternate process plans. In particular, for every product, the process planner can choose from a continuum of intermediate product dimensions (subject to certain process constraints) at each processing step. In Section 4, we provide a novel characterization of processing flexibility for tube drawing operations.

• The upstream and downstream stages are highly interdependent, and the choice of processing paths has a significant impact on the processing effort required at each stage of the manufacturing process. The selected process plan for each product determines both the total workload to deform the ingot to the required final dimensions, and the relative allocation of this workload between the upstream and downstream stages.

• Due to the significant equipment setup and changeover times and the fixed process scrap requirements, upstream operations strongly favor large lot sizes. Changing over from one product size to another requires changing the rolls, dies, and other tooling, as well as preheating ingots, and processing test runs (to stabilize and debug the process). The second factor contributing to the scale economies is the fixed scrap for each lot or ingot. For instance, a certain fixed length (independent of the ingot size) of the leading and trailing ends of a hot-rolled sheet must be scrapped because it does not meet material properties and dimensional specifications.

These three characteristics – processing flexibility, interdependence between successive stages, and economies of scale – create opportunities to improve the plant's effectiveness through principled process planning. In particular, we can exploit the process flexibility to achieve economies of scale in the upstream operation by limiting semi-finished stock to a few standard sizes even though the number of finished product sizes is very large. This standardization of sizes enables us to either produce to stock in the upstream operation (if the facility maintains semi-finished inventories) or consolidate multiple customer orders into a single lot at upstream operations. Both these strategies increase the upstream lot sizes, thus reducing the number of setups, increasing recovery, and lowering the unit cost of production. If the upstream operation produces to stock, limiting the number of standard sizes has the added advantage of reducing the safety stock levels due to greater commonality.

Observe that we have generalized the conventional notion of commonality. In discrete parts manufacturing, commonality refers to shared raw materials or
components, i.e., overlaps in the bills of material of two or more products (see, for example, Baker et al. [1986], and Gerchak and Henig [1989]). In the metal forming context, however, commonality refers to a shared set of processing steps from the first (ingot casting) stage to some intermediate hot or cold forming step. The degree of commonality between two products depends on the extent to which their process plans overlap. The inherent flexibility in metal forming operations permits us to introduce this commonality. However, selecting a very high degree of commonality might potentially increase the downstream effort significantly (since many different finished sizes must be produced from a common semi-finished size). This tradeoff between commonality and balanced workloads will be the main theme of our discussions in Sections 4 and 5.

One implication of this tradeoff is that, contrary to the prevailing practice of product-by-product process planning, we must simultaneously plan the processing steps for multiple products. We will argue that resolving the tradeoff effectively requires not only a principled planning model that judiciously selects process plans for multiple products, but also a good understanding of the processing constraints. Next, we describe the specific processing steps and constraints in tube manufacturing.

3. Tube Manufacturing: Background and Current Practice

This section specializes the previous generic two-stage process representation to tube manufacturing operations. Section 3.1 outlines the process flow and engineering principles underlying tube drawing and extrusion, and Section 3.2 describes current process planning practices and typical process engineering concerns.

The process of manufacturing hollow, seamless cylindrical tubes consists of an extrusion (hot forming) step followed by one or more tube drawing (cold forming) passes, possibly with intermediate annealing operations (see Figure 1). Each step of the process reduces the cross-sectional dimensions – outer diameter (OD) and wall thickness (WT) – of the workpiece, and increases its length. As we noted in Section 2, since extrusion is a hot process, it cannot achieve tight dimensional tolerances and is limited in terms of the minimum possible outer diameter and wall thickness. Subsequent tube drawing operations are, therefore, required to further reduce the cross-sectional dimensions, meet stringent quality standards, and achieve desired material properties. We refer to seamless extruded tubes produced by the extrusion press as blooms, and to drawn tubes,
i.e., blooms that have undergone one or more tube drawing passes, as tubes. Thus, blooms correspond to the semi-finished stock in our previous two-stage representation. Using simple schematics and a two-dimensional process plan representation, the next section describes the engineering principles and process limitations of extrusion and tube drawing. Subsequent sections then indicate how these constraints both interact with and are included directly within process planning.

### 3.1 Tube Drawing and Extrusion: Engineering Principles

We will concentrate on the tube drawing process, discussing extrusion only briefly at the end to indicate the types of processing constraints it imposes.

#### 3.1.1 Tube Drawing

Tube drawing involves pulling a tube, at room temperature, through a stationary die with an annular orifice that has a smaller cross sectional area than the tube. The tube consequently decreases in both wall thickness and diameter and increases in length. Repeated drawing steps therefore permits the manufacture of very small tubes from initially large bloom sizes. Normally, the annular space is formed by a inner mandrel and an outer die as shown in Figure 2. Figure 3 schematically illustrates various types of tube drawing depending on the relative sizes and location of the tube, die, and mandrel.

**Draw-bench operation**

Prior to its first drawing pass, each incoming bloom is crimped at one end, forming a "point". Each drawing pass starts by threading the tube over the mandrel, lubricating the inner and outer surfaces of the tube, and passing the crimped end of the tube through the opening of the die. A set of jaws on the other side of the die grabs the crimped end of the tube, and pulls the tube through the annulus formed by the mandrel and the die. The mandrel is held in place by a rod that extends past the opposite end of the tube. Notice that the diameter of the die and the mandrel respectively determine the outside and inside diameter of the drawn (output) tube.

**Graphical representation of tube drawing**

Before discussing the equipment and workpiece constraints that limit each tube drawing pass, let us first introduce a convenient graphical representation called the Tube Reduction Diagram. The Tube reduction diagram shows the changes in the tube's dimensions with each drawing pass. Since we are focusing
only on changes in the cross-sectional dimensions, we will consider a two-
dimensional version of the diagram as shown in Figure 4; the x-axis of this
diagram corresponds to the tube's outer diameter (OD), and the y-axis represents
its wall thickness (WT). A more comprehensive and accurate representation of
the tube drawing process would include other dimensions such as the tube's
length and eccentricity, as well as mechanical properties (e.g., hardness) that are
affected by the drawing operation. Tubes (and blooms) of varying sizes
correspond to distinct points on the tube reduction diagram. We visualize the
process plan for a product as a piecewise linear path from the bloom (starting
point) to the finished tube (target point), consisting of line segments connecting
the successive intermediate drawn tube sizes.

Observe that a process plan that successively reduces both OD and WT during
each drawing pass corresponds to a downward-sloping path from the north-east
bloom location to the south-west finished tube location. Also, for fixed bloom
and finished tube sizes, achieving greater reduction per draw corresponds to
permitting longer line segments, thus decreasing the number of drawing passes.
Conversely, if we limit the maximum length of each line segment, a bloom that
is farther away from the finished tube requires more drawing passes. Since the
total drawing effort (say, number of hours of drawbench setup and processing
time) varies directly with the number of drawing passes, we are interested in
constructing process plans that have fewer drawing passes (for a given set of
bloom sizes). Reducing the number of drawing passes also has the significant
added benefits of both eliminating materials handling steps which increases the
process yield, and decreasing lead times and inventories in the tube drawing
facility. Next, we will discuss the constraints that limit both the orientation and
the length of each line segment.

**Tube drawing constraints**

We wish to explore the limits of the drawing process so that we can later
"optimize" the process plans for all products while meeting the processing
limits. For tube drawing, these limits are determined by the following
considerations:

- **workpiece characteristics:** e.g., the tube must not break during the drawing
  process;
- **process yield and quality:** the drawn tube must have acceptable surface
  quality, must meet dimensional tolerances, and have the required
  metallurgical and mechanical properties;
- **equipment capabilities:** the equipment must sustain the load necessary to
  draw the tube, and tool wear must be controlled; and,
• **physical and operational considerations**: for example, the setup operations of threading the tube over the mandrel and through the die must be relatively easy.

To derive process constraints that satisfy these conditions, let us first understand the important physical principles underlying tube drawing, and describe the restrictions qualitatively. Later, we will show how to approximate and express these restrictions in terms of the dimensions of the input and drawn (output) tubes.

**Maximum reduction per draw**

First, drawing a tube requires a certain amount of force to pull the tube through the die. The amount of force depends on a number of factors including the required reduction in OD and WT, the deformation resistance of the input tube material, friction between the tube, die and mandrel, as well as processing conditions such as ambient temperature, lubrication, and tooling (for instance, die geometry). For a given set of processing conditions, the pulling force increases as the amount of required reduction increases, and is higher for input tubes that deform less easily. Higher pulling forces normally increase friction which in turn further increases the force necessary to pull the tube through the die-mandrel annulus. Beyond a certain limit, increasing the force might break the tube, degrade the surface quality, or cause excessive die wear. Furthermore, the maximum pulling force is also limited by the machine's power. To summarize, workpiece failure and equipment capabilities limit the maximum pulling force and hence the maximum amount of OD and WT reduction that can be achieved in a single drawing pass; the actual value of this limit depends on the deformation resistance of the input tube (which varies with the tube's processing history), its alloy, and the die-mandrel setup. We will refer to this restriction on the output tube's relative OD and WT as **max reduction/draw**.

**Work hardening**

The other important phenomenon to consider in tube drawing is **work-hardening**. As we reduce the cross-sectional area of a tube, the deformation within the die work-hardens the metal by increasing the average dislocation density. By definition, the amount of work-hardening corresponds to increasing the tube's resistance to deformation. The tube manufacturing process exploits this hardening while simultaneously limiting its cumulative effect. Work-hardening increases the strength of the metal, preventing the smaller output tube from immediately breaking as it pulls on the remainder of the tube during the drawing process. However, in a multi-draw process plan, where the output from one draw becomes the input to the next draw, the tube work-hardens
successively less with each draw, increasing the likelihood of breaking as it exits the die. To prevent tube breakage, we impose an upper limit on the total amount of deformation that a tube can experience before it must be annealed. Annealing consists of holding the tube at an elevated temperature for a specified time to resoften the metal, permitting additional work-hardening and hence additional draws to further reduce the cross-sectional dimensions. We refer to the restriction limiting the cumulative work-hardening before the tube must be annealed as the \textit{max work-hardening} constraint.

\textbf{Minimum cold work}

The third class of constraints stems from a product's temper specification. Some alloys require a certain minimum amount of work-hardening after the last annealing operation to assure adequate dislocation density so that the product achieves the desired microstructures during subsequent heat treatment. This requirement imposes a \textit{min final cold work} constraint.

\textbf{Sinking and ironing}

The \textit{max reduction/draw}, \textit{min final cold work}, and \textit{max work-hardening} constraints effectively restrict the \textit{length} of each line segment in the piecewise linear representation of a process plan on the tube reduction diagram. Other factors in tube drawing act to constrain the orientation (e.g., the angle) of these line segments. Notice that the orientation of the line segment connecting the input and output tubes of a drawing pass (in the tube reduction diagram) depends on the relative ratio of reduction in OD to reduction in WT. This \textit{OD-to-WT reduction ratio} impacts the process yield since it affects the surface quality and dimensional tolerances of the drawn tube. We will examine two extreme ratios corresponding to sinking and ironing operations.

A large diameter-to-wall reduction ratio corresponds to the \textit{sinking} operation. Figure 3a shows schematically a "pure" sinking operation that is unsupported by a mandrel. This mode of operation yields a large reduction in the outer diameter; the wall thickness may increase or stay the same depending on the state of stress that develops in the tube during drawing. Notice that sinking corresponds to a horizontal (or even upward sloping) processing path in the tube reduction diagram (Figure 4). Sinking operations can cause irregular surfaces on the inner tube diameter and are, therefore, usually avoided.

At the other extreme, the \textit{ironing} operation illustrated in Figure 3c has a relatively small OD-to-WT reduction ratio since it leaves the inner diameter unchanged while reducing the wall thickness. Ironing is difficult to achieve in
high volume production since it requires threading the tube's inner diameter over a mandrel with the same outer diameter. Difficulties with both sinking and ironing therefore impose lower and upper limits on the OD-to-WT reduction ratio for each drawing pass.

**Mathematical representation of constraints**

Our previous discussion identified and justified five different classes of tube drawing restrictions—max reduction/draw, max work-hardening, min final cold work, sinking, and ironing—due to equipment limitations and process yield or quality considerations. From a process planning perspective, we wish to translate these restrictions into limits on the length and the orientation of the line segments in the process plan's piecewise linear representation in the tube reduction diagram. Effectively, these limits specify the range of permissible input-to-output (dimensional) transformations during each draw, and between annealing steps. Using such limits, we can then easily check if a chosen plan is feasible, i.e., if it meets all five restrictions. To convert the original restrictions (e.g., max work-hardening) into equivalent mathematical constraints on the tube reduction diagram we will use certain approximations and surrogate metrics that are functions of the OD and WT of the input and output (or drawn) tube at each drawing pass, and the starting and ending dimensions between intermediate annealing operations. The subscripts in and out denote respectively the dimensions of the input and output tube for a given draw.

First, let us consider the max reduction/draw constraint. This restriction limits the drawing force and hence the amount of deformation in each drawing pass to prevent equipment and tube failures. We use the reduction in cross-sectional area (CSA) of the tube as a measure for the amount of deformation; the upper limit on CSA reduction is expressed as a proportion of the CSA of the input tube at each draw. Thus, the max reduction/draw constraint becomes:

\[
\frac{(CSA_{in} - CSA_{out})}{CSA_{in}} \leq \delta CSA_{\text{max}},
\]  

where the parameter \( \delta CSA_{\text{max}} \) has value between 0 and 1. Recall that the CSA of a hollow tube is \( \pi \times WT \times (OD-WT) \).

The amount of inelastic deformation, which dictates the amount of work-hardening, correlates approximately with the CSA reduction. Therefore, we can limit the amount work-hardening by specifying an upper limit on the total CSA reduction before an intermediate annealing operation is required. This
approximation gives the following max work-hardening constraint (again, we express the limit as a % of the CSA of the starting tube):

\[
\frac{\text{CSA}_s - \text{CSA}_e}{\text{CSA}_s} \leq \Delta \text{CSA}_{\text{max}},
\]

where \( \text{CSA}_s \) and \( \text{CSA}_e \) denote the tube's starting and ending cross-sectional areas between annealing steps. The parameter \( \Delta \text{CSA}_{\text{max}} \) has value between 0 and 1, and exceeds \( \delta \text{CSA}_{\text{max}} \).

The actual values of the limits \( \delta \text{CSA}_{\text{max}} \) and \( \Delta \text{CSA}_{\text{max}} \) in the max reduction/draw and max work-hardening constraints are determined by the process engineers through process understanding, experience and experimentation. They vary with the particular alloy being drawn. Certain alloys (such as the 6000 series of aluminum alloys) are designed specifically for large reductions per draw, and can withstand many draws before requiring an annealing treatment, while other alloys are significantly less easily drawn. The values of \( \delta \text{CSA}_{\text{max}} \) and \( \Delta \text{CSA}_{\text{max}} \) also depend on the equipment capabilities, the die geometry and setup, and processing conditions (e.g., lubrication, ambient temperature), as well as the dimensions of the input tube (e.g., "thin wall" tubes might permit only lower CSA reduction per draw) and its processing history.

Like the max work-hardening constraint, we can represent the min final cold work constraint in terms of CSA reduction as follows:

\[
\frac{\text{CSA}_s - \text{CSA}_f}{\text{CSA}_s} \geq \tau_{\text{min}},
\]

where the subscript \( s \) denotes the tube immediately after the last annealing step, and \( f \) represents the finished tube. The parameter \( \tau_{\text{min}} \) depends on the alloy and temper specification.

The sinking and ironing constraints limit the OD-to-WT reduction ratio; we will interpret and approximate them as lower and upper limits on the angle of the line segment connecting the output tube to the input tube (for each drawing pass) in the tube reduction diagram. We will refer to this angle as the drawing angle. Let us first consider the ironing constraint which specifies that the inner diameter or ID (\( = \text{OD} - 2 \text{WT} \)) of the output tube must be no greater than the ID of the input tube, i.e.,

\[
\text{OD}_{\text{in}} - 2 \text{WT}_{\text{in}} \geq \text{OD}_{\text{out}} - 2 \text{WT}_{\text{out}}.
\]
Effectively, the constraint imposes an upper limit on the drawing angle. Ironing corresponds to the limiting case with $\text{ID}_{\text{in}} = \text{ID}_{\text{out}}$. We must also provide for some clearance to thread the mandrel. Recall that the diameter of the bulb at the end of the mandrel, which equals the inner diameter $\text{ID}_{\text{out}}$ of the output tube, must be first threaded through the input tube before drawing. To conveniently thread the tube, we require a minimum clearance between the bulb and the inner diameter of the input tube. One way to approximate this bulb clearance requirement is by decreasing the upper limit on the drawing angle. In general, this upper drawing angle constraint has the following form:

$$\frac{\text{WT}_{\text{in}} - \text{WT}_{\text{out}}}{\text{OD}_{\text{in}} - \text{OD}_{\text{out}}} \leq \tan \theta_{\text{max}}.$$ (3.5)

Based on their experience, process engineers might choose a conservative (lower) value for the parameter $\theta_{\text{max}}$ to ensure good quality tubes.

Just as the ironing constraint imposes an upper limit on the drawing angle, the sinking constraint specifies a lower limit. The general form of the lower drawing angle constraint is:

$$\frac{\text{WT}_{\text{in}} - \text{WT}_{\text{out}}}{\text{OD}_{\text{in}} - \text{OD}_{\text{out}}} \geq \tan \theta_{\text{min}}.$$ (3.6)

Sinking corresponds to a drawing angle $\theta_{\text{min}}$ of $0^\circ$. Again, process engineers might specify a larger value for this lower drawing angle to ensure adequate process yield.

Calibrating the constraints

The constraints described above, with other constraints that capture limitations due to process equipment, lubrication, or operating conditions, can represent a set of rules or standard practices that serve to define bounding process conditions on the production of drawn tube. For a particular alloy and tube dimensions, the standard practice specifies the values of the various maximum and minimum CSA reduction and drawing angle parameters. Currently, these standard practices are primarily experience-based, and although they reflect actual physical processing constraints they have not been derived from first principles and engineering analysis.

To accurately calibrate these constraints (i.e., to determine the true values of the parameters $\delta_{\text{CSA, max}}$ and so on), we must use a combination of process
modeling and designed experiments. Process modeling techniques such as finite element methods provide a way to visualize the complicated deformations, die interactions, lubricant effects, and deformation heating effects associated with tube drawing. Nonlinear finite element programs such as ABAQUS (Hibbett, Karlsson and Sorenson, Inc. [1991]) and ALPID (Batelle Research Laboratories [1991]) are sufficiently powerful to model bulk deformation processes including the effects of large deformations, nonlinear constitutive behavior, and coupled thermomechanical deformations, although particular difficulties still exist with the modeling of three dimensional contact. While process modeling provides physical insight, it does not guarantee that all of the relevant factors affecting the process are correctly modeled or included within the model. Conversely, in many cases, designed experiments would be either impossible or too difficult to capture the range of conditions that modeling can simulate. Thus, process understanding and constraint characterization requires a combined approach employing modeling, designed experiments, and manufacturing experience. The manufacturing experience component is critical since without this experience models can often represent a process incorrectly, and process engineers might form incorrect conclusions from experiments.

This section has identified and formulated the deformation limits imposed by the tube drawing process. We defer discussion on how to use these constraints for process planning in order to first briefly outline the factors limiting the extrusion operation.

3.1.2 Extrusion

Extrusion consists of producing a long part with a given cross section by forcing a hot metal workpiece through a die with a cutout of that cross section (we consider only hot extrusion, although cold extrusion is also possible). Extrusion is a particularly efficient forming process since it can produce very complicated, and intricate geometries from large, simple starting workpieces. For the production of tubes, the starting extrusion workpieces are simple solid or hollow cylindrical ingots.

Like tube drawing, several processing constraints limit the bloom sizes that can be produced on a given extrusion press (we refer to seamless, cylindrical extruded tubes as blooms). First, as the differential between the cross-sectional areas of the ingot and the bloom increases, the required ram force to push the ingot through the die also increases. The press' capabilities therefore limit the maximum possible cross-sectional reduction. Extrusion engineers might specify
this constraint as an upper limit on the Extrusion Ratio which is defined as the ratio of cross-sectional areas of the ingot and the bloom. Notice that this constraint is analogous to the max reduction per draw constraint for drawing operations. The capacity of the extrusion press and its cylinder size also restricts the length and diameter of the starting ingot.

The extrusion process is also constrained by thermal considerations. The ingot must be preheated above a certain minimum temperature to achieve the required deformation with the available ram force. However, the heat dissipated during deformation increases the temperature of the metal. Above a certain alloy-dependent maximum temperature the metal becomes too soft and produces weak extrusions with poor surface quality. Furthermore, certain alloys can withstand only a limited amount of deformation before they develop internal defects that either weaken the extruded tube or develop later as surface defects. These factors again limit the amount of cross-sectional reduction.

Using these extrusion constraints we can develop standard practice rules analogous to the tube drawing guidelines. These rules then determine if a particular bloom size can be extruded from a specified ingot. We next describe current practices in process planning and process engineering for tube manufacturing operations.

3.2 Current Practice in Process Engineering and Planning

3.2.1 Process Planning

As we mentioned previously, neither the literature nor current practice adequately emphasize the strategic importance of process planning. Instead, process planning is viewed mainly as an operational function, with much of the emphasis on how to automatically generate the process plan for a given product specification (see, for example, Alting and Zhang [1989], CIRP [1985], Chang and Wysk [1985], van't Erve [1988], and Chang [1990]). In particular, the literature focuses on process planning for a single part (primarily non-prismatic parts), and considers mainly machining operations; thus, the literature addresses questions such as what is the appropriate computer representation of the product design, how to infer the required processing steps from this representation, and how to provide computer support for developing the detailed process plan. Researchers have identified two basic methods for process planning—generative methods that construct process plans from first principles based on the part geometry, tolerances, and material, and variant methods that identify similar parts that
were produced previously, and perturb the previous process plan to accommodate any differences in the design of the new part. Computational geometry and artificial intelligence methods have been proposed to provide decision support for process planning.

Unlike some machining operations, implementing generative process planning methods is relatively easy for metal forming operations. Given the sizes of the bloom and the finished tube, we can use the drawing constraints described in Section 3.1 to:

(i) verify if the bloom can produce the finished size, i.e., if the plan satisfies the drawing angle and min cold work constraints,
(ii) if feasible, determine the required number of drawing passes and intermediate annealing steps based on the max reduction per draw and max work-hardening constraints, and
(iii) appropriately space the drawing passes (i.e., select intermediate product dimensions) and annealing steps to meet the reduction and work-hardening constraints.

Indeed, this planning procedure is easy to visualize on the tube reduction diagram (Figure 4). Intuitively, the process plan consists of dividing the line connecting the two points representing the bloom and the finished product into multiple segments, each corresponding to a drawing pass. Each segment must satisfy the length and orientation constraints implied by the reduction per draw and drawing angle restrictions. Within the prescribed limits, the planner might choose a trajectory that ensures good process yield or optimizes drawing effort. The minimum number of draws required to produce the finished tube from the given bloom depends on the max reduction parameter $\delta \text{CSA}_{\text{max}}$. Specifically,

$$\text{Min number of draws} = \lceil (\text{CSA}_{\text{fin}} + \text{CSA}_{\text{bloom}}) / \log (1-\delta \text{CSA}_{\text{max}}) \rceil,$$

where $\text{CSA}_{\text{fin}}$ and $\text{CSA}_{\text{bloom}}$ are the cross-sectional areas of the finished tube and bloom, respectively, and $\lceil a \rceil$ denotes the smallest integer greater than or equal to $a$. (Recall that the value of the parameter $\delta \text{CSA}_{\text{max}}$ lies between 0 and 1, with higher values corresponding to greater reduction per draw.) We refer to the procedure for determining the process plan with the fewest number of draws to produce a finished product from a specified bloom as the draw planning method. The method is easy to implement both manually and using computer graphics support. The intermediate tube sizes determined by the draw planning method might require minor perturbations to accommodate available die and mandrel sizes.
We next describe a typical process planning scenario from practice. When the customer places an order, the process planner (who is not necessarily a process engineer) receives the product specifications, and is responsible for preparing a routing sheet containing the detailed processing steps—from ingot to finishing—to produce the item. In preparing this process plan, the planner must operate within certain prespecified boundaries. For instance, the extrusion plant might prespecify a set of standard bloom sizes that it can produce, thus limiting the available processing path choices from ingot to extrusion. In this case, the planner must

(i) select the best available bloom size (OD and WT) for the product,
(ii) specify the lot size (number and weight of ingots and blooms) after accounting for extrusion and tube drawing process yields,
(iii) determine the sequence (and intermediate dimensions) of tube drawing and annealing operations to reduce the selected bloom to the required finished size, and
(iv) specify the required finishing operations.

To select the appropriate bloom size for an order, planners often use the variant method since it utilizes a previously proven processing path with an acceptable recovery level. Thus, the process planner first identifies a previous product with identical or very similar specifications, and chooses the same bloom size for the new order. If the previous product had the same specification, the planner uses the same process plan; otherwise, he applies the draw planning procedure that we described earlier. Occassionally, the previous bloom size may no longer be available as a standard size, in which case the planner chooses a similar bloom from the available stock sizes. Thus, the process planner is generally more concerned with selecting a proven process plan, rather than one that explicitly considers drawing or extrusion effort. As one example, an analysis of actual process plans over a 3-month period in a tube manufacturing plant revealed that a potential savings in tube drawing effort (drawbench hours) of approximately 20% was possible by merely selecting, for each finished, the closest feasible bloom from the current list of standard bloom sizes (this analysis ignores the impact on extrusion effort). In part, the process planners' emphasis on process feasibility rather than manufacturing effort and complexity reflects the disadvantage of using standard practice rules that are not completely reliable because they are based on experience rather than deep process understanding. Furthermore, the standard practice rules do not provide any guidance on how the product quality varies as the draw parameters (e.g., drawing angle, CSA reduction per draw) vary.
How does the extrusion plant decide the standard bloom sizes? This decision is largely evolutionary. For instance, the extrusion facility is mainly concerned with maintaining large lot sizes in order to achieve its throughput and efficiency targets (e.g., number of pounds extruded per month, effective press utilization %, and so on). Therefore, if the plant receives a large order whose volume is high enough to justify introducing a new bloom size, the extrusion plant might agree to add this size to its standard list; conversely, sizes that are not active for a certain period of time are discarded from the list. Selecting standard bloom sizes in a principled way is the main theme of Section 4.

3.2.2 Process Engineering

Process engineers typically specialize in individual processes (e.g., extrusion or tube drawing), focussing on improving the efficiency or yield of that process. Thus, extrusion engineers are concerned with optimizing the extrusion speed and controlling the defect rate for a specified bloom. Similarly, for a given process plan, tube drawing engineers seek optimum die setups, lubricants, and drawing practices to improve quality.

Setting priorities is a challenge for process engineering since the engineer must address immediate process problems while still pursuing longer term improvements. The immediate problems frequently pertain to a specific lot or piece of equipment. Due to the pressure to find quick solutions to disruptions in daily production, the process engineer becomes preoccupied with "fighting fires", and lacks the guidance necessary to formulate a consistent plan of attack for long-term process improvement. Furthermore, the engineer seeks ways to incrementally modify the current practice without considering more fundamental changes in the process.

Even when the opportunity exists for broader process improvement, it may be difficult to select the most critical part of the process to address. What may appear to be the most difficult problem on a local process scale may not be economically the most important constraint on the process. To increase manufacturing flexibility without sacrificing either manufacturing efficiency or quality, the engineer needs to know both which process constraints, if relaxed, would offer the most benefit and which constraints are most amenable to relaxation. In many cases the rationale for one set of constraint parameter values has been lost due to changes in technology and product mix, and the standard practices specified for a given process may be the result of habit rather than
engineering knowledge. The next section provides a framework to incorporate manufacturing and process constraints directly into the planning process, which then feeds back sensitivity analysis to the engineering activity to identify "critical" constraints.

4. Tactical Process Planning for Tube Manufacturing

This section deals with how to use the inherent process flexibility of tube manufacturing wisely to mediate between the conflicting objectives of extrusion and tube drawing. We first develop a characterization of processing flexibility using the drawing and extrusion constraints described in Section 3.1, and briefly discuss the factors affecting the workload in extrusion and tube drawing. This discussion leads naturally to the tradeoffs and constraints of a medium-term planning model to select standard bloom sizes. We describe an iterative engineering-planning framework using this model, and identify several related research issues.

We define process flexibility as the opportunity to choose from a range of alternate process plans for each product. Tube drawing and extrusion operations permit wide flexibility, and the processing path choices greatly influence both the total workload and its relative distribution between extrusion and drawing. Exploiting process flexibility to balance and control the workload requires a medium to long-term systems view that closely coordinates engineering and planning activities. However, as we have noted, process planning is often treated as an on-line, operational function; it is typically myopic (i.e., considers one product at a time) and is largely based on past history. Correspondingly, process engineering efforts are mainly reactive, addressing current difficulties in individual processing steps. We motivate and formalize a medium-term process engineering and planning framework to systematically address extrusion-drawing tradeoffs.

4.1 Characterizing Tube Drawing and Extrusion Flexibility

Our discussion of tube drawing constraints in Section 3.1 had the implicit purpose of determining the set of finished tube sizes that can be produced from a given bloom. Indeed, current standard practice rules were developed primarily to support this "top-down" view, with planners using the rules to verify if they can produce a particular product from a specified bloom, and to construct a satisfactory process plan. Most computer aided process planning systems
reinforce this view of the planning function by automating the feasibility verification tasks. However, the mathematical representation of the constraints in the context of the tube reduction diagram is very powerful, and enables us to address the converse question, namely, "given a desired target point i.e., finished tube, what are the possible bloom sizes that can produce this tube?" This "bottom-up" view of processing constraints provides substantial latitude to the process planner in selecting an appropriate bloom size, and designing a tube drawing plan from first principles without the restriction of prior processing history.

To characterize the flexibility of the tube drawing process, we partition the feasible bloom sizes that can produce a given finished tube size according to the number of drawing passes and intermediate annealing operations they require. Consider, first, the subset of blooms that can produce the finished tube in a single draw. This subset, shown in Figure 5, consists of all sizes that satisfy the lower and upper drawing angle constraints (3.5) and (3.6), the minimum final cold work constraint (3.3), and the maximum reduction per draw constraint (3.1). We refer to the area contained within the Iso-CSA lines representing the minimum cold work limit $\tau_{\text{min}}$ and the maximum reduction per draw limit $\delta_{\text{SCS}}$ and the lines defining the upper and lower drawing angles as the 1-draw region. As shown in Figure 5, we can recursively construct the feasible areas for multiple draws, introducing intermediate annealing steps as necessary.

The feasible area representation conveniently characterizes the inherent flexibility of tube drawing operations, with larger areas denoting greater flexibility. (This type of flexibility is sometimes called range flexibility; see Slack [1983], Upton [1991].) Process engineers can increase flexibility in different directions by exploring the limiting values for each of the constraint parameters $\delta_{\text{SCS}}, \Delta_{\text{CSA}}, \theta_{\text{max}},$ and $\theta_{\text{min}}$. The feasible area representation also verifies our previous observation that blooms that are farther away from the finished tube require more number of draws, and hence more drawing effort.

The tube drawing feasible area is analogous to the concept of processing maps that have been popularized in materials manufacturing. Woodyatt et al. [1992] use a graphical representation to show the range of mechanical properties (e.g., tensile and yield strength) and chemistries (e.g., carbon, manganese, and sulfur content) corresponding to different grades of steel. Frost and Ashby [1982] and Ashby [1985] have developed deformation maps and hot-isostatic pressing maps that assist the process engineer to operate with a desired range of material behavior. Forming limit diagrams have been applied to the shaping of sheet
materials through stamping and sheet drawing (Wagoner, et al. [1989]). However, these concepts have not been extensively implemented within other metal working processes, such as tube forming or rolling. As we will show, they can be directly coupled to the planning process, providing both engineering input directly to the planning operation and economic information to the activity of process improvement.

Just as we used the tube drawing constraints to define feasible bloom sizes that can produce a given tube, we can also use the extrusion constraints to define feasible ingot sizes that can produce a given bloom size. Alternatively, given a set of standard ingot sizes, we can use the extrusion constraints to define the set or area of bloom sizes that the press can produce from these ingots. By overlapping this area with the feasible drawing area for a particular tube we can identify bloom sizes that are feasible for both extrusion and drawing.

We should note that the process flexibility demonstrated by feasible areas represents a double-edged sword. Exploited systematically, we can use this flexibility to enhance competitive advantage. Exploited piecemeal, the practice of incremental planning can result in contradictory process plans over time since similar tubes can have dramatically different processing paths. Furthermore, without a consistent set of practices, process engineers cannot rely upon either historical data or implement system-wide improvements.

4.2 Determinants of Extrusion and Tube Drawing Effort

Given the wide spectrum of bloom size and draw planning choices facing the process planner, we are interested in understanding the effects of these choices on extrusion and tube drawing effort in order to balance the workload.

**Extrusion effort**

To understand extrusion workload, we will focus on how the effective extrusion speed, a common performance metric for extrusion managers, varies with bloom dimensions and lot size. Effective speed is the number of "good" pounds extruded per hour of press usage (including batch setup time); it depends on the total processing time for a batch and its recovery rate. The total time to extrude a batch of blooms with specified length and CSA consists of: (i) the press setup or changeover time, which might be sequence-dependent, and includes the time to preheat the ingot, and change the tooling (dies, mandrel, and possibly the cylinder), and (ii) the actual extrusion time which equals the batch size.
(including scrap) times the extrusion rate. The extrusion rate (i.e., ram speed) decreases as the cross-sectional area of the bloom decreases.

The recovery rate (i.e., good pounds as a % of total pounds extruded) in extrusion also depends on the batch size and bloom dimensions. Planned scrap, consisting of fixed lengths (largely independent of bloom dimensions) from the leading and trailing ends of each bloom or batch of blooms, decreases as a % of total extruded weight when the batch size increases. The extrusion process also introduces random defects (surface defects and dimensional variations) which tend to increase as the bloom's CSA decreases.

Because effective extrusion speed and recovery increase with batch size, extrusion managers strongly prefer to produce fewer and preferably large CSA bloom sizes in large batches.

**Tube drawing effort**

*Tube drawing workload* increases directly with the number of drawing passes. The batch size and dimensions of the tube affect the time required for each drawing pass. We can broadly decompose the total time required for each drawing pass into two components: (i) batch setup time: consisting of the time to load and unload racks of tubes (using, say, cranes or forklift trucks), to change the die set on the draw bench, and to draw one or more trial tubes to validate and debug the drawing pass; and (iii) the processing time for each tube: consisting of the time to set up the tube on the draw bench (i.e., thread the mandrel through the tube, etc.), and the actual drawing time. The drawing time is proportional to the length of the output tube (which increases from one draw to the next), and the drawing speed; this speed depends on the draw bench's capabilities and the required CSA reduction per draw. In addition to draw-bench time, we must also consider the time required for materials handling and intermediate annealing operations. More importantly, materials handling and annealing introduce additional defects (e.g., surface defects) and might severely degrade process yield; the batch size (and hence workload) correspondingly increases to produce the required number of good finished tubes for a given process plan. To summarize, the various ingredients of tube drawing workload lead managers in the drawing facility to strongly prefer process plans that require very few drawing passes and no intermediate annealing steps so that they can improve their performance metrics such as recovery rate, throughput (e.g., total pounds or feet of good drawn tubes produced per month), and productivity.
4.3 The Bloom Sizing Problem

Flexibility in tube manufacturing impacts long-term capacity planning, medium-term tooling, and short-term lot planning decisions. In this section, we focus on a medium-term (say, annual) tactical planning decision, namely, the problem of selecting a set of standard bloom sizes. We refer to this decision as *bloom sizing*.

The interdependence and tradeoffs between extrusion and tube drawing workloads motivates the bloom sizing problem. The downstream (tube drawing) stage prefers to select a tailored bloom size for each of its finished tube sizes in order to minimize the number of draws and eliminate intermediate annealing steps. In terms of the tube reduction diagram, this *draw-effort minimizing* strategy would select, for each finished size, a bloom belonging to the 1-draw feasible region for that tube; yield variations and extrusion feasibility determine the exact bloom size within (or even beyond) this area. This strategy requires a large number of bloom sizes (in the worst case, as many bloom sizes as the number of finished products), with relatively low annual demand for each bloom size and possibly low extrusion rates.

Conversely, the extrusion plant prefers to exploit commonality in order to limit the number of blooms to a set of standard sizes that it can produce efficiently. Conceptually, the same bloom can "serve" k different finished products if it lies in the intersection of the feasible drawing areas for these k products. Even within this intersection, the drawing facility might prefer a size that has the smallest weighted distance (in terms of number of draws) to the k target points, while the extrusion plant might choose a different size to maximize effective extrusion speed. Observe that as k increases, the total volume of finished products served by the bloom increases, improving extrusion performance; however, the area of intersection decreases, possibly increasing the weighted distances and hence the total drawing effort.

This tradeoff between extrusion and drawing effort is the crux of the bloom sizing problem. Given the projected product mix and volumes, the *bloom sizing problem* consists of selecting a set of standard bloom sizes (and deciding the draw plan for each product) to "effectively" resolve the extrusion-drawing tradeoff. As we shall see later, we can represent the tradeoff in various alternative ways by including the extrusion and drawing effort either in the objective function or as constraints. Observe that selecting standard bloom sizes is a special case of the more general *commonality selection* problem of choosing
a set of standard "initial" flow paths that all end-products can use. An initial flow path refers to all the processing steps from ingot casting to an intermediate extrusion or drawing step. By focusing on bloom size standardization, we limit our attention to standard (common) processing paths that end at the extrusion (upstream) stage, assuming implicitly that cold drawing operations are tailored to individual orders. Besides simplifying our discussions, this restriction also reflects the practical concerns of avoiding inventories of semi-finished drawn tubes; the tube drawing facility therefore operates as a pure "make-to-order" facility. However, the concepts that we discuss also apply to the more general commonality selection problem.

We note that the bloom sizing problem can be viewed as a generalized, multi-dimensional version of the Assortment problem (Wolfson [1965] and Pentico [1974]) for selecting standard lengths to balance scrap costs against production economies and storage costs.

We decompose the problem of selecting standard bloom sizes into two interdependent tasks:

(i) characterizing the constraints, quality, and speed of the processing equipment using a process engineering model(s); and,

(ii) balancing extrusion and tube drawing effort using a planning model.

We will argue that effective bloom sizing requires iterative use of the planning and process engineering models. We first motivate and describe the elements of the planning model, and subsequently discuss its interactions with process engineering.

4.4 The Planning Model

The planning model takes as given the processing constraints and operating parameters specified by the process engineers. It explicitly incorporates extrusion and tube drawing effort to select a set of standard bloom sizes (bloom sizing), and a feasible assignment of each finished product (tube) to a selected bloom (tube-to-bloom assignment). We first make some simplifying assumptions, and describe a basic modeling approach called the total cost minimizing model. Later, we describe other modeling options, and Section 4.6 discusses open research issues.

The model requires three sets of inputs:

(i) Projected demand: For each possible product OD, WT combination, we require information on the expected total (say, annual) demand, specified in either feet or pounds, and the batch size or number of orders for that product during the year.
(ii) *Processing constraint* parameters: To select bloom sizes and drawing plans that are feasible, we require: (i) standard practice rules defining the feasible extrusion area, and (ii) the parameters (e.g., \( \delta \text{CSA}_{\text{max}}, \Delta \text{CSA}_{\text{max}}, \theta_{\text{max}} \) and \( \theta_{\text{min}} \)) defining the tube drawing restrictions. We assume that either these constraint parameters define preferred operating regions with acceptable yield, or the process engineer characterizes yield as a function of the drawing plan's parameters (e.g., variations of yield with draw angle and CSA reduction per draw).

(iii) *Processing effort* parameters: To quantify extrusion and drawing effort, we require formulae and procedures to calculate the total extrusion and drawing effort (expressed, say, in monetary values) as a function of the bloom dimensions and drawing plans. The extrusion and drawing effort models require speed and yield parameters based on process engineering.

The planning model has two sets of decisions: (i) selecting standard bloom sizes (OD and WT), and (ii) assigning each finished product to a selected bloom. These two decisions, *bloom selection* and *tube-to-bloom assignment*, together specify the preferred process plan for all the products in the medium-term. Given the projected demand for each finished size, the tube-to-bloom assignments determine the total required production volume of each standard bloom size, and hence the total extrusion effort. For each finished product, the drawing plan is a function of the bloom that is assigned to that product; hence, the tube-to-bloom assignment determines the total tube drawing effort. This assignment also determines whether the tube requires intermediate annealing steps.

### 4.4.1 Total cost minimizing model

First, let us describe a simplified planning model that minimizes the sum of the annual extrusion and drawing effort. For this model, we assume *lot-for-lot extrusion*, i.e., each extrusion lot produces blooms for a single drawing lot which, in turn, corresponds to a unique customer order. Our basic planning model also ignores the sequence-dependence of setup times in extrusion. Since cylinder changeovers are the most time-consuming activities, assuming that setup times do not depend on production sequence is reasonable if blooms and presses are partitioned (using, say, group technology) so that blooms requiring different cylinders are assigned to different presses (the plant that we studied followed this strategy). These two assumptions simplify the model by enabling us to express
both tube drawing and extrusion effort as separable, additive functions of the decision variables.

We assume that a set of m candidate bloom sizes is prespecified. The candidate blooms might consist of all sizes that can be produced using the current tooling (dies and mandrels), or might correspond to all the feasible (i.e., producible by extrusion) grid points of a rectangular grid superimposed on the tube reduction diagram. Alternatively, we might use a "feasible area overlapping procedure" (Loucks [1990]) that determines the intersection of the feasible drawing regions for closely clustered finished products to identify a list of promising bloom sizes that can serve several end products. We index the candidate bloom sizes from 1 to m, and the finished tube sizes from 1 to n.

Suppose we assign bloom j to tube i, i.e., the process plan for product i consists of extruding bloom j, and drawing it down to product i's OD and WT. Using the associated drawing plan, and the annual demand and average lot size for product i, we can calculate the total drawing effort to manufacture product i using bloom j. We convert this total drawing effort into an j-to-i drawing cost which we denote as \( d_{ij} \). Note that this drawing cost can readily incorporate the costs of material handling, annealing, and scrap reprocessing. For notational convenience, we assume that \( d_{ij} \) has a very large value if bloom j does not lie in product i's feasible area. Because we have assumed lot-for-lot extrusion, we can also compute the j-to-i extrusion cost \( e_{ij} \) to produce bloom j to meet product i's annual demand. This cost includes the cost of setup and actual usage of the extrusion press, as well as the cost of reprocessing extrusion scrap.

Using these costs, we can formulate the medium-term planning task as an assignment problem. For all \( i = 1,2,...,n \), and \( j = 1,2,...,m \), the model contains binary decision variables \( x_{ij} \) representing the tube-to-bloom assignment decisions. The variable \( x_{ij} \) takes the value 1 if we assign product i to bloom j, and the value 0 otherwise. In terms of these decision variables, the basic planning model has the following form:

\[
\begin{align*}
\text{minimize} \quad & \sum_{i=1}^{n} \sum_{j=1}^{m} \{ e_{ij} + d_{ij} \} x_{ij} \\
\text{subject to} \quad & \sum_{j=1}^{m} x_{ij} = 1 \quad \text{for all } i = 1,2,...,n, \text{ and } \quad (4.2) \\
& x_{ij} = 0 \text{ or } 1 \quad \text{for all } i=1,2,...,n, \ j=1,2,...,m. \quad (4.3)
\end{align*}
\]
The objective function (4.1) minimizes the total annual production cost which, under our assumption of lot-for-lot extrusion, is the sum of the drawing and extrusion costs over all the selected tube-to-bloom assignments. The constraints (4.2) and (4.3) ensure that every product is assigned to one bloom. Observe that, in this simple form, the planning model is easy to solve: assign each product \( i \) to the bloom \( j \) that minimizes \( (e_{ij} + d_{ij}) \).

**Model Variants, and Enhancements**

The assignment model (4.1)-(4.3) considers all products simultaneously, and captures the variations in extrusion speed and drawing effort with bloom size; however, its main disadvantage is the lot-for-lot extrusion assumption. Thus, the model ignores the savings in setup time at the extrusion press when we combine multiple customer orders requiring the same bloom into a single extrusion lot. To capture the setup savings, extrusion managers might either specify a surrogate constraint that limits the number of standard bloom sizes, or prefer to include an estimate of the annual setup cost in the objective function. We next outline some of these modeling options:

1. Impose an upper limit on the number of selected bloom sizes.
   To model this restriction, we introduce another set of binary variables \( y_j \) for each bloom \( j = 1, 2, \ldots, m \). The *bloom selection* variable \( y_j \) takes the value 1 if we select bloom \( j \), and the value 0 otherwise. The model contains two additional sets of constraints: (i) the forcing constraints:
   \[
   x_{ij} \leq y_j \quad \text{for all } i=1,2,\ldots,n, \ j=1,2,\ldots,m,
   \]
   which specify that we can assign a tube \( i \) to a bloom \( j \) only if we select bloom \( j \), and the upper limit, say \( p \), on the number of standard bloom sizes:
   \[
   \sum_{j=1}^{m} y_j \leq p.
   \]
   Note that adding constraints (4.4) and (4.5) to (4.1)-(4.3), transforms the assignment model into a p-median model (see, for example, Mirchandani and Francis [1990]).

2. Include an "implicit" fixed cost, say \( F_j \), for selecting each bloom size \( j \).
   For instance, if we assume that the facility produces each bloom size once every week, the fixed cost \( F_j \) might represent the annual setup cost to produce bloom \( j \). We drop constraint (4.5) from the p-median formulation, and change the objective function to:
\[
\text{minimize } \sum_{i=1}^{n} \sum_{j=1}^{m} (e_{ij} + d_{ij}) x_{ij} + \sum_{j=1}^{m} F_j y_j.
\] (4.6)

This formulation corresponds to a plant location model, with blooms representing plants, and products representing customers. By parametrically increasing the fixed costs \(F_j\), we can generate an entire family of solutions with decreasing number of standard bloom sizes.

(iii) Estimate of the number of setups for each bloom (as a function of its tube assignments) and include an explicit setup cost in the objective function. For instance, we might approximate the number of setups by assuming that the extrusion facility selects lot sizes using the economic order quantity formula (see, for example, Nahmias [1989]). In this square root formula, the annual demand for bloom \(j\), say, \(D_j\) depends on the tube-to-bloom assignments, i.e., \(D_j = \sum_{i=1}^{n} d_i x_{ij}\) where \(d_i\) denotes the known annual demand for tube \(i\). Observe that introducing the EOQ-based setup cost creates a non-linear, non-separable objective function in the optimization model.

Alternatively, we might assume that the extrusion facility accumulates all orders for a week, and consolidates all orders requiring the same bloom into a single extrusion batch; thus, the production frequency of each bloom depends on the demand pattern for the finished tubes it serves. To estimate the number of blooms produced each week (and hence the setups) under this policy we require additional information and assumptions regarding the order arrival process for each product. Again, a probabilistic model for estimating the number of setups as a function of the tube-to-bloom assignment variables introduces non-linearities in the objective function.

By redistributing the workload between extrusion and tube drawing operations, the bloom sizing and tube-to-bloom assignment decisions also affect the relative congestion in the two stages. The basic planning model does not capture the increased lead times and inventory costs due to this congestion. Thus, another useful model enhancement consists of including the effects of congestion using queueing approximations (see, for example, Bitran and Tirupati [1989]). Nof and Barash [1980] emphasize the role of the process planning function to judiciously select alternate routes and manage the relative congestion at different workcenters of a flexible manufacturing system.

4.4.2 Modeling Alternatives

The basic planning model that we have just described seeks feasible tube-to-bloom assignments that minimize the total extrusion and drawing cost. This
total cost minimizing model represents a centralized decision-making process, or assumes close cooperation between the extrusion and tube drawing facilities to achieve the global cost minimization objective. We can also extend this model to contexts where the extrusion and tube drawing facilities are separate cost or profit centers, and the extrusion facility specifies the "transfer price" (possibly with volume discounts) for each of its standard bloom sizes, or offers its capacity at a negotiated rate per hour of usage (including setup time). Indeed, we can even transform the model to a "profit maximizing" form if the extrusion facility is free to accept external orders, and the tube drawing facility can use external extrusion sources and also selectively reject customer orders that are relatively unprofitable.

Other organizational structures might require different models. We describe two alternative models—an extrusion constrained model, and a bloom coverage model—that represents more appropriately the tube drawing facility's viewpoint. We describe only the aggregate structure of the models instead of providing detailed mathematical formulations. For both models, the decision variables are the binary bloom selection and assignment variables ($y_j$ and $x_{ij}$).

**Extrusion-constrained model:**
This model has the following form:

\[
\text{minimize total drawing cost}
\]

subject to:

- Extrusion capacity constraint, i.e., upper limit on total extrusion hours to produce the required blooms, and
- Upper limit on number of standard bloom sizes chosen (optional).

Instead of including extrusion cost in the objective function, this model treats extrusion effort as a capacity constraint. By parametrically changing the available extrusion capacity, we can generate a set of pareto-optimal solutions that vary in their tube drawing and extrusion processing requirements. Figure 6 illustrates a tradeoff curve between extrusion and drawing effort; managers might find these tradeoff curves more appealing than the single solution generated by our previous total cost minimizing model.

**Bloom coverage model:**
This model focuses on selecting a subset of blooms that maximizes the total volume of drawn tubes that are "easy to produce". Since annealing costs are typically much higher than drawing costs, products that do not require
intermediate anneals might be considered easy to produce. Correspondingly, we define the \textit{coverage} of a bloom as the total annual volume (in pounds or feet) of finished products drawn from that bloom \textit{without any intermediate anneals} (or within a prespecified number of drawing passes). The model then becomes:

\begin{equation*}
\text{maximize total bloom coverage}
\end{equation*}

subject to:

\begin{equation*}
\text{Upper limit } p \text{ on the number of standard bloom sizes.}
\end{equation*}

The model does not explicitly consider extrusion and drawing effort, and does not require that every product should be assigned to some bloom. Instead the model seeks, say, the top 5 or 10 (i.e., \( p = 5 \) or 10) blooms that together cover, without intermediate anneals, a large portion of the total end-product demand; the annual demand (feet or pounds) of each product serves as its weight in computing this coverage. We can enhance the model, for instance, by including an extrusion capacity constraint similar to the previous model. Section 4.6 outlines other relevant modeling options including multi-objective problem formulations, and models to select a "robust" set of bloom sizes (to provide primary and secondary coverage, or to meet shifting demand patterns).

In summary, we can model the medium-term bloom selection problem in various alternative ways, depending on the organization structure, the relative costs and equipment utilization in extrusion and drawing, and the availability of data and methods to quantify production effort and inventory cost. The model might use various approximations or surrogate measures of performance, and capture extrusion and drawing considerations via either the objective function or constraints. Some models are well-known optimization problems (e.g., assignment, plant location, p-median) with proven solution methods, while others require new optimal or heuristic methods that exploit the problem structure. In Section 4.6, we describe some interesting generic optimization problems that are motivated by the medium-term planning model.

The models we have described in this section represent an improvement over current practice. They recognize the strategic value of \textit{simultaneously considering all products} that the facility expects to produce; by incorporating the impact on both upstream and downstream operations, these models provide a principled way to select common processing paths and balance the workloads. We remark that the underlying principle of exploiting process flexibility to establish upstream commonality extends to the ingot casting stage as well. Thus, we might consider a comprehensive long and medium-term planning model.
that selects standard ingot sizes and standard bloom dimensions, including not only standard OD–WT combinations but also standard lengths for each OD–WT combination, to effectively produce the projected mix and volumes of finished tubes. Vasko, Wolf and Stott [1989] and Vasko et al. [1989] address related problems of selecting optimal ingot sizes and choosing common metallurgical grades for steel rolling operations.

4.5. Planning-Engineering Iterations

The planning model requires the process constraints as input. These constraints impact both the feasibility of the tube-to-bloom assignments, and the calculation of drawing and annealing effort. Since the process planning decisions are sensitive to the accuracy of these constraints, the planning model requires good process understanding, i.e., an accurate representation of constraints, and characterization of processing speeds. This section explains the reverse effect, i.e., the planning model can provide valuable information to guide process engineering efforts. Thus, in addition to providing support for tactical planning decisions, the planning model plays an important role in continuous improvement efforts by generating sensitivity analysis information with respect to various constraints. This information identifies promising directions for process improvement or further refinement of process constraints.

As we noted in Section 3.2, the process variables controlling the tube drawing constraints are not well-understood currently, and the standard practice guidelines are based primarily on experience. To further explore and calibrate these constraints requires extensive experimentation and detailed process modeling. To effectively direct process modeling and improvement efforts, the process engineer requires some principled method to determine which constraints are critical in terms of improvement in manufacturing performance (e.g., cost, lead time, quality), and to prioritize the various improvement options for further exploration. Consider, for instance, two of the five tube drawing constraint classes—the max reduction per draw constraint, and the lower drawing angle—described in Section 3.2. Should the process engineer first explore the possibility of increasing the CSA limit ΔCSA_max in the max reduction per draw constraint, or should he study the effect on tube quality of reducing the lower drawing angle? Should the CSA limit be increased through better process understanding and more precise parameter estimation, or by redesigning the tooling (e.g., die geometry, lubrication)? What is the economic impact of increasing the CSA limit by, say, 1 percentage point, and how does this impact
compare to reducing the lower drawing angle by 1 degree? A parametric analysis using the planning model can provide some insights into these questions.

For a given set of standard bloom sizes and tube-to-bloom assignments, increasing the CSA limit might possibly reduce the number of draws for certain products, and hence reduce the drawing effort. Adding these savings in drawing effort over all the current tube-to-bloom assignments provides a lower bound on the total savings of increasing CSA reduction per draw limit. However, further savings might be possible by reassigning tubes to other blooms (e.g., if increasing the CSA limit does not reduce the drawing cost for the current assignment of product i to bloom j but reduces the j'-to-i drawing cost for another standard bloom j'), or even choosing alternate standard bloom sizes. To accurately estimate this total savings, we must re-solve the planning model with the new CSA limit. Figure 6 shows how we might represent the results of this type of sensitivity analysis to provide insights to process engineers and managers; the extrusion-drawing tradeoff curve shifts downward as the maximum CSA reduction per draw increases.

Similarly, consider the effect of reducing the lower drawing angle. Relaxing this constraint increases the feasible area for each product (see Figure 5). Consequently, a standard bloom size can now feasibly produce a larger set of products, possibly making alternate tube-to-bloom assignments more attractive in terms of reducing the total (extrusion + drawing) cost. Furthermore, since decreasing the lower drawing angle expands the intersection of feasible areas for a given subset of products, we might be able to select bloom sizes that are closer to the products, thus reducing total cost. As before, we must parametrically vary the lower drawing angle and re-solve the planning model in order to accurately evaluate the total savings obtained by reducing the lower angle.

These two examples illustrate the linkage, shown in Figure 7, between the planning model and process engineering and improvement efforts. The process model first determines a set of constraint parameters (possibly experience-based, and reflecting current wisdom in terms of preferred operating regions) that is known to provide feasible process plans. Using these parameters as the basis, the planning model performs sensitivity analysis to determine if the planning decisions are robust, and to evaluate the potential economic impact of refining the parameters. This sensitivity analysis prioritizes the next iteration of process understanding and improvement efforts, pointing to constraints that are worthwhile refining or improving (e.g., relaxing the parameters via changes in the process). We then re-solve the planning model using the new, improved
process constraint parameters. In effect, the planning–engineering iterations drive the plant's continuous improvement efforts. This complementary use of planning and process engineering models represents a new paradigm for both engineers and planners. This paradigm explicitly recognizes that good process understanding is a prerequisite for effective planning, while the economic considerations derived from planning must drive process engineering. We emphasize that the planning-engineering iterations are ongoing activities, and go well beyond the normal consultations between engineers and the model builder during the initial stages of validating and testing the basic planning model (e.g., Vasko et al. [1989]).

4.6 Research Issues

The tube manufacturing context we have described provides a rich set of management and engineering research opportunities. We limit our discussions to three promising areas for further research.

4.6.1 Planning (Bloom Selection) Methodologies:

We have described the elements of medium-term planning, and posed the essential tradeoffs in terms of an optimization problem that can be modeled in various ways. By discretizing the space of possible bloom and tube sizes, we can formulate the planning models as integer programs; however, these problems are difficult to solve optimally (most models we have considered are NP-complete or NP-hard). Developing effective solution methods for these models is important, especially since the iterative planning-engineering framework requires repeated application with varying constraint parameters. The basic planning model and some of its variants are well-known optimization problems (assignment, plant location, or p-median problems). However, other models such as the extrusion-constrained model and the bloom coverage models are new, and solving them effectively requires tailored optimization-based heuristic solution approaches that exploit their special structure using, say, decomposition techniques or polyhedral approaches (see, for example, Nemhauser and Wolsey [1990]).

Since bloom sizing decisions must consider multiple conflicting objectives, the formulation and solution of multi-objective models to identify a robust set of standard bloom sizes is another important area to investigate. Instead of combining all the factors into a single objective function or constraint, practitioners might prefer a multi-objective optimization framework that considers a hierarchy of metrics. For instance, the multiple objectives might
consist of (i) minimizing the total annealing effort or the weighted number of products requiring intermediate anneals, (ii) minimizing the number of standard bloom sizes, and (iii) minimizing the total tube drawing effort. Furthermore, selecting a "robust" set of blooms sizes might be an important practical consideration. Robustness of the bloom set might be defined as the ability to provide, for each product, both a good primary or preferred bloom as well as a feasible secondary or alternate bloom (for contingencies when the primary bloom is in short supply). Del Callar [1992] attempts to solve this model using genetic algorithms. Alternatively, we might define robustness in terms of the sensitivity of the bloom set's performance (i.e., total extrusion and drawing effort) to shifts in the product mix and demand pattern (e.g., increasing proportion of thinner wall tubes). Finally, all of our previous models implicitly assume that product quality is at an acceptable level as long as the drawing plan meets the max draw, work hardening, and draw angle constraints. In practice, however, the process yield and product quality vary even within this feasible region. If we can characterize this variation using engineering models (see Section 4.6.3), then we can either include an explicit quality cost in the planning model's objective function, or employ techniques such as fuzzy set theory (see, for example, Woodyatt et al. [1992] for an application to selecting metallurgical grades) to capture the quality and yield variations.

The medium-term process planning framework also motivates some generic optimization problems that might interest researchers in location theory and computational geometry. Consider, for instance, the following generic problem motivated by the feasible area overlapping procedure to identify promising candidate bloom sizes:

Given a set of points (tubes) on the plane, a weight (demand) for each point, and a feasible area (defined by inequality constraints) associated with each point, find a location (bloom) belonging to the intersection of these areas that minimizes the sum of the weighted distances to all the points.

We can extend this problem to select a prespecified number (say, p > 1) of locations to minimize total weighted distance assuming each original point is assigned to its closest location. Computer scientists and location theorists have analyzed simple versions of this problem for certain special metrics (see, for example, Shamos [1978], Francis, McGinnis and White [1992]).

The tube manufacturing context also motivates a new class of oriented location problems dealing with the optimal location of facilities that can serve demand only in a certain direction on the plane. Gopalan [1992] analyzes the performance of heuristics for certain special types of oriented location problems.
As an example, consider the bloom coverage model and assume, for simplicity, that the lower drawing angle is 0° and the upper drawing angle is 90° (i.e., a bloom can produce any tube lying to its south-west). The following question is one of many interesting oriented location problems related to bloom coverage:

Given a set of points on the plane and a prespecified radius \( r \), find the minimum number of locations that can cover all the given points assuming that each location can only cover points to its southwest and lying at most \( r \) units away.

Developing tailored solution methods to exploit the special structure of these oriented location problems is a promising research direction.

4.6.2 Inventory Policies for Systems with Commonality and Substitutability:

We have already discussed issues associated with approximating the benefits of commonality in terms of reduced setup times and improved recovery. Our previous model assumed, however, that both extrusion and tube drawing were make-to-order activities. The benefits of commonality become magnified when the upstream stage (extrusion) produces to stock since having fewer standard bloom sizes reduces the number of "part numbers" to monitor, and more importantly reduces the safety stock (of blooms) due to risk pooling. In such systems, the planning model requires an estimate or approximation of these benefits associated with managing bloom inventories as a function of the bloom selection and tube-to-bloom assignment decision variables. To identify a good approximation, we must first decide what inventory policy to use in the presence of commonality (i.e., multiple products using the same bloom size as starting stock) and substitutability (i.e., if the assigned bloom is not available in inventory, we can produce the tube from a different bloom size albeit with a possibly higher drawing effort). The tube manufacturing context introduces new dimensions that differentiate it from previous research on optimal inventory policies for systems having common components or substitutable products (see, for example, Collier [1982], Baker [1985], Baker et al. [1986], Gerchak and Henig [1986], [1989], Gerchak et al. [1988], Bitran and Dasu [1989], Bassok et al. [1991], and Ou and Wein [1991]).

Consider, for instance, the following short-term production/inventory policy for a given set of standard bloom sizes. Each tubular product has a prespecified "preferred" bloom size to use as starting stock. When an order arrives, if the corresponding preferred bloom is not available in stock, the planner has the option of selecting a prespecified "alternate" bloom that has a proven, but more expensive, drawing plan. If the alternate bloom is also not available in stock, the planner must expedite an extrusion lot for the preferred bloom. Given the
distribution of demand for various tube sizes, what is the optimal inventory policy and the expected inventory level for this system? Observe that optimality is defined with respect to a composite objective function that includes the expected expediting costs in extrusion, and the excess (extrusion, drawing, and scrap) costs of using alternate blooms, in addition to the conventional ordering, setup and inventory carrying costs. We must first understand the performance of this system, for a given set of standard bloom sizes and tube-to-bloom assignments, before we can incorporate the related economies of scale and scope in the medium-term planning model.

4.6.3 Long-term Process Development and Improvement

We have argued for a proactive process engineering strategy to realize long-term process improvements with potentially greater economic benefit compared to the current reactive mode of problem-solving that is driven by day-to-day process difficulties. The short term solutions normally encompass a narrower scope of change, accomplished through incremental adjustments and experimentation around the current operating parameters. Longer term improvements offer greater economic advantage, but require deeper process knowledge and understanding of the underlying principles. In particular, our integrative approach requires the definition and calibration of explicit constraints that can be then relaxed or tightened to investigate the effect on processing flexibility. As we mentioned in Section 3.2, the true underlying constraints are currently not well-understood, presenting a rich set of research opportunities. Designed experiments and accumulated experience can provide guidance on the variables governing a particular process, but the results are often phenomenological rather than fundamental. We must complement experimentation and experience with methods such as finite element analysis to better understand the effects of die geometries, machine setups, material properties, and process path. We also require guidelines or a structured approach to combine these model-based and experimental methods so that process engineers can identify and address process improvement opportunities efficiently. Dorah [1992] uses a finite element model to study the tube drawing process, and reports results from a set of designed experiments to investigate the effects of die setups and geometry on process yield.

The current state-of-the-art in modeling metal working processes still has limitations associated with interfaces. Friction, lubrication, and the effect of surface condition are all imperfectly accomodated within simulation models, and computational capabilities limit the model size if we wish to represent three dimensional forming operations. Another arena concerns the formulation of
inexpensive lubricants that provide maximum lubrication with the least environmental impact. Large quantities of lubricants are required for cold rolling and tube drawing, and their handling, cleaning, and disposal in many cases is not resolved. These lubricants also depend on and influence the surface condition of the metal, and appropriate practices to produce the proper surface condition are still in their infancy.

Characterizing product quality within the feasible area is an important and very promising research thrust. Tube quality varies as a function of the process parameters (e.g., drawing angle), but current understanding of this relationship is imperfect. Indeed, the constraint approximations and limiting parameters specified in current standard practice rules are often conservative estimates that supposedly ensure a prespecified yield and acceptable quality level. However, the iterative planning-engineering framework that we have proposed can explicitly account for the cost of quality; it does not require a prior specification of a quality target, but instead determines the appropriate level by formally incorporating the tradeoffs between poor quality and lower effort. For instance, the planning model might possibly choose a processing path with lower recovery if this path has lower manufacturing complexity or requires less effort (inspite of the lower recovery). Therefore, to exploit this feature, we require a functional description of how recovery varies as we select different trajectories within the tube drawing feasible area (e.g., recovery as a function of drawing angle, and CSA reduction per draw). Notice that, contrary to many process improvement efforts that deal primarily with parameter optimization, we require a characterization of the response surface rather than a point solution. An important research issue concerns how to design parsimonious experiments to provide adequate accuracy for the response function, similar to recent schemes that reduce the number of experiments required to optimize a process' operating parameters (see, for example, Alkhairy and Staelin [1992a], [1992b]).

Finally, a fruitful and novel research direction concerns the direct integration of the cost of process engineering analysis within the planning model. In Section 4.5 we proposed an iterative framework, with sequential engineering and planning activities. Consider instead an integrated model where, instead of specifying hard process constraints in the planning model, we permit relaxing the constraints at a cost. For instance, instead of specifying a fixed lower drawing angle of 25° we permit even lower values for this constraint; reducing the value, however, entails additional costs for process analysis and experimentation, and lower yield. Using this type of model, we can evaluate the return on investment for different types of process engineering efforts, and directly identify constraints
that are critical or cost effective from a planning perspective. We are not aware of any similar effort to include quality metrics and process engineering cost within a planning model.

Next, we briefly describe process planning and engineering issues associated with rolling operations, and identify similar opportunities for modeling and collaboration.

5. Short-term Planning of Rolling Operations: Combining Aluminum Sheet Orders

This section briefly describes the interactions between upstream and downstream operations in aluminum sheet and plate rolling operations, and addresses short-term decisions concerning how to plan production for a set of confirmed orders. Our main purpose in this section is to illustrate how the previous concepts—simultaneously planning the processing paths for multiple orders, accounting for the impact of these decisions on both the upstream and downstream stages, and integrating process engineering activities with planning—apply to other metal forming operations besides tube manufacturing, and are also relevant for short-term planning. We discuss the process flow in rolling operations and outline the underlying engineering principles, describe the economics of rolling and market characteristics, and identify short-term planning issues and modeling requirements. Our description is based on a rolling facility that largely produces specialty sheet products to order (the industry distinguishes between flat plates and sheets based on thickness, with plate products being much thicker).

5.1 Process Flow and Engineering Principles of Rolling

Rolling operations consist of processing a rectangular ingot of the required alloy and dimensions at synchronized hot rolling mills, followed by one or more cold rolling passes. The process flow might also include preprocessing steps, intermediate thermal operations, and a final finishing stage. We focus on the interactions between hot rolling (the upstream stage) and cold rolling (the downstream stage). Each step of the process successively decreases the workpiece's gauge (i.e., thickness) and increases its length; normally the width remains relatively constant. Thus, the % decrease in cross-sectional area equals the % decrease in gauge.
A hot "line" consists of one or more reversing or multi-stand rolling mills arranged in series and operating synchronously (to avoid the need to reheat the metal between stations). The release of ingots into the line and the workload distribution among the mills is controlled so that the workpiece does not wait between stations (due to blocking). Like extrusion, hot rolling can achieve greater reduction per unit energy input (since the metal deforms more easily at elevated temperatures) compared to cold rolling, but is limited in terms of the smallest possible gauge and the dimensional tolerances it can achieve. Sheets from the hot line are processed at cold mills to further reduce the gauge, and meet temper and other material properties. If the hot rolled sheet has a significantly larger gauge relative to the required finished product, the material must undergo multiple cold rolling passes since the process constraints that we describe next limit the amount of reduction in each pass.

5.1.1 Engineering principles

The dominance of rolling in the production of formed metal products has led to more advanced technical understanding of the rolling process relative to, say, tube drawing. Wusatowski [1969] and Roberts [1978] provide comprehensive reviews of the fundamentals of rolling processes. Both hot and cold rolling processes impose constraints similar to extrusion and drawing. This section introduces a few essential characteristics of metal rolling, and describes one class of process constraints, namely, the maximum gauge reduction per rolling pass.

The amount of gauge reduction in each pass depends on the type and power of the rolling mill, the thickness and width of the incoming plate, the type and condition of the metal alloy, the work rolls' diameter, the metal temperature, the nature of the lubricant, and the desired manufacturing tolerances. Because rolling deformation induces dynamic recrystallization (Sakai and Jonas [1984]), the amount of reduction achieved by hot rolling also influences the evolution of microstructure in the metal, and hence its properties and subsequent processing. This tight coupling between the complicated deformation fields and distribution of mechanical properties throughout the plate makes the design of hot rolling practices particularly challenging.

To increase productivity, rolling mills attempt to achieve as large a gauge reduction as possible, as quickly as possible, during every pass of the metal through the rolls. However, the amount of reduction in each pass is limited by several factors including:
(i) the load capacity of the rolling mill, since excessive reductions cause commensurately excessive loads, preventing control of the plate thickness uniformity, and causing excessive wear or damage to the rolling mill;

(ii) the width of the plate, since the width controls the load required for deformation;

(iii) the metallurgical characteristics and processing history of the workpiece. Excessive reduction can cause cracks on the leading edges and sides of the plate, defects on the interior of the plate due to deformation-induced microstructural changes, and surface damage; and,

(iv) the quality of the ingot. Inhomogeneities or defects in the ingots reduce the strength of the plate. Consequently, the amount of reduction is restricted to prevent the defects from causing the plate to fracture.

These factors impose an upper limit on the reduction in cross-sectional area (i.e., reduction in gauge) during each rolling pass; this constraint is similar to the max reduction per draw constraint for tube drawing. The maximum possible gauge reduction in each pass depends on the alloy, the dimensions of the workpiece, and the mill's design and capabilities.

Like tube drawing, cold rolling also introduces additional constraints similar to the max work hardening restriction (for certain materials), and a minimum cold work requirement to meet temper specifications.

One of the important features of some cold mills is their ability to change the spacing between the rollers while processing a sheet. By using this facility during the last cold rolling pass we can produce segments with different output gauges, to satisfy different customer orders, in the same sheet. The sheet segments are then separated (by cutting the sheet at the appropriate lengths) before performing order-specific finishing operations. The maximum gauge reduction constraints together with other process limitations impose upper and lower limits on the possible output gauges that can be combined into a single lot. Note that changing the gauge during a rolling pass entails additional scrap, requires slowing down or stopping the mill, and demands special operator attention.

The range of permissible gauge reductions and combinations give process planners wide flexibility. The planners can redistribute the workload between hot and cold rolling by choosing different gauges for the semi-finished (i.e., hot rolled) sheet; in particular, choosing a thicker gauge improves the productivity of hot rolling operations but increases the cold rolling workload. We next discuss these productivity issues before describing a model to exploit the available flexibility in rolling operations for short-term production planning.
5.2 Rolling Economics and Market Characteristics

Like extrusion, rolling operations are characterized by considerable economies of scale. The per pound production cost is lower for larger rolling mills than smaller rolling mills, and decreases as the ingot size increases (due to reductions in scrap and setups). Typical performance metrics for rolling operations include recovery rate (good pounds as a percentage of total pounds rolled) and productivity or output rate (good pounds produced per hour of press operation).

The processing speeds and costs of rolling vary widely depending on the size of the mill, the required reduction in gauge, and the alloy and temper. The productivity of rolling operations is determined by the rolling speed, machine setup and changeover time, and planned scrap. Rolling speed decreases as the amount of required gauge reduction increases, and total setup time increases as the product diversity (number of different gauges and widths) increases. Finally, the planned scrap—due to ingot scalping, head and tail scrap (material removed from the leading and trailing ends of a sheet due to quality considerations), and side trim—as a percentage of ingot weight decreases as the ingot weight increases, i.e., larger ingots have better recovery rate. Scrap also has a direct impact on total production cost since it represents unproductive use of rolling capacity, and entails a reprocessing cost (the energy cost and vapor loss during the melting operation) to recycle the metal for ingot casting.

Thus, considerations of cost, productivity, and recovery drive rolling mills towards developing the capability for processing larger ingot sizes. In contrast, the market forces are driving in the opposite direction. In particular, because many of the customers are moving towards just-in-time manufacturing in their fabrication and assembly operations, they prefer to place smaller but more frequent orders instead of maintaining large quantities of sheet stock as raw material inventories. In one instance, we observed that the average order size (in terms of pounds of each product ordered) roughly halved over the last 5 to 10 years, while the facility concurrently upgraded its processing capabilities to handle ingot sizes that were approximately 70% larger than before. Consequently, the maximum ingot size is currently about three times as large as the average order size. The next section describes an order combination strategy to mediate between these two opposing trends.
5.3 Short-term Planning Issues

We wish to exploit the inherent flexibility of rolling operations to simultaneously address the rolling mills' preference for processing large ingots while meeting small customer orders. To accomplish this objective, we must assign multiple (say, two or three) incoming orders to a single ingot. If the orders have similar characteristics, i.e., same alloy, and similar widths and gauges, we can choose process plans that have a high degree of commonality which permit common initial rolling operations on a single, large workpiece before it is cut for order-specific final operations. We also exploit the cold mill's capability to produce different exit gauges within a single coil during the final combined cold rolling pass. Figure 8 presents a schematic for the "combined" processing plan for two orders (see Ventola [1991] for more details).

The Order Combination Model

Let us further explore the short-term order combination problem. We are given a set of confirmed or anticipated orders, each specifying the alloy, dimensions (width, gauge, length of each roll, tolerances), total weight, and due date. We are also given a set of standard ingot sizes, as well as the processing and recovery parameters (rolling speeds, gauge reduction constraints, planned scrap requirements, and so on) for each operation as function of the process plan.

Consider the planning process for the available orders for a particular alloy. The order combination model seeks to combine "compatible" orders (say, 2 or 3 orders per ingot) to minimize the total production cost while meeting customers' quantity requirements, specifications, and due dates. A group of orders is said to be compatible if we can develop a feasible process plan satisfying all the processing constraints to produce the selected orders using a single ingot. Thus, determining the compatibility and cost of jointly producing a combination of orders implicitly requires (i) selecting an appropriate ingot, (ii) developing the common process plan, and (iii) verifying feasibility of that plan with respect to processing constraints such as the maximum gauge reduction per pass, and maximum gauge differential between the orders in the group. Notice that the process plan requires determining the relative allocation of workload (i.e., reduction in gauge) between hot and cold rolling operations; this allocation depends on the efficiencies, capabilities, and congestion in the upstream and downstream operations.

We refer to any group of compatible orders as a feasible order combination. The production cost of each combination must include the cost of processing the
workpiece at each workstation (including setup times), and the cost of reprocessing scrap (including gauge change scrap, and trim loss when we combine orders of different widths). A particular order can belong to numerous alternative feasible combinations. Given the set of all feasible order combinations (including "single" combinations, that dedicate an ingot to a single order) and their associated costs, the order combination model must select a subset of these combinations to "cover" all the orders, i.e., each order must belong to exactly one selected combination. The model resolves tradeoffs between combining orders (and hence incurring additional processing costs) and dedicating ingots to single orders (thus decreasing productivity). Since the number of feasible combinations is exponential in the number of orders, solving the order combination problem manually is very time-consuming, and will likely result in suboptimal solutions. Balakrishnan and Gopalan [1992] develop an integer programming approach to find near-optimal order combinations. Their approach extends to the following enhanced versions of the order combination problem.

Our previous description implicitly assumes that the order combination model would be used, say, once a week to plan the production for orders that are due that week. In practice, the plant's order books might contain confirmed orders that have later due dates; accounting for the additional feasible combinations containing these orders might potentially decrease unit production cost even further. However, producing these orders before their due dates entails additional (finished goods) inventory holding costs. This observation leads to an enhanced order combination problem that minimizes the sum of production and inventory holding costs, where we now permit early production of orders in order to exploit the economies of scale in production costs due to order combination. Another model enhancement stems from the plant's (limited) leeway in deciding the shipped weight of each order. Customers normally specify a nominal weight, but permit a limited variance (say, ±10%) around this nominal weight. The plant can exploit this flexibility during the order combination phase. In particular, if the nominal weights for a particular combination of orders does not consume the entire ingot, the plant can increase the individual order weights up to the upper limit or vice versa. Making these choices in a principled way requires a profit maximizing order combination model (instead of our previous cost minimizing model) that does not specify exactly the shipped weight but instead imposes upper and lower limits on the shipped weight for each order.
Strategic uses of the Order Combination Model

We can use the short-term order combination model in an iterative planning-engineering framework similar to our proposed approach in Section 4.3. The model assumes prespecified values for the process parameters and constraints. In particular, the model requires as input the maximum gauge reduction parameters, and guidelines regarding the permissible gauge combinations. Often these parameters are conservative and experience-based, but can be refined through process modeling and experimentation. By solving the order combination model for various values of, say, the maximum gauge reduction and combination parameters, we can determine the sensitivity of the objective function (total costs or net profits) to these parameters. This analysis can then assist the process engineer in prioritizing various process analysis, modeling, and improvement opportunities.

The model can also be used to determine standard ingot sizes. Recall that the model requires prespecified ingot sizes (width, thickness, and total weight). By iteratively varying the ingot sizes, we can select cost-effective standard sizes that are appropriate for the projected mix of products.

In summary, this section has described the process flow and constraints in rolling operations, and identified an opportunity to improve operations by exploiting the inherent process flexibility for short-term production planning. The flexibility permits rolling facilities to continue processing large ingots in spite of the decreasing order sizes. Although we focused on a short-term problem, the principles that we discussed in this section are very similar to our previous discussion of tube manufacturing. Indeed, we can formulate a medium-term gauge and width standardization model (to determine standard sheet sizes produced by the hot line) similar to the medium-term bloom sizing model that we described in Section 4. Likewise, we can develop short-term order combination models for tube manufacturing.

6. Concluding Remarks

We believe that there are substantial opportunities for planning and process improvement within the metal working industry by taking advantage of the inherent process flexibility, and by coupling engineering activities with the short, medium, and long-term planning efforts. We have demonstrated these opportunities for both tube drawing and flat rolling of metal plate and sheet.
The improvements we propose derive from some fundamental characteristics of continuous metal forming, namely, the highly interdependent upstream (hot working) and downstream (cold working) operations with wide latitude in selecting both the upstream and the downstream process paths. In spite of this strong interdependence, the upstream and downstream operations frequently function independently, pursuing local objectives without consideration for overall efficiency. As a result the two operations develop conflicting objectives to the detriment of the overall process. The methodology we propose couples the upstream and downstream processes to exploit the process flexibility while developing global process plans.

We believe that the strategic implications of this inherent process flexibility has been unappreciated and certainly unexploited. Upstream flexibility allows process plans that produce disparate products using common process paths through much of their processing history. We have presented methods to incorporate this process commonality within optimization models that simultaneously develop process plans for multiple products. The models incorporate the primary parameters required for a global process plan: projected demand, processing constraints, and processing effort and cost. This approach can have dramatic and far-reaching benefits including operations streamlining, inventory and flow time reductions, reduced setup times, and improved quality. The model can impose constraints on both the upstream and downstream processes, including limits on the number of upstream products through direct constraints or their cost implications. The exact structure of the model depends on organization structure, cost models, capacity limits, and ease of application. We have proposed several model variants and discussed the implications of each form.

Much of our work has also revealed the benefits of close linkage between engineering and planning models. The engineering models provide the manufacturing constraints that dictate the extent of commonality that is possible between product process plans. The feasible area representation of tube drawing illustrates how engineering constraints dictate a region of process flexibility that a planning model can subsequently operate within. Similarly, the rolling constraints dictate the limits on order combination within a single ingot. This combination of engineering and systems modeling is particularly powerful, for the planning process does not have to rely upon either previous process history or a single process plan. The incorporation of engineering within the planning process also provides feedback to the engineering process. Parametric analyses of the sensitivity to different engineering constraints indicate which constraints
have the greatest influence on process performance. Engineers can therefore pursue improvements that provide the greatest benefit.

We believe that the interdisciplinary approaches described in this article provide fertile ground for significant research—new optimization models and algorithms, inventory management paradigms and problems, and process modeling opportunities. Our work has additionally indicated several areas of process improvement, in both tube drawing and rolling, that could have a significant influence on performance. Not included in this list are other challenging and important research issues relating to incentive and performance evaluation systems for decentralized, but closely coupled, manufacturing stages (e.g., what metrics to use to evaluate the performance of extrusion and tube operations, what is the appropriate transfer pricing scheme), the design of appropriate cost accounting systems, and questions of technology choice, capacity expansion, and cellular manufacturing.

We believe that other industries will also benefit from the concepts presented here, particularly those involving large capital equipment and continuous processing. Prototype industries beyond metal working include the paper industry, the food industry, and the structural polymer product industry. Large benefits can be accrued by exploiting process flexibility to determine the appropriate level of commonality for multiple products, and integrating engineering and planning activities. The field is rich with both economic and technical opportunities.

Acknowledgments
We are very grateful to Duane Dunlap, Ricardo Kamenetzky, Brian Landis, Randy Martin, Brian McConville, Robert Pahl, Scott Roodvoets, and Charles Straface for sharing their insights and providing us extensive support, access, and feedback. Three Leaders for Manufacturing students—Timothy Loucks, David Ventola, and Michael Dorah—played an important role by helping us better understand metal forming principles, practices and performance metrics. Finally, Ram Gopalan and Joseph Del Callar provided invaluable assistance during all phases of our collaborative research projects.
References


Batelle Research Laboratories. 1991. ALPID, Columbus, Ohio.


Figure 1: Tube Manufacturing Process Flow

UPSTREAM
Extrusion

Extrusion Press

Semi-finished stock (blooms)

Inventory

Draw Bench

Annealing Furnace

Draw/Annealing

DOWNSTREAM

Finished Tube
Figure 2: Draw Bench Schematic

- Drawing direction
- Ending cross-section
- Drawn (output) Tube
- Die
- Mandrel Bulb
- Input Tube
- Rod
- Starting cross-section
Figure 3: Types of Tube Drawing Operations

(a) Typical drawing conditions

(b) Pure Sink – Unsupported Draw

(c) Ironing
Figure 4: Tube Reduction Diagram

- **Wall Thickness**: WT
- **First draw**: Starting point
- **Second draw**: Intermediate tube
- **Finished tube**: Final product
- **Bloom**: Raw material

**Graph Parameters**:
- **Y-axis**: Wall Thickness
- **X-axis**: Outer Diameter OD
Figure 5: Tube Drawing Constraints

- Max. Drawing Angle
- Max. work hardening (anneal)
- 3 or more draws
- 2-draw region
- 1-draw region
- Min. cold work
- Max reduction per draw
- Finished Tube
- Outer Diameter
- Wall Thickness
Figure 6:
Extrusion versus Draw effort Tradeoff Curves and Sensitivity Analysis

Current practice

Increasing values of Max reduction per draw parameter

Extrusion effort

Tube Drawing effort
**Figure 7: Planning–Engineering Iterations**

**Process Engineering**
- Process models
- Designed Experiments
- Engineer's/operator's experience

**Process Planning**
- Planning Model
- Planner's experience
- Past practices

**Process Constraints**
- Drawing constraint parameters
  - max/min draw angle
  - max reduction/draw
  - max work hardening
- Extrusion constraint parameters

**Processing Effort**
- Drawing speed
- Extrusion speed
- Planned scrap

**Sensitivity Analysis**
Impact on performance of:
- drawing constraints
- extrusion constraints
- speeds, recovery rates

Select area for improvement, refinement, further investigation

- Standard bloom sizes (OD, WT)
- tube-to-bloom assignment

Projected Product Mix
Available Equipment

Standard bloom sizes (OD, WT)
tube-to-bloom assignment
**Figure 8: Combined Process Plans for Two Sheet Orders**

**Order 1:**
- Width = 48"
- Gauge = 0.063"
- Weight = 10,000 lbs
- Temper = T4

**Order 2:**
- Width = 48"
- Gauge = 0.05"
- Weight = 5,000 lbs
- Temper = O

1. **Ingot Casting**
   - 50" wide, 25,000 lbs ingot

2. **Hot Rolling**
   - 50" wide, 0.25" gauge, coiled sheet

3. **Cold Rolling**
   - 50" wide, 0.063" gauge
     - 12,000 lbs coiled sheet
   - 50" wide, 0.05" gauge
     - 7,000 lbs coiled sheet

4. **Tempering**

5. **Finishing**

ORDER 1

ORDER 2
<table>
<thead>
<tr>
<th>Date</th>
<th>Due</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUG 16</td>
<td>1993</td>
</tr>
<tr>
<td>SEP 6</td>
<td>1995</td>
</tr>
<tr>
<td>S.P.: 2</td>
<td>1997</td>
</tr>
<tr>
<td>OCT 5</td>
<td>1998</td>
</tr>
</tbody>
</table>

Lib-26-67