

第三章 MRP 运动

不像其它很多方法和技巧，物料需求计划有着“实际而有效的作用”，而这正是它的最可贵之处。

——约瑟芬·奥利基，1974

3.1 物料需求计划——MRP

到 20 世纪 60 年代早期，许多公司都开始使用数字式计算机来进行日常的会计活动。考虑到制定计划和进行库存控制的复杂性和单调性，尽力把计算机也延伸到这些功能活动中就是很自然的事情了。这个领域的早期试验者之一就是 IBM，在那里约瑟芬·奥利基和其它人开发了**物料需求计划（material requirements planning, MRP）**。尽管早期发展很慢，但是当 1972 年美国生产与库存控制协会（APICS）发起了“MRP 运动”来推动 MRP 的使用时，MRP 开始取得巨大的发展。从那时候开始，在美国 MRP 就开始变成基本的生产控制范式。到 1989 年，MRP 软件以及运行支持的销售额就超过了 10 亿美元。

因为它是这么的普遍，使得训练有素的制造经理都必须精通一些 MRP 是怎么运行的。因此，在这一章里面，我们会描述 MRP 范式和接踵而至的**制造资源计划（manufacturing resources planning, MRP II）**范式，以及它的现代版本，**企业资源计划（enterprise resources planning, ERP）**。同时我们将突出 MRP 所代表的基本思想和它所留下的一些没有解决的困难。

3.1.1 MRP 的主要思想

就像我们在第二章里提到过的，在 MRP 之前，大多数生产控制系统都是基于一些统计性的再订货点变量。从本质上讲，这意味着任何零件、成品、或部件的生产都是由它们低于某个特定水平的库存状态触发的。奥利基和其它一些 MRP 的开创者认识到，相对于部件而言，这种方法更加适合成品。理由是，对于成品的需求是来源是系统外部的，因此它受不确定性的影响。然而，由于部件是用于生产成品的，对于部件的需求是成品的需求的函数，因而，对于任何给定的最终装配计划，部件的需求都是已知的。（109|110）像统计再订货点法中所做的那样，把这两种需求同等处理，会忽略部件需求对成品需求的依赖性，并因此会导致生产调度中的低效率。

任何产生于系统外部的需求都称为**独立需求（independent demand）**，它包括所有对成品的需求和可能的一些对于部件的需求（如它们被当成备件出售时）。**相关需求（Dependent demand）**是对于组成独立需求产品的部件的需求。运用这些术语，MRP 的主要视角可以阐述如下：

相关需求不同于独立需求。对用于满足相关需求的生产应该制定排程，以便能明确地识别出它与用于满足独立需求的生产的联系。

就像我们将要看到的一样，MRP 的基本机理正是做这些事情。通过向后追溯，从一个独立需求物件的生产排程推出一个相关需求物件的生产排程，MRP 把统计再订货点法中所

缺失的独立与相关需求之间的联系加入到物料需求计划中了。由于 MRP 基于需求对那些应该首先开始生产（推入）的物件计算出生产排程，它因此也被称为**推式（push）**系统。这与**拉式（pull）**系统相反，譬如丰田的**看板（kanban）**系统，当库存消耗时它才授权生产。我们在第四章中会更加详细地讲解看板，并在第十章中提供推式系统与拉式系统之间更加全面的对比。

3.1.2 MRP 概览

MRP 的基本功能正如它的名字所揭示的一样通过对物料的需求进行规划，MRP 可以对来自工厂内外的订单进行协调。外部的订单被称为**采购订单（purchase orders）**，而来自内部的订单则被称为**加工任务（job）**。MRP 的主要焦点就在于对加工任务和采购订单进行调度，以满足由外部需求产生的物料需求。

MRP 处理生产控制的两个基本维度是数量和时间。这个系统必须确定从用于出售的成品到用于制造成品的部件，到作为物料购买进来的投入品，这些所有类型物件的合适的生产数量，它也必须确定满足订单期限的生产时间（即加工任务开始时间）。

尽管有一些系统使用的是连续性时间，而在许多 MRP 系统中，时间被分成许多**时段（buckets）**，一段就是一个用于把时间和需求分隔成不连续块的间隔。在这个时间间隔（时段）过程中所积累的需求都被看成是该时段开始时到达。因此，如果一个时段的长度是一个星期，而在第三个星期期间，星期一的需求是 200 件，星期二是 250 件，星期三是 100 件，星期四是 50 件，星期五是 50 件，那么第三个时段内的需求就是 950 件，并且都在星期一早上到达。在过去数据处理比较麻烦的时候，典型的时段大小都是一个星期或更长一点。现在，大多数现代 MRP 系统都使用日时段，尽管仍然还有很多系统使用星期作为时段。

MRP 不仅处理成品，或者叫**最终物件（end items）**，而且还处理成品的组成零件，我们称之为**低层物件（lower-level items）**。**物料清单（bill of material, BOM）**描述了最终物件与低层物件之间的关系，如图 3.1 所示。对最终物件的需求引发出对于对低层物件的相关需求。就像我们在上面提到的一样，所有对于最终物件的需求都是独立需求，而大多数对于低层物件的需求都是相关需求。然而，对于以备件、研究用零件和质量测试等形式存在的低层物件的需求也可以是独立需求。（110|111）

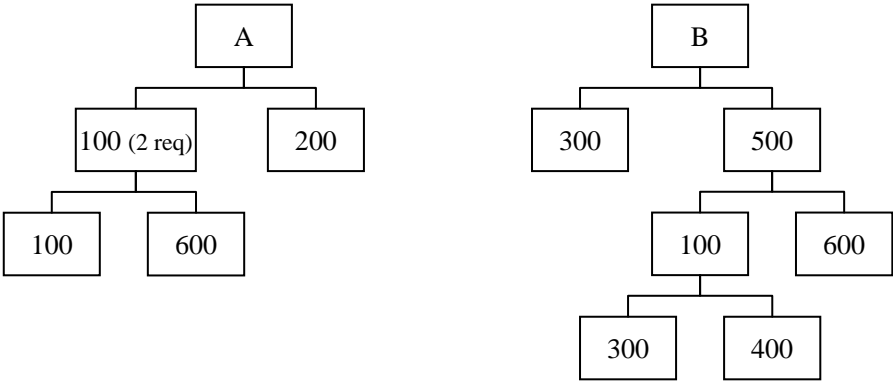


图 3.1 两份物料清单

为了方便MRP的处理过程，物料清单中的每个物件都被赋予一个**低层码（low-level code, LLC）**。这个代码显示了某个特定零件在物料清单中的最低层数。¹最终产品（就是那

¹ 不幸的是，低层码的性质是，物料清单中的零件的层级越低，它的低层码数值就越高。

些不是任何其它物件组成部分的物件)的低层码是零。一个只被最终产品使用的组件的低层码是一。一个只为组件使用的零件的低层码是二,依此类推。例如,在图 3.1 中,零件A和B都是低层码为零的最终产品。对于这些零件的需求都是独立需求。初看上去,似乎零件 100 的低层码应该是一,因为它直接为A使用。然而,因为它同时也是零件 500 (它的低层码是 1) 的一个组成部分,它的低层码被指定为二。同样地,由于零件 300 既要用于低层码为零的零件B的制造,又要用于低层码为二的零件 100 的制造,它的低层码为三。

大多数商业 MRP 软件包都包括一个 **BOM 处理器 (BOM processor)**,它主要是用于维护物料清单,并且自动指定低层码。物料清单处理器的其它功能包括产生“导入 (goes into)”列表(在零件被用到的地方)和物料清单打印。

除了物料清单的信息外,MRP 还需要与来自**主生产计划 (master production schedule, MPS)**的独立需求有关的信息。主生产计划包含**毛需求 (gross requirements)**,称为**现有 (on-hand)**库存量的当前库存状态,已及称为**计划接收量 (scheduled receipts)**的正在执行中的订单(购买而来的和正在制造中的)的状态。

基本的 MRP 步骤很简单。我们将逐一详细讨论其中的每一步。但是,简单地说,对物料清单的每一级,从最终产品开始,MRP 为每个零件主要是做下述工作:

1. 净值计算 (*Netting*): 通过从毛需求中减去持有库存和所有的计划接收量算得**净需求 (net requirements)**。0 层的物件的毛需求来自于主生产计划,而对于那些低一层级的物件的毛需求则来自于之前 MRP 运算的结果。
2. 批量确定 (*Lot sizing*): 把净需求分成合适的**批量 (lot sizes)**而形成加工任务。
3. 时间分段 (*Time phasing*): 用**提前期 (lead times)**来预估加工任务的到期时间,以决定开始时间。
4. **BOM 展开 (BOM explosion)**: 通过使用开始时间,批量和物料清单形成下一层级所需要部件的毛需求。
4. 迭代 (*Iterate*): 重做这些步骤,一直到所有的层级都处理完。(111|112)

当物料清单中的每个零件被处理完后,对下面的层级的需求也就形成了。在开始下一层级之前,MRP 会把这一层的所有零件都处理完。这样做会在一个低层零件被处理前形成它的毛需求。我们会在 3.1.4 节中详细描述其中的每一步。MRP 系统的基本输出是计划投入量 (**planned order releases**)、变更通知 (**change notice**) 和例外报告 (**exception notice**)。这些我们将会 在 3.1.3 节中进行定义。图 3.2 示意了所有的过程。

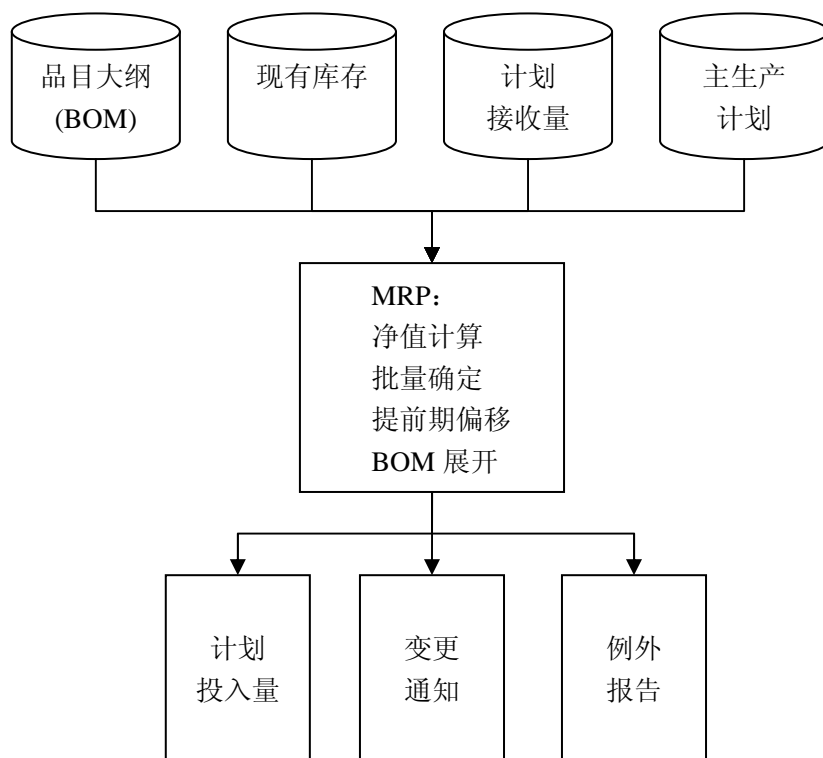


图 3.2 图解 *MRP*

现在我们将用一个简单的例子来说明。假如对零件 A 的需求是由下面的主生产计划中的毛需求给定的：

| 零件 A | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|------|----|----|----|----|----|----|----|----|
| 毛需求 | 15 | 20 | 50 | 10 | 30 | 30 | 30 | 30 |

进一步假设没有计划接收量（这些假设其实有一点欺骗性，我们在后面会讨论它们），而且持有库存量为 30 个单位。我们还假设零件 A 的批量为 75 个单位，并且提前期为 1 个星期。MRP 的处理过程如下：

净值计算。 手头上的 30 个单位将可以满足第一个星期的需求，并且剩余 15 个。剩余的 15 个不能全部满足第二个星期的需求，而留下 5 个单位的空缺。因此，净需求如下所示：（112|113）

| 零件 A | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|-------|----|----|----|----|----|----|----|----|----|
| 毛需求 | | 15 | 20 | 50 | 10 | 30 | 30 | 30 | 30 |
| 持有库存量 | 30 | 15 | -5 | - | - | - | - | - | - |
| 净需求 | | 0 | 5 | 50 | 10 | 30 | 30 | 30 | 30 |

批量确定。 首次未被满足的需求发生在第二个星期。因此，第 1 个**计划产出量**(planned order receipt) 为第二个星期 75（批量）个单位。由于第二个星期只需求 5 个单位，就会有 75 个单位一直持有到第三个星期，而这时的需求为 50。这样对第四个星期就只剩下 20 个，而这

时的需求为 10。当满足星期四的需求之后，剩余的对于满足第五个星期 30 个单位的需求就是不够的。因此，在第五个星期的开头，我们需要另外一批 75 个单位的到达。减去 30 个单位后，我们就有 55 个单位来应付第六个星期，而这时的需求是 30 个，这样，到第七个星期还剩余 25 个单位。而这 25 个单位对于满足 30 个单位的需求是不够的，所以在第七个星期我们需要另外一批 75 个单位的到达。这个批量能够满足第七个星期的剩余需求量（5 个）和第八个星期的 30 个单位的需求。在下面的表格里，我们将表示出上述计算的结果：

| 零件 A | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|-------|----|----|----|----|----|----|----|----|----|
| 毛需求 | | 15 | 20 | 50 | 10 | 30 | 30 | 30 | 30 |
| 持有库存量 | 30 | 15 | -5 | - | - | - | - | - | - |
| 净需求 | | 0 | 5 | 50 | 10 | 30 | 30 | 30 | 30 |
| 计划产出量 | | | 75 | | | 75 | | 75 | |

时间分段。为了确定什么时候启动生产加工任务（如果是在内部制造）或采购订单（若是从其它地方采购），我们只需要从计划产出量的时间中简单地减去提前期，就可以得到计划投入量。使用一个星期的计划提前期，结果如下所示：（113|114）

| 零件 A | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|-------|----|----|----|----|----|----|----|----|----|
| 毛需求 | | 15 | 20 | 50 | 10 | 30 | 30 | 30 | 30 |
| 持有库存量 | 30 | 15 | -5 | - | - | - | - | - | - |
| 净需求 | | 0 | 5 | 50 | 10 | 30 | 30 | 30 | 30 |
| 计划产出量 | | | 75 | | | 75 | | 75 | |
| 计划投入量 | | 75 | | | 75 | | 75 | | |

BOM 展开。一旦我们已经决定了零件 A 的开始时间和数量，产生出对它所有的部件的需求就是一件非常简单的事情了。例如，每单位的零件 A 需要两个单位的零件 100。对于生产零件的零件 100 的毛需求就可以通过简单地加倍零件 A 的计划投入量来进行计算。由零件 A 产生的零件 100 的毛需求必须与其它零件（例如零件 500）产生的零件 100 的毛需求相加，以便得到零件 100 的总毛需求。只要我们按照零件的低层码的顺序（从低到高）进行处理，我们就将在处理每个零件之前得到它的累计毛需求。

3.1.3 MRP 的输入和输出

MRP 的基本输入包括对最终物件的需求预测，相关的物料清单，现在的库存状况，以及明确生产政策所需要的数据。这些数据来源于三个途径：（1）品目大纲文件、（2）主生产计划和（3）库存状态文件。

主生产计划。主生产计划是 MRP 系统的需求来源。它给出了有独立需求的所有零件的数量和到期时间方面的数据。这其中包括对所有最终产品的需求以及对较低层零件的外部需求（例如对备件的需求）。

主生产计划中至少包含：需求数量，每一个采购订单的到期时间。MRP 使用这些信息而取得启动 MRP 程序所需要的毛需求。MPS 使用零件代号与包含其它处理信息的品目大纲进行联系。

品目大纲文件。品目大纲通过零件代号组织起来，并且至少包括零件的描述、物料清单信

息、批量信息以及计划提前期。

一个零件的物料清单文件只列出与制造这个零件直接相关的部件及其相应数量。BOM 处理器使用这些信息可以显示任何物件的完整 BOM，尽管 BOM 处理中不需要如此详细的信息。

通过使用低层码，MRP 可以保证在处理一个零件之前可以把它的需求收集完整。为了证明这样做的必要性，我们可以试想不这样做会发生什么。在我们的例子里，MRP 可能会在处理完零件 A 和 B 之后，在处理零件 500 之前就先处理零件 100。如果是这样，对零件 100 的需要就会缺少由零件 500 所产生的那部分需求。如果我们回头增加一个零件 100 的生产计划，那么结果可能就会有太多的零件 100 的小单生产加工任务，而这若干个小生产加工任务很可能有相同的到期日。因而就很可能无法通过共享某些关键设一次准备来获得规模经济了。使用低层码可以防止这种事情的发生。

实施 MRP 生产需要其它两方面的信息：**批量规则（Lot-sizing rule, LSR）**和**计划提前期（planning lead time）**。批量规则确定加工批量，而合适的加工批量可以平衡降低库存（需要小批量）和增加能力（需要大批量来避免频繁的生产准备）之间的矛盾。如第二章所述，EOQ 和 W-W 都是可能可以作为批量规则。我们将稍后在本章中讨论它们以及其它的规则。

计划提前期是用来决定加工的开始时间。在 MRP 中，这一程序是很简单的：开始时间等于到期日减去计划提前期。因此，如果提前期总是等于计划提前期，MRP 的实施将使工件在需要时恰好准备好（即 JIT）。（114|115）可是，实际提前期是不断变化的而且从来不可能提前知道。因此，在 MRP 系统中决定使用什么样的提前期将是一个非常难的问题，这一点我们将在本章和第五章中进一步讨论。

持有库存。 持有库存数据的储存是按照零件编码编排的，包含的利息有零件存储位置和持有数量。持有库存包括原材料库存，半成品库存（即从原材料开始已经被处理过并储存在库的库存），以及装配库存。持有库存也可能包含关于预备为以后生产预留数量的分配信息。

计划接收量。 该文件包括下达的**采购订单（purchase orders）**和**加工任务（jobs）**。**计划接收量（scheduled receipt, SR）**是已经下达的**计划投入量（planned order release）**。对采购的部件来说，要做的就是生成采购订单并将其送达供应商。对生产的零部件来说，这需要收集所有必须的工艺路线和生产信息，给所需生产的工件配备必须的库存，并将工令送达工厂开始生产。一旦采购订单或者工令送达，计划投入量就从数据库中删除，计划接收量也就随之产生。因此计划接收量就是之前的 MRP 运行之后产生的、正在生产或者还未从供应商那里收到的零件和订单。还未到达库存地点的零件被认为是**在制品（work in process）**的一部分。当零件生产完成（即已经完成工艺路线而转为库存），计划接收量就从数据库中删除，并且持有库存在此前的基础上加上已经完成的部分。向供应商下达采购单后的处理程序和上述基本相同。

每一张计划接收量单据至少包括如下信息：识别码（采购订单号码或者工令号）、到期日、下达日、单位、质量要求和现有数量，当然也可能包含其它信息如价格或成本、工艺路线数据、供应商数据、原材量要求、特殊搬运事项、预期最终数量以及预期完成日期等等。

持有库存信息和计划接收量单据对决定净需求是非常重要的，这一程序通常称为“**需求满足分析（coverage analysis）**”，这涉及到确定持有库存、采购订单和生产工令可以满足多少需求。

如果需求一直不变，并且工件都按时完成，所有现存的计划接收量就会与最终的需求完全对应。遗憾的是，需求是变化的，工件也不总是会按时完成，因此计划接收量有时需要调整，此类调整在**变更通知（change notices）**中说明，我们将在下面进行讲解。

MRP 输出。 MRP 系统的输出包括计划投入量、变更通知和例外报告。计划投入量最终转化成工厂生产的工件。

计划投入量 (planned order release, POR) 至少包括三方面的信息：(1) 料号 (可能只有一个)、(2) 需要的数量、(3) 加工任务的交期。工件和计划投入量不必而且通常也不会与顾客订单一致。实际上，在需要很多普通零件的情况下，计划投入量通常会包括许多不同的生产线，更不用说顾客了。尽管如此，如果所有的生产都按时完成，那么顾客订单也会在相应时间完成。这在 MRP 生产中是自动完成的，我们将在以后详细讲述。(115|116)

变更通知表示对现有工作的改动，如到期日或优先序的变化。将到期日前移称为“**提前 (expediting)**”，而移后则称为“**延期 (deferring)**”。

在大型管理信息系统中，“**例外报告 (exception reports)**”被用来通知使用者期望与现状之间存在差异。这类报告可以用来表明工件数量差异、库存差异以及紧急加工等。

3.1.4 MRP 程序

尽管 MRP 的基本概念很简单，但是细节却非常繁琐。本节我们将详细讲解 MRP 的各个细节以使读者对大多数商业 MRP 系统的基本工作原理有一个系统认识。鉴于此，我们将运用下面一些符号，其定义为：

D_t = 时段 t (如一周) 内的毛需求

S_t = 当前计划在时段 t 内完成的数量 (即计划接收量)

I_t = 时段 t 结束时的预计库存，而持有库存设为 I_0

N_t = 时段 t 的净需求

有了这些符号，我们接下来介绍 MRP 的四个基本步骤：净值计算、批量确定、时间分段和 BOM 展开。

净值计算。 净需求计算 (需求满足分析) 实现了两个重要功能：(1) 调整计划接收量，即把预计会延迟的提前，把预计会提前的延迟；(2) 计算净需求。

大多数 MRP 系统假设所有的计划接收量会在其它新产生的工作完成之前收到。这是现实的：既然计划接收量已经在执行当中，其它新的计划投入量不可能越过已有的计划接收量得以实施。如果计划接收量还未交付给销货方，那么把现有的订单提前要比开始一个新的订单更容易。同样，已经在车间里的计划接收量肯定比现在才开始的订单要先完成。因此，我们可以假设生产所需物料首先来自于持有库存，其次是计划接收量 (忽略其到期日)，最后是新的计划投入量。为了计算第一批计划接收量什么时候达到，我们首先得确定将来什么时候持有库存会减少为零。算式如下：

$$I_t = I_{t-1} - D_t$$

从 $t = 1$ 和 I_0 持有库存开始计算。随着 t 的增加不断计算 I_t ，直到 I_t 小于零。此时计算出的

t 即是第一批计划接收量应当到达的时间。如果第一批计划接收量的到期日与上面计算的 t 不同，那就需要调整。这时就需要发出一个变更通知：计划接收量需要延迟或提前。²一旦计划接收量改变，计划的持有库存就应当反映出这一变化。即：(116|117)

$$I_t(\text{变动后}) = I_t(\text{变动前}) + S_t$$

S_t 是计划接收量移动到 t 时段内的数量。如果 I_t 仍然小于 0，下一批计划接收量也应当移动到 t 时段，继续重复这一过程直到 I_t 非负或者没有其它计划接收量。

一旦计划的持有库存在 t 时段非负，我们继续前面的程序， t 继续增加来计算 $I_t = I_{t-1} - D_t$ ，直到 I_t 又小于 0。我们继续重复这一过程直到计划接收量结束或者达到时间轴的末端。如果发生后一种情况而计划接收量还有剩余，就需要发出变更通知删除多余的生产订单或者推迟到以后，因为现在没有对该计划接收量的需求。更多的情况是在需求完结之前持有库存和计划接收量用完。短缺的库存和计划接收量就是**净需求(net requirements)**。

一旦计划接收量改变，净需求就很容易确定。我们假设 t^* 为计划接收量妥善调整后计划持有库存第一个出现负值的时间段，³那么净需求在 t^* 时刻之前为零，在 t^* 时段等于首次出现负数的计划持有库存，而在 t^* 以后等于毛需求。我们用如下算式表示：

$$CT_{best} = \begin{cases} 0 & \text{若 } t < t^* \\ -I_t & \text{若 } t = t^* \\ D_t & \text{若 } t > t^* \end{cases}$$

然后得到的净需求将用于确定批量。

在进入批量大小确定阶段之前，我们思考一个例子来形象地说明上述需求满足分析程序。表 3.1 中包含了主生产计划关于 A 部件的毛需求、3 批计划接收量和当前持有库存的数量。(117|118)

表 3.1 输入数据示例

| 零件 A | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|-----------|----|----|----|-----|----|----|----|----|
| 毛需求 | 15 | 20 | 50 | 10 | 30 | 30 | 30 | 30 |
| 计划接收量 | 10 | 10 | | 100 | | | | |
| 调整后的计划接收量 | | | | | | | | |
| 计划持有库存 | 20 | | | | | | | |
| 净需求 | | | | | | | | |
| 计划产出量 | | | | | | | | |
| 计划投入量 | | | | | | | | |

² 当然，这种到期日的自动变更只发生在数据库中，知道有人将其付诸实施。变更通知的用途是将这个信息传递给负责保证工作按时完成的人。这在理论上都十分简单，但是许多情况下都可能把工作提前到一个不可能完成的时间。这就看出了 MRP 数据库中的数据并不总是能反映车间的真实情况。

³ 注意如果我们不先调整计划接收量，出现这种情况的时间段可能就不止一个。即若干个净需求之间夹了一个计划接收量。

我们从计算计划持有库存开始。假设初始库存为 20 件，减去阶段 1 的毛需求 15 件，剩余 5 件持有库存。注意这里因为我们总是先使用持有库存再使用计划接收量，因此我们不考虑阶段 1 的计划接收量 10 件。

往后移动到阶段 2，我们可以看到毛需求 20 大于持有库存 5，因此我们发出变更通知，将阶段 1 的 10 件计划接收量推迟到阶段 2。但是，就算如此我们还是只能提供 15 件产品，与需求的 20 件短缺 5 件。所以我们将阶段 2 的第二批计划接收量 10 件加上，总量就达到 25 件。注意这里计划接收量已经在阶段 2 使用，因此不需要发出变更通知。在将前两批计划接收量调整至阶段 2 并减去毛需求后，剩余的持有库存为 5 件。由于这一数量远远少于第三阶段的需求 50 件，我们需要发出提前通知将第三批计划接收量的到期日从阶段 4 提前到阶段 3，此时剩余持有库存为 55。在有些系统中，加工批量可以被分割，从而只提前生产需要的数量。但本例中我们将整个批量的生产都提前。提前后的计划接收量满足阶段 4 的需求 10 后剩余 45，满足阶段 5 的需求 30 后剩余 15，而阶段 6 的需求超过了计划持有库存，并且此时没有计划接收量可供调整。因此，第一次没有满足的需求发生于阶段 6 并等于 15 件。表 3.2 总结了通过需求满足分析的计算过程所得到的计划持有库存。

现在净需求计算起来就容易了。如表 3.2 所示，因为计划持有库存大于 0，阶段 1 的净需求为 0，而阶段 6 净需求则为 15，即此时计划持有库存的负数。而阶段 7 和 8 的净需求等于毛需求，都是 30。

表 3.2 调整后的计划接收量、计划持有库存和净需求

| 零件 A | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|-----------|----|----|----|-----|-----|----|-----|----|----|
| 毛需求 | | 15 | 20 | 50 | 10 | 30 | 30 | 30 | 30 |
| 计划接收量 | | 10 | 10 | | 100 | | | | |
| 调整后的计划接收量 | | | 20 | 100 | | | | | |
| 预计持有库存 | 20 | 5 | 5 | 55 | 45 | 15 | -15 | - | - |
| 净需求 | | | | | | | 15 | 30 | 30 |
| 计划产出量 | | | | | | | | | |
| 计划投入量 | | | | | | | | | |

形成批次。 当我们计算出了净需求，就需要安排生产来满足需求。因为 MRP 假设需求是确定的，且随时间而变化，这恰恰就是我们在第二章中用 Waner-Whitin 算法取得最优解的那个问题。我们将在 3.1.6 节中对这个问题及其他一些确定批量的技术进行讨论。为了清楚地说明基本的 MRP 计算过程，我们在这里集中讲述两个非常简单的批量规则。（118|119）

最简单的批量规则也就是我们熟知的**批对批法（lot for lot）**，在该法则下，某段时间内应当生产的产品数量等于该时段的净需求。该规则用起来要比例 3.1.2 中所使用的固定批量法容易，这也与 JIT 思想（仅生产需要数量的产品）相符合。

另一种简单的批量规则称为**固定订货期法（fixed order period, FOR）**，又称为**周期订货量法（period order quantity）**。该法则试图通过合并 P 个时间段的净需求来减少准备次数。当 P = 1 时，定期用量法与批对批法相等。

表 3.3 计划产出量和计划投入量

| 零件 A | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|-----------|--|----|----|-----|-----|----|----|----|----|
| 毛需求 | | 15 | 20 | 50 | 10 | 30 | 30 | 30 | 30 |
| 计划接收量 | | 10 | 10 | | 100 | | | | |
| 调整后的计划接收量 | | | 20 | 100 | | | | | |

| | | | | | | | | | |
|--------|----|---|---|----|----|----|-----|----|----|
| 预计持有库存 | 20 | 5 | 5 | 55 | 45 | 15 | -15 | - | - |
| 净需求 | | | | | | | 15 | 30 | 30 |
| 计划产出量 | | | | | | | 45 | | 30 |
| 计划投入量 | | | | | 45 | | 30 | | |

再回到我们前面的例子，假设部件 A 和 B 的批量规则是固定订货期法且 $P=2$ ，而其他部件则采用批对批法。那么，我们计划在阶段 6 和阶段 8 分别接收 A 部件 45 件（包括阶段 6 和 7 的净需求）和 30 件（我们不能将超出计划范围的合并进去）。根据批量规则计算后的结果如表 3.3 所示。

时间分段。 几乎所有的 MRP 系统都假设制造一个零件的时间是固定的，尽管有一些系统允许将计划提前期作为工件大小的一个函数。但是，不管具体情况如何，MRP 将提前期看成零件或工件的一个属性，而不是车间工作的状态。我们在后面会看到这会引起很多问题。

如果我们回到前面的例子并假设零件 A 的计划提前期是两个时段，我们就可以计算出计划投入量，如表 3.3 所示。

BOM 展开。 表 3.3 显示了生产零件 A 的最终结果，而零件 A 又是由两个零件 100 和 1 个零件 200 组成，因此 A 零件的计划投入量连带又会产生相应的对零件 100 和 200 的毛需求。具体的说，在阶段 4 和 6 我们分别需要 90 件（每个零件 A 需要两个）和 60 件零件 100，同样，还分别需要 45 件和 30 件的零件 200。这些需求应当加到已经计算好的零件需求计划当中（例如在 MRP 中已经处理了其他某些需要它们作为组成部分的零件）。为了对此更为形象地加以说明，我们继续使用上面的例子。

下一步是处理其他低低层码的零部件，我们选取零件 B 的生产为例。假设零件 B 的主生产计划如下：(119|120)

| t | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|----|----|----|----|----|----|----|----|----|
| 需求 | 10 | 15 | 10 | 20 | 20 | 15 | 15 | 15 |

除此以外，我们还假设了零件 B、100、300 和 500 的库存和零件数据（为了简便起见，我们在此忽略了对零件 200、400 和 600 的 MRP 处理）：

| 零件号 | 持有库存 | 计划接收量 | | 批量规则 | 提前期 |
|-----|------|-------|-----|---------|-----|
| | | 到期日 | 数量 | | |
| B | 40 | 0 | | FOP 2 周 | 2 周 |
| 100 | 40 | 0 | | LFL | 2 周 |
| 300 | 50 | 2 | 100 | LFL | 1 周 |
| 500 | 40 | 0 | | LFL | 4 周 |

因为没有对零件 B 的计划接收量，该部分的 MRP 计算很简单，表 3.4 显示了计算结果。

表 3.4 零件 B 的 MRP 处理

| | | | | | | | | |
|------|----|----|----|----|----|----|----|----|
| 零件 B | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| 毛需求 | 10 | 15 | 10 | 20 | 20 | 15 | 15 | 15 |

| | | | | | | | | |
|-----------|----|----|----|---|-----|----|----|----|
| 计划接收量 | | | | | | | | |
| 调整后的计划接收量 | | | | | | | | |
| 预计持有库存 | 40 | 30 | 15 | 5 | -15 | - | - | - |
| 净需求 | | | | | 15 | 20 | 15 | 15 |
| 计划产出量 | | | | | 35 | | 30 | 15 |
| 计划投入量 | | | 35 | | 30 | | 15 | |

我们已经完成了对低层码为零的零件（即零件 A 和 B）的处理，余下的只有零件 500 的低层码为一，因此我们接下来处理零件 500。

对零件 500 的需求仅来自于零件 B（即零件 A 不需要零件 500），每个零件 B 需要一个零件 500，因此零件 B 的计划投入量也就是零件 500 的毛需求。同样，这里也没有计划接收量，MRP 的处理过程见表 3.5。

表 3.5 零件 500 的 MRP 计算

| 零件 500 | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|-----------|----|-----|----|---|-----|---|----|---|---|
| 毛需求 | | | 35 | | 30 | | 15 | | |
| 计划接收量 | | | | | | | | | |
| 调整后的计划接收量 | | | | | | | | | |
| 预计持有库存 | 40 | 40 | 5 | 5 | -25 | - | - | - | - |
| 净需求 | | | | | 25 | | 15 | | |
| 计划产出量 | | | | | 25 | | 15 | | |
| 计划投入量 | | 25* | 15 | | | | | | |

*表示延迟投入生产

因为零件 500 的提前期为 4 周，在第四周之前没有足够的时间来完成第一批的 25 件，因此在第一周（越早越好）安排计划投入量并在例外报告中注明该批生产可能来不及按时完成。（120|121）

接下来我们来看低层码为 2 的零件 100，零件 100 需要满足两方面的需求：每个 A 零件需要 2 个零件 100，每个零件 500 需要 1 个零件 100，没有计划接收量。其 MRP 处理如表 3.6 所示。

表 3.6 零件 100 的 MRP 计算

| 零件 100 | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|--------------|----|----|----|---|-----|---|----|---|---|
| 来自 A 的需求 | | | | | 90 | | 60 | | |
| 来自零件 500 的需求 | | 25 | 15 | | | | | | |
| 毛需求 | | 25 | 15 | | 90 | | 60 | | |
| 计划接收量 | | | | | | | | | |
| 调整后的计划接收量 | | | | | | | | | |
| 预计持有库存 | 40 | 15 | 0 | 0 | -90 | - | - | - | - |
| 净需求 | | | | | 90 | | 60 | | |
| 计划产出量 | | | | | 90 | | 60 | | |
| 计划投入量 | | | 90 | | 60 | | | | |

低层码为三的只有零件 300，对它的需求来自于零件 B 和 100，而且在第二周还有 100 件的计划接收量，由于恰好是在第一个没有满足需求的计划期到达，因此不需要进行调整。其 MRP 处理如表 3.7 所示。

表 3.7 零件 300 的 MRP 计算

| 零件 300 | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|--------------|----|----|-----|----|-----|----|----|---|---|
| 来自 B 的需求 | | | 35 | | 30 | | 15 | | |
| 来自零件 100 的需求 | | | 90 | | 60 | | | | |
| 毛需求 | | | 125 | | 90 | | 15 | | |
| 计划接收量 | | | 100 | | | | | | |
| 调整后的计划接收量 | | | 100 | | | | | | |
| 预计持有库存 | 50 | 50 | 25 | 25 | -65 | - | - | - | - |
| 净需求 | | | | | 65 | | 15 | | |
| 计划产出量 | | | | | 65 | | 15 | | |
| 计划投入量 | | | | 65 | | 15 | | | |

现在我们已经对所有相关的零件进行了 MRP 处理（零件 200 和 400 的处理过程与上述过程完全相同）。(121|122) 表 3.8 总结了 MRP 系统经过上述计算后的输出结果。在每一份变更通知中，系统报告了受影响的零件号及其数量、改变前的到期日、改变后的到期日以及是延迟还是提前。而新的计划投入量则报告了下达时间、（新的）到期日、下达的数量以及是否可能迟到。

表 3.8 MRP 输出总结

| 处理 | 零件号 | 改变前到期日 或下达日 | 新到期日 | 数量 | 通知 |
|-------|-----|----------------|------|-----|----|
| 变更通知 | A | 1 | 2 | 10 | 延迟 |
| 变更通知 | A | 4 | 3 | 100 | 加快 |
| 计划投入量 | A | 4 | 6 | 45 | OK |
| 计划投入量 | A | 6 | 8 | 30 | OK |
| 计划投入量 | B | 2 | 4 | 35 | OK |
| 计划投入量 | B | 4 | 6 | 30 | OK |
| 计划投入量 | B | 6 | 8 | 15 | OK |
| 计划投入量 | 100 | 2 | 4 | 90 | OK |
| 计划投入量 | 100 | 4 | 6 | 60 | OK |
| 计划投入量 | 300 | 3 | 4 | 65 | OK |
| 计划投入量 | 300 | 5 | 6 | 15 | OK |
| 计划投入量 | 500 | 1 | 4 | 25 | 迟到 |
| 计划投入量 | 500 | 2 | 6 | 15 | OK |

3.1.5 MRP 系统中的特殊问题

到目前为止，我们主要集中于 MRP 处理过程的机理。现在我们来考虑影响 MRP 运行效果的一些技术性问题，特别是当系统运行出现意外时如何改善绩效。(122|123)

更新频率 (Updating Frequency)。 影响 MRP 系统有效性的一个关键因素是更新频率。如果

更新太过频繁，车间就会出现例外报告泛滥以及经常改变计划投入量的情况。⁴而另一方面，如果更新频率太低，我们就常常会按照过时的计划进行生产。因此，在设计MRP系统时应当平衡及时性和稳定性这两方面的需要。

固定计划订单 (Firm Planned Orders)。 频繁改变生产计划可能会使生产系统变得非常不稳定，这使得管理者难以有效安排工人换班和生产准备。因此，最大限度地降低变化对生产计划的影响是非常必要的，一种方法就是使用**固定计划订单 (Firm Planned Orders)**。固定计划订单是指对计划投入量实施冻结，即令该计划量按期执行而忽略系统的变化。结果，在MRP 处理中总是把固定计划订单当作计划接收量来看待（即在需求满足分析中必须包括固定计划订单量）。通过将一定时间内的计划投入量转换成固定计划订单，生产计划就可以变得更加稳定。这在短期对于管理控制尤其重要。固定计划订单对于减少系统**紧张 (nervousness)** 同样有用，这方面内容我们将在后面更为详细地讲解。

MRP 中的故障检修 (Troubleshooting in MRP)。 一个叫墨菲的哲人曾经说过：“如果事情可以变糟的话，它往往就真的会变糟。”在 MRP 系统中，有很多事情可能变糟。生产任务可能延迟，零件可能报废，需求可能变化等等。因此，多年以来 MRP 系统已经获得了许多特性来帮助计划者应对变化的情况。这样的例子包括挂钩 (pegging) 和从下而上再计划 (bottom-up replanning)。

挂钩 (pegging) 使得计划制定者可以看到现有计划投入量所对应的需求来源。办法就是在每个零件的毛需求与其所有的需求来源之间建立关联。我们以表 3.7 中显示的第三周对零件 300 的 65 件计划投入量为例，挂钩技术会将这一需求连接到第四周对 60 件零件 100 和 30 件零件 B 的需求，反过来这些需求又会连接到它的需求来源，即零件 B 对应主生产计划，零件 100 对应第六周生产零件 A 的需求（见表 3.6）。

挂钩技术可以应用在**从下而上再计划 (bottom-up replanning)**。我们以下面的例子来解释。假设我们发现第二周到期的 100 件零件 300 的计划接收量无法到达了（有人在文件柜后面发现了早就应当发给供应商的采购订单），当然恰当的做法是马上下订单并打电话给供应商看是否可以加快交付订单。如果回答是不可能，我们就可以运用从下而上再计划技术来考察延迟交货的影响。

从表 3.7 我们可以看出对毛需求的影响是第二周的 125 件，如果计划接收量不会到达，那么我们就只有现有的 50 件产品来满足需求，剩余 75 件没有满足。在这 125 件中，35 件是生产零件 B 所需（低层码为零），90 件是零件 100 所需（低层码为二）。如果优先供应最低层次的零件（理由是最低层次的产品有可能造成最大的扰动），那么我们可以看到只能供应第二期所需的 90 件零件 100 中的 50 件。通过挂钩技术可以进一步发现这些需求来源于零件 A 的 90 件需求，而现在只能向客户提供 50 件产品 A。这个时候我们可能就会想联系客户看是否可以先交付 50 件 A 产品，稍后再交付剩余的 40 件。（123|124）

除此以外，我们也可以将现有的 50 件库存先供应 B 零件的生产（这里的理由是先供应能够产生利润的产品）。这样做的话，可以满足生产 B 零件所需的 35 件需求，此外还剩余 15 件可以供应生产 90 件零件 100 的需要。同样将这些需求连接到其源需求，可以看到第 4 期所需的 90 件 A 零件还有 75 件不能供应。如果主生产计划中对零件 B 的需求是实际客户的需求而零件 A 只是预测需求的话，那么我们可以考虑先满足 B 零件。当然还有一种选择就是将这 50 件分别供应给零件 B 和零件 100。怎样选择取决于客户及其接受延迟供货的意愿。

⁴ 在过去，当计算机系统的存储空间很小、运算速度也很慢的时候，计算机处理的成本也会制约更新频率。然而，随着最近几年计算机性能急剧提升，处理能力对于决定更新频率已经不那么重要了。

除了利用挂钩，我们还可以消除 100 件零件 300 的计划接收量并更新 MRP。这将产生第一周的计划投入量以及一个例外通知，告知该项生产预计会迟于计划完成。尽管如此，对 MRP 的更新不能决定哪些客户订单将因此而被延迟。从下而上再计划和挂钩技术能帮助计划制定者做到这一点。利用固定计划订单可以使计划制定者绕过标准的 MRP 过程以制定一个补救计划。

3.1.6 MRP 中的批量规则

为了展示基本的 MRP 处理过程，我们已经介绍了两种简单的批量规则：固定订货期法和批对批法。在这一节里，我们将围绕批量规则问题展开讨论，并介绍其它更为复杂的批量规则。

批量规则问题解决的是小批量和大批量之间的平衡问题，小批量倾向于增加准备成本（材料、跟踪和人力成本等）并减少生产量，而大批量倾向于增加库存。

回顾一下第二章中的解决批量问题的 Wagner-Whitin (W-W) 方法，该方法假设了固定的生产能力以及准备和库存会增加成本。在这些假设下，我们说明了可以运用 W-W 算法取得批量问题的最优解。当然，该方法存在的问题在于是否 *可以* 知道准备成本和库存持有成本的大小以及生产能力是否会受到约束。就像有人评论准备成本说：“我居然还要给一台机器开支票。”在许多情况下，准备“成本”被用作生产能力的限制。这里的关键在于设计一个批量规则，以使更高的准备成本产生更大的批量（例如 EOQ）。因为更大的批量需要更少的准备次数，这就节约了生产能力。相反，当生产能力不紧张的时候，更小的准备成本可以减少批量规模（进而减少库存），代价是更多次数的生产准备。因此，通过调整准备成本，计划制定者可以用库存来换取生产能力。

遗憾的是，当生产能力受限时，所谓的 W-W 性质即只有当库存降至 0 时才开始生产不是最佳的。不仅如此，许多提到的批量规则都有 W-W 性质，并且在评估其绩效时通常将其和 W-W 算法进行比较。因此，尽管许多假设在现实情况下可能无效，但是大多数批量规则的设计者还是接受了 W-W 这个范例。有趣的是，我们知道还没有哪个商业 MRP 软件包使用了 W-W 算法，理由常常是它太过复杂或太慢。但是随着快速计算机的出现，速度不再是问题，有效的 W-W 算法在现代个人计算机上可以运行得很快。一个更可能的原因可以见之于有人观察得出的结论：“人们宁愿要接受一个他们解决不了的问题，也不愿要接受一个他们并不了解的答案。”不管理由是什么，现在又有人提出了许多替代的批量规则，并且可以在大多数商业 MRP 系统中以多种形式来运行。我们这里讨论一些使用更为普遍的方法。（124|125）

批对批法 (Lot-for-Lot)。 如我们前面所说，LFL 是最为简单的批量规则——在时段 t 只生产该时段需要的产品。因为该规则在任何时段末都不会产生库存（在 MRP 假设下），因此该方法可以最大限度地降低库存（假设在每一个时段都可以生产所需的产品）。然而，在 W-W 范例中，因为在每一期都有对应需求的生产准备，因此该方法也就带来了最大的准备成本。尽管有这一缺陷，LFL 在许多方面还是很有吸引力的。首先它简单；其次它与 JIT 只在需要的时候生产需要的产品的思想相符（见第四章）；最后，因为其程序不会把有些时段需求累积到一起而在许多时段不生产产品，所以它更倾向于得到平滑的生产排程。在准备时间（成本）达到最小的情况下，这一方法也许是最好的。

固定批量法和经济订货批量。 第二种非常简单的方法是不管订单什么时候下达都按照一个预先确定的批量。我们在第一个例子中使用固定批量这一方法。运用这一方法通常有两个简单的原因。

首先，当有固定容量的转载工具来运输工件时，生产与这些运载工具容量相应数量的工件就很有意义。在有些情况下，工厂在不同时候使用不同大小的装载工具。例如挡板用的转载工具的搬运数量通常就比火花塞小。为了避免剩余，协调装载工具容量与工件数量就很有必要，一种方法就是批量大小选择 2 的幂级数（1、2、4、8、16 等）。

其次，确定加工数量会影响准备次数。因为基本的权衡在于准备成本和库存持有成本，选择一个合适的固定订货批量的问题就与第二章中讨论的经济批量问题非常相似。主要的差别在于经济批量模型的前提假设是有恒定的需求速率，而在 MRP 系统中，需求有可能是不恒定的。尽管如此，我们仍然可以通过用估计的平均需求代替该模型中的恒定需求来使用 EOQ 模型。那么，如果用 A 来代表准备成本，h 代表年平均库存成本，我们可以运用第二章中的 EOQ 公式

$$Q = \sqrt{\frac{2AD}{h}}$$

来计算固定订货批量 Q。就像之前讨论的，我们可能想将该数量取为最接近的 2 的幂级数。可以通过调整 A/h 的比率来产生需要的准备频率。将 A/h 调高可以减少准备频率，而减少 A/h 则可以增加准备频率。经过数次实验之后，我们就可以找到一个与生产线的生产能力相容的 A/h 值。当然，因为该值取决于实际的订单，所以它可能会频繁地变化。

与批对批法不同，固定批量法（不管使不使用 EOQ 来得到订货数量）不会拥有仅当库存降为零时才生产 W-W 性质。这意味着可能导致相应的库存成本并且还不能避免生产准备，很显然这是很没有效率的（在 W-W 假设下）⁵。（125|126）

尽管如此，我们可以稍微改变一下规则，即只考虑与一个或多个时段的需求相等的加工数量，然后选择与想要的固定加工数量最相近的那个。这样做就重现了 W-W 性质。思考下面这个例子：假设固定订货批量为 50 件，而净需求为：

| | | | | | | | | |
|-----|----|----|----|----|----|----|----|----|
| 净需求 | 15 | 15 | 60 | 65 | 55 | 15 | 20 | 10 |
|-----|----|----|----|----|----|----|----|----|

然后，为了保持 W-W 性质，计划产出量应为：

| | | | | | | | | |
|-------|----|--|----|----|----|----|--|--|
| 计划产出量 | 30 | | 60 | 65 | 55 | 45 | | |
|-------|----|--|----|----|----|----|--|--|

在第 1 期，30 比 15 更接近于 50，所以我们订购两期的需求量而不是一期的。在时段 3，60 比 125 更接近 50，所以我们订购一期的需求而不是两期的，等等。

固定订货期法。 固定订货期法（FOP）在 3.1.4 节关于 MRP 处理过程的举例中使用过。其操作很简单：如果你准备在时段 t 生产，那么就生产 $t, t+1, \dots, t+P-1$ 这些时段所有的需求量，此处 P 是该策略的一个参数。如果 $P = 1$ ，该策略就是 LFL，因为只生产当前期的需求。因为所有的生产量都是给定一组时段的确切需求，该策略就具有了 W-W 性质。

该策略尽管很简单，却有一些微妙之处。该策略没有说每隔 P 期就会生产，如果有些时段没有需求，就跳过。我们来看 $P = 3$ 时下面的例子：

| | | | | | | | | | |
|-----|---|----|----|---|---|----|----|----|----|
| 时段 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 净需求 | | 15 | 45 | | | 25 | 15 | 20 | 15 |

⁵ 当然，事实上我们可能不会计划到恰好下一个订单量到达的时候库存用完。不仅如此，我们甚至可能会使用安全库存（在下一部分讨论）来提供缓冲而对于周期库存（即预计使用的库存）则坚持 W-W 性质。

| | | | | | | | | | |
|-------|--|----|--|--|--|----|--|--|----|
| 计划产出量 | | 60 | | | | 60 | | | 15 |
|-------|--|----|--|--|--|----|--|--|----|

既然第 1 期没有需求我们就跳过它，第一次需求发生在第 2 期，因此我们合并第 2、3、4 期（注意第 4 期没有需求）的需求得到第 2 期的计划投入量 60 件。我们再次跳过第 5 期，合并第 6、7、8 期的需求得到第 6 期的计划投入量 60 件。最后第 9 期为 15 件，然后由于已经到了时间轴的末端就不用再往下考虑了。（126|127）

确定“最优”P 值的一种方法是运用 EOQ 公式和平均需求量，这与固定批量规则的用法类似。在之前的例子中，9 期的总需求为 135 件，可得其平均需求为 15 件。假设准备成本和每个时段的运输成本分别为 150 元和 2 元，我们可以计算 EOQ 公式得出：

$$Q = \sqrt{\frac{2AD}{h}} = \sqrt{\frac{2 \times 150 \times 15}{2}} = 47.4$$

然后就可以计算得出：

$$p = \frac{Q}{D} = \frac{47.4}{15} = 3.16 \approx 3 \text{ 期}$$

当然，运用这种方法计算的 P 值的有效性受到 EOQ 限制条件的局限，这部分内容我们在第二章中已经提到过。

部件-时期平衡法（Part-Period Balancing, PPB）。 PPB 是一种将 W-W 范例的假设与 EOQ 机制结合在一起的策略。EOQ 对于批量问题的解有一个性质就是它令平均库存持有成本等于准备成本。

PPB 思想即是平衡（即使之相等）库存成本和准备成本。为了描述这个思想，我们需要将**部件-时期**的概念定义为一批生产的零件数量乘以其作为库存持有的期数。例如说，不管是存储时间为 10 期的一个零件，还是存储时间为 2 期的 5 个零件，或者存储时间为 1 期的 10 个零件，都表明其零件-期数为 10，也就具有相同的库存成本。零件-期数平衡力求使库存成本与准备成本尽可能接近。我们可以借助前述例子的数据来说明这个问题。

我们只考虑能够保留 W-W 性质的那些数量，这样我们选择的范围就相对要小得多。由于第 1 期没有需求，因此第 1 期不生产，第 2 期可选择的产量为 15 件（只生产第 2 期所需）、60 件（生产第 2、3 期的需求）和 85 件（生产第 2、3、6 期的需求）等。下表显示了零件-期数和相应的成本：

| 时段 2 生产数量 | 准备成本 | 零件-期数 | 库存成本 |
|-----------|------|--------------------------|------|
| 15 | 150 | 0 | 0 |
| 60 | 150 | $45 \times 1 = 45$ | 90 |
| 85 | 150 | $45 + 25 \times 4 = 125$ | 290 |

因为 90 元是现有选择中与 150 元最接近的数值，所以我们选择在第 2 期生产 60 件。第 3、4、5 期没有需求不生产。第 6 期的选择为 25、40、60 和 75 件，同样我们将计算过程用表显示出来：（127|128）

| 时段 6 生产数量 | 准备成本 | 零件-期数 | 库存成本 |
|-----------|------|-------------------------|------|
| 25 | 150 | 0 | 0 |
| 40 | 150 | $15 \times 1 = 15$ | 30 |
| 60 | 150 | $15 + 20 \times 2 = 55$ | 110 |

| | | | |
|----|-----|----------------------|-----|
| 75 | 150 | $55+15 \times 3=100$ | 200 |
|----|-----|----------------------|-----|

第 6 期生产 60 件时的库存成本与 150 最为接近，因此生产数量为 60 件，这个数量覆盖了第 6、7、8 期的需求，剩下的只有第 9 期的 15 件需求。注意由此得出的结果恰好与 FOP 策略得出的生产计划相同。

其他方法。 研究人员还提出了其他的批量规则方法，大多数都是根据 W-W 准则试图提供一个接近最佳的解决方案。正如我们所讨论的，W-W 准则是否恰当仍然是一个有争议的话题。贝克（Baker）于 1993 年对许多现存的批量规则作了很好的综述。

最后我们要注意的，尽管 W-W 算法在有些情况下是最佳的，在实际应用中其他规则可能会表现得更佳。例如，Bahl 等（1987）在一篇关于批量规则的文献综述中说道：如果没有对其进行修改以赋予 W-W 性质，固定批量规则在具有生产能力限制的多级系统中确实要比那些拥有 W-W 性质的批量规则要表现得更好。他们总结道：“强加的 W-W 性质在现实生产中可能并不实用，原因在于其他大多数批量规则力求避免的物料剩余反而有利于最终产品的准时交付，因而变成了一种优势。”这一点的确是非常有意义的，因为这些剩余物料成为了安全库存的一种形式。关于这一问题我们将在接下来的一节进行探讨。

3.1.7 安全库存与安全提前期

运营管理研究人员对 MRP 系统中安全库存和安全提前期的作用已经进行了长期的争论。奥利基（Orlicky）认为它们在 MRP 系统中没有立足之地，只有对于最终产品可能还有意义。他相信 MRP 系统的运行已经完全能够应付低层物料了。自从奥利基所处时代以来，许多研究人员都对其观点不予赞同，因为 MRP 处理的是确定性情况，而逻辑上来说实际生产中还必须处理不确定的和随机的情况。

不确定性的来源多种多样。首先，除纯粹接单生产（make-to-order）的系统外，在所有的系统中，不管是需求量还是需求的时间都不是事先可以确切预知的。其次，由于机器故障、质量问题以及人员变动等问题的存在，生产时间几乎总是受制于一些变量。第三，生产数量具有不确定性，因为产出损失（yield loss）或者是质量问题（fallout），产出零件的数量总要比预先投入的物料要少。

安全库存（safety stock）和安全提前期（safety lead time）可以用来应对这些问题。Vollmann 等（1992）表示：安全库存应当用来应对生产和需求量的不确定性，而安全提前期则应当用来应对生产和需求时间的不确定性。（128|129）

在 MRP 系统中建立安全库存是非常方便的。假设我们希望为零件 B（参见表 3.4）保有 10 件安全库存，那么我们跟以前一样计算最初的净需求时就应当减去这额外的 10 件安全库存。从表 3.9 可以看出，预计持有库存减去安全库存首次出现负数是在第 3 期（与之之前的第 4 期对应）。

因此，最初的计划投入量是 5 件，以达到保有的安全库存水平，然后再加上 20 件以满足实际需求。

在 MRP 计算中引入安全提前期则稍微有点不同。如果名义提前期是 2 周并且我们需要 1 周的安全提前期，那么可以用两步来进行偏移：第一步是根据计划产出量的时间（即到期日）偏移得到安全提前期，第二步是运用通常的 MRP 方法来获得计划投入量的时间。我们用前例中表 3.10 中的数据来展示如何运用 1 周的安全提前期。（129|130）

表 3.9 零件 B 的安全库存的 MRP 计算

| 零件 B | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|-----------|----|----|----|----|----|----|----|----|----|
| 毛需求 | | 10 | 15 | 10 | 20 | 20 | 15 | 15 | 15 |
| 计划接收量 | | | | | | | | | |
| 调整后的 SR | | | | | | | | | |
| 预计持有库存 | 40 | 30 | 15 | 5 | - | - | - | - | - |
| 预计持有库存-SS | 30 | 20 | 5 | -5 | - | - | - | - | - |
| 净需求 | | | | 5 | 20 | 20 | 15 | 15 | 15 |
| 计划产出量 | | | | 25 | | 35 | | | |
| 计划投入量 | | 25 | | 35 | | 30 | | | |

表 3.10 零件 B 的安全提前期 MRP 计算

| 零件 B | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|-----------|----|----|----|----|-----|----|----|----|----|
| 毛需求 | | 10 | 15 | 10 | 20 | 20 | 15 | 15 | 15 |
| 计划接收量 | | | | | | | | | |
| 调整后的 SR | | | | | | | | | |
| 预计持有库存 | 40 | 30 | 15 | 5 | -15 | - | - | - | - |
| 净需求 | | | | | 15 | 20 | 15 | 15 | 15 |
| 计划产出量 | | | | | 35 | | 30 | | 15 |
| 调整后的计划产出量 | | | | 35 | | 30 | | 15 | |
| 计划投入量 | | 35 | | 30 | | 15 | | | |

除通常的 MRP 计算之外，还增加了一个步骤即“调整后的计划产出量”，这一步根据 1 周的安全提前期将接收日期相应偏移 1 周。注意这与简单地延长计划提前期对计划投入量的影响是一样的。尽管如此，在拥有安全提前期的系统中，生产的到期日比没有安全提前期的系统要早。安全提前期对单个零件的影响是非常简单的，将零件提前一周准备好意味着除非运送时间耽误一周以上的时间，否则生产就不会耽搁。但是如果我们考虑多种零件和生产线的情况比我们想象的要更为复杂。

例如，假设一个工厂生产一种零件需要 10 种组件同时放在组装线旁，再假设可以运用均值为 3 周、标准差为 1 周的正态分布合理接近实际的制造提前期。为了保证良好的服务水平，我们需要组装线在 95% 情况下都能准时开始生产，如果 s 为各组件的服务水平（按时交货的概率），那么 10 个组件全部按时到货的概率为（假设相互独立交货）：

$$\text{Prob}\{\text{准时启动组装}\} = s^{10}$$

因为我们需要该概率等于 0.95，可以解得：

$$s = (0.95)^{1/10} = 0.9949$$

由于制造提前期呈正态分布，这就表示大约均值之上 2.6 个标准差，即 5.6 周，大约是计划提前期均值的两倍。

当然，该分析假设这 10 个组件是彼此独立地运送到组装线的，但如果所有的组件都是在同一个车间生产出来的话这显然就不成立。然而此处的问题至少说明：如果我们想要保证一条组装线的服务水平，那么组件的服务水平应当更高。

总之，尽管安全库存和安全提前期在 MRP 系统中可能有用，但我们必须认识到所有的程序都是依存于 MRP 系统这样一个事实。安全库存需要在顾客需求之外增加一定数量的生

产，而安全提前期则要求到期日比实际需要的要早。这两者都会使可供销售量（available-to-promise）计算结果（用来向顾客允诺交货，将在后面讨论）的准确性降低。过量的安全库存和过长的安全提前期都会导致顾客认为生产计划不可行而转寻他处，尽管生产计划实际上是可行的。此外，总会存在这样一种风险：一旦安全库存或者安全提前期被使用者发现，就会产生一套非正式的“真实”数量和到期日。这种行为可能颠覆正式的生产系统并降低其绩效。

3.1.8 应对产出损失 (Accommodating Yield Losses)

上面的讨论和例子阐述了如何应对需求和时间的不确定性。然而，如果要应对生产过程中的随机零件报废即产出损失却需要另外的计算方法。假设净需求是 N_t 件，平均产出率是 y ，再假设对于这里讨论的例子， N_t 是一个很大的数，因为我们不必担心数量取整的问题。

因此，如果我们按照 $N_t(1/y)$ 件产品的量投入物料进行生产，平均来讲我们就会得到 N_t 件产品，也就是净需求的量。但是如果 $N_t(1/y)$ 是一个很大的数，那么我们很有可能不能恰好得到 N_t 件产品。（130|131）我们有大致相同的概率得到比净需求多或少的产品。得到多余的产品意味着我们必须保持多余的库存直到将来有变成新的净需求。如果产品是高度定制化的，这就会是一个难题。另一方面，如果我们的产品比预计要少，那就需要再次开始生产以弥补差额，这就可能造成订单不能按时送达客户。

在这一点上安全库存可以提高客户服务水平和反应速度。跟以前一样我们将计划的生产数量扩大至 $N_t(1/y)$ 并保持安全库存来应对实际产量少于平均产出的情况。另一个策略是不保持安全库存但是将生产量扩大 $1/y$ ，这样最后我们得到的产品可能将多于净需求，剩余的产品将会变成库存。这两种方法在本质上是一样的，因为产生的结果都是以额外库存为代价来换取更好的服务。

最后，我们必须指出：任何产出策略的有效性都取决于产出本身变动性的大小。例如，如果一项生产任务需要生产 100 件产品并且每一件产品独立完成的概率为 0.9，那么得到的产品数量的均值和标准差分别为 90 和 3。因此如果准备的生产原料为 120 件产品的量（即 $100/0.9 + 3 \times 3$ ），我们就有超过 0.99 的概率（超过均值 3 个标准差）得到至少 100 件产品。然而，如果产出情况是共同完成的形式，即要么全部完成要么一件都不完成（正如成批加工），那么我们就需要下达两次 100 件产品的生产任务，从而使得准时完成 100 件产品的概率为 0.99。在第一种情况（独立），平均库存增加量为 8 件（ $120 \times 0.9 - 100$ ），而在第二种情况（批量），平均库存增加量为 80 件（ $200 \times 0.9 - 100$ ）。我们看到的结果就是平均产出率不足以决定一项有效的产出策略，加工过程的机制和变动性所造成的产出损失也必须一并考虑进去。

3.1.9 MRP 的问题

虽然在奥利基之前 MRP 已经有热心的支持者了——但是奥利基的书的副标题还是写《一种新的生活方式》，然而 MRP 的一些问题确实早前就已经被发现了。其中三个最严重的问题是：（1）MRP 排程的能力不可行，（2）提前期过长，（3）系统“紧张”。这些问题与其他问题首先产生了新的 MRP 程序和新一代的 MRP 系统，称为制造资源计划或者 MRP II，随后又演变成企业资源计划（ERP），我们会在接下来的一节中对其进行讨论。

产能不可行 (Capacity Infeasibility)。 MRP 的基本工作模式是针对具有固定提前期的生产线。由于这个提前期不依赖于车间有多少工作量，这就存在一个隐含的假设：不管工作量多大，生产线总是拥有足够的生产能力。换句话说就是 MRP 假设所有的生产线都有无限的产能。这样当生产水平处于或接近产能时就会产生问题。

解决这个问题的一种方法是确保满足系统需求的主生产计划在产能上可行。一种叫做粗能力计划 (RCCP) 的程序可以提供这样一种检查，这我们会在后面看到。正如其名称所显示的，RCCP 是一种近似估算。对于产生的 MRP 计划还有一种更为精确的能力评估方法，即**能力需求计划 (Capacity requirements planning, CRP)**。RCCP 和 CRP 都是 MRP II 中常见的模块。

过长的计划提前期 (Long Planned Times)。 正如我们在之前的安全提前期的讨论中所看到的，在 MRP 系统中有很多方面的压力促使延长计划提前期。(131|132) 在本书第二篇中，我们会看到过长的提前期肯定会导致过高的库存。但是，只要延迟交货的成本高于过量库存带来的成本（这是很显然的事情，因为库存不会对你大喊大叫，但不满意的客户会），生产管理者就会倾向于更长的提前期。

另外，由于 MRP 使用固定的提前期，而实际生产时间总是变化，这还会进一步激化过长提前期所造成的问题。为了弥补这一缺点，管理者在计划提前期上通常会选择悲观的预测。例如，假设生产的平均提前期是 3 个星期，标准差为 1 个星期，为了保持好的客户服务水平，计划提前期被设为 5 个星期。由于实际提前期是随机的，有些就会低于 3 个星期的均值，而另外一些会高些。如果这是一个标准正态分布的话，最有可能的提前期就会是 3 个星期，这样最有可能的库存持有时间就是 2 个星期，其结果是造成大量的库存。

计划提前期越长，零件在进行下一步生产之前就要等待越长的时间，系统就会产生越多的库存。因为将计划提前期设为等于平均的制造时间就只能使零件得到 50% 的服务水平（因而由其构成的零件的服务水平就会更低），管理者事实上总是会选择比平均制造时间长得多的提前期。这会降低响应速度同时又带来高库存水平。

系统紧张 (System Nervousness)。 当主生产计划的一个微小变动造成计划投入量大幅度变动时，MRP 系统就会出现紧张。这会导致奇怪的结果。例如，正如我们下面的例子表明的，需求的降低可能导致正常情况下可行的 MRP 计划变得不可行。

以下的例子摘自 Vollmann 等 (1992) 的论文。我们考虑两种零件，零件 A 的提前期为 2 周，运用固定订货期法确定批量，设定订单间隔为 5 周。每一件零件 A 需要一个组件 B，而组件 B 的提前期为 4 周，并运用定固定订货期法确定批量，设定订单间隔为 5 周，表 3.11 和 3.12 给出了两种零件的 MRP 计算结果。(132|133)

表 3.11 需求变化前零件 A 的 MRP 计算

| 零件 A | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---------|----|----|----|----|---|----|---|----|
| 毛需求 | 2 | 24 | 3 | 5 | 1 | 3 | 4 | 50 |
| 计划接收量 | | | | | | | | |
| 调整后的 SR | | | | | | | | |
| 预计持有库存 | 28 | 26 | 2 | -1 | - | - | - | - |
| 净需求 | | | 1 | 5 | 1 | 3 | 4 | 50 |
| 计划产出量 | | | 14 | | | | | 50 |
| 计划投入量 | 14 | | | | | 50 | | |

表 3.12 需求变动前组件 B 的 MRP 计算

| 组件 B | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---------|---|----|----|---|---|---|-----|---|---|
| 毛需求 | | 14 | | | | | 50 | | |
| 计划接收量 | | 14 | | | | | | | |
| 调整后的 SR | | 14 | | | | | | | |
| 预计持有库存 | 2 | 2 | 2 | 2 | 2 | 2 | -48 | - | - |
| 净需求 | | | | | | | 48 | | |
| 计划产出量 | | | | | | | 48 | | |
| 计划投入量 | | | 48 | | | | | | |

表 3.13 需求变动后零件 A 的 MRP 计算

| 零件 A | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---------|----|----|----|----|----|---|---|---|----|
| 毛需求 | | 2 | 23 | 3 | 5 | 1 | 3 | 4 | 50 |
| 计划接收量 | | | | | | | | | |
| 调整后的 SR | | | | | | | | | |
| 预计持有库存 | 28 | 26 | 3 | 0 | -5 | - | - | - | - |
| 净需求 | | | | | 5 | 1 | 3 | 4 | 50 |
| 计划产出量 | | | | 63 | | | | | 50 |
| 计划投入量 | | | 63 | | | | | | |

现在我们将第 2 期的需求从 24 降到 23，看起来很显然的情况应该是：能够满足第 2 期 24 件零件的生产计划当然也应当满足同期 23 件零件的生产计划。然而确定批量过程中对需求的加总却产生了一套完全不同的计划投入量。在零件 B 的生产计划中（见表 3.14），计划投入量变得不可行了。

表 3.14 需求变化后组件 B 的 MRP 计算

| 组件 B | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---------|---|----|-----|---|---|---|----|---|---|
| 毛需求 | | | 63 | | | | | | |
| 计划接收量 | | 14 | | | | | | | |
| 调整后的 SR | | | 14 | | | | | | |
| 预计持有库存 | 2 | 2 | -47 | - | - | - | - | - | - |
| 净需求 | | | 47 | | | | 48 | | |
| 计划产出量 | | | 47 | | | | 48 | | |
| 计划投入量 | | | 47* | | | | | | |

减少系统紧张的方法有多种，其中一种是合理运用批量规则。很显然，如果我们使用批对批法，计划投入量的变化就不会大于主生产计划的变化。但是，批对批法可能造成过多的准备时间，所以我们需要寻找其他的解决办法。

Vollmann 等（1992）推荐我们在不同的物料清单水平上使用不同的批量规则，如在中等水平上使用固定批量法或批对批法，在低水平上使用固定订货期法，而在最终零件生产上使用固定批量法。因为订货批量在高水平上保持不变，生产批量的变化就会减少系统紧张。当然，在确定固定批量的数量时务必要格外注意。（133|134）

除了运用合适的批量规则可以减少系统紧张，其他一些措施也可以减轻系统紧张的影响，其中一种方法是直接减少输入本身的变化。要做到这一点可以通过冻结主生产计划（MPS）的早期计划部分，这样就减少了 MPS 中可能出现的变化，继而减少了计划投入量的变化。既然早期计划投入量最容易受到生产计划变动的影响，因此在主生产计划最早的几期内设定一个禁止变更的**冻结期（frozen zone）**就可以极大地减少系统紧张带来的问题。

一些企业通常将主生产计划的最初 X 周设定为冻结期，然而在大多数系统里，“冻结（frozen）”这个词可能太过于严格了，因为变化可以限制却不能完全禁止。（也许“阻滞期（slushy zone）”这个词可能更为确切。）“时间栅栏”这一概念将这种行为正式化了。最早的时间栅栏，比如说是四个星期，那就意味着绝对的冻结，即决不能有任何变动。随后的时间栅栏，可能是 5 到 7 周，这期间的变动受到限制但相对不那么严格了，在模型中如果存在选择的话变动是可以接受的，但是可能会对客户进行财务上的惩罚。再接下来的时间栅栏，可能是 8 到 12 周，变动就更加没那么严格了。在这种情况下，如果所有组件都是现存的，那么零件数量上的变动则是可以接受的。在最后的时间栅栏里，即 13 周以上，就不存在限制了。

另一种减少系统紧张所带来损失的方法是使用**固定计划订单（firm planned orders）**。与冻结期和时间栅栏不同，固定计划订单下达固定的计划投入量。通过将早期的计划投入量转化为固定计划订单，可以消除生产计划早期的系统紧张，正是在这段时期系统紧张的破坏性最大。考虑一下如果我们在需求发生变动之前将表 3.11 中的第一批计划投入量转化成固定计划订单将发生什么情况，这有可能导致 MRP 处理过程将其当成计划接收量来对待。经过这种转化后就没有系统紧张了，如表 3.15 和表 3.16 所示。

表 3.15 固定计划订单下零件 A 的 MRP 计算

| 零件 A | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|--------|----|------|----|------|---|---|----|---|-----|
| 毛需求 | | 2 | 23 | 3 | 5 | 1 | 3 | 4 | 50 |
| 计划接收量 | | | | | | | | | |
| 固定计划订单 | | | | 14 | | | | | |
| 预计持有库存 | 28 | 26 | 3 | 14 | 9 | 8 | 5 | 1 | -49 |
| 净需求 | | | | | | | | | 49 |
| 计划产出量 | | | | [14] | | | | | 49 |
| 计划投入量 | | [14] | | | | | 49 | | |

表 3.16 固定计划订单下组件 B 的 MRP 计算

| 零件 B | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---------|---|----|----|---|---|---|-----|---|---|
| 毛需求 | | 14 | | | | | 49 | | |
| 计划接收量 | | 14 | | | | | | | |
| 调整后的 SR | | | | | | | | | |
| 预计持有库存 | 2 | 2 | 2 | 2 | 2 | 2 | -47 | - | - |
| 净需求 | | | | | | | 47 | | |
| 计划产出量 | | | | | | | 47 | | |
| 计划投入量 | | | 47 | | | | | | |

当然，使用固定计划订单和时间栅栏意味着生产计划的冻结部分对需求变动的响应变弱，但同时的缺点是固定计划订单代表着计划制定者必须进行冗繁的手工输入。（134|135）

3.2 制造资源计划——MRP II

物料需求计划提供了一种计划和获得物料从而支持生产的系统方法。该方法通过电脑实施会相对简单和容易，但是仍然存在一些问题。

正如我们已经提到过的，产能不可行、过长的计划提前期、系统紧张和其他问题的存在会损害 MRP 系统的运行效果。随着时间的发展，人们又设计出了其他办法来解决上述这些问题，这些解决方法就包括在我们所说的**制造资源计划**，也就是 **MRP II**。

MRP II 除了解决 MRP 失效的问题外，同时还给我们带来了其他很多功能，这使得 MRP II 成为真正意义上的综合性制造管理系统。MRP II 带来的其他功能包括需求管理、预测、能力计划、主生产计划、粗能力计划、能力需求计划、人员排班以及输入/输出控制。(135|136) 本节我们将介绍 MRP II 各种功能赖以发挥作用的 MRP II 层级体系以及与之相关的模块。我们的介绍从某种意义上是简略的，理由有两点：第一，MRP 和 MRP II 的内容本身就可以写一整本书，我们在此推荐 Vollmann 等（1992）的著作作为一本优秀的综合性参考书。第二，我们将在第十三章中介绍层级生产计划问题（在拉式生产的背景下），到时还将详细介绍与时间度量、预测以及需求管理等计划层级相关的一类问题。

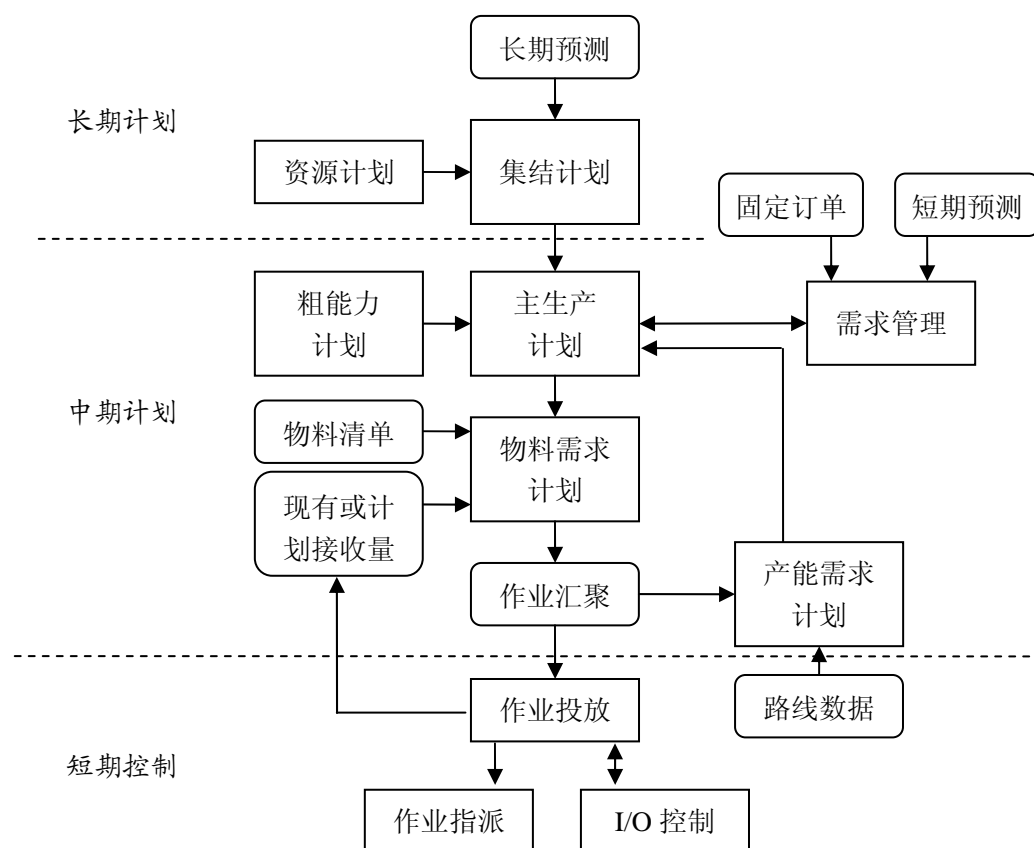


图 3.3 MRP II 层级

3.2.1 MRP II 层级

图 3.3 描述了 MRP II 层级的一个例子。我们在此用例子这个词是因为 MRP II 可能有成千上万种层级，就像 MRP II 软件有成千上万的生产商一样。（尽管大多数生产商现在称自己

为 ERP 或者“企业”软件生产商)。

3.2.2 长期计划

MRP II 层级的顶端是**长期计划 (long-term planning)**，包括三项功能：资源计划、集结计划和预测。长期计划的时间跨度从大约 6 个月到 5 年不等。更新计划的频率从一月一次到一年一次不等，通常是每年 2 到 4 次。详细程度通常是零件族 (part family) 的级别 (即具有相似需求和生产特性的一类最终产品)。(136|137)

预测 (forecasting) 功能的目的在于预测未来的需求。长期预测对于决定产能、设备和人员需求是非常重要的，而短期预测则是将零件族的长期预测转化成单个最终产品的短期预测。这两种预测的数据都输入到中期层级的**需求管理 (demand management)** 功能。我们将在第十三章中详细介绍预测技巧。

资源计划 (resource planning) 是指在长期内决定能力需求的过程。关于是否新建工厂或扩建工厂的决定就属于产能计划功能的一部分。资源计划的一个重要的输出是在长期计划层面上的预计可用产能。这一信息将作为一个参数输入到集结计划中。

集结计划 (aggregate planning) 用来决定长期的生产、人员、库存和加班加点等的水平，其详细程度通常在月和零件族的层面上。例如，集结计划将决定在预期需求会增加时 (通过预测得知) 是否通过建立库存或加班加点，或者两者兼行并施，从而随着需求的变化而改变产能。线性规划之类的最优化技术通常被用来帮助制定集结计划。我们将在第十六章中详细介绍集结计划和相关的支持模块。

3.2.3 中期计划

在中期层级上有许多生产计划功能模块。其中包括需求管理、粗能力计划、主生产计划、物料需求计划以及产能需求计划。

当跟踪单个客户订单时，将长期集结预测转化成详细预测的过程我们称之为**需求管理 (demand management)**。需求管理模块的输出是一套实际的客户订单再加上预期订单的预测，然后随着时间的推移，预期订单转化成实际订单。

这是通过一种称为**可供销售量 (available to promise, ATP)** 的技术来实现的。这一技术使计划制定者清楚地知道主生产计划中哪些订单已经生产完毕以及多少产品可供销售给新的客户，将 ATP 与一个能力可行的主生产计划结合在一起可以便于协商出可实现的到期日。如果接收的订单超过预期，以至于先前制定的提前期延长，这时就需要额外的生产能力。另一方面，如果接收的订单少于预期，那就需要折价销售或其他刺激措施来扩大需求。无论发生哪一种情况，都可能需要修正预测甚至是集结计划。

主生产计划 (master production scheduling) 从需求管理模块获得需求预测和固定订单，并使用集结能力限制来产生预期的详细生产计划，其详细程度处于最高层级。这就是 MRP 中使用的“需求” (即，料号、数量和到期日)。因此，主生产计划包括每一期、每一个计划日对具有独立需求的每一种零件的订单数量。对于大多数企业，这些是面向**最终产品 (end item)** 层的。但是，在有些情况下，为每一个产品大类作计划比为每一种最终产品做计划更有意义。我们可以从汽车工业中找到这样的例子，在汽车工业里，只有到组装线上的最后一分钟我们才能决定汽车确切的制造方式和规格。在这种情况下，通常用**最终装配计划 (final assembly schedule)** 来确定什么时候生产最终产品，而主生产计划则只是用来对每一类产品预先做出计划。这种类型的主生产计划有一个关键的输入——**超级物料清单 (superbill of material)**，它包含用百分比的形式对每一类产品中不同型号的数量所作的预测。如果了解有关最终装配计划中超级物料清单的完整讨论，我们建议读者参考 Vollman 等 (1992)。(137|138)

粗能力计划（rough-cut capacity planning, RCCP）的作用在于对一些关键资源进行快速的能力检验，以确保主生产计划的可行性。粗能力计划的详细程度介于集结计划 and 能力需求计划之间，能力需求计划是在 MRP 处理之后进行能力检验的又一工具。能力需求计划必须使用主生产计划中每一个最终产品的**资源清单（bill of resource）**，资源清单显示的是为了生产一件最终产品所需花在每一项关键资源上的时间。该时间不仅包括最终产品本身的时间，而且包括由此引起的其它相关需求的时间。例如，假设零件 A 由两个组件 A1 和 A2 组成，零件 A 需要 1 个小时的组装时间，而组件 A1 和 A2 分别需要半个小时和 1 个小时的生产时间，因此，零件 A 的资源清单应当显示每个 A 零件需要 2 个半小时。另外我们还假设零件 B 不包含任何组件，需要 2 个小时的生产时间。

在上面的例子中，我们还假设在零件 A 和 B 的主生产计划中包含如下信息：

| 周次 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|------|----|----|----|----|----|----|----|----|
| 零件 A | 10 | 10 | 10 | 20 | 20 | 20 | 20 | 10 |
| 零件 B | 5 | 25 | 5 | 15 | 10 | 25 | 15 | 10 |

零件 A 和零件 B 的资源清单如下：

| 加工中心 | 零件 A | 零件 B |
|------|------|------|
| 21 | 2.5 | 2.0 |

那么，零件 A 和零件 B 的粗能力计划计算如下所示：

| 周次 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|-------------|----|-----|----|-----|----|-----|-----|----|
| 零件 A（小时） | 25 | 25 | 25 | 50 | 50 | 50 | 50 | 25 |
| 零件 B（小时） | 10 | 50 | 10 | 30 | 20 | 50 | 30 | 10 |
| 总时间（小时） | 35 | 75 | 35 | 80 | 70 | 100 | 80 | 35 |
| 可用时间 | 65 | 65 | 65 | 65 | 65 | 65 | 65 | 65 |
| 剩余（+）/短缺（-） | 30 | -10 | 30 | -15 | -5 | -35 | -15 | 30 |

如果我们只考虑这 8 个时段的总体情况，那么我们可以得出的结论是：我们拥有 520 小时的生产能力但只有 510 小时的生产需求，因此生产能力够用。（138|139）但是，在进行粗能力计算后，我们可以发现在好几期中生产能力不够用，而在有些期内生产能力是过剩的。现在就有待计划制定者来决定如何才能改变这种不均衡的情况，通常有如下选择：（1）通过改变到期日来调整主生产计划，（2）通过加班加点生产或转包部分生产来调整产能。

粗能力计划并不作任何偏移（offsetting），因此使用的时段必须足够长，以使零件及其组件能够在单期内完成。粗能力计划同样假设不管生产在加工中心内如何安排（即产生空余时间）需求都能够满足。在这种假设下，粗能力计划可能出现的情况作出了乐观的估计。

另一方面，RCCP 也不进行任何净值计算。这对最终产品（它们的需求对最终产品库存进行净值计算相对要容易些）来说也许是可以接受的，但对构成最终产品的组件来说就不那么容易让人接受了，尤其是有许多共用组件而且 WIP 水平很高时。RCCP 的这方面特性使得它有些保守。

RCCP 的这两个特性使 RCCP 的行为难以确切的计量，通常来讲第一个特性所作的近似决定了第二个特性，这使 RCCP 可以乐观的估计可能出现的情况，但是并不总是这样的。

结果就是粗能力计划实际上可能确实是非常粗略的。

能力需求计划（Capacity requirements planning, CRP）对MRP的生产计划提供一个比RCCP更为详细的能力检验。它所必需的输入包括所有计划投入量、现存WIP水平、工艺路线数据以及所有加工中心的生产能力和提前期。与其名称有点出入的是，能力需求计划并不给出有限能力分析，相反，CRP处理的是**无限前向负荷（infinite forward loading）**。CRP通过确定的提前期预测每一个加工中心加工任务的完成时间，然后计算出预期负荷完成时间。这些负荷然后被用来与可用的生产能力进行比较，但对于超负荷的情况并不做出任何修正。

6

为了形象地说明 CRP 是如何运作的，我们举一个简单的例子。一个加工中心的提前期为 3 天，生产能力为 400 件/天，开始当天刚好有 400 件待加工零件已经投入到该加工中心，之前已经有 500 件等待了 1 天，另 300 件已经等待了 2 天，接下来的 5 天里计划投入量如下表所示：

| 第几天 | 1 | 2 | 3 | 4 | 5 |
|-------|-----|-----|-----|-----|-----|
| 计划投入量 | 300 | 350 | 400 | 350 | 300 |

通过已知的 3 天提前期，我们可以计算出有多少零件何时生产完成离开加工中心。如果我们之前预期一天会有 400 件完成，该加工中心就处于超负荷状态。由此产生的**负荷状态图（load profile）**如图 3.4 所示。第一天负荷为 300 件（这是在该加工中心已经等待了 2 天的那 300 件零件，在第一天结束时离开）；第二天负荷为 500，这同样是开始时已经等待了一天的那一批。由于 500 大于该加工中心的生产能力 400 件/天，这就产生了超负荷的情况。（139|140）

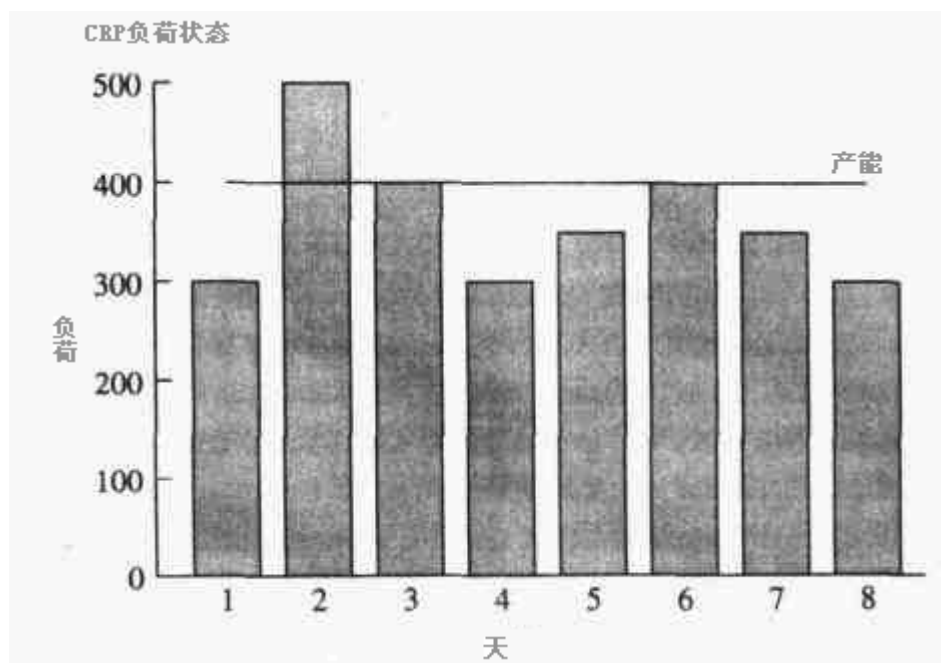


图 3.4 CRP 负荷描述

⁶ 不像 MRP 和 CRP 那样，真正的有限能力分析并不假设固定的提前期，工件通过制造环节的时间取决于该环节中已经具有的工件数量及其优先序。大多数有限能力分析的软件包都会对通过工厂的物流流进行某种程度的确定性仿真。因此，有限能力分析要比 CRP 复杂得多。

需要注意的是，即使当负荷超过生产能力时，CRP 仍然假设加工中心生产产品的时间不发生变化。我们知道超负荷生产时所耗的时间当然比低负荷时所耗的时间要多，因此 CRP 在超负荷情况下的估计都将出错。所以，CRP 通常不是估计负荷状态的很好方法，除非是在非常接近的短期内。CRP 的另一个问题是它仅告知计划制定者出了问题，但提供不了是什么导致了问题的出现以及如何来解决问题。为了知道应该采取什么措施，计划制定者首先必须获得一份分解负荷的报告，从而知道是哪些工作导致了问题的出现，然后必须通过挂钩技术追踪到主生产计划的需求。这一过程是非常冗繁的。

和 MRP 自身一样，CRP 的一个基本缺陷在于它预先假设产能是无限的。这一假设来自于固定提前期的假设，而固定提前期并不取决于加工中心的负荷。同样是这个加工中心，如果起始时并没有等待完成的加工任务以及相应的计划投入量，生产过程遵循的批量规则是对需求分组以减少生产准备，其计划投入量如下：

| 第几天 | 1 | 2 | 3 | 4 | 5 |
|-------|------|---|---|------|---|
| 计划投入量 | 1200 | 0 | 0 | 1200 | 0 |

运用 CRP 我们可以从负荷状态图看到超负荷的情况出现在第三天和第六天。如果我们用的是有限产能负荷，我们将得到一个非常不同的结果。在其中两天会没有产出（第一批投入量需要这两天的时间生产完），接下来的六天每天都会有 400 件的产出，而在第四天的第二批投入量到达时第一批的最后部分正好进入加工中心生产。我们将在第七章中介绍产能、在制品和穿越加工中心的时间之间的关系。

因此，尽管 CRP 的引进让人满怀希望，其目标也非常有价值，但它存在着根本性的问题。首先，这需要大量的数据，并且输出非常巨大和冗繁。其次，对于超负荷状态它不能提供补救的办法。最后，由于 CRP 用的是无限负荷而大多数现代系统处理的是现实中的有限产能负荷，使用 CRP 的公司越来越少了。（140|141）

MRP II 的所有早期版本中的物料需求计划和很多现代 ERP 系统与我们之前描述的 MRP 过程是非常相似的。MRP 的输出是**工件池（job pool）**，包含计划投入量。这是由**作业投放（job release）**职能完成的。

3.2.4 短期控制

上述的中长期计划功能模块产生的计划是通过短期控制模块来执行的，短期控制包括**作业投放（job release）**、**作业指派（job dispatching）**和**输入/输出控制（I/O control）**。

作业投放将计划产出量转换成计划接收量。作业投放的一个重要职能是零件**分配（allocation）**。当有多种高层级物料需要同一低层级零件，而现有该零件的数量又不足所有时，就有可能产生冲突。这时作业投放通过将有限的零件合理地分配给某一项加工就能化解这些冲突。假设有两个计划投入量同时需要组件 A，但是该组件只够满足其中之一的需求；此外，第一个计划投入量还需要组件 B，第二个计划投入量还需要组件 C，而且持有库存中有足够的组件 B 但没有足够的组件 C，在这种情况下作业投放就会将现有组件 A 分配给第一个计划投入量，因为其所需的两种组件都有足够的库存。这时还会产生一个短缺通知，告知第二个计划投入量尚停留在工件池还未生产。（141|142）

一旦加工任务或者采购订单被投放，就需要对其继续控制以确保产品数量和规格与要求相符。如果是采购组件，还需要对订单进行跟踪，也就是对订单到达的时间和延误订单进行监控。如果是内部生产，下一步就是**车间作业控制（shop floor control, SFC）**功能的范畴，它也称为**生产活动控制（production activity control, PAC）**。由于 SFC 更为传统而且运用更为广泛，本书中我们一律用这个名称。在车间作业控制下有两项子职能：**作业指派（job**

dispatching) 和输入/输出控制 (I/O control)。

作业指派。 作业指派的基本思想非常简单, 即开发一种合适的规则来安排每一个工作站前的工件加工队列, 既能使到期日不变又能保持尽可能高的机器利用率和尽可能少的生产时间。人们为此提出了许多规则。

最简单的一个指派规则就是**最短生产时间 (shortest process time, SPT)**。在 SPT 下, 加工中心的工件队列被重新分类, 并将生产时间最短的工件置于队列最前面, 这样就能确保生产时间最短的工件总是最先被处理, 使得小型工件能够快速通过加工中心。SPT 的应用通常能够降低平均生产时间并提高机器使用率, 而且即使在定制过程中不考虑到期日, 平均到期日的效果通常也会非常好。

但是当队列中有生产时间相当长的工件时 SPT 就会出现問題, 这时会有工件等待很长时间却得不到处理。因此尽管 SPT 的平均到期日效果不错, 但是生产延迟的差异却可能非常大。避免这一问题的方法之一就是运用 SPT^x 规则, 其中 x 是一个参数。根据这一规则, 得到处理的将是生产时间最短的工件, 除非某一工件已经等待了 x 单位或者更长的时间, 这时这一工件将优先得到处理。这一规则在很多情况下都会表现出很好的效果。

如果队列中的工件在尺寸和工艺路线上都非常接近, 那么**最早到期日 (earliest due date, EDD)**就可以作为一个很好的指派规则。在 EDD 下, 距到期日最近的工件将最先处理。EDD 在上述情况下都能有很好的效果, 但是在更为普遍的情况下通常并不比 SPT 更优。

以下是其他三种常见的规则:

最小松弛时间 (Least slack): 一个工件的松弛时间即到期日减去剩余的生产时间 (包括准备时间) 再减去当前时间。该值最小的工件将被赋予最高的优先权, 即最先生产。

剩余加工平均最小松弛时间 (Least slack per remaining operation): 该规则与上一规则相似, 不同之处在于: 我们将松弛时间除以工艺路线上尚未进行的加工次数。同样, 该值最小的工件拥有最高的优先权。

临界比率 (Critical ratio): 将剩余时间 (即到期日减去当前时间) 除以未完成加工的小时数计算出一个指数, 根据这个指数对工件进行排序。指数大于 1 的加工将会提前完成, 而指数小于 1 的加工则会来不及, 如果指数为负数则表明已经来不及了。与前两种规则一样, 临界比率最小的工件将拥有最高的优先权。

生产管理方面的著作给我们提供了至少 100 种指派规则, 其中 Blackstone 等 (1982) 对许多规则作了很好考察, 他们利用一个模拟工厂对许多种规则进行了测试。

当然, 任何指派规则都不能适用于所有情况, 因为从本质上讲, 所有的指派规则都是针对短期的。制定出良好计划的唯一方法就是将车间看成一个整体。这样做的问题在于: (1) 车间调度问题是一个极其复杂的问题, 而且需要大量的计算时间; (2) 得到的排程通常是不直观的。我们将在第 15 中对排程问题作更为详细的讲解。

输入/输出控制 (I/O)。 输入/输出控制最早是于 1970 年作为一种控制提前期的方法由怀特提出的。I/O 控制是通过如下方法进行的:

1. 监控每一个加工中心的在制品水平。
2. 如果在制品超过某一水平, 则表明现有投放率过高, 应当降低。
3. 如果在制品低于某一特定水平, 则表明现有投放率过低, 应当提高。

4. 如果在制品介于上述两个水平之间则表明当前投放率正常。降低或提高投放率必须通过改变主生产计划来进行。

I/O 控制提供了一种检测投放与可用能力情况的简易方法。然而，等到在制品水平出现异常时，系统往往在许多方面已经无法控制了。这可能就是所谓的拉式系统（例如丰田的看板系统）优于 MRP、MRP II 和 ERP 之类的推式系统的一个原因。与推式系统通过主生产计划控制投放、通过 I/O 控制衡量在制品水平所不同的是，看板系统直接控制在制品并每天测量产出率。因此，看板系统不允许在制品水平过高并能快速发现问题（即生产不足）。我们将在第四章和第十章中分别详细探讨看板系统和拉式、推式系统的基础。（142|143）

3.3 超越 MRP II ——企业资源计划（ERP）

继 MRP II 发展起来之后的几年，软件供应商和咨询公司提出了许多种可能的继承者。其中 MRPIII 从来就没有真正进入过人们的视野，而令人费解的缩略词 BRP（业务需求计划）同样也没有被广泛接受过。最后企业资源计划（ERP）成功脱颖而出，尽管它的首字母缩写同样是让人没胃口。

ERP 的推广很大程度上因为一些软件商的成功，其中就有著名的 SAP，SAP 不只是定位于生产运作领域，而是瞄准了所有运作过程（如制造、分销、会计、财务和人事）。因此，它提供的系统是设计用来控制整个企业的。

SAP 的 R/3 软件是一个典型的交互式综合性 ERP 系统，根据 1997 年《商业周刊》上 Edmonson 的一篇文章所说，该系统是一个功能强大的网络，通过单击鼠标就能够加快决策制定、削减成本以及使管理者得以控制遍布全球的商业帝国。该文章的华丽词藻所言不虚，ERP 系统通过将各类信息连接起来使公司的高层管理者得以更为容易地实时掌握全球业务。

这种大型综合系统的优点包括：

1. 集成化的功能模块
2. 一致的用户界面
3. 集成化的数据库
4. 单一的软件供应商商和合约
5. 统一的结构体系和工具箱
6. 统一的产品支持

但该系统也存在如下缺点：

1. 与现有系统不兼容
2. 实施昂贵且时间长
3. 与当前管理惯例不相容
4. 失去了运用各种战术系统的柔性
5. 产品开发和实施周期长
6. 回收期长
7. 缺乏技术创新

正如我们下面所说的，尽管存在以上缺点，ERP 在市场上还是获得了显著的成功。

3.3.1 ERP 的历史与成功

ERP 的成功至少部分是由于它与之前的三个潜在趋势不谋而合。一是**供应链管理 (supply chain management, SCM)** 获得了承认。在许多方面, 供应链管理将传统的库存控制方法拓宽到一个更为广泛的层面, 包括分销、仓库管理和多区位生产。更重要的是, 对供应链管理的定义引出了对物流问题的重视, “物流管理理事会”这类贸易组织的扩大即反映了物流领域的重要性, 该理事会的成员数从 1990 年的 6256 家增加到 1997 年 14000 家。(143|144)

另一个激发了市场对 ERP 接受的趋势就是**业务流程重组 (business process reengineering, BPR)** 运动 (见 Hammer 和 Champy 1993)。90 年代之前, 很少有企业愿意大幅度改变其管理结构来支持一个新的软件包, 但是 BPR 的兴起指导管理者们考虑改变其管理结构。今天, 许多管理者认为实施 ERP 的一个好处即在于帮助他们对业务进行重组。

第三个趋势是分布式处理和微型计算机处理能力的爆炸性增长。1960 年在一台 100 万美元的计算机上需要一个星期运行时间的 MRP 现在在一台笔记本电脑上只需要几秒钟。以前需要在一个中心储存器上储存的公司数据, 现在可以储存在生产用的个人电脑或工作站上。所有的个人电脑数据通过跨公司网络相连并可以被所有职能部门共享。ERP 销售商的最新成果就是根据这一思想设计的 (Parker 1997)。

ERP 系统销量的增长表明了它被市场接受的程度。1989 年 MRP II 的总销售额为 12 亿美元, 仅占整个美国软件销售市场的三分之一 (《工业工程》1991), 而 1995 年仅 ERP 系统的前 10 大销售商就创下了 28 亿美元的销售额, 1996 年和 1997 年更达到 42 亿美元和 58 亿美元。1997 年 SAP 一家公司就销售了超过 32 亿美元的 ERP 系统软件。

但是在 ERP 销售的大幅增长之外还存在着一些问题。许多公司抱怨该系统的实施带来的巨额成本。一项对实施 ERP 系统的财富 1000 强公司的调查显示, 44% 的公司表示他们花费了至少 4 倍于软件本身的费用来帮助实施 ERP 系统 (如咨询费用等)。据我们所知有些公司在花费了数百万美元之后已经取消了实施计划, 原因在于他们不想一遍又一遍地往里面投钱。

尽管 ERP 系统的实施耗资甚巨, 一些公司也声称获得了生产率的大幅提升。在 1996 年完成了 SAP 的会计模块系统安装后, 孟山都 (Monsanto) 公司副总裁鲍勃·巴雷特 (Bob Barrett) 表示, 该系统的实施将计划周期从 6 周减少到了 3 周, 除此以外还降低了库存、运营成本, 并提升了公司跟供应商谈判的能力, 这些成果每年给公司节省下 200 万美元。

3.3.2 案例: SAP R/3

SAP 公司销售的 R/3 软件是一个典型的 ERP 系统。R/3 运用客户/服务器技术提供一个他们所称的**数据仓库 (data warehouse)**, 该数据仓库允许所有使用者使用一个单一的数据系统, 并让使用者通过一个通用的界面与其他软件共享数据。

与大多数 ERP 系统一样, R/3 是一个大型的、事务处理型的软件包, SAP 将其设计成四个应用套装, 包括财务、人力资源、生产和物流、销售和分销, 每一个套装包含多个应用程序。有趣的是, 在它们的中心是一个极其简单的物料需求计划模块, 在逻辑上与 30 年前奥利基所描述的完全一样。

SAP 的 R/3 系统还可以通过附加模块而不断得到更新, 包括针对各种特定工业领域的模块。关键的意图是要建立起最佳惯例 (best practice) 并将其加入到软件当中。许多使用该系统的公司必须大幅改变公司的管理流程来适应该软件及其包含的最佳惯例, 事实上, 这种对公司管理者个性的大幅限制最终将是 SAP 成功给社会带来的最大影响。但是由于将多年的惯例程式化不太可能是保持公司竞争优势最佳战略, 这有可能导致很多 ERP 的使用者容易

受到更具创新性的竞争者的挑战。(144|145)

3.3.3 制造支持系统

制造支持系统 (manufacturing execution system, MES) 是指车间作业控制 (SFC) 的自动执行。但与 SFC 不同是, MES 自动跟踪在制品, 记录生产、产出和质量数据, 执行计划并将新工件送进系统等等。如果要问 MES 是否是 ERP 的一部分的话, 恐怕咨询公司和软件供应商之间发生激烈的争论了, 但是随着像 SAP 的 R/3 系统那样的软件将越来越多的业务功能整合到一起, 就很难说 MES 还可以继续独立保持多久了。

3.3.4 先进计划系统

ERP 系统将公司数据整合在一起, 而**先进计划系统 (advanced planning system, APS)** 的作用就是对贯穿组织上下的数据进行分析。APS 的功能随供应软件的销售商的不同而不同, 大多数 APS 是基于记忆的算法, 包括有限能力排程、预测、可供销售量、需求管理、仓库管理、分销和交通管理等。很多时候, ERP 销售商与其他更为专业化的软件开发商一起合作提供这类功能模块。有趣的是, 这类附加模块的方法越来越类似于早前的 MRP II 系统的方法, 即通过重新计算排程以弥补 MRP 中无限后向排程的问题。

3.4 总结

物料需求计划 (MRP) 是从对独立和非独立需求之间区别的认识发展而来的, 也是现代计算机在生产控制领域的首次大规模应用。MRP 提供了一个基于需求订购原材料的简单方法, 即依靠主生产计划和物料清单。因此, 它非常适合用于控制零部件采购, 但是在生产控制中却仍然存在问题。

随后制造资源计划 (MRP II) 被用来解决 MRP 中存在的问题, 并进一步将业务功能整合到一个统一的框架上来。MRP II 提供了一个高度通用的控制结构, 将生产控制问题分解成一个基于时间维度和产品集合的层级结构。要运用数百种工具、协调成千上万的订单、生产成千上万的由上千个零部件构成的产品, 没有这样一套层级方法实际上是不可能的。最近, ERP 已经将这种层级方法整合到一个强大的管理工具当中, 来维护和跟踪庞大数量的数据。

尽管 MRP、MRP II 和 ERP 对生产领域做出了极大的贡献, 这些系统下的基本模块仍然存在基本的问题 (即无限能力和固定提前期的假设, 即使在最复杂的 ERP 系统中这些假设仍然存在)。正如我们将在第五章中进一步讨论的, 对于长期的一个关键问题是如何在保留简单性和广泛适用性的前提下解决 MRP 的基本难题。我们将在第三篇中介绍这一问题, 在此之前我们要在第四章了解 JIT 运动所带来的思想以及在第二篇对工厂行为建立一些基本的认识。(145|146)