

一种线性规划问题的自动化建模方法研究^{*}

陈雪龙¹ 韩胜菊² 王延章¹ 许永涛¹

(大连理工大学信息与决策技术研究所 大连 116024)¹ (大连水产学院信息工程系 大连 116023)²

摘要 提出了一种基于自然语言理解及 XML 的线性规划问题自动化建模方法,并给出实例。实践证明,该方法具有灵活性、通用性及可行性。

关键词 线性规划, 自动化建模, XML

The Research of an Automatic Modeling Method of Linear Programming Problem

CHEN Xue-Long¹ HAN Sheng-Ju² WANG Yan-Zhang¹ XU Yong-Tao¹

(Institute of Information and Decision Technology, Dalian University of Technology, Dalian 116024)¹

(Department of Electronic and Information Engineering, Dalian Fisheries College, Dalian 116023)²

Abstract In this paper, an automatic modeling method of linear programming problem based on the understanding of natural language and XML is proposed. Meanwhile, an application example is analyzed. It is proved by practice that this method is flexible, general and feasible.

Keywords Linear programming, Automatic modeling, XML

1 引言

线性规划(LP, Linear Programming)问题是管理决策中最基本的问题,LP方法是应用最多效果最好的管理优化方法。LP模型描述简单、直观,便于建立,具有非常有效的单纯形等求解方法,而且大多数优化决策问题可以简化建立或间接建立为LP问题,借助LP方法求解。

在DSS中,如果只是建立一个适用于某个企业的、固定的、难以修改的LP模型,则必然难以适应决策问题的多样化、半结构化、多变性等特点,更难以实现系统的通用性。所以,为了给决策者及决策分析人员针对综合经济发展研究、分析问题提供一个便于建立模型、修改模型及求解模型的计算机环境,必须提出一种具有通用性、友好性及交互性的LP问题自动化建模方法。

由此,本文提出了一种基于自然语言理解及XML的LP问题自动化建模方法。实践表明,该方法具有科学性与可行性。

2 LP问题建模过程分析

LP问题的数学模型描述如下:

$$\begin{aligned} & \text{目标函数: } \max Z = C_1 x_1 + C_2 x_2 + \dots + C_n x_n \\ & \text{约束条件: } s. t. \begin{cases} a_{11} x_1 + a_{12} x_2 + \dots + a_{1n} x_n \geq b_1 \\ a_{21} x_1 + a_{22} x_2 + \dots + a_{2n} x_n \leq b_2 \\ \vdots \\ a_{m1} x_1 + a_{m2} x_2 + \dots + a_{mn} x_n \geq b_m \\ L_1 \leq x_1 \leq U_1 \\ \vdots \\ L_n \leq x_n \leq U_n \end{cases} \quad (1) \end{aligned}$$

其中,对 $i=1,2,\dots,n$, x_i , L_i 和 U_i 分别表示决策变量及其下界限制值和上界限制值; C_i 和 a_{ij} 分别表示目标函数系数和约束方程系数, $j=1,2,\dots,m$,这里的约束可以是大于、小于或等

于型的任何一种。上述模型描述的是如何确定决策变量 x_1, x_2, \dots, x_n ,使得它们在满足各个约束方程和上下界限值条件下,获得最大的目标收益。

由(1)式可见,这样一个较为复杂的模型要交给决策者直接使用是比较困难的,因为决策者对实际问题很清楚,但不一定懂模型。由于实际问题的复杂性,使得要建立和调整这样一个模型很麻烦。因此,非常需要一个方便、实用的LP问题自动化建模方法,以便将一个实际问题转化为LP模型。

模型是对客观现实的一种描述。对模型的表示一般都是采用符号语言,这种符号模型是实际系统的一种抽象描述。从对实际问题描述的角度来看,数学表达精炼、准确、符号化,有利于在计算机上求解,但数学模型一般都较抽象;从实际使用的角度来看,用户往往更熟悉实际问题,并习惯使用自然语言对问题加以描述。因此,需要在数学模型与使用者之间架起一座桥梁,使使用者能从实际问题很容易地走向数学模型,这就是数学模型的另一种描述方法,即自然语言符号化的问题描述方法。

因此,LP问题的自动化建模关键在于将自然语言描述的LP问题自动地转换成计算机可识别的符号语言描述的LP模型。文章下面的章节将详细探讨其关键环节与方法。

3 LP问题的语义框架生成

认知科学对人的问题求解形式的研究表明,人在问题求解的初期阶段,为了理解他所面临的问题,首先要形成一个由问题的基本概念、假设等构成的结构化语义框架,用以表现与问题相关联的主要元素及其之间的相互关系,然后才能对问题进行更为深入的分析^[1]。

又由文^[2]提出的E-R-P知识表示体系可知,一个决策问题 p 可以表示为:

$$p = \langle O, I, X, R_p, C \rangle \quad (2)$$

其中 O 为目标实体属性集, I 为条件实体属性集, X 为关联实

^{*}国家自然科学基金资助项目(编号:70271045)与国家“十五”重点科技攻关项目(编号:2002BA105A-01-02)。陈雪龙 博士生,主要从事地理信息系统、决策支持系统与数据仓库的研究;韩胜菊 硕士,主要从事计算机信息系统、数据结构的研究;王延章 博士生导师,主要从事系统集成和优化理论与方法、决策支持系统、经济系统分析的研究;许永涛 博士生,主要从事电子政务、资源整合理论与方法的研究。

体属性集, R_p 为所涉及的问题关系集, C 为问题对应的属性集。这里关于实体、关系的概念均来自哲学观点。实体是客观实在的一种抽象, 而关系是实体间的联系, 二者均具有层次关系。这里的实体、关系与数据库理论中的实体、关系有本质的区别, 这里的实体、关系是对事物本质的刻画, 而数据库理论中的实体、关系则是面向数据的组织与表示的。

因此, LP 问题在问题求解初期阶段建立的语义框架可以通过实体集、实体属性集、问题属性集及问题关系集来刻画, 如图 1 所示。其中框表示实体和属性的集合, 菱形框表示实体与实体、实体与属性之间的关系。

由 LP 问题的自然语言描述生成问题的语义框架过程涉及到自然语言的理解问题, 包括词法、句法、语义等分析。这部分内容不是本文的重点, 相关细节请参看文[3]。

每个实际问题都可以形成类似于图 1 的一个语义框架。语义框架本身是一种知识表达方式, 可以作为模型描述知识存储于知识库或模型库中。它所描述的是实际问题中实体、属性及其之间的联系, 具有可操作性, 完全可以使用一种算法对其进行操作, 以达到从问题描述到模型生成的目的。

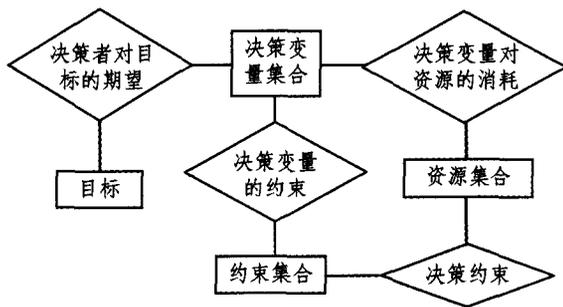


图 1 LP 问题语义框架

4 符号语言描述的 LP 模型的生成

4.1 LP 模型符号语言的 XML 表述

LP 模型不同于经济预测或仿真分析模型, 可用特定格式的数据文件加以描述。LP 模型描述要综合考虑模型求解、调整等多种因素。本文把前述 LP 模型拆分成 3 个 XML 文件加以描述。

(1) UL.xml 文件

此文件为变量上下界约束文件, 存放 L_i 和 U_i 信息。该文件大小依赖于决策变量的个数 n 。文件格式如下:

```
<xml>
.....
<变量描述>
  <下界值>/</下界值>
  <上界值>/</上界值>
<变量标注>/</变量标注>
</变量描述>
.....
</xml>
```

文件中的“下界值”、“上界值”元素分别表示决策变量 x_i 的 L_i 、 U_i 值。这里要求每一个决策变量必须给定其上下界约束值, 若对应的实际问题相应决策变量没有界约束时, 下上界可分别给以最小值和最大值, 如‘0’和‘1.0e16’。这里的最小值可以为负。“变量标注”元素是作为变量的说明与注释, 可有可无。

(2) SB.xml 文件

此文件为约束方程符号与右端项数据信息文件, 主要存放约束方程的小于、大于或等于符号信息和对应方程的右端约束值。文件最后一个“约束描述”节点的意义与其它元素不同, 用来标识目标函数是取极大还是极小的标志。该文件大小依赖于约束方程的个数 m 。文件格式如下:

```
<xml>
.....
<约束描述>
  <约束方程编号>/</约束方程编号>
  <约束类型>/</约束类型>
  <右端 b 值>/</右端 b 值>
  <约束说明>/</约束说明>
</约束描述>
.....
</xml>
```

这里约束方程编号只是代码, 不要求为连续值, 但相互之间要区别开来, 约束类型为一代码值, 等于类约束为‘0’, 大于类为‘1’, 小于类为‘-1’; 右端 b 值为相应约束方程对应的约束值。文件中最后一个“约束描述”节点为目标函数状态, 相应类型极大化为‘1’, 极小化为‘-1’, 右端值记为‘0’。“约束描述”元素中的“约束类型”与“右端 b 值”元素是必需的有效内容。“约束说明”为注释型说明, 可有可无。

(3) XC.xml 文件

此文件为约束方程系数文件, 存放决策变量在约束方程及目标函数中对应的系数值。文件格式如下:

```
<xml>
.....
<变量系数描述>
  <约束方程编号>/</约束方程编号>
  <变量编号>/</变量编号>
  <系数值>/</系数值>
  <说明>/</说明>
</变量系数描述>
.....
</xml>
```

这里的“约束方程编号”与 SB.xml 文件中的相同, 并且要求系数按一个方程一个方程地顺序给出, 顺序要与 SB.xml 文件中约束顺序一致。“说明”元素可有可无。目标函数的系数类同约束, 在文件最后给出。

采用 XML 进行模型的符号语言表示主要有以下优点:

- (1) 实现了表现形式和数据的分离, 这允许它能够对来自不同系统的异构模型进行集成和处理。
- (2) 有很多成熟的 XML 解析技术 (如 DOM、SAX 和 XPP^[4]), 因此 XML 形式的模型处理效率很高。
- (3) 由于 XML 文件是纯文本格式, 因而可以对其所描述的模型进行灵活的编辑与修改。
- (4) XML 文件是严格结构化的, 便于检查模型的正确性及有效性。

4.2 XML 描述的 LP 模型生成过程

在 LP 模型的自动化生成过程中, 对问题描述过程中产生的语义框架, 可以通过定义一组模型结构控制参数、模型参数以及若干算法对其进行操作来完成将 LP 问题的语义框架描述转化成 XML 形式的符号语言表示, 以至最终生成 LP 模型。

模型结构控制参数为决策变量个数 n 和资源种类数 m , 因而模型中决策变量约束条件个数为 n , 资源约束条件个数为 m , 总体约束条件个数为 $n+m$; 模型参数为 C_i 、 a_{ij} 、 L_i 和 U_i , $i=1, 2, \dots, n$, $j=1, 2, \dots, m$ 。一旦这些参数被确定, LP 模型的结构也就随之确定。反过来, LP 问题的变化, 即问题语义框架的修改, 将导致模型结构控制参数与模型参数的改变, 模型结构也随之改变。

由语义框架形式的模型描述知识, 可得模型控制参数及模型参数 (m 、 n 、 C_i 、 a_{ij} 、 L_i 和 U_i), 调用相应的 XML 文件生成程序模块, 即可完成 XML 描述的 LP 模型生成。对于 XML 文件的生成, 目前已有非常成熟的处理方式, 这里不再赘述。

综合上述的讨论, 并结合第 2 节对 LP 问题建过程的分析, 给出如图 2 所示的 LP 问题自动建模方法流程示意图。

图 2 说明了从自然语言到符号语言以至模型生成的过程。决策分析人员只需通过人机交互界面, 利用自然语言进行 LP 问题的描述, 模型生成的其它过程均可由系统自主完

成,体现了模型生成的自动化特性。

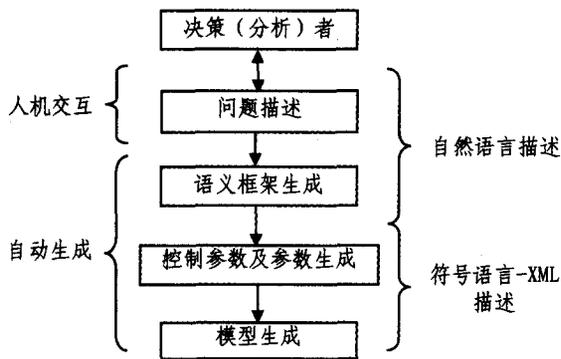


图2 LP问题通用建模方法流程示意图

5 应用实例分析

本节通过一个简单的LP问题建模过程验证本文提出的方法的科学性与可行性。

某厂拟生产甲、乙两种适销产品,每件利润分别为300元500元。甲、乙产品的部件各自在A、B两个车间分别生产,每件甲、乙产品的部件分别需要A、B车间的生产能力1、2工时;两种产品的部件最后都要在C车间装配,装配每件甲、乙产品分别需要3、4工时。A、B、C3个车间每天可用于生产这两种产品的工时分别为8、12、36。问题是:应如何安排生产这两种产品才能获利最多?

这是一个典型的产品配比LP问题。经过问题的自然语言理解过程,生成的语义框架描述如图3。

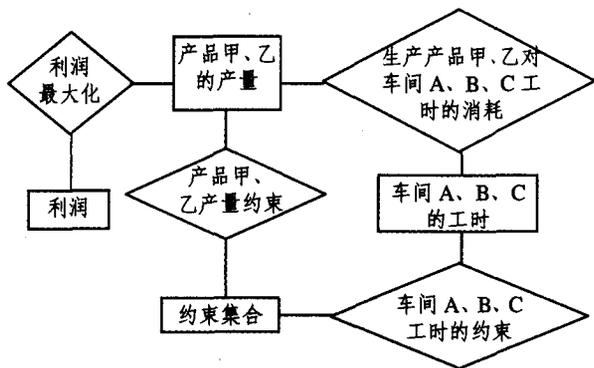


图3 一个实际LP问题语义框架

由问题的语义框架,可得到模型结构控制参数及模型参数,据此可以给出问题模型的数学描述:

$$\begin{aligned} & \text{目标函数: } \max Z = 3x_1 + 5x_2 \\ & \text{约束条件: } s. t. \begin{cases} x_1 \leq 8 \\ 2x_2 \leq 12 \\ 3x_1 + 4x_2 \leq 36 \\ 0 \leq x_1 \\ 0 \leq x_2 \end{cases} \end{aligned} \quad (3)$$

调用XML文件生成程序模块,生成对应的UL.xml文件、SB.xml文件以及XC.xml文件如下:

```
(1)UL.xml
<xml>
<变量描述>
<下界值>0</下界值>
<上界值>1.0e16</上界值>
<变量标注>产品甲产量</变量标注>
</变量描述>
<变量描述>
<下界值>0</下界值>
```

```
<上界值>1.0e16</上界值>
<变量标注>产品乙产量</变量标注>
</变量描述>
</xml>
(2)SB.xml
<xml>
<约束描述>
<约束方程编号>1</约束方程编号>
<约束类型>-1</约束类型>
<右端b值>8</右端b值>
<约束说明>A车间工时</约束说明>
</约束描述>
<约束描述>
<约束方程编号>2</约束方程编号>
<约束类型>-1</约束类型>
<右端b值>12</右端b值>
<约束说明>B车间工时</约束说明>
</约束描述>
<约束描述>
<约束方程编号>3</约束方程编号>
<约束类型>-1</约束类型>
<右端b值>36</右端b值>
<约束说明>C车间工时</约束说明>
</约束描述>
<约束描述>
<约束方程编号>4</约束方程编号>
<约束类型>1</约束类型>
<右端b值>0</右端b值>
</约束描述>
</xml>
```

```
(3)XC.xml
<xml>
<变量系数描述>
<约束方程编号>1</约束方程编号>
<变量编号>1</变量编号>
<系数值>1</系数值>
</变量系数描述>
<变量系数描述>
<约束方程编号>2</约束方程编号>
<变量编号>2</变量编号>
<系数值>2</系数值>
</变量系数描述>
<变量系数描述>
<约束方程编号>3</约束方程编号>
<变量编号>1</变量编号>
<系数值>3</系数值>
</变量系数描述>
<变量系数描述>
<约束方程编号>3</约束方程编号>
<变量编号>2</变量编号>
<系数值>4</系数值>
</变量系数描述>
</xml>
```

模型求解时,只需读出3个XML文件的内容,生成相应的模型控制参数及模型参数,调用相应的求解方法(如单纯形法、二阶段法、误差处理技术等)求解即可。

结束语 LP问题是管理决策中最基本的问题。在管理决策中LP问题并不确定,它随着用户相关的管理决策问题的目标和依据条件的不同而不同,这就要求给出一种带有通用性、自动化的LP问题建模方法。由此,本文给出了一种基于自然语言理解及XML的LP问题自动化建模方法。本方法是决策用户从繁琐的数学模型构造中解脱出来,利用自然语言,通过人机交互界面即可完成LP模型的构造工作,具有用户友好的特性;同时,模型的表示采用了XML文件形式,具有结构良好、易于集成、调用灵活等特性。实践表明,本文提出的方法具有科学性与可行性。

参考文献

- 1 Belew R K. Representation for Decision Support Systems. In: Methliie I B, Sprague R H, eds, Evolutionary DSSs, New York: North-Holland, 1985
- 2 Wang Yanzhang, Yang Bai, Wang Zhongtuo. ERP System of Knowledge Representation for DSS [C]. In: Proceedings of International Conference on Information & Systems, International Academic Publishers, Hangzhou, 1991. 852~856
- 3 向阳. 基于管理问题理解的模型智能构造系统研究: [学位论文]. 哈尔滨工业大学, 1999
- 4 刘芳,肖铁军. XML应用的基石: XML解析技术[J]. 计算机工程与设计, 2005, 26(10): 2823~2839