

基于决策问题形式化的模型构造方法

陈雪龙¹, 王延章¹, 李 强²

(1. 大连理工大学管理学院信息与决策技术研究所, 大连 116024; 2. 大连保税区经济发展局, 大连 116600)

摘 要: 在对目前决策模型构造方法的局限性进行分析的基础上, 给出决策问题的形式化表示, 提出一种基于决策问题形式化的模型构造方法, 从描述模型与问题模型两个角度探讨了该方法构造模型的过程。并以实例论证了该方法的科学性与可行性。

关键词: 决策问题形式化; 模型构造; 描述模型; 问题模型

Model Construction Method Based on Formalization of Decision Problem

CHEN Xue-long¹, WANG Yan-zhang¹, LI Qiang²

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

2. Economic Development Bureau, Dalian Free Trade Zone, Dalian 116600)

【Abstract】 The paper analyzes the limitation of the existing model construction methods. And the formal representation of decision problem is given. The construction method of decision model based on formalization representation of decision problem is proposed. The model construction process of the method is discussed in detail in terms of description model and problem model. An example of the method is given to prove the scientificity and the feasibility of the method.

【Key words】 decision problem formalization; model construction; description model; problem model

1 概述

自 20 世纪 70 年代以来, 决策支持系统(DSS)理论研究取得了巨大的进展, 但其自身存在的弱点^[1]在某种程度上制约了它的应用发展。目前绝大多数 DSS 的问题在于 DSS 辅助决策者建模的功能较弱, 不能很好地支持解决问题模型的构造, 在模型构造过程中需要用到大量的构模专业知识, 而大多数决策者并不是构模专家, 因此要求他们面对具体问题构造模型进行决策是难以实现的。没有一个有效的模型辅助构造过程, 要缩小在模型构造者和决策制定者之间的差距是非常困难的^[2]。

近年来, 为有效克服 DSS 这个弱点, 研究人员提出了多种构模方法, 文献[3]中将这些方法划分为 3 个类: 结构化构模^[4], 逻辑构模^[5]和图语法构模^[6]。结构化构模和逻辑构模方法的突出优点是描述简洁, 简化问题描述规模能够识别模型的基本组成成份, 以及这些成份之间的关系, 易于正确性校验等。但是, 这些方法的模型部件与数据部件特性不匹配, 为了获得构模所需的输入数据, 必须借助外部应用程序来实现, 并且在模型构造过程中, 还需要决策者具有相当好的数学基础。图语法构模较易于被决策者接受和理解, 用图形来描述模型, 能使决策者对模型结构有更加直观清晰的了解, 但模型中的有些关系无法用图形描述, 而且即使是稍复杂的模型, 用图形描述也显得相当庞大。同时也存在计算机储存与处理速度上的技术问题。因此, 该方法不适用于构造复杂模型。

为了解决上述构模方法所存在的问题, 基于知识的模型构造方法被广泛提出^[7-9], 尤其是面向对象方法, 反映了模型的自然层次结构, 是一种非常有前途的方案。但由于这些建

模方法没有严格遵照问题的形式化表示, 使得构造的模型在应用于具体问题求解时, 常有模型不适用或过于复杂等问题。

鉴于此, 本文从决策问题的形式化出发, 提出了基于决策问题形式化的模型构造方法。由于该方法的构模过程严格遵照了决策问题的形式化表示, 因此可有效解决构模结果的不适用性以及复杂性, 并支持决策问题的求解过程。

2 决策问题的形式化

一个客观事物总是可分成单元、元素或下一个层次的事物。文中将这些单元、元素、下一层次事物称为实体(entities)。其形式化表示如下:

设 e 表示一个实体概念, A_e 表示实体对应的属性集合, 即 $A_e = \{a_1, a_2, \dots, a_q\}$ 。其中, $a_i (\forall i \in \{1, 2, \dots, q\})$ 为实体相应的第 i 个属性; q 为对应的属性个数。这样, 一个实体就是下列的一个二元组 (e, A_e) 。实体集是由实体构成的某一类实体, E 是一个实体集, 则 $E = \{e | A_e \supseteq A\}$ 。其中, e 表示一个实体; A 是一个给定属性集, 一般是实体集或类的属性规范, 同时也是类实体的属性集。由于实体的概念和属性是相对的, 某些实体在另一事务体系中可以是相应的一些属性, 因此实体具有层次结构。

从事物普遍联系观点, 构成一个客观事物的所有实体都是相互联系的。这种联系体现在实体之间相互依赖、相互作

基金项目: 国家自然科学基金资助项目(70271045)

作者简介: 陈雪龙(1978 -), 男, 博士研究生, 主研方向: 地理信息系统, 决策支持系统, 数据仓库; 王延章, 博士生导师; 李 强, 博士研究生

收稿日期: 2007-03-10 **E-mail:** chenxl_dg@hotmail.com

用的关系(relations)。给定一个系统 S 的实体集 E 和属性集 A ，那么 E 上的一个关系 r 定义为 $r = \{O, I, B, C\}$ 。式中， C 是关系对应的属性集； $O \subseteq E \times C, O \neq \Phi, C \subseteq A, O$ 是被关联实体属性集； $I \subseteq E \times C$ 且 $I \neq \Phi$ ， I 是施联系实体属性集， B 是关联矩阵，即 $B = [b_{ij}]$ ，其中

$$b_{ij} \triangleq \begin{cases} 1 & \text{当 } x_i \text{ 与 } y_j \text{ 关于给定关系属性是关联的} \\ 0 & \text{否则} \end{cases}$$

其中， $x_i \in I, \forall i \in \{1, 2, \dots, m\}, y_j \in O, \text{对 } \forall j \in \{1, 2, \dots, n\}, m = |I|, n = |O|, B$ 为 $m \times n$ 布尔矩阵。

上述关系用图 1 加以描述，它类同于一般的输入输出关系。

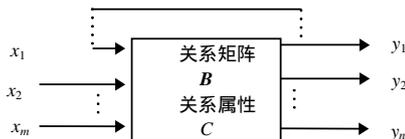


图 1 关系的图表示

一般情形， $O \cap I \neq \Phi$ ，即在一个关系中存在这样的实体，它既联系于其他实体而又被其他实体所关联，如图 1 中实体 $x_1, x_1 = y_1$ 。

上述关系很容易推广到用于模糊或概率型关系的直接表述，即令关联矩阵 B 中元素 $b_{ij} = [0, 1]$ ，则可以用于表述隶属度或概率值。

这样，任一个客观事物的运动或任一领域的知识，都可以分解成实体和关系加以描述。

问题是人的主观愿望或要求与客观事物系统状态的差异，因此，问题(problems)是关于人们主观目标(目标的概念来源于客观事物系统)和客观事物系统的集合。由此可以给出基于实体、关系的问题形式化表示：

给定客观事物系统 $S = \{E, R\}$ ，问题 p 定义为 $p = \{O, I, X, R_p, C\}$ ，其中， $O \subseteq E \times C, O \neq \Phi$ 为目标实体属性集； $I \subseteq E \times C, I \neq \Phi$ 为条件实体属性集； $X \subseteq E \times C$ 为关联实体属性集； R_p 为所涉及的问题关系集； C 为问题对应的属性集。

决策问题的形式化表示，可有效指导决策模型的构造，并且构造的模型可有效支持决策问题的求解。

3 基于决策问题形式化的模型构造

DSS 中的模型按其性质可分为描述模型和问题模型。下面给出两类模型的定义。

定义 1 描述模型是对客观系统组成部分及其联系机制的一种描述。

定义 2 问题模型是对决策问题的一种抽象描述，是问题的一种同态映射。

3.1 描述模型的构造

描述模型是对客观系统的一种描述，决策问题形式化表示中的实体 E 和关系 R 就是系统的一种最基本的描述，即

$$S = \{E, R\}$$

但这种描述不够深入，一般需要进一步确定实体的某些属性状态，以及这些属性状态的具体联系机制，这是描述模型构造的主要任务。

描述模型以知识为基础，是关系知识的进一步深化。给定 S 的一个关系 $r(r \in R)$ ，对应的输出实体集 O 确定一组属

性状态变量 Y ；对应的输入实体集 I 确定一组属性状态变量 X 。那么描述模型构造就是构造一组映射 F ，使得状态变量 X 通过 F 映射到状态变量 Y ，即 $F = X \rightarrow Y$ 或 $Y = F(X)$ 。

关于 F 的形式可由关系 r 的属性确定。如果 r 为逻辑型关系，则 F 为逻辑型映射；如果 r 为线性的，则 F 为线性映射；如果 r 为非线性的，则 F 为非线性映射，等。这样， F 的构造可变成一般化的系统辨识问题。对于系统辨识的具体方法可参看文献[10]。

这样对 $\forall r \in R$ ，可建立描述模型 $M_r, M_r: Y = F(X)$ 。

3.2 问题模型的构造

问题模型是问题的一种同态映射，因此一般要求是可解得，至少是可模拟的。依据决策问题的形式化表示，使得问题模型的构造变得规范化。

给定问题 $p = \{O, I, X, R_p, C\}$ ，其中， $O, I, X \subseteq E; R_p \subseteq R$ ，对 $\forall r \in R_p$ ，已知 M_r ，则对应问题 p 的模型可以以 M_r 为基础进行构造。

通常，一个问题模型可以表示为如下的四元组：

$$M = \{I, O, OP, K\}$$

其中， I, O 分别代表模型的输入输出属性数据； OP 代表模型的操作，包括属性数据的存取、模型的求解算法等； K 代表关于模型的知识。

因此，问题模型的构造主要包括输入输出数据(或模型变量)的构造以及模型操作(或输入输出映射关系)的构造。对于模型知识，作为决策分析人员对模型维护、调用的参考信息，可以通过多种手段获取。

对于模型变量的构造，可抽取描述模型中与问题模型相关的变量(包括自变量与因变量)，分别作为问题模型的输入变量与输出变量。即与问题模型目标相关的变量作为输出变量，决定模型输出的变量作为模型的输入变量。

对于模型映射关系的构造，依据问题 p 的类属性，如模拟问题、优化问题或评价问题等，可以确定模型类别。通常情况下，对应一类问题(是可解的)具有一种或多种通用方法或求解器，如线性规划问题、投入产出问题等。因此，在确定了模型类别之后，通常可以得到带有未知参数的函数，即参数化函数，通过系统辨识操作，确定未知参数，即可完成问题模型映射关系的构造。对于无法给出参数化函数的问题，可以借助具有学习功能的自适应网络，完成映射关系的构造。具体方法可参看文献[10]。

实际上，问题模型是描述模型的一种有机组合。这样，如果对应一个客观系统 S ，其对应的关系集 R ，若对 $\forall r \in R$ ，都能有相对应的描述模型 M_r 存在，则有关系统的各种问题模型的建立相对变得简单化。

4 应用实例分析

本节以畜牧业发展规划的模型构造为例，具体探讨基于决策问题形式化的模型构造过程。在探讨模型构造过程之前，首先给出畜牧业发展规划问题的形式化表示。

在分析畜牧业生产过程时，可将其实体分为畜种实体集 E_c 和环境实体集 E_e 两大类实体，其对应的属性集分别为 A_c 和 A_e ，如下：

$E_c = \{\text{猪, 肉牛, 奶牛, 奶羊, 马, 驴, 役牛, 绒山羊, 蛋鸡, 鸭, 鹅, 兔, 水貂, 蚕, \dots}\}$

$E_e = \{\text{粗饲料, 粗蛋白, 精饲料, 土地, 电, 劳动力, 技术人员, 物耗成本, 劳动力价值, \dots}\}$

$A_c = \{\text{饲养量, 价格, 产值, } \dots\}$

$A_e = \{\text{供应量, 价格, 价值量, } \dots\}$

依据上述实体,可建立反映畜牧业内部各实体及其属性间联系的关系集,包括: r_1 为品种猪状态方程关系; r_2 为品种肉牛状态方程关系; r_3 为品种奶牛状态方程关系; r_4 为品种奶羊状态方程关系; r_5 为单组品种状态方程关系; r_6 为资源消耗关系; r_7 为出栏产品关系; r_8 为畜产品关系; r_9 为畜牧业线性组合关系。

在以上实体与关系的基础上,可建立畜牧业系统的问题描述系统,包括: p_1 为品种猪状态分析问题; p_2 为品种肉牛状态分析问题; p_3 为品种奶牛状态分析问题; p_4 为品种奶羊状态分析问题; p_5 为单组品种状态分析问题; p_6 为资源消耗问题; p_7 为出栏产品问题; p_8 为畜产品问题; p_9 为畜牧业产值分析问题; p_{10} 为畜牧业优化问题。

对于以上问题,可采用如下形式化表示:

$$p_1 = \{O, I, X, R_p, C\}$$

其中, $O \in E_c$; $I \in E_c$; $X = \Phi$; $R_p = \{r_1\}$; $C = \{\text{模拟问题, } \dots\}$ 。对于 $p_2 \sim p_9$,同样可采用上述方式进行表示。对于优化问题 p_{10} ,可表示为

$$p_{10} = \{O, I, X, R_p, C\}$$

其中, $O \in E_c$; $I \in E_c$; $X \in E_c$; $R_p = \{r_1, r_2, \dots, r_9\}$; $C = \{\text{优化问题, } \dots\}$ 。

(1)畜牧业发展规划的描述模型构造

在已输入的畜牧业实体集、关系集的基础上,可构造畜牧业的描述模型。如上所述,描述模型要给出关系所描述的实体及其属性状态之间的具体联系机制,要确定一组状态变量 X 到 Y 之间的映射。对于畜牧业品种猪状态方程关系 r_1 ,可以以施联系实体集的实体,如“ $t-1$ 年猪繁殖组存栏数”、“ $t-1$ 年猪育肥组存栏数”等,取其数量值的属性为自变量 X 。以施联系实体集的实体,如“ t 年猪繁殖组卖出数”、“ t 年猪繁殖组购入数”等,取其数量值的属性为控制变量 U 。这样就完成了施联系实体集 I 到自变量 X 的映射。同样,以被关联实体集的实体,取其数量值的属性为因变量 Y ,这样就完成了被关联实体集 O 到因变量 Y 的映射。由于自变量 X 和因变量 Y 分别为相同实体的不同年份值,因此,可用 X_{t-1} 和 X_t 分别表示自变量和因变量,可以得到描述畜牧业品种猪状态方程关系的描述模型,具有如下形式:

$$X_t = AX_{t-1} + BU$$

畜牧业模型中的 A, B 参数可直接由畜牧业专家给出或通过其他参数算出,也可以利用线性回归、最小二乘法等方法进行辨识。

以上给出的描述模型是针对属性为差分方程的关系,对于关系 $r_2 \sim r_5$,因为其属性都属于差分方程,所以其构造方法类同。

对于关系 $r_6 \sim r_8$,其属性为投入产出关系,可以用类似的方法给出其描述模型,具体形式为

$$Y = A_t X$$

其中,参数 A_t 一般为畜牧业专家所熟悉的、经济含义明显的参数,通常可由专家直接给出或有其他参数间接计算出。

同样,对于属性为线性组合方程关系的关系 r_9 ,给出其描述模型为

$$V = A_t X$$

其中,参数 A_t 一般为各种产品的价格,可直接给出。

(2)畜牧业发展规划的问题模型构造

在以上生成的描述模型的基础上,以畜牧业优化问题 p_{10} 为例,探讨问题模型的构造。这里, p_{10} 对应的问题模型是一个动态线性规划模型。

1)问题模型的变量设置

在构造问题模型时,以描述模型的自变量和因变量为问题模型的模型变量,即上述描述模型中的 X, U 和 Y 等。

2)问题模型的约束设置

在构造问题模型时,对于差分方程关系所对应的描述模型,以其因变量与自变量和的差为零为问题模型的约束方程,即以 $X_t - AX_{t-1} - BU = 0$ 为约束方程。对于线性组合方程关系所对应的描述模型,以其因变量与自变量和的差为零为问题模型的约束方程,即以 $V - A_t X = 0$ 为约束方程。对于投入产出关系所对应的描述模型,以其自变量的和小于给定值为问题模型的约束方程,即以 $A_t X \leq b$ 为约束方程。

3)问题模型的目标函数

对于畜牧业发展规划,往往要考虑为数众多的目标,如总产值最大、成本最小、提供的大牲畜要达到一定水平等。但为了使问题处理方便,一般只取其中几个进行加权处理,作为目标函数,而使其他目标以约束方程的形式出现在模型中。所形成的问题模型可用多目标的求解软件进行优化分析。

对于 $p_1 \sim p_9$,可用类似方法给出其问题模型。至此,完成了畜牧业发展规划的模型构造。

5 结束语

本文首先论述了目前模型构造方法所存在的局限性,给出了决策问题的形式化表示,并提出了一种基于决策问题形式化的模型构造方法。该方法将决策模型按照其性质分为描述模型和问题模型,利用决策问题形式化表示中的E知识和R知识实现描述模型的辅助构造;在描述模型的基础上,并利用决策问题形式化表示中的P知识实现对问题模型的辅助构造。由于该方法的提出是建立在对决策问题形式化表示的基础上,因此模型构造变得规范化,并且所构造的模型可有效适用于决策问题求解过程。文章给出的建模实例,从实践角度论证了该方法科学性及其可行性。

参考文献

- [1] Zopounidis C, Doumpos M, Atsatsinis N F. On the Use of Knowledge-based Decision Support System in Financial Management: A Survey[J]. Decision Support Systems, 1997, 20(3): 259-277.
- [2] Yeom K, Lee J K. Logical Representation of Integer Programming Models[J]. Decision Support Systems, 1996, 18(3): 227-251.
- [3] Daniel R D. An Introduction to Model Integration and Integrated Modeling Environment [J]. Decision Support Systems, 1993, 10(3): 249-254.
- [4] Groffrion A M. An Introduction to Structured Modeling[J]. Management Science, 1987, 33(5): 547-585.
- [5] Kimbrough S O, Lee R M. Logical Modeling: A Tool for Management Science[J]. Decision Support Systems, 1988, 4(1): 3-16.
- [6] Jones C V. An Integrated Modeling Environment Based on Attributed Graph-grammars[J]. Decision Support Systems, 1993, 10(3): 255-275.

(下转第56页)