

# 税务稽查选案专家系统的研究

赵 璇 (中国科技大学研究生院 100039)

史忠植 (中国科学院计算所 100080)

**摘要:**本文在分析稽查选案业务的基础上,探讨了稽查选案专家系统的总体结构、设计思想及实现方式,提出了符合目前选案业务需求,适宜当前选案工作特点的专家系统解决方案。并以中科院计算所研制的“面向对象知识处理系统”为工具,开发了一个选案专家系统原型,目前已在沈阳市国家税务局稽查分局试运行成功。

**关键词:**税务 稽查 选案 专家系统

## 1. 引言

税收是国家财政的命脉,为保证国家税收及时、足额地入库,国家税务总局于1994年工商税制改革后确立了“以纳税申报和优化服务为基础,以计算机网络为依托,集中征收、重点稽查”的征管模式,税务稽查成为整个税收征管工作的“重中之重”。

税务稽查分为选案、实施、审理和执行四个环节,选案是稽查工作的首要环节,起着为后续工作提供目标和方向的作用。这项工作专业性强,不仅需要税收法规、税收会计等多方面知识的掌握,还需要实践和经验的大量积累。这些经验是整个税务系统的宝贵财富,将它们积累起来,用结构化、形式化、计算机可以识别、处理的方式存储起来,并利用计算机高效处理这一优势应用到实际工作中,必将发挥巨大的作用,并从根本上解决税务系统管户多,稽查人力匮乏这一难题。

## 2. 专家系统设计思想用于稽查选案的意义

在我国,将计算机用于税务稽查尚属新生事物。现有的业务需求也不尽完善,选案方法、选案指标、选案参数的设定主要凭实践经验,由于经验的局限性,以及市场经济的不断发展,税收法规政策的不断变化,偷逃税款的手段方式也不断翻新,这就需要选案参数等不断调理优化,立求选案的准确性。为此要求选案系统具有良好的灵活性、适应性、易操作性,能便捷的产生各种选案指标,并方便地对其进行维护。

现有的稽查选案系统大都采用传统的数据处理方式,存在以下弊端:

(1)传统数据处理方式的特点是“数据+算法”,将问题求解的知识隐含地编在程序中,维护起来需要在对源程序及各块接口细节全部了解的基础上,通过修改源程序才能实现,且需要重新编译;一个地方的修改可能导致

与其他地方的冲突,且这种冲突是不可预见的,这在某种程度上制约了维护的方便性、灵活性、简易性;

(2)传统数据处理方式作用于数据库或文件,其中仅包含明确的事实(实体和关系),而选案知识除事实外,大多是根据专家知识和经验总结出来的规则及解决问题的方案等内容,且往往带有不确定性、不精确性,有些是在不确定的信息的基础上凭直觉、经验作出的决定,针对这类知识很难用“非此及彼”的二值逻辑来表达。

将专家系统的设计思想用于稽查选案具有以下意义:

(3)传播成本高的专门知识。由于稽查人力有限,稽查专家的数量远不能满足实际需要,通过建造专家系统可以起到汇集、复现专家知识的作用,从而解决税务系统管户多,稽查人力匮乏这一难题;

(4)专家系统的设计思想是把知识从处理流程中独立出来,组成一个叫知识库的实体,知识库的处理是通过独立于知识库的、易识别的控制策略来进行的,因而相对于传统数据处理方式来说,它是一种“开放式”的系统,具有较好的开放性、灵活性、易维护、易扩展等优势,在税务稽查选案领域有着极大的应用潜力;

(5)专家系统作用于知识库,其中除事实外,还可以对根据人类专家知识和经验总结出来的规则、解决问题的方案等进行表示和推理。

## 3. 计算机选案业务简介

所谓计算机选案(以下简称选案),是指依据科学的选案标准,利用计算机对现有的内外部数据资料1进行加工处理、打分排序、分析筛选,以发现对税收影响异常的纳税人。

选案主要围绕“选案方法”和“选案指标”建立,选案方法是提供给用户使用的选案单位,它是若干选案指标

的有机组成。选案指标是基本的选案元素,它是从某一方面评价纳税异常与否的标准。所有的选案指标并非重要程度都一样,它们通过权值来体现。为了将纳税人的纳税异常程度进行量化,采用“打分”的方法进行选案,打分结果越高,异常程度越严重。

常用的选案概念、术语的层次结构及相互之间的关系可参见图1。图中的选案结果总分即打分结果,它是某次选案涉及的所有选案方法分数累加得到的总分值,选案方法分数又由它下属的所有选案指标分数加权求和而来。

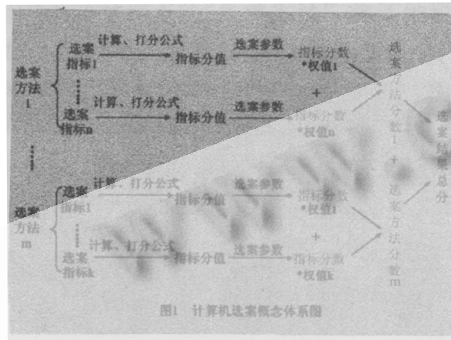


图1 计算机选案概念体系图

下面以销售税金负担率指标为例来说明选案指标的打分过程(参见图2)。

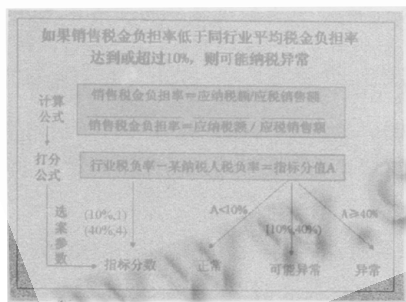


图2 选案指标打分示意图

#### 4. 选案专家系统总体结构

根据专家系统设计思想,并结合选案业务需求,构造如下系统总体结构(参见图3),各部分功能如下:

(1)数据采集、预处理模块:根据用户的选案要求,从所有纳税人资料中采集与本次选案任务相关的数据进行预处理,并将结果存入综合数据库中备用;

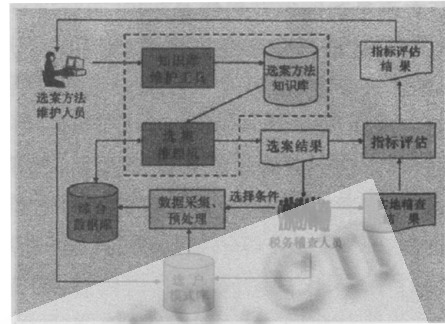


图3 选案专家系统总体结构示意图

(2)知识库维护工具:选案方法维护人员可通过它对选案方法知识库进行创建、维护和管理,它提供知识的浏览、增加、修改、删除等功能;

(3)选案推理机:从综合数据库中提取与本次选案相关的数据,利用选案知识库中的知识完成选案推理工作,并将选案结果反馈给稽查人员;

(4)指标评估模块:将计算机选案结果与实地稽查反馈结果进行比较、分析,验证选案指标的准确性,找出偏差原因,为维护选案方法知识库提供依据;

(5)选案方法知识库:存放与选案领域相关的各种规则、知识,包括稽查专家的实践经验和各种相关的税收法律、法规。选案方法知识库是税收业务知识的具体反映,在选案中起着主导作用;

(6)综合数据库:存放与每次选案相关的纳税人的各项涉税信息及选案中间结果;

(7)选案模式库:存储常用的选案模式。有时在进行选案前,需要先从不同角度对纳税人的各项涉税信息进行全面分析,找出稽查重点,再利用选案方法进行选案。选案模式就是由这些常用的分析方法归纳、提取而成的,它可以辅助选案人员对纳税人的资料进行分析,从而缩小选案范围,提高工作效率;

(8)选案结果:待查纳税人清册及各项选案指标结果分析表,通过该结果稽查人员可掌握所有异常纳税人各项选案指标的得分情况及税收异常程度;

(9)实地稽查结果:稽查人员将选案结果用于实地稽查得到的结果;

(10)指标评估结果:根据指标评估算法得出的对各项选案指标实际应用成果的评价,它是对选案方法知识

库进行修改、完善的参考依据。

### 5. 选案专家系统的设计思想及实现

、受篇幅所限,这里只介绍专家系统的核心——知识库和推理机的设计和实现。

#### (1) 选案方法知识库

①知识表示方法——面向对象知识表示法。所谓知识表示就是研究表示知识的原则和有效的表示方法或模式,以把人类知识表示成计算机能处理的数据结构。对于每条选案指标,它的结构既包括属性信息,如指标权值、选案参数、指标分数,又包括与属性相关的具体操作,如指标计算公式、打分公式。为此可把选案指标视为包括属性及其相关操作的对象,因而可采用面向对象的知识表示方法对其描述。

面向对象的知识表达方法以对象为中心,将对象属性、动态行为特征、相关领域知识、数据处理方法等有关知识“封装”在表达对象的结构中,不仅与人的认知习惯相近,还具有模块性强、知识单元较独立、代码可复用等优点。

具体到选案方法知识库,它由若干选案方法构成,可将选案方法视作一个对象类,它有自己的属性和行为,将它进行实例化即为每种具体的选案方法,其属性包括它所用到的选案变量及用该方法选案得到的选案方法分数;其行为是用此方法选案所要做的各项具体工作。另外,例属于某个选案方法的选案指标可以作为该选案方法的子类,它除了继承其父辈的属性和行为外,还有自己的属性,如指标权值、指标分数。这样,对每个选案方法而言,它是一个由选案方法根节点和若干选案指标子节点组成的知识树。

为了实现多种方法综合选案的目的,把每个选案方法视为一个子知识库,并将它们的父节点作为根节点——姑且称之为选案方法根节点。由此不难想象,整个选案方法知识库就是一棵层次结构的树——根节点下面并列所有选案方法,每个选案方法的子节点为该方法下的选案指标。

②选案方法知识表示示例。选案专家系统采用中科院计算所史忠植、谭宁等同志研制的面向对象知识处理系统(Object-Oriented Knowledge Processing System,简称OKPS)作为开发环境。OKPS是一个专家系统开发平台,它包括两部分:可视化知识管理工具和面向对象的推理机。OKPS把知识看作对象类,将客观事物和规律

的属性及其行为属性封装起来,通过属性槽和侧面描述事实性知识,用行为表示推理规则,知识对象之间通过语义联系和对象约束关系构成知识树,由推理机控制搜索求解;知识处理系统内部各个部分是相对独立的,它们被封装在相应的类中,用对象的消息控制机制进行各个部分间的接口。推理机可针对求解问题的不同性质,实施不同的工作机制,如提供先序、中序、后序三种不同搜索策略。

图5是用OKPS开发的“选案方法知识库维护工具”界面。

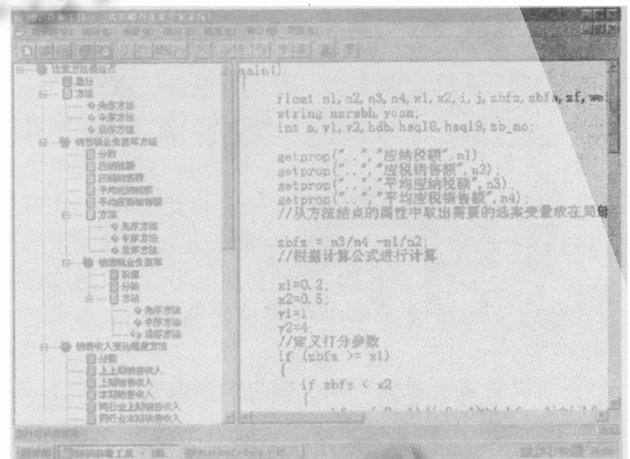


图5 选案方法知识库的知识表示直观图

上图的左侧窗口显示了存储选案方法的知识库结构,它是一棵以选案方法根节点为根的知识树,根节点的下一层结点(如销售税金负担率方法)的名字就是所有选案方法的名称,每个选案方法结点下属的子结点与此选案方法涉及的选案指标相对应。每个结点可视为一个对象,它有自己的属性和行为(图中表示为方法)。属性用来保存数据,行为用于描述它要完成的操作。上图的右侧窗口即为销售税金负担率指标后序方法的控制脚本。

(2)知识的利用——推理机。在选案系统中,推理机的任务就是利用选案方法知识库中的知识,并根据纳税人的各项涉税信息完成选案(打分)的过程。

由于某次选案的结果总分由它涉及的所有选案方法分数累加而来,选案方法分数又通过它下属的选案指标分数加权求和而来,为此要实现每个纳税人的综合打

分,首先要对各选案方法包含的所有选案指标打分。因而从打分的角度来说,推理机制采用自底向上的方式,而对各结点的访问则采用自上而下的方式,即先序方式。

推理机运行选案时,所有处理过程是严格按照下列次序执行的——就像出场次序一样:

①执行选案方法根结点的“先序方法”,为开始选案做打开数据库之类的准备工作;

②对每一个纳税人重复步骤3至步骤9

③对此次选案涉及的每一种选案方法,依次执行步骤4至步骤9

④执行方法结点的“先序方法”,将该纳税人的有关选案变量值读到方法结点的相应属性中;

⑤对当前结点的每一个子结点依次执行步骤6至步骤8;

⑥执行它的“先序方法”,进行变量有效性检查之类的操作;

⑦如果它还有子结点,递归的调用步骤5至步骤8;

⑧执行它的“后序方法”,进行指标分数计算,向选案结果-指标分数表中写数据等操作;

⑨执行方法结点的“后序方法”,进行选案方法分数计算,向选案结果-方法分数表写数据,并将此选案方法分数累加至选案结果-总分表;

⑩执行选案方法根结点的“后序方法”,即完成关闭数据库等操作。

## 6. 小结

将专家系统的设计思想引入税务稽查选案,并采用面向对象方法构造知识库,有着传统数据处理方式不可比拟的优点:

(1)即专家系统优于传统数据处理系统最本质的一点,它把知识从处理流程中独立出来,组成知识库,知识

库的处理是通过独立于知识库的控制策略来进行,使系统具有较好的开放性、易维护性;

(2)系统采用面向对象的知识处理方式构造知识库,将每种选案方法视为一个对象,这种构造方式接近人类思维和会话的自然形式,易于理解;

(3)专家系统能够更好地描述和处理选案专家的各种经验知识;

(4)系统具有良好的模块性、灵活性。每种选案方法将与其相关的属性及操作行为封装起来,具有相对独立性,对一个知识单元的修改不会影响其他单元;另外,每种选案方法都具有相同的结构和格式,增加新的选案方法时,只需做简单的修改即可,有较好的可扩充性;

(5)对某一地区、某一稽查部门来说,使用构造一致的选案方法知识库进行选案,选出的待查户具有可比性,便于下达稽查计划。

目前该系统已在沈阳市国家税务局稽查分局试运行成功。实践证明将专家系统的设计思想用于稽查选案是可行的,也是大有潜力的。随着选案业务需求的不断完善,以及系统功能的不断扩充,它必将成为稽查人员选案的有力工具。

## 参考文献

- [1] 税收征管软件业务规程,国家税务总局,1998年;
- [2] 石纯一,黄昌宁:人工智能原理,清华大学出版社,1993;
- [3] 史忠植:高级人工智能,科学出版社,1997;
- [4] 张全寿 周建峰:专家系统建造原理及方法,中国铁道出版社,1992。

(来稿时间:1999年7月)