

矿床统计预测单元划分的方法与程序^{*}

李新中¹ 赵鹏大² 肖克炎³ 胡光道²

(1 北京大学地质学系, 北京 2 中国地质大学, 北京 3 中国地质科学院区划室, 北京)

提 要: 在矿床统计预测中, 如何确定最佳的统计预测单元, 目前尚无通用的准则和算法, 它的划分往往取决于地质学家对研究区控矿条件的认识, 矿化的实际分布形态及统计预测所采用的评价模型等多种因素。在系统总结和研究单元划分的领域专家知识的基础上, 提出了用产生式规则语言来表达单元划分的知识结构和实现方式, 探讨了用 C 语言设计和实现单元划分专家系统的方法和途径。系统功能齐全, 基本上满足了资源评价的各种需求, 实现了预测单元划分的自动化和智能化。

关键词: 矿床统计预测单元 规则 模拟 领域专家知识

矿床统计预测单元的划分是矿产资源评价过程中最关键的一个环节, 其目的是为了确定地质变量的观测尺度和取值范围, 提高评价结果的准确性, 而单元类型和大小, 尤如样品采集和分析那样, 其取样方法及大小不同, 获得的结果对地质现象描述的精确程度也不同, 从而直接影响资源评价的效果。预测单元划分得太小, 则明显地扩大了无矿单元和单一控矿单元的数目, 增加了预测的工作量, 不利于资源评价模型的建立, 而预测单元划分得太大, 则歪曲了实际有矿单元的分布形态, 使误判有矿的地区面积增大, 不利于找矿工作的进行, 并使预测的靶区可信度降低。因此, 如何确定最佳预测单元, 并不是一件轻而易举的事, 它必须结合实际资料水平和采用的评价模型, 选择合理的单元划分方法。一般情况下, 当地质条件变化多样, 区内的地质及物化探资料齐全时, 采用不规则单元为宜; 当需要对地质变量进行统一观测和取值范围时, 应采用规则单元; 而对于预测区内地质构造变化, 特别是对矿带内岩层、矿层与线型褶皱构造紧密有关的矿带及矿田范围内进行评价时, 应采用垂直矿带走向, 按一定间距穿越整个矿带划分成长条形单元, 即穿透单元作为变量取值或评价的基本单位。故此, 根据不同的地质信息水平及矿产资源评价方法的要求, 本文分别设计了面向初级用户的单元划分专家咨询法, 以及面向高级用户的模拟试验法和仿真 CAD 法。

1 系统的总体功能结构^[1]

单元划分咨询系统是用 C 语言及 GURU 语言写成, 它是矿床统计预测专家系统的一个子系统^①, 部分功能需要主系统的支持, 如图件数据的采集模块及图件编辑模块^[2]等, 但该系

^{*} 地矿部科技司高科技项目《矿产资源地质环境专家系统》资助 (编号: 8505705-1)

李新中, 男, 32岁, 北京大学地质学系博士后, 数学地质专业博士, 现主要从事地质专家系统、地质过程的计算机模拟, GIS 软件的开发及在地质学中的应用等研究工作。邮政编码: 100871

① 李新中, 1994, 中大比例尺矿床统计预测专家系统的研制 (博士论文), 中国地质大学 (武汉)

1997-06-09收稿, 1998-01-12修改回

统由于遵循结构化程序设计的原则，因此单元划分系统自成体系，它与主系统各功能模块及外部应用程序具有方便的接口，从而便于系统的扩充和维护。图1直观地表达了本系统的主要功能。各功能在系统中以命令菜单的形式调用。

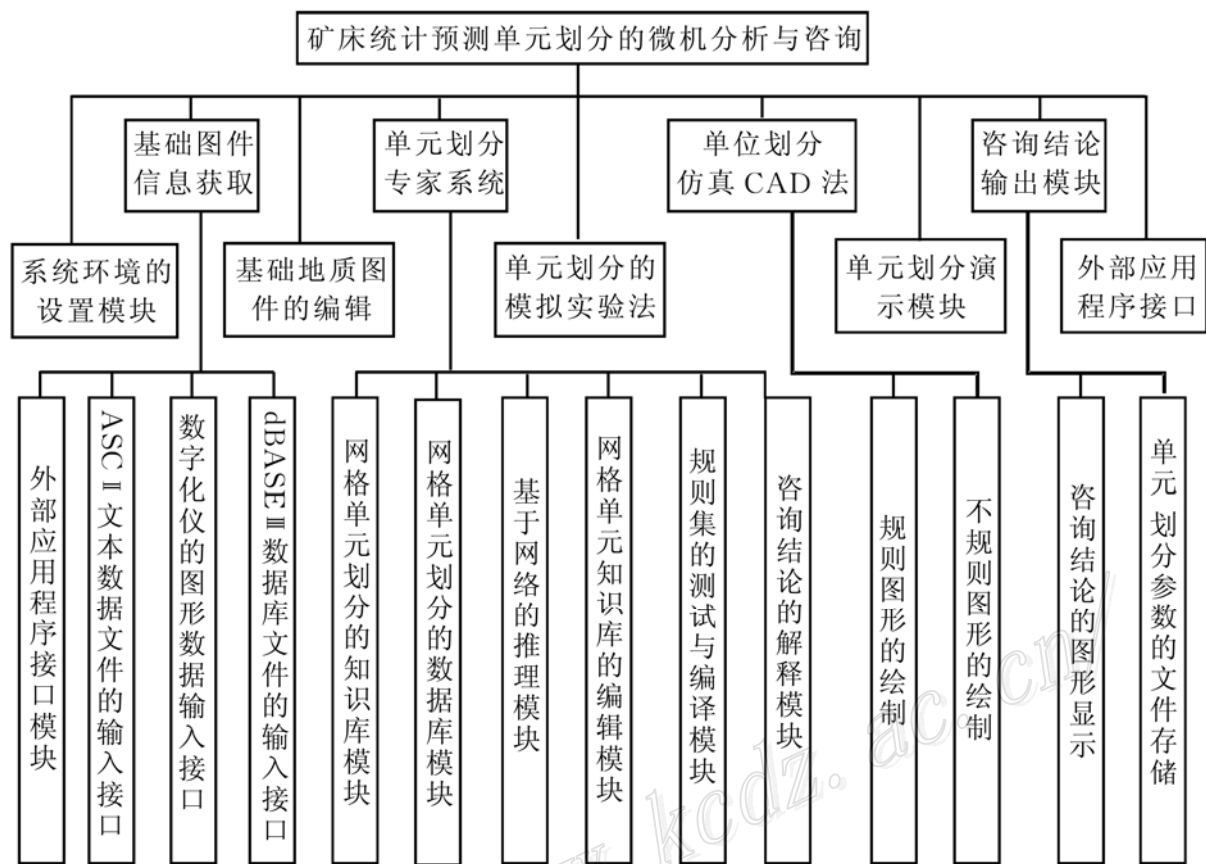


图 1 单元划分的结构功能框图

Fig. 1. Function architecture block diagram of the unit partition system.

2 网格单元划分的专家咨询法

2.1 单元划分的领域专家知识的抽取^[3,4]

基于网格单元的资源预测方法主要是采用概率统计的分析方法，其基本思想就是选择一定大小的网格将整个研究区划分为面积相等形状相同的单元，用作统一观测和取值的基本单位，通过样本的观测来描述总体，并遵循抽样的随机性及样品的代表性原则。因此，单元的大小，对建立的资源评价模型的合理性和预测的效果有直接的影响。为此，本文将一个有经验的矿床统计预测专家对预测单元划分所考虑的因素归纳为以下几点。

(1) 统计预测的比例尺及精度要求：从事矿产资源评价的地质学家经过大量的实践，将地质图件的比例尺与预测单元的大小关系总结为：当预测比例尺为1:5万时，单元大小应为0.25~1 km²为宜；1:20万，单元大小应为4~16 km²；1:50万，单元大小应为25~100 km²。因此，在同比例尺精度条件下，模型区和预测区的单元大小应根据相应的地质图比例尺而确定。

(2) 研究区地质条件复杂程度：合理的单元划分应该能使模型区建立的评价模型反映代表区域地质构造背景和矿床地质条件变化的信息。如当地质条件变化复杂，信息量大时，在不增加预测工作量的前提下，应采用较小的预测单元；而地质条件单一，单元信息量少时，在

既不会漏矿又能满足评价方法所要求的地质变量个数的前提下, 应该采用较大面积的单元。

(3) 研究区出露的矿点数及其空间分布特征: 不同大小面积的单元, 其矿点分布模型也不同。从矿点分布的均方差 $\sigma(x)$ 与其数学期望 $E(x)$ 的比值变化可知, 当单元面积越大, 则单元矿点的分布模型越接近具有丛集型分布特征的负二项分布, 单元面积越小则越接近具有随机型分布特征的波松分布。因此, 在用理论分布模型拟合时, 应以有矿单元的分布形态尽可能地与实际矿化的分布形态吻合为最佳的单元划分方案。

(4) 研究区内矿点数及研究范围大小: 在实践中, 地质学家根据研究区内矿点数量以及预测范围的大小, 提出最优单元面积的经验计算公式: $s_1 \times 2/n$, s_1 为研究区总面积, n 为矿点总数, 赵鹏大教授提出, 在单元大小能保证当矿床点的分布为随机型时, 用二项分布模型 $P_n(x) = C_n^x p^x q^{n-x}$ 的均方差 $\sigma(x)$ 和数学期望 $E(x)$ 的比值关系, 将以落入单元内的期望矿点数等于或小于实际落入单元矿点数标准差的 3 倍为准则。即: $\sigma(x)/E(x) = \sqrt{((1-s)/(n \times s))} = 1/3$, 其中 $E(x) = ns$ 为落入单元内的矿点数的数学期望; $\sigma(x) = ns(1-s)$ 为实际落入单元内矿点数的标准差; n 为预测区矿点总数; s 为单元面积百分比 ($s = 9/(9+n)$); s_1 为预测区总面积 (km^2), 则统计预测单元面积为 $9/(9+n) \times s_1/100 (\text{km}^2)$ 。

(5) 保证统计分析所必需的单元数: 单元的划分必须使有矿单元的数目符合统计分析的最小数目要求, 即含矿单元数要大于变量数, 如在应用判别分析、因子分析及回归分析等评价方法时, 必须遵守此规则。

(6) 地质特征的空间变异性: 利用区域化变量理论研究单元内地质变量的空间结构性。当地质变量的结构变化表现为各向同性时, 我们采用正方形单元; 当结构变化表现为各项异性时, 则采用长方形单元, 并以变化性最大的方向为长方形的短边。

2.2 单元划分领域专家知识的表达方式^[5~7]

知识表示是一组用于描述知识对象的语法和语义约定, 是知识的形式化, 符号化过程。根据单元划分领域专家知识的特点, 我们采用基于规则知识表示的 GURU 专家系统工具提供的产生式规则语言, 以“属性—值”的形式来表示单元划分的事实和规则, 其规则的知识表示形式为:

```

PRIORITY : 〈数值〉
COST : 〈数值〉
IF : 〈事实变量1〉〈运算符〉〈值〉
    〈事实变量2〉〈运算符〉〈值〉
    .....
    〈事实变量 n〉〈运算符〉〈值〉
THEN : 〈结论变量1〉〈运算符〉〈值〉
    〈结论变量2〉〈运算符〉〈值〉
    .....
    〈结论变量 n〉〈运算符〉〈值〉
COMMENT : 〈文本注解〉
READY : 〈GURU 命令序列〉
NEEDS : 〈变量表〉
CHANGES : 〈变量表〉
  
```

REASON：〈规则理由〉
VAR.：〈变量名〉
FIND：〈提问文本〉
LABEL：〈变量注解〉

其中 PRIORITY：指明该规则做匹配测试时的优先次序；COST：规则的成本大小，激活规则时，首先激活代价较小的规则；COMMENT：用于对该规则的文本注解；READY：在激活该规则时所需执行的 GURU 命令序列；NEEDS：指明在测试该规则之前应寻找的一些变量，若该项省略，则推理机自动寻找仅在前提中出现的变量；CHANGES：指明该规则将对其操作的那些变量，若该项省略，则推理机认为操作只与结论中所列的变量有关；REASON：该规则对目标存在的理由；VAR：变量名；LABEL：该变量的文本注解；FIND：激活该规则时，对出现的未知变量的自动提问文本，在每条规则中前提或结论中的变量均可赋予一个从 0（表示未知）到100（表示全真）间的整数作为其确信因子，变量间可进行“与（and）”、“或（or）”、“非（not）”及“异或（xor）”的操作运算，前提和结论中允许多个结点共存，规则集激活时，它们之间形成一种网络关系，有利于表达复杂的地质经验知识。

2.3 知识的获取及编辑^[9]

对单元划分领域专家知识的获取和编辑，系统支持两种方式，一种是菜单驱动的知识库编辑器，其命令菜单见图2，系统采用多级菜单命令和窗口引导开发者输入关于单元划分的有关信息，用户无须了解系统关于规则编辑的有关知识即可进行知识库的建造，其命令菜单见图2。它比较适合初级规则集开发管理人员，规则集的建立分为：①初始化；②目标；

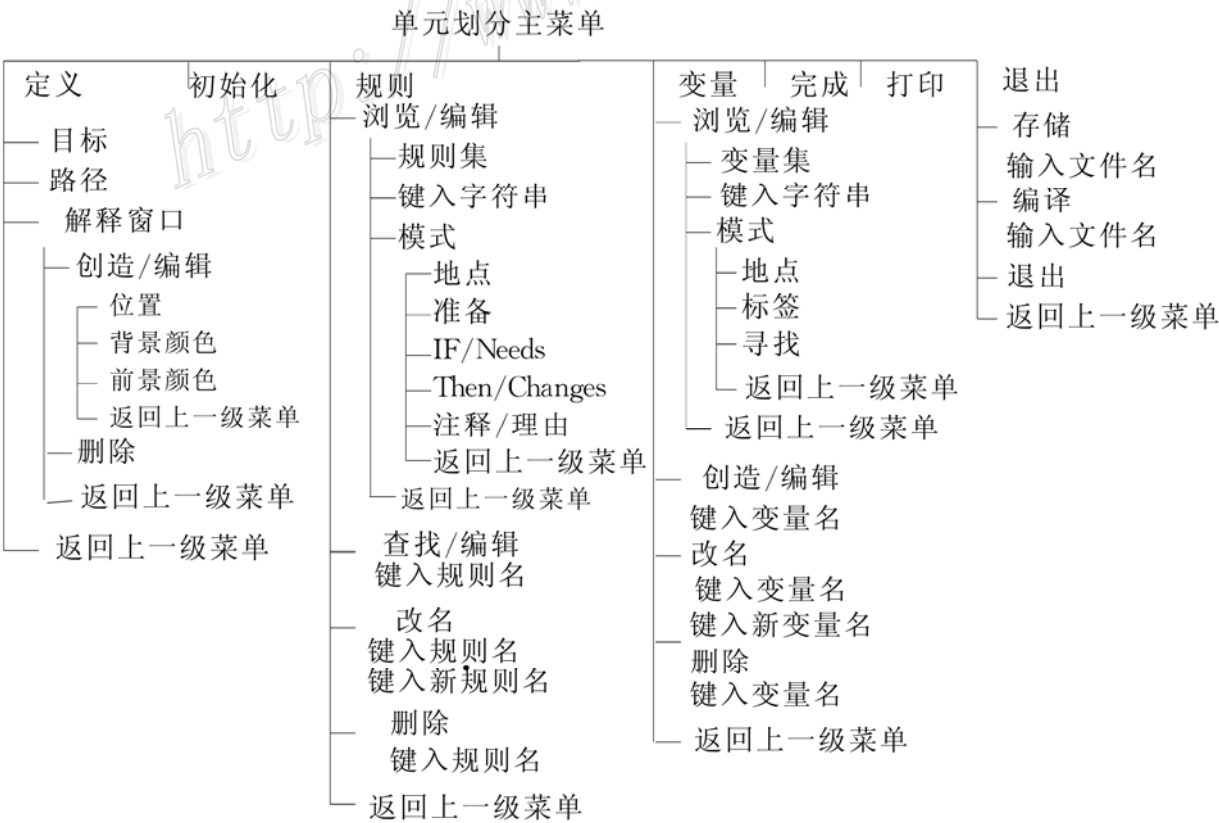


图 2 单元划分的知识获取与编辑命令菜单
Fig. 2. Instruction menu of rule set editing.

③ 规则；④ 变量；⑤ 完成；⑥ 窗口；⑦ 访问 7 个部分。另一种方式是对于有经验的规则集开发人员系统允许开发者利用 GURU 内部或外部提供的文本编辑器进行知识库的建造。开发者在进行知识获取时，虽然需要规则集的语法知识，但比菜单引导能够进行非常灵活的规则集查询和编辑。规则集经编译后即可用于问题的咨询。单元划分的规则集见脚注文献^①。

2.4 专家系统的咨询方式^[6,8]

GURU 是一个协同式集成化的专家系统开发工具，除了一个专家系统“外壳”以外，系统与外界的各种系统程序 DOS, Lotus 1-2-3 及 dBASE III 等都有方便的接口，并提供了互相调用的功能。根据这一功能，系统用 C 语言设计了单元划分的图形输出咨询接口，将专家系统的字符咨询转换为图示语言的咨询输出，从而使专家系统的咨询更为灵活，咨询的结论形象、直观易于用户的理解和接受。系统咨询的方法采用菜单提问的方式，通过对定制的四钟推理接口即正向、正向追踪、反向、反向追踪推理的选择来激活相应的规则。当用户选择追踪推理时，系统在咨询的同时动态地显示当前激活的规则及其对当前问题的解释；当用户选择正向推理时，可以按 CTRL+Y 键来得到解释，用户也可以在咨询后，通过定制的菜单接口，选择 How 或 why 命令序列来解释咨询的结论。经推理计算，先输出单元划分的最佳参数表列，然后利用外部应用程序接口转换成相应的网格图形叠加于地质图上。其咨询的过程以陕西勉略宁地区 1:5 万矿产预测为例，其咨询方式如下：

输入预测区比例尺 Enter the Sale of Prediction Field : 1:50000
输入预测区面积 Enter the Area of the Prediction Field AX: AY;
横向距离 AX (km) : 40; 纵向距离 AY (km) : 15
输入预测区出露的矿点数 Numbers of Mineral Deposit Points : 100
地质条件的复杂程度 Complex Degree of Geological Environment (0.1~1.0) : 0.6
地质条件变化的最大方向 Select Maximum Variation Direction (0 / 1) : 0
系统正在思考，请等待 MILASP is Thinking Away , Please Wait
最佳单元面积 The Unit Area (km²) = 1.00
研究区单元总数 The unit total number = 600
单元的横向边长 X side of the unit (cm) = 2.0
单元的纵向边长 Y side of the unit (cm) = 2.0
预测区横向单元数 Number of X side = 40
预测区纵向单元数 Number of Y side = 15

通过对以上参数表列的图形输出，可以看到有矿单元的分布形态与原有的矿床分布形态吻合较好，因此，2×2 的网格单元是该区 1:5 万矿产预测单元划分的最佳方案。说明该系统的解题水平已达到专家级的水平，其知识库的组织是合理的。

3 单元划分的模拟试验法模块

模拟试验法主要是根据用户直接提供的单元划分参数，在微机终端进行网格单元的划分

① 李新中，1994，中大比例尺矿床统计预测专家系统的研制（博士论文），中国地质大学（武汉）

和显示,单元大小可根据用户的经验和知识,随意缩放和调整,直到用户认为有矿单元的分布形态与实际矿床点的分布形态一致为止。该模块操作界面友好,划分方法形象直观。用户每次启动系统时,首先自动弹出一窗口提问输入参数,待用户输入后,窗口自动消失,并在地质矿产图上显示叠加该参数的网格图形。以此往复,直到用户选择弹出窗口中的退出选择项,则系统存贮参数后返回主系统。

4 应用仿真 CAD 技术的非网格单元的划分^①

前述两种单元的划分都是以满足抽样的随机性和代表性为前提而进行的规则网格单元的划分,其优点是单元在统一观察和度量的前提下,它可以把众多的地质变量所包含的矿产资源信息最大限度地反映出来,有利于矿与非矿地质特征的判断,虽然此法给矿产预测的计算机化带来了方便,但它也存在一些问题,如二个等大的单元中,提取的花岗岩的出露面积相同,但其中一个单元的花岗岩是由后期构造作用的外来推覆体,而另一个单元的花岗岩则是真正的侵入体,它是有根的。因此,两个单元的地质意义与控矿意义是相差极大的。再者,在众多的等大单元中,通常研究程度和观察尺度水平是不尽一致的,而单元中各种物探、化探和重砂异常也具有不同的等级。因此有的地质学家提出了以地质体为单位进行预测单元的划分,这样可以揭示不同性质、不同等级地质体的综合信息特征,便于控矿信息的充分提取和方法的应用。

根据地质体单元边界形态的不规则特点,本文设计了一个人机交互的应用仿真 CAD 橡皮带式绘图技术的非网格单元划分模块,可以绘制圆、矩形、直线、不规则多边形等图形。因此,地质人员根据微机终端所显示的地质图面信息,可直接用鼠标凭自己的地质认识和经验直接在屏幕上以地质体为单位的预测单元划分。对于圆形地质体,用户只需移动鼠标于当前地质体中心按键定位,随后屏幕光标随着用户拖动鼠标自动形成一个以地质体中心到当前屏幕光标位置为半径的圆,该圆的大小随着屏幕光标的移动始终拉伸着一个变化的圆,光标走到那里,圆就拉伸到那里,以此对图形实行动态跟踪,当用户认为研究对象与圆扣合时,只须再按一次鼠标键,则当前图形即可固定于地质图面上,系统同时自动记录当前图形单元的坐标并存贮于相应数据文件中,并为主系统的调用提供单元边界坐标数据^①。同理,对于长方形、矩形和不规则多边形等形状的地质体系统也以同样的方式绘制预测单元,并记录其图形边界坐标值于相应的数据文件中。其工作功能框图详见^①。

参 考 文 献

- 1 赵鹏大,胡光道,李新中. 基于语义网络知识表示的专家系统的设计. 地球科学, 1994, 19 (4): 391~401
- 2 李新中. 地质图件的微机编辑及信息的智能取值系统. 见: 岩石圈地质科学. 北京: 地震出版社, 1996, 162~172.
- 3 赵鹏大,胡旺亮,李紫金. 矿床统计预测. 北京: 地质出版社, 1982, 16~46, 94~176.
- 4 赵鹏大. 地质勘探中的统计分析. 武汉: 中国地质大学出版社, 1990, 171~181, 257~262.
- 5 Nilsson N J. 人工智能原理. 北京: 科学出版社, 1984, 49~116, 171~241
- 6 何新贵. 知识处理与专家系统. 北京: 国防工业出版社, 1990, 32~88

① 舒青, 1990, 微型计算机高级图形程序设计技巧实例, 中科院希望高级电脑技术公司, 第118~197页

- 7 黄可鸣. 专家系统导论. 南京: 南京大学出版社, 1988: 2~40.
- 8 李新中, 赵鹏大, 胡光道. 基于规则知识表示的模型单元选择专家系统的实现. 地球科学, 1995, 20 (2): 173~178.
- 9 MDBS, Inc. GURU reference manual. Vol. 1~2, Lafayette: MDBS, Inc, 1985, 1~2: 1~300.

THE METHOD AND PROGRAM FOR THE UNIT PARTITION IN THE STATISTICAL PREDICTION OF MINERAL DEPOSITS

Li Xinzhong¹, Zhao Pengda², Xiao Keyan³, Hu Guangdao²

(1 *Department of Geology, Peking University, Beijing 100871*)

(2 *China University of Geosciences, Wuhan 430074*)

(3 *Division of Minerogenic Regionalization, Chinese Academy of Geological Sciences, Beijing 100037*)

Key words: statistical prediction unit of mineral deposits, rule, simulation, field expert knowledge

Abstract

In statistical prediction of mineral deposits, there have been no general criteria and formulae as to how to determine the most suitable unit. Its partition usually depends on geologists' recognition of ore-controlling conditions, distribution shapes of mineralized occurrences as well as appraisal models used. Based on systematic summarization and research of field expert knowledge of unit partition, this paper suggests that knowledge structure and realization method with IF~THEN rule language is used to express unit partition. It also discusses the methods and ways of using C programming language to design and realize the expert system for unit partition. The program system has perfect functions. These methods can satisfy various requirements of resources assessment, and realize automation and intellectualization of the prediction unit partition.