

知识表达语言 (NFA) 及其在石油测井 解释专家系统中的应用

宋国宁 齐继光 张士杰 杨淮清

(中国科学院沈阳自动化研究所)

伊广林 吴付东 吴金丽 吕健儒

(石油部胜利油田测井公司, 山东)

提要 石油测井解释是一项逻辑推理和数值计算交错进行的复杂过程。为了描述测井解释专家的这种知识、经验并模拟其思维方式, 在扩充纯产生式规则的基础上, 我们开发了知识表达语言——NFA, 它把逻辑推理和数值计算综合成统一的形式。

石油测井解释专家系统LIX先后在INTERDATA-85机和PE-3230机上实现, 现场(胜利油田)运行近两年, 解释了130余口井, 符合率94%以上。

LIX实质上是NFA语言的解释系统, 它的研制成功, 说明了NFA语言的有效性和实用性。

1 概述

石油测井解释的任务是: 首先把测井数据(如地层的电阻率、自然电位、声波时差, 等等)转化为地质信息(如岩性、泥质含量、孔隙度、渗透率、含油饱和度等等), 再参照地质观察, 最后, 给出油、气、水的综合评价, 即对地层内的油和气的储量进行判断。在解释过程中, 解释人员不仅要针对不同的岩性, 选择数百个表示地层特性的解释参数, 而且要使逻辑推理和数值计算交错进行: 前者为后者选择合适的计算公式(或算法), 后者为前者提供判断依据。可见, 测井解释是非专家所难以胜任的智力活动。这正是我们研制石油测井解释专家系统LIX的目的。

产生式系统被广泛地用于AI程序中, 但是它有一个明显的缺点: 与其他程序设计语言不同, 它没有函数、子程序及其调用, 故不能表达算法知识。为此, 我们用专门设计的关联函数来对纯产生式规则加以改造和扩充, 从而把逻辑推理和数值计算综合成统一的形式——NFA语言。在关联函数中, 引入三值逻辑概念(即真、假和未知), 以表达缺乏信息时的推理。

本文着重介绍NFA语言, 并以LIX系统的成功运行实例来说明这种知识表达语言的有效性和实用性。

2 测井解释知识模型

本知识模型内包含两部分知识: 一部分是成熟的经验, 它们以解析式或算法等形式写进了专著; 另一部分是各个专家所特有的非解析的直觉知识, 即逻辑推理知识。

整个模型是模块化的多层次结构, 它的主要特点有: ①对测井信息进行必要的校正, 以提取有效信息; ②纳入多种测井信息和解释方法以便进行综合评价, 提高了模型的适应性和评价的可靠性; ③模块化结构有利于模型的局部增、删、改。

3 知识表达语言——NFA

3.1 扩充的产生式规则和NFA语言

在专家系统中, 描述领域专家的知识有多种方法, 其中最普遍使用的是产生式规则, 它的一般形式为

$$a, b, c, \dots \implies r$$

其中 a, b, c, \dots 是前提, 一般为基本事实, 也可能是由其他规则推导出的结论, r 是结论。这种形式描述纯逻辑推理过程是很有趣

的, 可是, 对于大量的实际应用, 特别是在仪表型 (instrument-based) 和实时 (real-time) 专家系统中, 例如, 在石油测井解释中, 逻辑推理和数值计算交叉混杂, 逻辑推理时, 需要知道某些量的值, 而计算这些值时, 又要通过逻辑推理和判断来选择合适的计算公式或选取恰当的参数, 可见, 纯产生式规则不能很好地描述某些领域的专家知识。

为此, 我们将纯产生式规则加以改造和扩充, 用关联函数 F 把它的前提和结论等两个部分有机地“关联”起来。这时产生式规则

$$a, b, c, \dots \implies r$$

改造成为

$$F(a, b, c, \dots) \implies r$$

即前提条件 a, b, c, \dots 不能仅仅通过逻辑与的关系导致结论了, 而必须借助于关联函数 F 所规定的综合分析或计算来得出结论 r 或数值 r 。

例如, 产生式 $a, b, c, \dots \implies r$ 可改写为

$$\text{AND}(a, b, c, \dots) \implies r$$

它表示: 仅当 a, b, c, \dots 都满足时, 才有结论 r , 这就是原来的产生式规则。

又如 $\begin{cases} a \implies r \\ b \implies r \\ c \implies r \end{cases}$ 可改写为

$$\text{OR}(a, b, c) \implies r$$

它表示: a, b, c, \dots 中只要有一个满足条件时, 就可推导出 r 。

各种计算公式, 如

$$\frac{v-a}{b-a} \implies r$$

也可写成 $\text{SCALE}(v, a, b) \implies r$

它表示: r 是变量 v 在区间 a, b 上所占的百分比, 这是纯数值计算过程, SCALE 为算术函数。

关联函数一方面使逻辑推理过程多样化 (除了上述的 and, or 外, 还可以设计多种其他的逻辑型关联函数), 另一方面使逻辑推理与数值计算统一于这种扩充的产生式形式之中。

一条扩充的产生式规则箭头右端的结论与另一条规则中的某个前提条件匹配时, 便构成一次推理。如

$$F_1(a, b, c) \implies d$$

$$F_2(d, e, f) \implies g$$

构成的一次推理如图 1 所示。

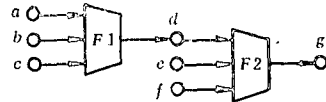


图 1 一次推理

把表达某个领域知识的全部扩充的产生式规则 (即各个一次推理) 都如图 1 所示的连接起来, 便形成完整的推理网络, 它就是知识库。

上述的每条扩充的产生式规则中都含有节点 (变量)、关联函数和变元, 故我们将这种扩充的产生式规则简称为 NFA 语言 (即 Node-Function-Arguments 的缩写)。

$F(a, b, c) \implies r$ 又可以写成 $r = F(a, b, c)$, 这时, NFA 语言的语法用 BNF 来表示是十分简单的, 即

$$\langle \text{规则} \rangle ::= \langle \text{函数名} \rangle (\langle \text{变元} \rangle ; ?)$$

3.2 关联函数

在 NFA 语言中, 关联函数是核心部分。设计各种不同功能的关联函数, 才能满足表达专家知识的需要。

NFA 具有如下几类关联函数。

3.2.1 纯逻辑函数 (如 SELECT, SWITCH, AND, OR, H-AND, H-OR, 等)

例如, $\text{SWITCH}(P, A, B) \iff R$

- 当 $P = T(\text{True})$, 选择推理路线 $A \implies R$;
- 当 $P = F(\text{False})$, 选择推理路线 $B \implies R$ 。

在 AND, OR, ... 等函数中, 采用三值逻辑 (除 T, F 外, 引入 X 代表 “unknown”)。

例如, $\text{OR}(A, B, C) \implies R$

- 当 $A = X$, $\text{OR}(B, C) \implies R$;
- 当 $A = B = X$, $C \implies R$;
- 当 $A = B = C = X$, $X \implies R$ 。

AND, OR 本质上是搜索函数, 一般情况下是按变元的先后顺序搜索的。为了加速求

解, 可采用启发搜索函数(如H-OR, H-AND等)。

例如, $EVAL(a, b, c) \implies H$;

$H-OR(H, A, B, C) \implies R$

其中 EVAL是计算启发信息H值的评价函数, 它同领域知识密切相关, 由领域专家提供, 它通过 a, b, c 来估计A, B, C中哪条推理路线成功的可能性更大。

• 当 $H=2$, 则 H-OR 选择 B 路线进行搜索, 若失败, 则改选 A 依次往下搜索;

• 当 $H=3$, 则 H-OR 选择 C 路线进行搜索, 若失败, 也改选 A 依次往下搜索。

3.2.2 纯算术函数

这类函数多数是某问题领域中的经验公式。例如, 阿尔奇 (ARCHIE) 公式

$ARCHIE(a, b, m, n, R_i, R_w, \phi) \implies S_w$

代表
$$\sqrt[n]{\frac{a \cdot R_w}{b \cdot R_i \cdot \phi^m}} \implies S_w$$

3.2.3 混合比较函数 (如 GE, LE, EQ, IN, OUT, 等等)

例如, $GE(v, a, b) \implies r$

- 当 $v = X, X \implies r$;
- 当 $v \geq a + b, T \implies r$;
- 否则, $F \implies r$ 。

又如, $IN(v, a, b) \implies r$

- 当 $v = X, X \implies r$;
- 当 $v > b, F \implies r$;
- 当 $v < a, F \implies r$;
- 否则, $T \implies r$ 。

3.2.4 非线性饱和函数 (如 SATM, SATN, SATMN, 等等)

例如, $SATMN(v, a, b) \implies r$

- 当 $v = X, X \implies r$;
- 当 $v \leq a, a \implies r$;
- 当 $v \geq b, b \implies r$;
- 否则, $v \implies r$ 。

3.2.5 各种专用处理函数 (如 FUZZY, CORR, CROSS, 等等)

例如, 模糊评价函数 $FUZZY(a, b, c, \dots) \implies (EV, CF)$

该函数的右端有二个变量, 其中 EV 表示评价结论, CF表示结论 EV的可信度。

3.2.6 评价函数

例如, 加权函数 $AWT(v, w) \implies u$

- 当 $v = T, w \implies u$;
- 当 $v = F, -w \implies u$;
- 当 $v = X, 0 \implies u$ 。

又如, 综合评价函数 $CWT(p, u_1, u_2, \dots, u_n) \implies r$

令 $CF = \frac{\sum u_i}{\sum |u_i|}$, 则

- 当 $CF \geq p, CF \implies r$;
- 否则, $F \implies r$ 。

3.2.7 辅助函数

例如, 输出函数 $RECORD(a, b, c, \dots) \implies r$ 表示将 a, b, c, ... 等结论输出到存储介质上, 此处 r 是虚设的, 它代表推理的某一阶段。

3.3 推理网络和关联函数在推理中的作用

图 2 是推理网络的一部分。它是由下列的扩充的产生式规则所构成的。

- $F1(a1, b1, c1) \implies r1$
- $F2(a2, b2, c2) \implies r2$
- $F3(a3, b3, c3) \implies (r31, r32)$
- $F4(r1, r2) \implies r4$
- $F5(r1, r2, r31, r32) \implies r5$
- $F6(r4, r5) \implies (r61, r62)$

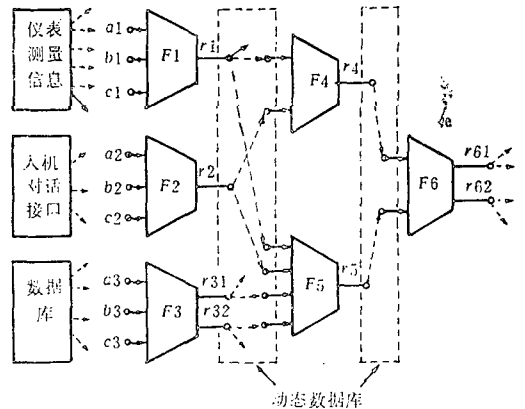


图 2 局部推理网络

图 2 中的 a_i, b_i, c_i 等是基本事实, 它们可能来自: ①仪表直接测到的信息; ②人机

对话接口；③专用数据库。

推理过程根据数据库和知识库的具体内容来接受控制策略（如，逆向推理）和关联函数的控制。从推理的角度看，关联函数有如下两个重要作用

- 控制推理路线。关联函数的变元的求值顺序由已求值的那些变元来决定；
- 规定求值规律。按一定的规律建立起结点值和各变元值之间的关系。

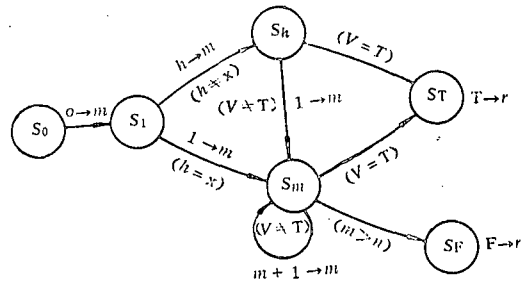


图 3 H-OR(H, a, b, c, d) => r 的状态图

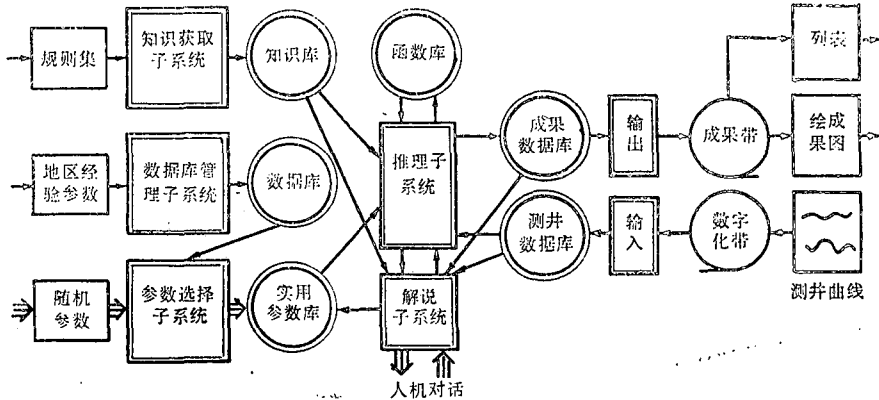


图 4 LIX 的结构框图

可以用有限自动机的状态图来描述关联函数的解释过程。例如，启发搜索函数 H-OR 的状态如图 3 所示。

图 3 中的 m 代表当前选择的推理路线序号， n 代表可能的推理路线数， v 代表第 m 条推理路线的返回值， h 是启发信息。

4 测井解释专家系统 LIX 的实现

4.1 系统的结构

图 4 是 LIX 的结构框图。它是由知识获取子系统、参数选择子系统、数据库管理子系统、推理子系统和解说子系统等所组成。

知识获取子系统的任务是：把领域专家用 NFA 语言写的规则转换成机内代码，即知识库，供推理子系统及解说子系统等使用。该子系统能检查出知识库内的词法、语法和逻辑错误，并提供灵活方便的修改手段，还具有优化规则的功能，例如，可将某些规则归纳、整理成新的更合理的规则，供领域专家参考、决策。

数据库管理子系统和参数选择子系统的任

务分别是构造数据库和实用参数库，以适应多油田、多套含油层系及多变的随机参数等复杂情况。数据库内以 20 个地质层系为组合单元，存放 100 多个油田（或区、块）的地区经验参数，数据库管理子系统可以根据不同时期的勘探与开发任务，更换其内容。参数选择子系统根据具体油田、地质层系以及输入的随机参数来形成当前需要的实用参数库。

解说子系统的任务是：回答用户提出的八种问题（why, what, how...），将系统的推理过程、得出结论的依据等形象地显示出来，供用户或领域专家来评判结论是否正确，依据是否充分，推理过程有何差错，等等，以便改进系统的设计和知识库。它还可以帮助新手学习并掌握本系统。

推理子系统是 LIX 的核心部分，它的任务是：使用知识库求解问题，最后以成果图的形式给出油、气、水层的评价结论，并能列表输出 29 种描述储集层性质的地质参数。

本子系统采用逆向推理的控制策略，即从

总目标 TASK 出发,沿知识库规定的推理链,逆向前进。

TASK 又划分为若干个子任务,通过人机对话来选择

TASK = SELECT(TASP, TASK1,
TASK2, ..., TASK4)

TASP = ASK(TASP = ?)

若选择了子任务 TASK1,它又把地层分成三大模式:泥岩层 SHZ,致密层 DNZ,储集层 REZ。对这三种模式利用上一个采样点作为启发信息来进行启发式搜索

TASK1 = H-OR(H_1 , SHZ, DNZ, REZ)
 H_1 = LAST(SHZ, DNZ, REZ)

若 $H_1 = 2$,则从 DNZ 出发,重复上述的推理过程,直至达到推理网络的最底层,并返回最高层的总目标 TASK 为止。

推理过程是深度优先的搜索过程,但同时受到关联函数的控制。在每个结点上推理的总规律是相同的,故推理子系统是一个递归程序。

4.2 应用效果

4.2.1 NFA语言易学易懂

这种语言的语法如前所述,极为简单,实际上只有一条语法规则: $N = F(A_1, A_2, \dots, A_n)$ 。胜利油田测井公司的解释专家通过几个小时的学习,就十分熟练地用 NFA 语言构造用于测井解释的实用知识库,目前该知识库有 400 余条规则。

4.2.2 LIX 先后在 INTERDATA-85 机和 PE-3230 机上用 FORTRAN 语言编程实现(采用了一些技巧,使 FORTRAN 语言能递归),现场(胜利油田)运行近两年,解释了 130 余口井的测井曲线,评价的符合率为 94.2%,达到了解释专家的水平。

5 结束语

本系统 LIX 不仅能连续接收和处理测井曲线上的大量数据,而且能以成果图的形式或列表形式输出油、气、水的综合评价,并附以可信度,所以,它具有仪表型(instrument-based)专家系统的特点。

NFA 语言是产生式规则的扩充,它克服了纯产生式规则不能表达算法知识的缺点,将逻辑推理和数值计算综合成统一的表达形式,因此 NFA 具有很强的知识表达能力。

LIX 是 NFA 语言的一种机器实现,实质上,它是 NFA 语言的解释系统,因此,可以把 LIX 看成一种专家系统的建造工具。LIX 的实现和成功地运行,充分地说明了 NFA 语言的有效性和实用性。我们期望 NFA 能用在其他问题领域。

参 考 文 献

1. Nilsson N J, Principles of Artificial Intelligence. Palo alto:Tiaga Publishing Co, 1980
2. Bachanan B G, Shortliffe E H, Rule-based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Project. Reading, Massachusetts: Addison-wesley, 1984
3. Barr A, Feigenbaum E A (eds.), Handbook of Artificial Intelligence (Vol 1), Los Altos, CA: Kaufmann, 1981
4. Hayes-Roth F, Waterman D, Lenat D (eds.), Building Expert System, Reading, MA: Addison-Wesley, 1983
5. Gao S, A Design and Implementation of the Tool for Building Expert Systems. Master Dissertation, 1985
6. Mo X, INLOG — an Expert System for Log Interpretation. Master Dissertation, 1984
7. Georgeff M P, Lansky A L, Bessiere P, A Procedural Logic. In: Proceedings of the Ninth IJCAI, Los Angeles, California, USA, 1985
8. Vilain M, The Restricted Language Architecture of a Hybrid Representation System. In: Proceedings of the Ninth IJCAI, Los Angeles, California, USA, 1985
9. Smith R G, Baker J D, The Dipmeter Advisor System: A Case Study in Commercial Expert System Development. In: Proceedings of the Eighth IJCAI, Karlsruhe, West Germany, 1983
10. Bonnet A, Dahan C, Oil-well Data Interpretation Using Expert System and Pattern Recognition Technique. Proceedings of the Eighth IJCAI, Karlsruhe, West Germany, 1983
11. Davis R, Buchanan B G, Shortliffe E H, Production Rules as a Representation. Artificial Intelligence 1977, (8):15-45
12. SONG G N et al, A Knowledge Representation Language NFA. 1986 (Submitted to the IJCAI-87)
13. 石油部胜利油田测井公司,中国科学院沈阳自动化研究所.砂泥岩地层测井解释专家系统研制报告.1985.12
14. 石油部胜利油田测井公司,中国科学院沈阳自动化研究所.砂泥岩地层测井解释专家系统技术报告.1985.12

Knowledge Representation Language NFA and Its Application in Expert System for Oil Log Interpretation

SONG Guoning et al.

In order to describe oil log interpretation experts' knowledge and experiences and simulate their modes of thinking we developed a new knowledge representation language NFA, which is an extension of pure production rules and integrates the numerical calculation with the logical reasoning in the unified form.

The expert system LIX for oil log interpretation, developed by us, is in essence an interpreter of NFA language. Its rate of coincidence in interpreting oil-well data is more than 94% as compared to human experts.

(上接第63页)

连续键处理技术, 仪表的键盘显示程序使用户能在仅有五个键及五只数码管的面板上实现编程和控制所需的各种操作。诸如读取和修改各种参数, 输入各种命令, 手控调节器输出等等。而且由于对键盘实行中断扫描, 所以全部键盘操作均能立即得到响应, 没有任何延误。异常复位处理程序和自检程序提高了仪表的可靠性。而数字滤波和线性化程序则提高了仪表的测量精度。所有这些, 都充分利用了微处理器用于仪表的优点, 使仪表质量达到了较高的水平。

广东三水啤酒厂采用 DTL-812 型控制器控制锥形发酵罐的温度。控制系统如图2所示。啤酒发酵工艺要求发酵罐内的温度按一定的时间曲线变化, 发酵周期为十几天至几十天不等。发酵过程温度控制的好坏对产品质量有直接影响。发酵过程中罐内液体放出热量使温度升高, 因而须在夹套中通入冷媒以控制温度。每个发酵罐有上中下三个夹套, 三个测温点。过去发酵罐的控制是由操作工根据温度指示表控制阀门, 或是用普通调节器控制阀门, 而调节器的给定则由操作工根据时间改变。这样控制往往不能满足工艺要求。现在采用一台 812 型控制器控制 3 个点的温度, 由于控制温度曲线可编程, 保证了温度可按工艺要求变化, 从而稳定了产品质量。该厂采用的控制器规格为: 热电阻输入, 分度号 BA2, 量程

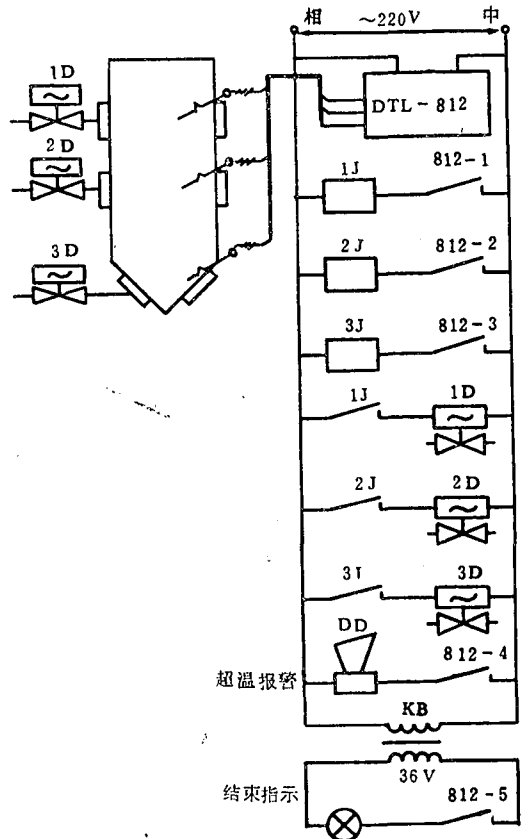


图 2 控制系统框图

- 10—+40℃, 双位式控制, 记录输出 0—10mA, 曲线时间 0.1~999.9 小时。

DTL-812 型控制器输入输出可采用标准信号, 除了专用于温度控制外, 还可代替三台一般 PID 调节器, 适用于各种场合。所以说, 它的应用前景是非常广阔的。