



句法知识系统的推理方法¹⁾

杨光正

(中国科技大学自动化系,合肥 230026)

摘 要

本文在文法产生式表达知识的基础上讨论了句法知识系统的推理方法。Earley 算法是一种高效的句法分析算法,它可成功地用作句法知识系统的搜索策略。本文还讨论了句法知识系统的启发式搜索策略,并且提出了一种高效的深度优先搜索策略。

关键词: 基于知识的模式识别,句法分析,搜索策略,启发式方法。

一、引 言

对于复杂的模式识别问题,常规方法的效果都是有限的。为了提高识别能力,基于知识的模式识别方法被提了出来^[1,2]。一般认为,句法识别本身就具备专家系统的基本性质^[3]。在句法模式识别中,产生式集、句法分析方法、文法推断相应地就是专家系统中的知识库、搜索策略和知识获取。

对于一个知识系统,首先应解决两个关键问题:1)这个知识系统如何表达知识;2)这个知识系统的搜索策略或推理方法。

对于第一个问题,国际上流行的观点是:一个文法产生式 $[A \rightarrow \beta]$ 可以用来表示知识 $[IF A THEN \beta]$ ^[2,3]。这是不正确的。因此国外迄今未能用句法方法建立一个基于知识的模式识别系统。

作者提出了自己的观点,证明了形如 $[A \rightarrow \beta]$ 的文法产生式应该表示知识 $[IF \beta THEN A]$ 。并在这个观点指导下,初步建立了一个用句法方法的基于知识的模式识别系统^[4,5],从而沟通了人工智能系统与句法模式识别系统。

本文着重讨论在句法知识系统中如何应用句法分析方法来解决推理问题。要指出的是,假如没有正确解决知识表达问题,知识系统的推理问题是无法处理的。

本文于 1991 年 2 月 25 日收到。

1) 本课题得到 1990 年国家自然科学基金的支持。

二、搜索策略的考虑

在句法知识系统中,可能派生的所有句型 α 是状态的自然表示.于是推理过程就是从 S 出发,搜索状态空间中目标状态 x .或是从状态 x 出发,搜索目标状态 S .本文中利用高效的 Earley 算法来建立一个新的推理方法.

对于复杂的识别问题,启发性搜索是必要的.有如下两类启发信息:

1) 逻辑性启发信息.这类启发信息指出当前推理过程哪些规则可以应用.这类启发信息是离散的,应尽可能在算法中加以考虑.

2) 经验性启发信息.这类启发信息是在总结专家经验的基础上建立的.在一定区间中连续取值.可以建立一定的数学模型来研究它们.

在句法系统中,属性文法是有力的表达知识的工具.其产生式的语义规则部分可以包含启发式函数来协助推理过程、计算结果的可信度和扩展文法的表达知识的能力.在本文提出的句法知识系统中,概率和模糊隶属度是相应的属性文法产生式的属性.

Earley 算法是一个有效的句法分析算法^[3,6].每个分析表中列入当前分析阶段的数据,包括各个应有的项目以及其它有关的信息.它记录各项当前子剖析的数据,并提出新的子剖析.这些子剖析就是知识源所产生的局部解释(假设).此外,它还可以记录由于算法或文法的要求而新增的附加数据和其它有关信息.

作者对 Earley 算法作了改进.对每个项目的添加都作向前看一个符号的预测.这是一种逻辑性启发信息.这种预测效果很好并且不致引起算法的过分复杂化.

一个简单的例子如表 1 所示.表中给出文法 G 、待分析的句子 x 并用 Earley 算法进行分析.并给出相应的状态搜索图(见图 1).可以看出

1) 每次预测运算的结果是搜索一个新的状态.

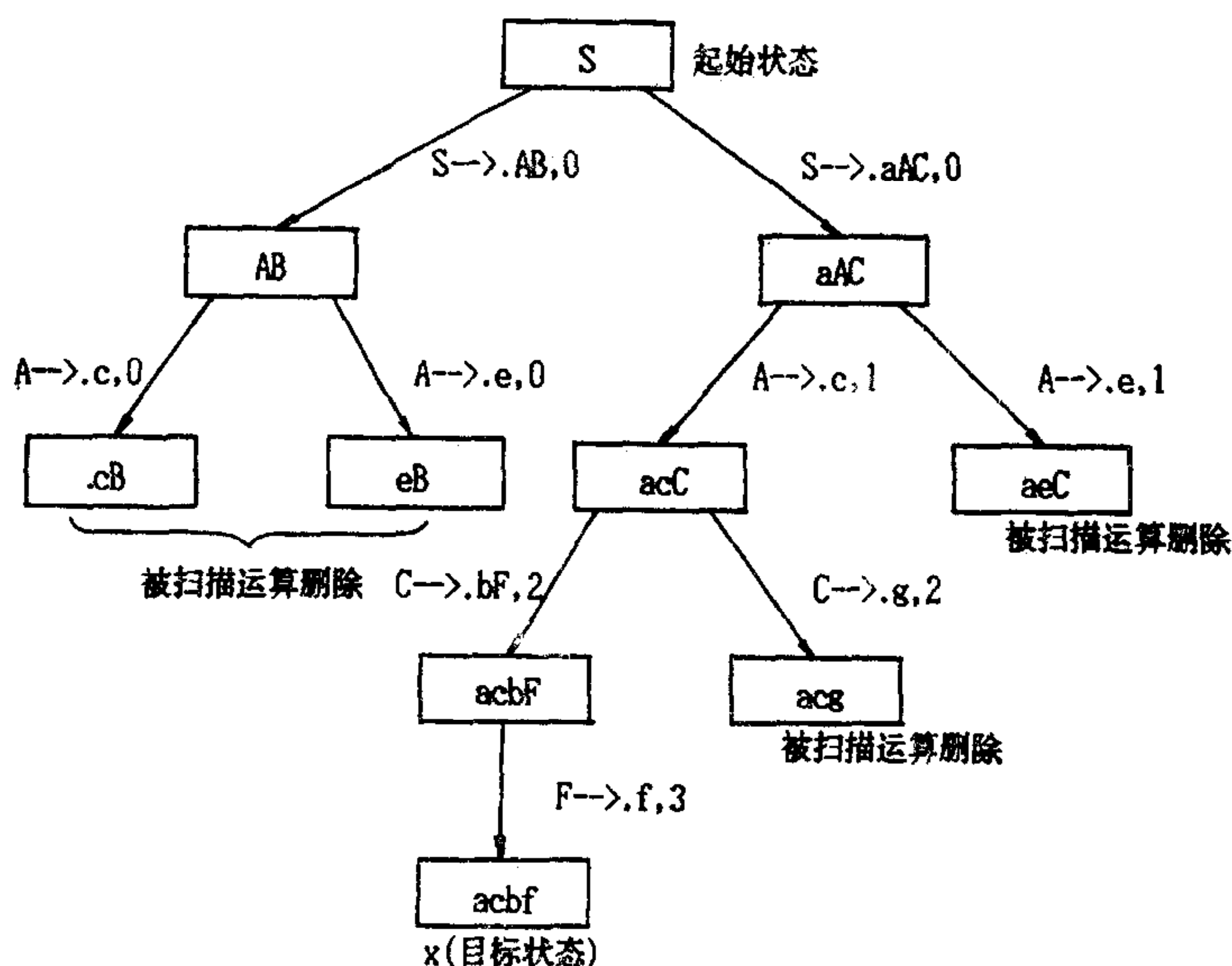


图 1

2) 扫描运算和完成运算仅用于验证已有搜索的有效性, 本身并不提出新的状态. 可见, 将 Earley 分析算法用于搜索策略时, 它是一个宽度优先的算法.

给定文法:

$$G = (V_N, V_T, P, S), \quad V_N = \{S, A, B, C, D, F\}, \quad V_T = \{a, b, c, d, e, f, g\},$$

$$P: S \rightarrow AB \quad S \rightarrow aAC \quad A \rightarrow c \quad A \rightarrow e \quad B \rightarrow bD$$

$$C \rightarrow bF \quad C \rightarrow g \quad D \rightarrow d \quad F \rightarrow f$$

要求分析 $x = acbf$

用 Earley 算法构造分析表如下:

表 1

I_0	I_1	I_2	I_3	I_4
$S \rightarrow \cdot AB, 0$	$S \rightarrow a \cdot AC, 0$	$A \rightarrow \cdot c, 1$	$C \rightarrow \cdot bF, 2$	$F \rightarrow \cdot f, 3$
$S \rightarrow \cdot aAC, 0$	$A \rightarrow \cdot c, 1$	$S \rightarrow aA \cdot C, 0$	$F \rightarrow \cdot f, 3$	$C \rightarrow bF \cdot, 2$
$A \rightarrow \cdot c, 0$	$A \rightarrow \cdot e, 1$	$C \rightarrow \cdot bF, 2$		$S \rightarrow aAC \cdot, 0$
$A \rightarrow \cdot e, 0$		$C \rightarrow \cdot g, 2$		

本文提出了一个深度优先的句法分析算法, 使之满足启发式搜索的要求, 称为 ED 算法, 基本思想是

1) 按输入句子符号构造分析表. 分析表中的项目有形式 $[A \rightarrow \alpha \cdot \beta, i, j, t]$. 其中 i ——扩展表号; j ——左边为 A 的产生式子集中, 由启发函数所决定的序号; t ——扩展指示符, $t = 1$ 扩展, $t = 0$ 其它.

2) 根据启发函数所确定的序列, 每次预测运算只添加当前项, 并且令 $t = 1$.

3) 扫描运算和完成运算的作用仍如 Earley 算法, 它们产生的项目有 $t = 0$.

4) 当新增加的分析表为空时, 执行删除运算, 方法是反向删除全部 $t = 0$ 的项目, 直至 $t = 1$ 为止.

(1) 如果这时预测项目已搜索完, 则删除此项目, 继续反向删除 $t = 0$ 的项目.

(2) 如果尚有剩余的预测项目, 则将其中顺序最高的项目添上, 令 $t = 1$.

5) 检查 I_n 中无形如 $[S \rightarrow \alpha \cdot, 0, j, 0]$ 的项目. 如有, 句子被接受; 否则转删除运算.

6) 如执行删除运算时在 I_0 中无新项目 $[S \rightarrow \alpha, 0, j, 1]$ 可以添加, 则判 x 不可接受而停止.

对于表 1 所示的例子, 当用 ED 算法时, 自 S 开始首先扩展状态 AB , 然后扩展到 cB . 它被扫描运算所删除. 返回 AB 后顺序扩展到 eB 状态, 仍被扫描运算所删除. 返回 S 状态后再按顺序扩展到 aAC 状态. 如此进行直至搜索到目标状态 $acbf$ 为止. 过程不经历状态 acg 和 aeC .

关于 ED 算法的讨论.

1) 关于计算复杂度. 因为 ED 算法与 Earley 算法具有相同的搜索空间, 在最不利的情况下它的计算复杂度与 Earley 算法的复杂度相同.

2) 派生树的分析. 在 ED 算法中如果句子能被接受, 则保存在分析表中全部预测运算都是有效的, 所以派生树分析的时间复杂度为 $O(n)$.

3) ED 算法会产生左递归的问题。但这在应用于知识系统时问题不大。

4) ED 算法同样可采用向前看一个符号的预测,如同在 Earley 算法中所采用的那样,它能大大提高搜索效率。

三、结 论

1. 由于具有更强的表达能力和多样的描述方法,句法知识系统较之一般的人工智能产生式系统可以提供更好的知识表达方式,例如可采用随机文法、模糊文法、以及功能很强的属性文法等,并且可以采用深度优先的搜索算法,以适应不同应用领域的要求。

2. 句法方法有形式语言这个很强的理论背景,可以用来支持人工智能的研究和理论的发展。句法分析算法在句法知识系统搜索策略中的应用就是一个明显的例子。

参 考 文 献

- [1] Nandhakumar, N. & Aggarwal, J. K., The Artificial Intelligence Approach to Pattern Recognition——A Perspective and an Overview, *Pattern Recognition*, 18(1985), 383—389.
- [2] Bunke, H., Hybrid Method in Pattern Recognition, *Pattern Recognition Theory and Applications*, Springer-Verlag (1987).
- [3] Fu, K. S., *Syntactic Methods in Pattern Recognition*, Academic Press, New York(1974).
- [4] 杨光正, A Syntactic Approach for Building Knowledge-Based Pattern Recognition System, Proc. 9th ICPR, Rome (1988), 1236—1238.
- [5] 杨光正,句法模式识别与专家系统的关系,模式识别与人工智能,2(1988)
- [6] Aho, A. V. & Ullman, J. D., *The Theory of Parsing, Translation and Compiling*, Prentice-Hall Inc. (1972).

ON THE REASONING METHOD OF SYNTACTIC KNOWLEDGE SYSTEM

YANG GUANGZHENG

(Dept. of Automation, University of Science and Technology of China, 230026)

ABSTRACT

In this paper the reasoning method for the syntactic knowledge system is discussed on the basis of knowledge representation by means of grammar productions. In the syntactic pattern recognition the Earley parsing algorithm is an efficient analysis algorithm. In the syntactic knowledge system it can be used as a search strategy successfully. For applications of syntactic knowledge system in different fields the heuristic search strategy is also discussed, and an efficient bottom-up search strategy the ED algorithm is presented.

Key words: Knowledge-based pattern recognition; syntactic knowledge system; syntactic analysis; search strategy; heuristic approach.