

一种识别线划图形的方法

戴汝为

(中国科学院自动化研究所)

摘 要

把模式识别的统计方法和句法方法统一起来成为一种有效的方法,已经有过一些研究,取得了进展。本文在作者用属性文法来统一两种方法的工作基础上,借鉴认知心理学关于记忆和语言的一些论点,讨论了高维属性文法的句法和词意之间的折衷关系,对于线划图形,给出句法、词义以及模块之间关系信息的距离度量,从而可以用最小距离准则来进行线划图形(尤其是汉字)的识别。

一、引 言

统计方法和句法方法是模式识别的两种一般性方法^[1]。两种方法各有优缺点,前者的弱点是不能描述复杂模式的结构以及模式与子模式之间的关系,后者在处理具有噪声的模式及利用数字的词义信息方面又往往无能为力。近几年来曾经探讨用属性文法把两种识别方法统一起来,并已经取得了一些进展^[2-7]。统计模式识别已经研究得很充分,这里不多说。回顾句法模式识别的形成,把模式的产生和构成与语言的产生和构成加以类比的想法是富有启发性的,这样一来数理语言(其中乔姆斯基的短语结构语言是一个重要组成部分)就成为以往句法模式识别的基础。然而只从工程技术的角度来考虑识别问题,有一定的局限性。人进行模式识别的本领是非常高明的,一些认知心理学的试验表明,“人们说话或理解别人的话都不能离开记忆,在人们的记忆中是概念,是概念和概念之间的关系”,也就是说人们的记忆中是概念,是词义而不是句法起重要作用。参照这种说法,在文法中引入词义及词义规则,把表示句法的文法限制为简单的有限状态形式,也许可以说是一种合理的途径。另外,人的记忆不是以 bit 为单位,而是以模块(Chunk)为单位,研究模块间相互关系有积极意义,这一点对于汉字这类线划图形的识别更为明显。汉字中的合体字就是由两个以上的构件组成,每个构件可以看成是一个模块。例如构件口和木,可以因其相对关系和大小的不同而构成呆、困和杏三个不同的字。构件(或模块)之间的关系,对于组成汉字具有相当大的重要性。

本文在作者原有工作基础上^[4,5,7],着重考虑词义信息,进一步阐述了不仅在鍊文法中,而且在树状文法中,句法和词义的复杂性之间可以有折衷关系,以使词义复杂而句法简化,反之亦然。在引进较复杂的基元属性和连接关系后,一些用树状属性文法描述的图形,可以用有限状态属性文法描述。这就为描述与识别线划图形,特别是由线段组成的汉

字提供了一种新的方法。

二、有限状态属性文法

线划图形是二维平面上的图形,识别二维图像往往需要采用高维模式文法,其中树状文法由于能较有效地进行句法分析,所以是一种采用得较多的方法,但对于某些线段间互相交叉的线划图形(如图 1 所示的心脏轮廓^[9]及汉字里),用通常的树状文法来加以描述是颇不方便的。利用文[7]中所定义的关联属性文法,引入适当的连接属性,并把基元考虑得复杂一些,那么参照鍊文法中句法和词义之间所存在的折衷关系,就可以把句法部分限制为简单的有限状态形式,就能够充分利用词义信息进行模式识别。例如汉字中的日和日,结构完全相同,但意义完全不同。这可以利用词义信息来判别。

如果线划图形中的所有基元都是直线段,那么就可以用一个基元的起始连接点到终止连接点的向量作为连接向量,至于两个向量交叉时,则连接关系的句法符号用 CX 表示^[5]。这样表示既简便又不致发生含混。下面用关联属性文法^[7]来描述图 1 中的汉字。

如果考虑基元有两个, a_1, a_2 , 每个基元有三个连接点。 a_1 的属性 $A(a_1) = (l_1, l_2)$, a_2 的属性 $A(a_2) = (l'_1, l'_2)$, $(l_1, l_2), (l'_1, l'_2)$ 分别为连接点 1 至 2, 2 至 3 之间线段的长度。那么用下述属性文法 G 可以产生图 1 的汉字

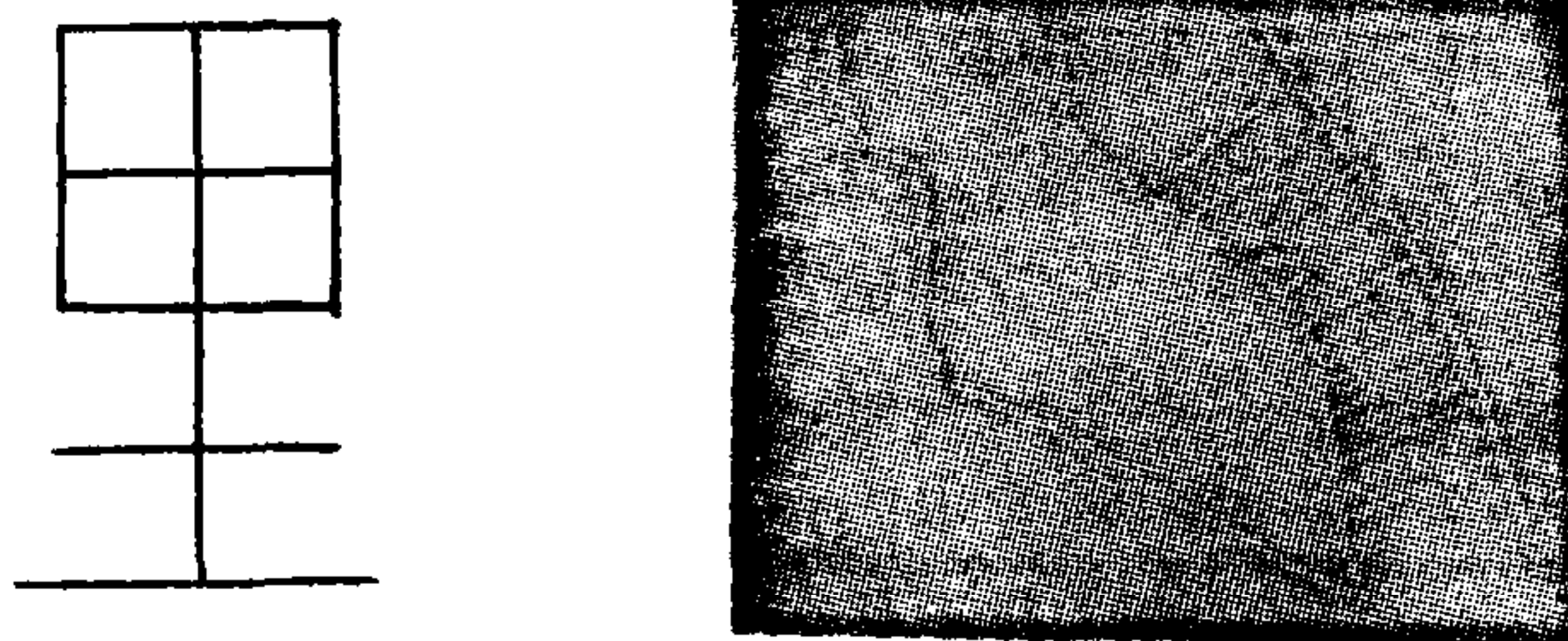
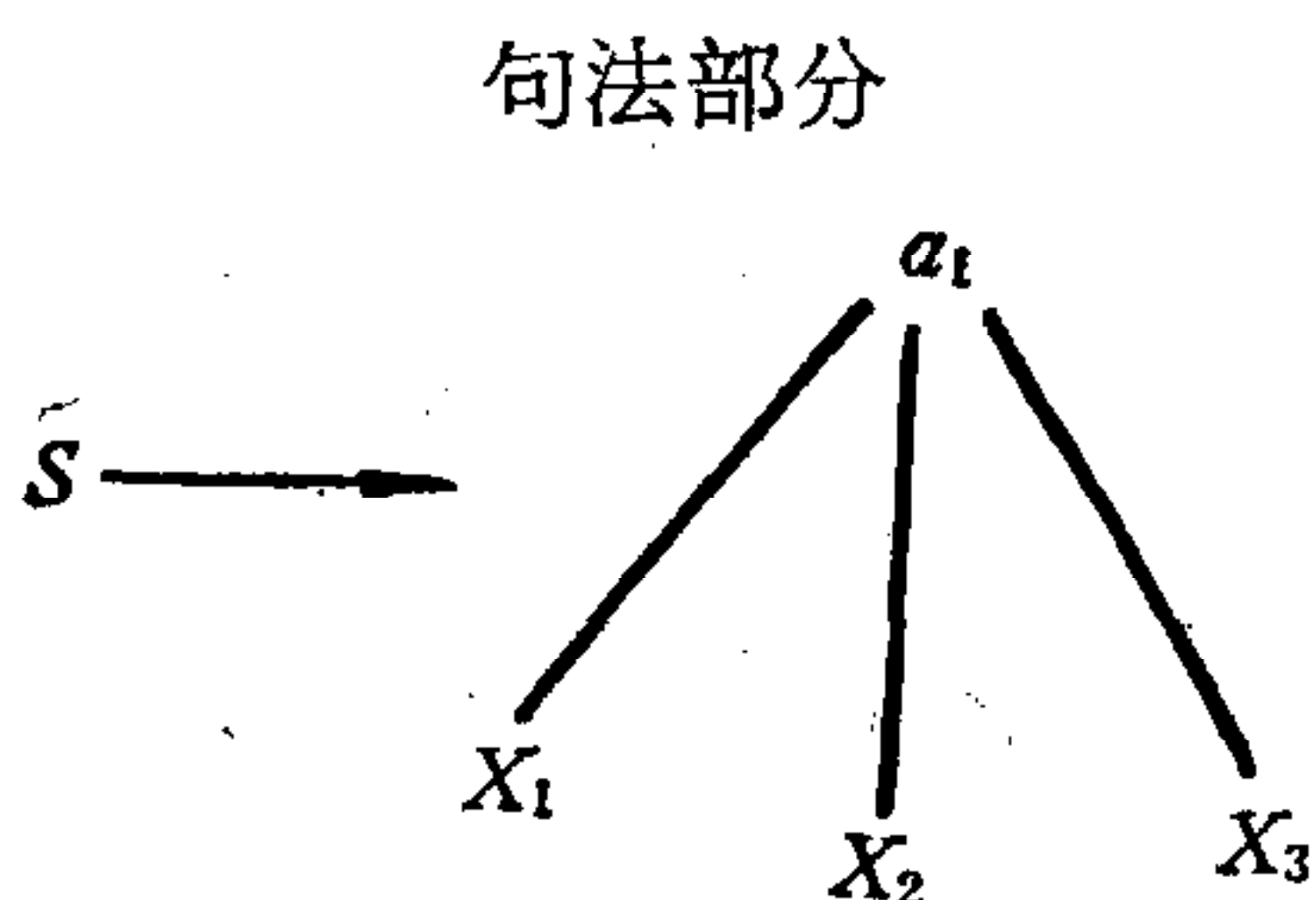


图 1

$$G = (N, V, S, P). \text{ 其中 } N = \{X_1, X_2, X_3, Y_1, \dots, Y_r, S\},$$

$$V = \left\{ a_1: \begin{matrix} 1 & 2 & 3 \\ \rightarrow & \rightarrow & \rightarrow \end{matrix}, a_2: \begin{matrix} \downarrow 1 \\ \downarrow 2 \\ \downarrow 3 \end{matrix}, a_1, a_2 \text{ 的阶分别为 } r(a_1) = \{3, 0\}, r(a_2) = \{2, 1, 0\} \right\}.$$

产生式 P 的集为 P :



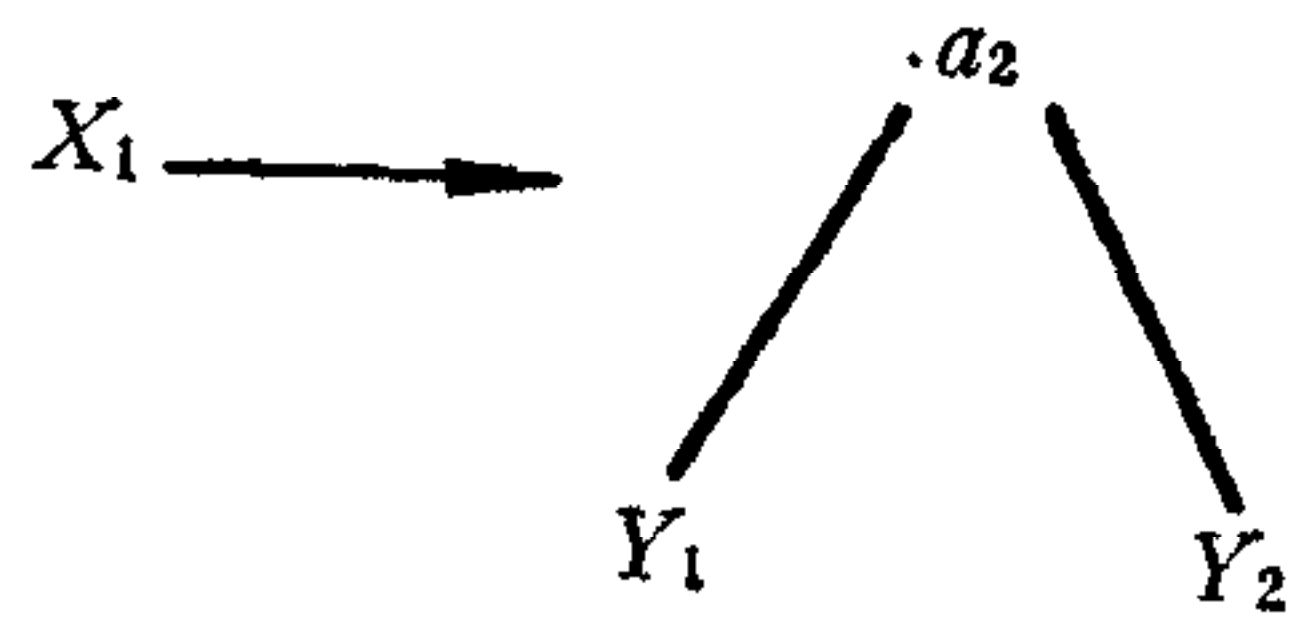
词义部分

$$A(a_1) = (1, 1)$$

$$CAT(a_1, X_1) = CAT(a_1, a_2) = ((1, 1)(x, -\pi/2))$$

$$CAT(a_1, X_2) = CAT(a_1, a_2) = ((2, 1)(CX, -\pi/2))$$

$$CAT(a_1, X_3) = CAT(a_1, a_2) = ((3, 1)(+, -\pi/2))$$



$$A(a_2) = (1, 1)$$

$$CAT(a_2, Y_1) = CAT(a_2, a_1) = ((2, 1)(CX, \pi/2))$$

$$\bar{C}AT(a_2, Y_2) = CAT(a_2, a_1) = ((3, 1)(+, \pi/2))$$



$$A(a_1) = (1, 1)$$

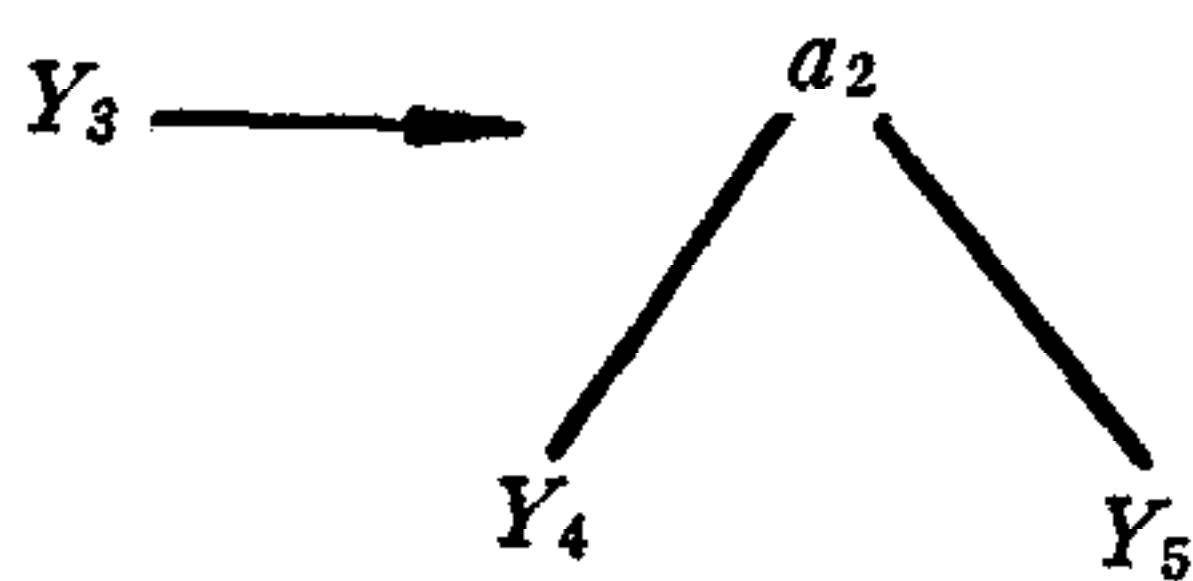


$$A(a_1) = (1, 1)$$



$$A(a_2) = (1, 1)$$

$$CAT(a_2, Y_3) = CAT(a_2, a_2) = ((2, 1)(+, -\pi))$$



$$A(a_2) = (1, 1)$$

$$CAT(a_2, Y_4) = CAT(a_2, a_1) = ((2, 2), (CX, \pi/2))$$

$$CAT(a_2, Y_5) = CAT(a_2, a_1) = ((3, 2)(CX, \pi/2))$$



$$A(a_1) = (1, 1)$$



$$A(a_1) = (1.2, 1.2)$$



$$A(a_2) = (1, 1)$$

与通常的树状文法相比较,属性文法具有明显的优点,基元的属性及连接属性起了作用。这种属性分别表示笔划的长度以及笔划之间的相对关系,如果考虑笔划长度及笔划间夹角为随机量,或者是一些满足给定条件的参量,那么上述文法实际上描绘了大小、形状均不同的一类汉字。

图 2 给出了由 G 产生并带有属性的树。

在上述分析中,有的笔划本来是一完整的笔划,由于基元的选取,分成了两段。如果使基元连接关系复杂一些,那么可以把文法的句法部分变得简单。另外,在书写汉字时,

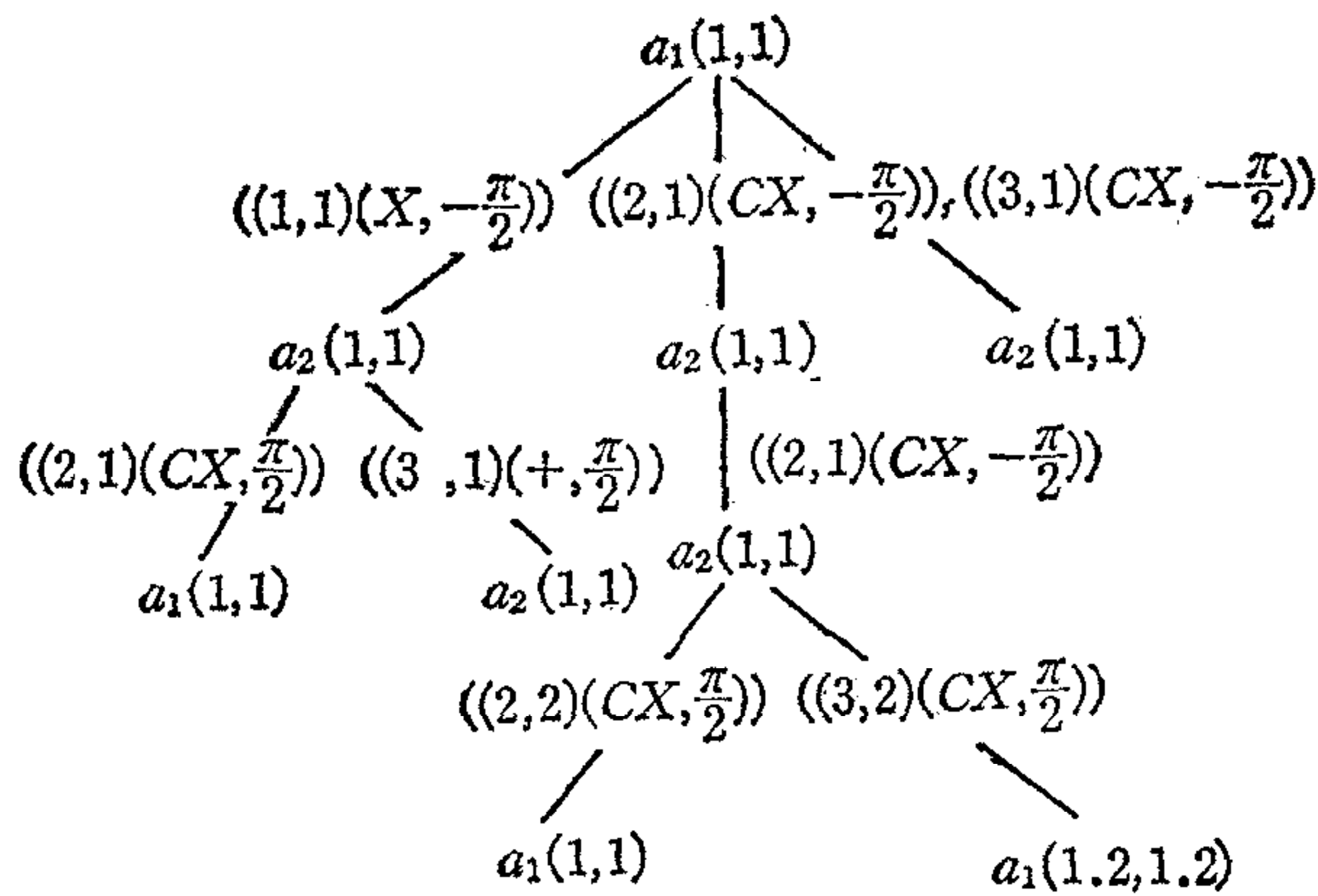


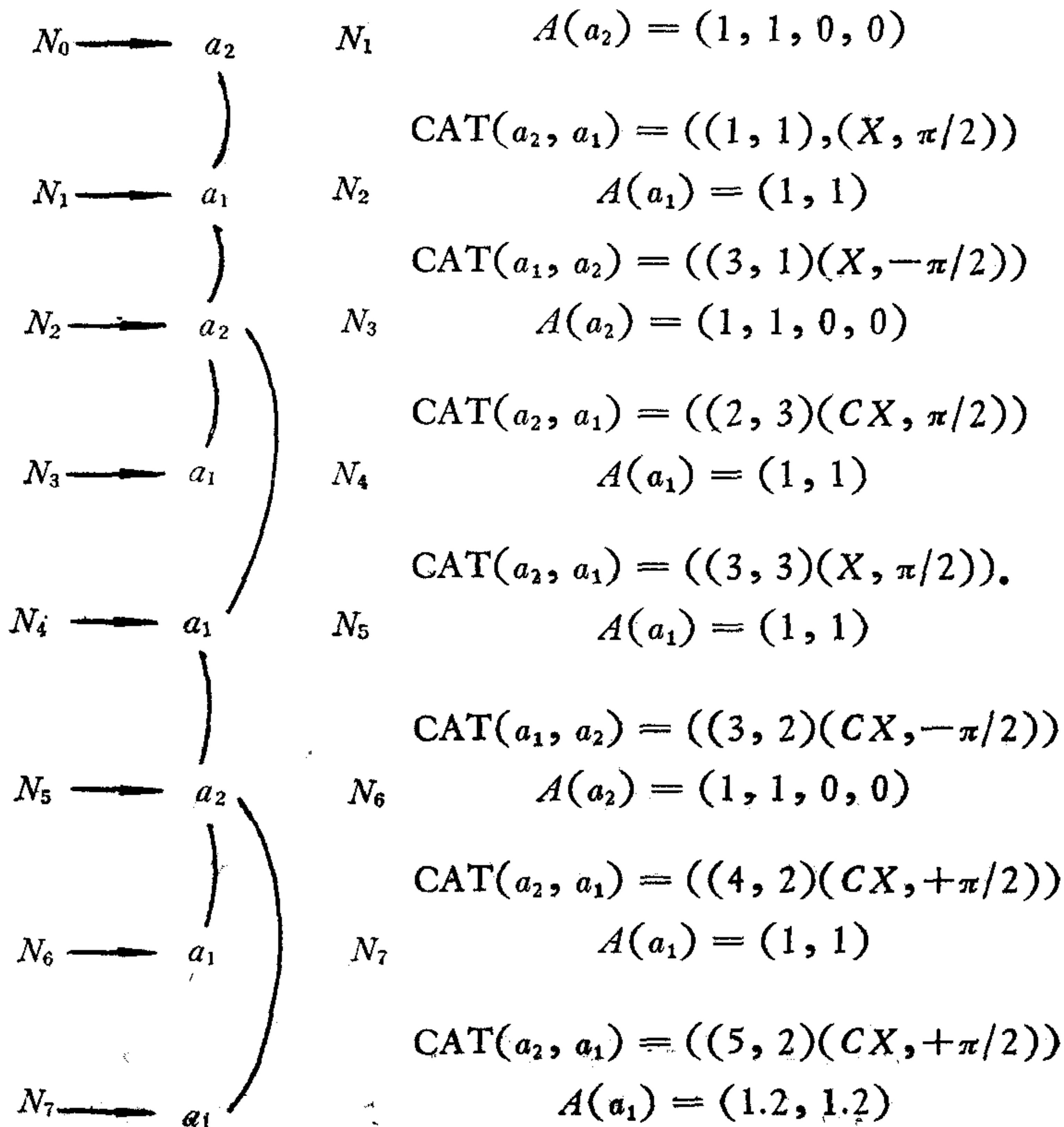
图 2

有笔划的顺序,在实时手写汉字的识别中,笔划的顺序是很重要的信息。根据某种笔划的顺序,很自然地可把句法形式变为有限状态形式。下述有限状态文法同样描述图 1 所示的汉字。

$$G = (N, \Sigma, S, P),$$

$$N = \{N_0 = S, N_1, N_2, \dots, N_7\}, V = \left\{ a_1: \begin{matrix} 1 & 2 & 3 \\ \cdot \rightarrow \cdot \end{matrix}, a_2: \begin{matrix} \downarrow 1 \\ \downarrow 2 \\ \downarrow 3 \\ \downarrow 4 \\ \downarrow 5 \end{matrix} \right\}.$$

P:



于是可以用图 3 所示的具有连接属性的链来表示所描述的字符。

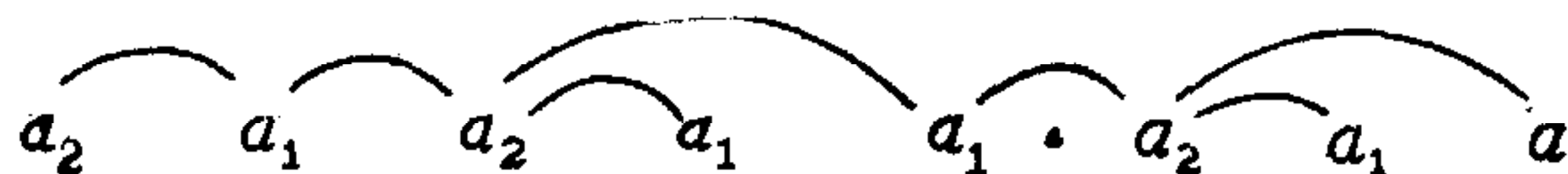
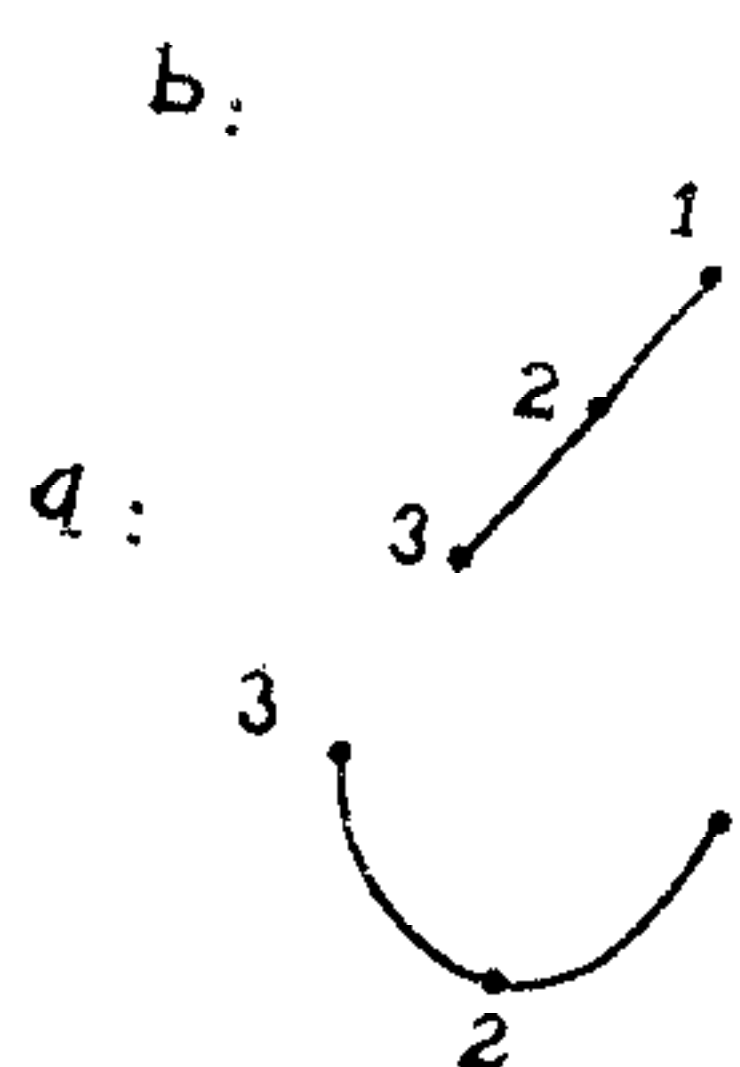


图 3

为了讨论起来简单起见，上面的例子是直线段构成的图形。对于曲线段构成的图形也可以类似地加以处理。主要在于找到由基元构成图像的一种排列次序。如果采用下述两种基元(实际上可概括为一种基元,通过基元的属性来加以表别即可),那么也可以用有限状态属性文法来描述图 1 所示的心脏轮廓图形。



有三个连接点接近于一直线段，两个连接向量分别为 p_1 和 p_2 ，并且 2, 3 可能蜕化为同一点,那么连接向量为 $p.b$ 的属性为

$$A(b) = (L, \Phi, R).$$

两个连接向量分别为 $p_1, p_2, A(q) = (L', \Phi', R')$.

这里采用光滑连接符号 $((m, n), \phi)$ ，即在 (m, n) 连接点处，从曲线上连接点一端的任一串点趋近于连接点时,切线的极限与从另一端任一串点趋向连接点的切线的极限相等。

下述文法 $G_1 = (N, V, P, S)$ 产生心脏轮廓图，其中 $N = \{N_0 = S, N_1, \dots, N_{19}\}$, $V = \{b, q\}$ (基元属性从略)。

$P:$

句法部分

词义部分

$N_0 \rightarrow bN_1$	$\left((2 = 3, 1) \left(+, \theta_{pp_1} = \frac{\pi}{2} \right) \right)$	
$N_1 \rightarrow qN_2$	$\left((1, 1) \left(X, \theta_{pp} = \frac{\pi}{2} \right) \right)$	CAT(q, q) = ((2, 3), ϕ)
$N_2 \rightarrow bN_3$	$\left((2 = 3, 1) \left(+, \theta_{pp_1} = \frac{\pi}{2} \right) \right)$	
$N_3 \rightarrow qN_4$	$\left((1, 1) \left(X, \theta_{p_1p} = \frac{\pi}{2} \right) \right)$	
\vdots		
$N_{16} \rightarrow bN_{17}$	$\left((2, 1) \left(+, \theta_{p_1p_1} = \frac{\pi}{2} \right) \right)$	CAT(q, q) = ((2, 3) ϕ)
$N_{17} \rightarrow qN_{18}$		
$N_{18} \rightarrow qN_{19}$	$((2, 1), \phi)$	
$N_{19} \rightarrow b$	$((3, 1), \phi)$	

从上述例子可以看出,对于基元之间相互有交叉的图形,用关联属性文法来加以描述是有效的,且比一般不带属性的树状文法要方便有效。另外,如果能找到基元构成图形的一种次序,那么就可以用有限状态属性文法来描述。由此带来的问题是由于连接点增加,

连接关系变得复杂,不仅仅是相邻两个基元之间的连接关系.可以认为,有限状态属性文法不仅可以作为属性鍊文法的标准型,而且可以作为属性树状文法的标准型.由于这种办法简化了句法部分,且又不使词义部分过份复杂,这就有利于实际应用.

下面着重讨论线划图形中具有代表性的一类,即汉字的描述问题.

三、模块与模块之间的关系

汉字从形体结构上看由构件组成,每个构件也可以看成一个模块.汉字可分为只有一个构件组成的独体字以及由两个或两个以上构件组成的合体字.这类字中大部分是形声字.随着社会经济文化的日益发展,汉字由表意向表音方向发展.产生形声字,是汉字发展史上的一件大事.形声字大部分是合体字.下面讨论用关联属性文法来描述汉字中的构件以及构件之间的相互关系.

如果 C_1 和 C_2 分别表示两个模块,取 C_1 的一个参考点 $P_1 = (x_1, y_1)$,例如取 P_1 为 C_1 的几何中心.另外取 C_2 的一个参考点 P_2 ,用 $v = re^{i\theta}$ 表示连接 P_1 到 P_2 的向量. C_2 对 C_1 的关系可以用下述二元式表示:

$$\text{REL}(C_2, C_1) = (P, v), P \in \{AR, A, R, B, L\}.$$

其中 P 是一个句法符号, v 表示 REL 的词意信息. AR 表示围绕, A 表示在上方, R 表示在右方, B 表示在下方, L 表示在左方(见图 4).

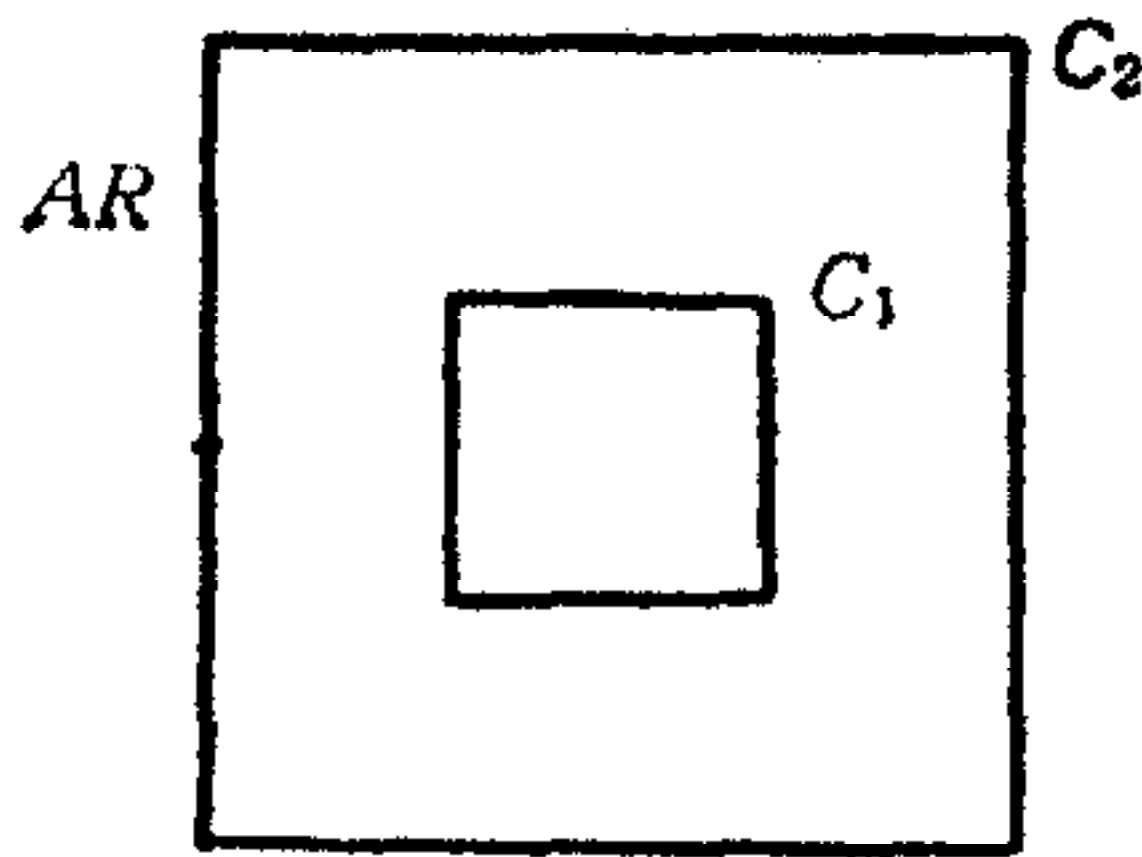


图 4

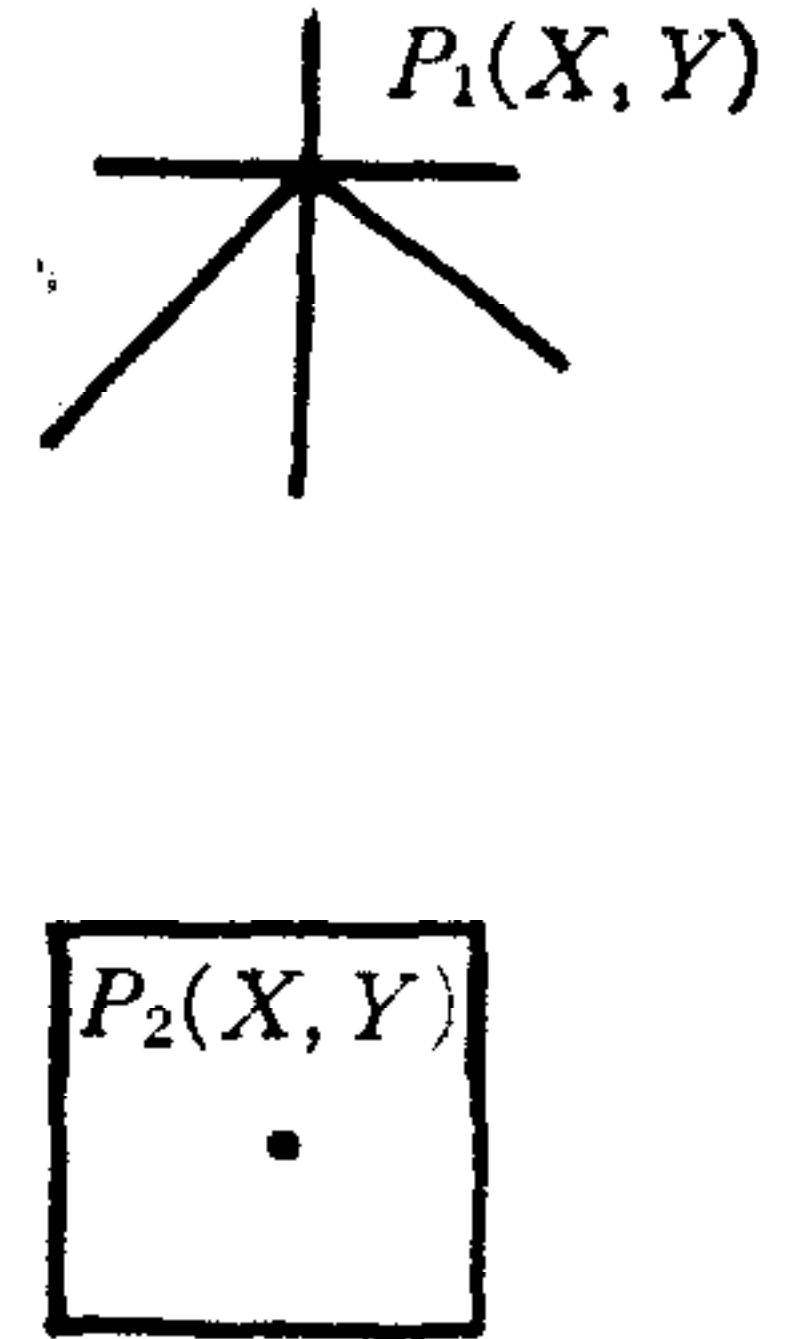
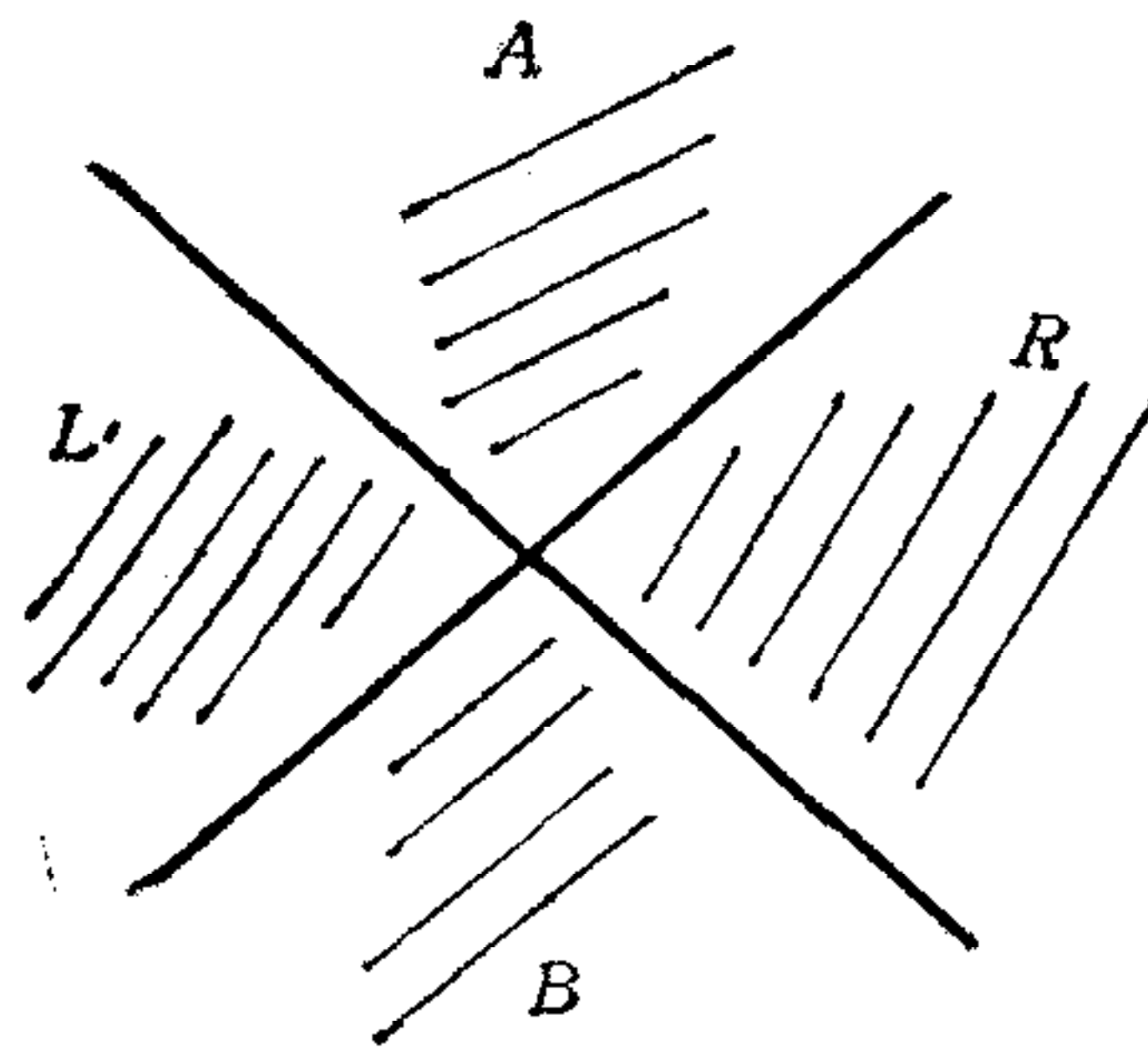


图 5

除了围绕的关系外,当 $r > r_0$ 时, $C_1 \cap C_2 = \phi$. r_0 为给定阈值.

考虑如图 5 所示的两个模块.

C_1 与 C_2 之间三种不同关系就形成三个不同的合体汉字,如图 6 所示.

(1) $\text{REL}(C_2, C_1) = (AR, v)$. 其中 $v = re^{i\theta}$; $r_0 \leq r \leq r_1$;

$$-\varepsilon - \frac{\pi}{2} \leq \theta \leq \varepsilon - \frac{\pi}{2};$$

r_0, r_1, ε 是表示阈值的正整数.

(2) $\text{REL}(C_2, C_1) = (A, v)$.

(3) $\text{REL}(C_2, C_1) = (B, v)$.

对于 n 个模块 ($n > 2$) 之间的关系 $\text{REL}(C_1, C_2, \dots, C_n)$, 可以转换为二元关系:

$$\text{REL}_1(C_1, \text{REL}(C_2 \cdots \text{REL}_{n-1}(C_{n-1}, C_n))),$$

或者转换为两个两个模块之间的二元关系：

$$\text{REL}_1(C_1, C_2) \wedge \text{REL}_2(C_1, C_3) \wedge \cdots.$$

四、句法、词义及模块间关系的距离度量

对于一个与独体字对应的模块 C ，可用一棵树来表征它的结构。这里所说的树是指带有方向的树。例如图 2 就是描述汉字里的句法结构的一棵树。两个模块 C_1 和 C_2 之间的结构距离 $d_{syn}(C_1, C_2)$ 表示 C_1 结构的树 α 与 C_2 结构的树 β 之间的距离 $d_{syn}(\alpha, \beta)$ 。这个距离是用一系列变换 T (包括代换、除去、插入三种运算，每种运算赋予加权量) 按一定准则把树 α 转变成 β 的花费^[8]，可以用已有的算法来求这个距离。如果两棵树的距离为零，则称所对应的两个模块具有相同的结构。对于两个结构不同的模块，经过一系列变换可变成结构相同的模块。

对于结构相同的两个模块 C_1 及 C_2 ，以同样的笔划顺序，分别用两条属性鍊表示为

$$C_1 = a_1 a_2 \cdots a_n,$$

$$C_2 = b_1 b_2 \cdots b_n.$$

把 a_i, b_i 的长度作为词意，表示为 $A(a_i) = l_i, A(b_i) = \tilde{l}_i$ 。对于某些情形，只要考虑相邻两个基元之间的连接关系即可。如果

$$a_i \text{ 与 } a_{i+1} \text{ 的连接关系为 } \text{CAT}(a_i, a_{i+1}) = (P_i, \theta_i), i = 1, \cdots, (n-1),$$

$$b_i \text{ 与 } b_{i+1} \text{ 的连接关系为 } \text{CAT}(b_i, b_{i+1}) = (P_i, \tilde{\theta}_i), i = 1, \cdots, (n-1).$$

那么 C_1 和 C_2 之间的词意距离 $d_{sem}(C_1, C_2)$ 定义为

$$d_{sem}(C_1, C_2) = \left[\sum_{i=1}^n \lambda_i \{ (l_i - \tilde{l}_i)^2 + (\theta_i - \tilde{\theta}_i)^2 \} \right]^{\frac{1}{2}}.$$

其中 λ_i 是表示加权量的正数。上述距离反映了两个模块相应笔划的长短及形状不同的程度。

下面考虑合体字。一个合体字往往由若干个模块，例如 N 个模块构成，表示为

$$C_1 = f_1(C_1^1, C_2^1, \cdots, C_N^1), C_2 = f_2(C_1^2, \cdots, C_N^2).$$

其中 f_1, f_2 表示某种结构关系。设 N 个模块中间有 k 个两两之间具有二元关系。其中 C_i^1 与 C_j^1 的关系为 $\text{REL}(C_i^1, C_j^1) = (P_{ij}^1, v_{ij}^1)$ ， C_i^2 与 C_j^2 的关系为 $\text{REL}(C_i^2, C_j^2) = (P_{ij}^2, v_{ij}^2)$ ，则 (P_{ij}^1, v_{ij}^1) 与 (P_{ij}^2, v_{ij}^2) 之间的关系距离为

$$d_{rel}[(P_{ij}^1, v_{ij}^1), (P_{ij}^2, v_{ij}^2)] = \begin{cases} \mu |v_{ij}^1 - v_{ij}^2|^2, & \text{如果 } P_{ij}^1 = P_{ij}^2; \\ \delta \gg 1, & \text{如果 } P_{ij}^1 \neq P_{ij}^2. \end{cases}$$

其中 μ, δ 为正数，表示加权量。上述距离主要反映出相应的关系是否一致。于是可以定义两个模块 C_1 和 C_2 之间的关系距离 $d_{rel}[C_1, C_2]$ 如下：

$$d_{rel}[C_1, C_2] = \sum_k d_{rel}[(P_{ij}^1, v_{ij}^1), (P_{ij}^2, v_{ij}^2)].$$

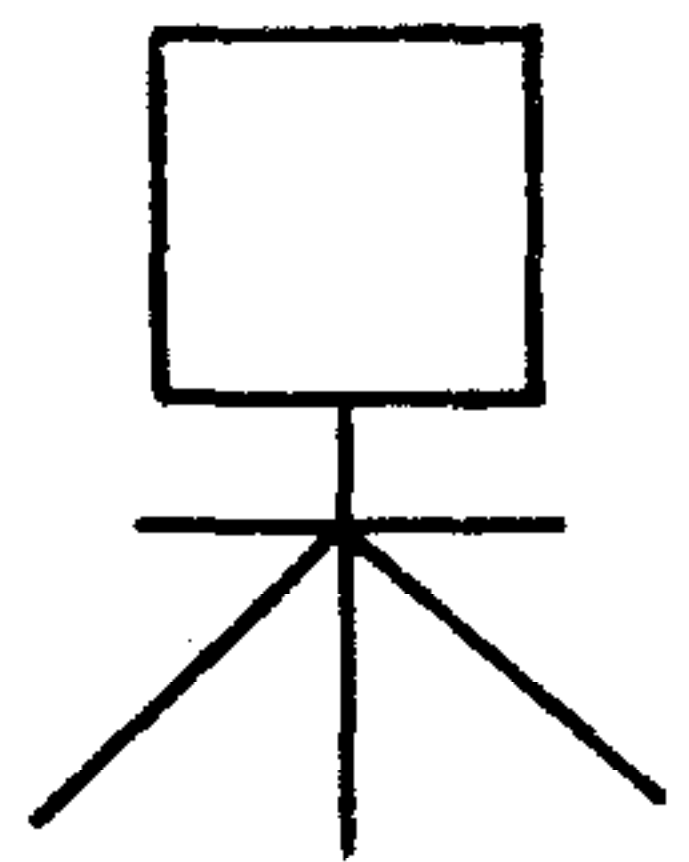
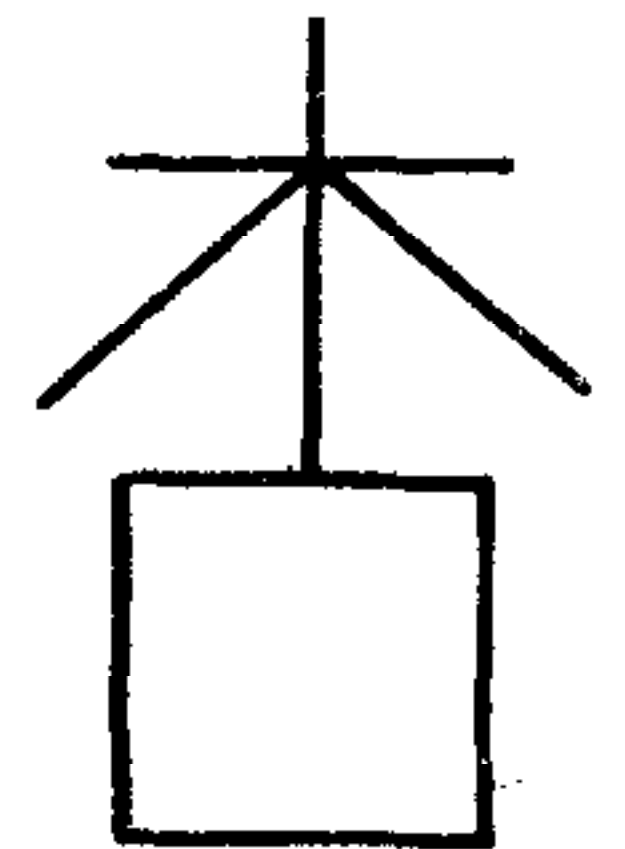
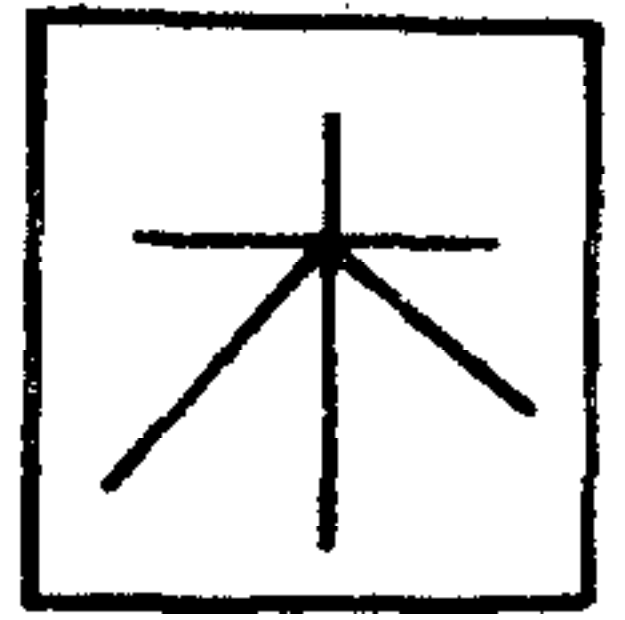


图 6

总而言之,可以利用模块之间的结构距离、词义距离以及模块间的关系距离来表征两个模块之间的相似程度,用最小距离准则来进行汉字识别。例如有 m 个汉字类,对每一类汉字建立一个标准模板,这个模板包括构成这类汉字的各模块间相互关系,用树式鍊表示句法结构以及笔划之间的连接属性,然后首先进行粗分类,再进行细分类。对于一个输入的汉字,通过计算输入与模板之间的关系距离及各个相对应的构件之间的句法距离,以这一距离最小且为零作为准则,进行粗分类。在此基础上,再根据词义距离最小作为准则来进行更为精确的分类。因为在汉字中存在着结构相同但并非一类的汉字,例如日,曰;干,土;肚,肝。所以引入词义信息是有意义的,可以根据词义距离来正确地进行区分。

五、结 束 语

人进行模式识别的本领是很高明的,一个人在听别人含混不清的讲话或者进行复杂图形识别过程中,不仅仅是用到有关句法的信息,还要依靠词义信息,以及各种知识而且词意信息往往比句法信息更为重要。至于模式识别的工程应用,一般说来,如果方法简单,用起来比较方便,就容易被接受。在以往句法模式识别的基础上引入词意信息有可能对以往的统计法和句法法取长补短,有机地结合起来,使识别方法简单、方便。

在与钱学森教授进行有关模式识别方法的讨论中,受到了启发与鼓励,谨表示衷心的感谢。

参 考 文 献

- [1] 傅京孙著,模式识别原理及其应用(戴汝为、胡启恒翻译、整理),科学出版社,1983年。
- [2] You K. C. and Fu K. S., A Syntactic Approach to Shape Recognition Using Attributed Grammars. *Trans on System Man and Cybernetics*, SMC-9 (1979), No. 6.
- [3] Tsai W. H. and Fu K. S., Attributed Grammar—A Toll for Combining Syntactic and Statistical Approach to Pattern Recognition, SMC-10 (1980) No. 12.
- [4] Tai J. W. Attributed Parallel Tree Grammars and Automaton for Syntactic Pattern Recognition. *Proc. 5-th ICPR Miami Beach, Dec. 1980.*
- [5] Tai J. W. and Fu K. S., Semantic Syntax-Directed Translation for Pictorial Pattern Recognition. *Purdue University Tec Rep. TR-EE 81—38 1981*
- [6] Shi Q. Y. and Fu K. S., Parsing and Translation of (Attributed) Expansive Graph Language for Scene Analysis. *6-th ICPR Munich, Germany. Oct. 1982.*
- [7] 戴汝为,模式识别的一类属性文法,自动化学报,第9卷第2期,1983年。
- [8] Fu K. S., *Picture Syntax. Pictorial Information System.* Edited by Chang S. K. and Fu K. S. Springer-Verlag, 1980.
- [9] M. Yichida, M. Ikeda and Tsuji, A plan-Guided Analysis of Cineangiograms for measurement of Dynamic Behavior of Heart Wall. *IEEE Trans. on Pattern Analysis and machine Intelligence*, PAMI-2 No 6, Nov. 1979.

A LINE DRAWING PATTERN RECOGNITION METHOD

TAI JUWEI

(Institute of Automation, Academia Sinica)

ABSTRACT

The unification of syntactic and statistical approach has been investigated so as to get an efficient method for pattern recognition. In this paper, based on the work done by the author, and referred to a point of view in cognitive psychology, the recognition of line drawing patterns is discussed. The tradeoff between syntax and semantics in high dimensional attributed grammar is shown by an example. In addition, the distance measures of syntax, semantics, and relationships between chunks are also given for line drawing pattern recognition. As a matter of fact, the minimum distance criterion can be applied for line drawing pattern (especially for Chinese character) recognition.