

统计和规则相结合的汉语组块分析

李素建 刘 群 白 硕

(中国科学院计算技术研究所 北京 100080)
(lisujian@ict.ac.cn)

摘 要 从文本中获得的组块对机器翻译、信息检索等很多领域都非常有用. 介绍了规则和统计进行组块分析的处理策略, 提出了规则与统计相结合的处理方法. 并且结合组块分析的实际情况改进了一般评价系统性能指标, 通过封闭测试和开放测试验证, 与单纯规则组块划分相比较, 组块识别的精确率和召回率都得到了提高, 组块划分错误率降低了 7%.

关键词 组块分析, 部分分析, 语料库

中图法分类号 TP18

CHINESE CHUNKING PARSING USING RULE-BASED AND STATISTICS-BASED METHODS

LI Su-Jian, LIU Qun, and BAI Shuo

(Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100080)

Abstract To acquire chunks from running texts is useful for many applications, such as machine translation, information retrieving, etc.. Described in this paper are the schemes of rule-based chunker and statistics-based chunker. Also proposed is a method to combine rule-based processing with statistics-based processing. According to the practical situation the mistake recall is introduced to rate the performance of the system. Compared with the rule-based system, the precision and recall are enhanced to identify chunks, and the error rate is reduced about 7%. The performance of the whole system has been improved greatly.

Key words chunk parsing, partial parsing, corpus

1 引 言

句法分析是自然语言处理中的重点和难点, 目前针对完整的句法分析在自然语言处理领域中遇到困难, 提出了部分句法分析, 它是近年来出现的一个新的语言处理策略. 部分句法分析 (partial parsing), 又叫组块分析 (chunk parsing), 是和完全句法分析相对的. 完全句法分析则着眼于充分分析整个句子的语法特点, 从而最大限度地揭示句子所反映的主题内容. 而

组块分析只限于把句子解析成较小的单元, 而不揭示这些单元之间的句法关系. 这样就降低了分析的难度.

组块分析并不是完全摒弃了完全句法分析, 而是充分引入和利用完全句法分析的理论和技术, 先把局部和小处的问题实际解决. 当前存在着两大研究分支:

① 统计方法, 如 Church^[1]应用 HMM 方法把短语边界识别化为在词类标记对间插入左右边界的问题;

② 规则方法, 其中 Abney^[2]使用的有限状态叠

原稿收到日期: 2000-10-23; 修改稿收到日期: 2002-01-23

本课题得到国家“九七三”重点基础研究项目基金资助 (G1998030507-4)

(finite-state cascade)简化了句子的语法分析.

近年来中文信息处理者也开始探索使用组块分析方法,孙宏林^[3]使用组块分析方法得到“V+N”序列的汉语短语,此外识别各种名词短语^[4~7]的组块分析也都取得了令人满意的结果.

不管使用统计方法还是规则方法,国内外利用组块分析的重点都放在了名词短语的识别上.结合汉语的特点,根据词组本位论,发现理解汉语语句时是从一个个语块开始着手的,但这些组块不仅仅局限于和名词有关的组块上,Abney^[8]提出了英语中的各种基本组块(形容词组块,动词组块等),有关划分汉语各种组块的工作还很少.本文提出了一个基于统计和规则相结合的组块划分器获得汉语中的基本组块(base chunk, BaseCK),并采用《人民日报》的真实语料进行测试,精确率能达到 96%以上,召回率约为 92%.

2 统计方法模型

在介绍统计模型之前,首先对组块的相关概念明确定义一下.根据 Abney^[8]对于组块的定义,我们结合汉语的特点定义如下.

定义 1. 组块(chunk)是一种结构,是符合一定句

[句法/n 分析/v]	[是/v]	[自然/n 语言/n 处理/v]	中/f 的/u	[重点/n]	./w
$\langle \emptyset, n \rangle$	$\langle n, v \rangle$	$\langle v, v \rangle$	$\langle v, n \rangle$	$\langle n, n \rangle$	$\langle n, v \rangle$
0	3	2	2	3	3
					$\langle v, f \rangle$
					$\langle f, u \rangle$
					$\langle u, n \rangle$
					$\langle n, w \rangle$
					1

对于任意的一个有 n 个词的汉语语句 $w_0/\emptyset, w_1/t_1, w_2/t_2, \dots, w_n/t_n, w_i(1 \leq i \leq n)$ 表示一个汉词, t_i 表示 w_i 所对应的词性标注, w_0/\emptyset 可以看做是固定标志,代表一个语句的开始.令 $t_i^* = \langle t_{i-1}, t_i \rangle (1 \leq i \leq n)$, 并设组块状态集 $CS = \{0, 1, 2, 3, 4, c_1, c_2, \dots, c_n (1 \leq i \leq n, c_i \in CS)\}$ 是对于上面汉语语句的组块标注.因此要找到最好组块标注序列 C^+ 满足如下条件:

$$C^+ = \arg \max_{C_i} P(C_i | w_1^n, t_1^n), \quad (1)$$

C_i 表示一种可能的组块标注序列.因为常用的汉词个数就有 6000 左右,不可能对每个汉词一一加以考虑,在这种情况下我们主要考虑语句中的词性标注.式(1)便可简化为

$$C^+ = \arg \max_{C_i} P(C_i | t_1^n) =$$

$$\arg \max_{C_i} \frac{P(C_i)P(t_1^n | C_i)}{P(t_1^n)} =$$

$$\arg \max_{C_i} P(C_i)P(t_1^n | C_i). \quad (2)$$

法功能的非递归短语.每个组块都有一个核(head),组块内所有成分都围绕着核进行扩展.

在汉语中定义的组块类型有名词组块、动词组块、形容词组块、副词组块和介词组块.组块是严格按照语法定义的,而不是按照语义和功能进行定义,因此它与关系紧密的短语(根据搭配关系、语义一致性、互信息等形成)是有区别的.

定义 2. 基本组块也可以称为最小组块,是构成语句的最小句法功能单位,不能为其它组块所包含.

Church^[1]使用的统计技术把 HMM 标注器的输出作为输入,通过在词之间插入左右方括号 ‘[’, ‘]’ 来标识名词组块(noun chunks).本文中的统计模型通过对以上标注器进行合理吸收和适当改进,设定汉词与汉词之间的标记有 5 种情况:[,],][, I(no brackets inside a chunk), O(no brackets outside a chunk).所以组块标注可以看做是每个词对间的状态标记,可能处于以下 5 种状态之一: 0([), 1(]), 2(][), 3(I), 4(O). 在建立的隐马尔科夫模型中,把组块标注看做为不可观察的过程,两个相邻词的词性标注对序列作为可观察的过程.

例 1. 一经过组块划分汉语句子,我们可以看到每个词所处的组块状态如下:

为了使式(2)适于计算,我们假设每个词标注间加入组块状态的事件是独立的,同时使用二元模型,则实际的工作模型是

$$C^+ = \arg \max_{C_i} P(C_i)P(t_1^n | C_i) =$$

$$\arg \max_{C_i} \prod_{i=1}^n P(c_i | c_{i-1}) \prod_{i=1}^n P(t_i^* | c_i), \quad (3)$$

这样最后得到的组块标注序列 C^+ 满足式(3).

此模型的两组参数: $P(c_i | c_{i-1})$ 和 $P(t_i^* | c_i)$, $i \in [0, n]$ 可以从以正确标注组块的语料中进行参数训练得到,一般的参数估计法为最大似然估计,

$$P(c_i | c_{i-1}) = P(c_i, c_{i-1}) / P(c_{i-1}) \approx$$

$$f(c_i, c_{i-1}) / f(c_{i-1});$$

$$P(t_i^* | c_i) = P(t_i^*, c_i) / P(c_i) \approx$$

$$f(t_i^*, c_i) / f(c_i); \quad (4)$$

其中 $f(c_i, c_{i-1})$ 是语料库中 c_i 和 c_{i-1} 共现的频度, $f(t_i^*, c_i)$ 是语料库中 t_i^* (即词对 $\langle t_i^*, t_{i-1} \rangle$) 间赋予组

块标记 c_i 的频度, $f(c)$ 为语料库中组块标记 c 出现的频数.

在该模型中对于数据稀疏问题的解决方法是, 对于 t_i^* 取值为 c_i 的情况, 赋值 $f(t_i^*, c_i) = 1$. 寻找一个组块标记的最优路径采用动态规划算法, 该路径的概率值最大. 对于输入语句长度为 t 的语句(状态数为 $n=5$), 其时间复杂度为 $O(n^2t)$.

3 规则方法

由于式(3)需要大规模的经过组块标注的语料库, 但目前这种语料很难获得, 我们只手工标注了一部分语料进行训练, 语料规模小对于组块标注是个极大的障碍. 为了解决这个问题, 同时因为基于统计的方法在分析语言结构方面先天不足, 不能系统化地表示语言. 结合规则方法对统计得到的结果进行进一步处理, 即由统计总结出规则, 并由系统进行学习改进后再继续应用到系统中去.

定义 3. 基本组块规则. 语句中如果存在 $w_1/t_1, \dots, w_m/t_m$ 构成一个基本的组块, 则词性标注序列 $t_1t_2 \dots t_m$ 称为一条基本组块规则.

由第 2 节介绍的统计方法语料进行组块划分, 对每一个组块构成提取出一条基本组块规则. 统计基本组块规则的数目及发生的频数.

定义 4. 合式的基本组块规则. 首先选定一个阈值 θ , 假设某条基本组块规则 r 的出现频数 $f(r)$ 满足 $f(r) > \theta$ 时, 则称规则 r 为合式的基本组块规则.

定义 5. 基本组块规则集 G . 所有合式的基本组块规则构成基本组块规则集 G .

汉语中一些常用的字经常嵌入到组块中形成固定结构, 因此在基本组块规则中会出现一些诸如

- (1) $nc \rightarrow a$ '的' n
- (2) $nc \rightarrow a$ '的'
- (3) $nc \rightarrow n$ '的' n
- (4) $nc \rightarrow m^* q n$
- (5) $nc \rightarrow m^* q$
- (6) $vc \rightarrow [d^*] v$
- (7) $vc \rightarrow d$ '地' v

“所”“的”之类的字. 例如, “数学家/ n 所/ u 认为/ v 的/ u ”, 这里规则中的 u 必须加以限定, 也就是我们要把规则改为: n 所 v 的. 否则就会极大地影响精确率. 同时我们采用了北京大学计算语言学研究所的汉语词性体系(见表 1).

表 1 汉语词性标注所用到的符号(共 26 类)

词性符	含义	词性符	含义
a	形容词	o	象声词
h	前接成分	J	简称略语
c	连词	q	量词
d	副词	r	代词
e	叹词	s	处所词
f	方位词	t	时间词
g	语素	u	助词
b	区别词	v	动词
k	后接成分	w	标点符号
i	成语	x	字
p	介词	y	语气词
l	习用语	z	状态词
m	数词	^	开始
n	名词	\$	结束

构成组块的几个词性标注序列构成一条合式的基本组块规则, 加入到基本组块规则集中去. Abney^[2]提出了有限状态叠来进行组块分析, 从而减少分析中的复杂度. 根据这个原则, 我们由基本组块规则集 G 来构造有限状态自动机, 对输入语句进行分析. 分析实质上是一个模式匹配过程, 在这个过程中, 如遇到冲突, 则按最大匹配原则选择合适的模式. 例如, 图 1 是基本组块规则集的一个子集及其对应的有限自动机.

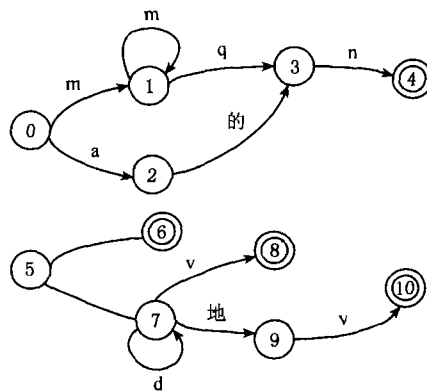


图 1 基本组块规则集的一个子集及其对应的有限自动机

规则子集中, * 表示一个或多个; [] 表示可以省略; nc (noun chunk) 表示名词组块; vc (verb chunk) 表示动词组块; 词性标注参见表 1. 图 1 中圆圈表示自动机的状态, 双圆圈表示一个组块的终止状态. 输入已经正确分词和标注的汉语语句, 输出经过组块标记的汉语语句.

例 2. “-/m 个/q 好/a 的/u 开端/n 是/v 成功/n 的/u 一半/n ./w”.

表 2 描述了有限自动机进行组块分析的过程. I 代表输入, S 代表一个特殊的状态栈, 在每次进行一个新的组块分析前, 存放所有自动机的初始状态, 其内部可以自动调整, 分析开始时把需要的自动机状态放在栈顶. 状态 s_i 代表自动机移进到状态 i , r_j 代表根据基本组块规则 j 得到了一个名词组块(nc)或动词组块(vc). 每一次都是从某个自动机的初始状

态出发寻找匹配, 当达到该自动机的终止状态时则产生一个组块. 若遇到不能形成组块的词不影响组块划分的进行, 跳过从下一个词继续进行处理. 最后输出经过组块标注的语句“[-/m 个/q] [好/a 的/u 开端/n] [是/v] [成功/n 的/u 一半/n] ./w”.

由于在组块标注过程中发生冲突的时候采用最大匹配方法, 极大地降低了组块划分的精确率. 因此我们可以根据使用频度规定规则的优先级, 发生冲突时优先考虑使用频度大的规则, 从而使划分更为准确.

通过统计语料库中的所有组块长度, 假设每个组块的平均长度为 $avg(|ck|)$, 则每个组块分析的时间复杂度为 $O(avg|ck|)$, 对于一个长度为 t 的语句, 平均有 $t/avg(|ck|)$ 个组块, 因此该语句分析的时间复杂度为 $O(t)$.

表 2 规则划分组块过程

S	I									
	-/m	个/q	好/a	的/u	开端/n	是/v	成功/n	的/u	一半/n	./w
5, 0	s1	s3	nc, r5							
5, 0			s2	s3	s4	nc, r1				
0, 5						s6	vc, r6			
5, 0							s2	s3	s4	nc, r3

4 统计和规则相结合

从以上两节的分析中可以看出, 规则处理的优势在于充分利用语言学知识, 由有限自动机对组块进行快速地划分. 而统计处理的优势在于全部知识是通过大规模语料库的参数训练得到的, 有很好的一致性和较高的覆盖率^[9]. 所以最好的方法是考虑把它们结合起来, 充分发挥二者的优势.

统计和规则相结合体现在两个方面: 一是在形成和完善规则以及统计模型过程中, 两者相辅相成; 二是对新的语料进行分析时规则和统计方法并举, 新语料由预处理器分别送到规则或统计组块划分器中进行处理.

图 2 中, 椭圆表示所用到的或生成的资源, 矩形表示一种处理的过程. 虚线框内是对规则模型和统计模型的完善过程. 最初由于没有一个大规模的、带准确组块标注标记的语料库, 因此先以粗规则集标注一部分语料. 通过人工校对, 发现和改正其中的错

误, 对统计模型进行参数训练, 并调整规则库内容, 然后使用规则和统计方法标注一部分新语料, 对所得到的更多的正确标注语料, 重新进行参数训练和规则调整, 如此循环. 随着语料库数量的扩大, 规则描述会更准确, 统计信息更全面, 从而充分发挥规则和统计相结合处理优势, 提高自动组块标注的准确性. 在收集组块标注语料时, 当新语料到来, 首先经过预处理器处理一下, 适合于规则分析的语句进入规则标注器进行组块划分, 适合于统计分析的语料进入统计标注器进行组块划分, 这样提高了组块划分的正确率, 降低了手工校正的工作量.

上面两节对规则和统计方法的时间复杂度进行过分分析, 考虑效率因素我们采用规则划分组块为主, 并辅以统计方法. 在标注过程中, 规则方法由于采用最大匹配方法进行确定性划分, 某些规则被屏蔽, 从而使组块划分的效果变差, 严重降低了精确率. 例如统计发现, 由于名词和动词在汉语使用中相当普遍和灵活, 它们在句中出现的频率较高, 规则模型机械按照语法的定义取出最长的可能组合形式进行组块

划分,经常出现错误,而使用统计方法标注的效果比较理想.因此由预处理器过滤出这些语句送入规则

标注器进行处理,这样并不会对系统的速度造成很大的影响.

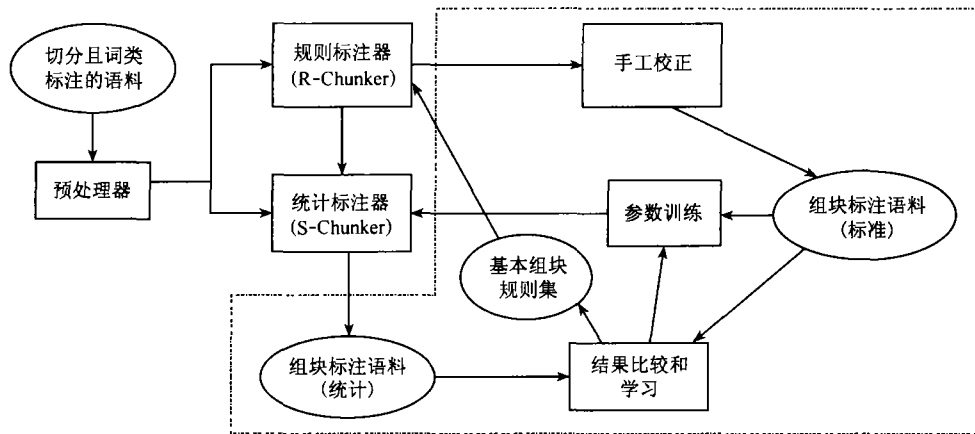


图2 统计和规则相结合处理的结构图

5 实验结果及分析

根据上面介绍的规则和统计相结合的方法,我们实现了一个面向汉语真实文本的汉语组块划分系统.本文使用1996年电子版《人民日报》(约计2400万字)已进行词性标注的文本作为语料,利用规则模型和统计模型组块标注文本的同时对模型进行修改、训练和完善.最终得到有价值的组块标注语料,规则模型获得常用的组块规则有160多条;统计模型中的两组参数(见式(4))也得到估值.同时手工对一汉语短文(约计60000汉字)进行词性标注和组块标注(约11200个组块),作为测试评价语料.使用组块划分器对该文进行自动组块标注,所得结果与测试评价语料相比较,采用以下几个指标:块划分精确率(*precision*)、召回率(*recall*)、错误率(*error*)和漏召率(*leakage*)来衡量本文系统的性能,定义如下:

a = 正确识别的组块总数;

b = 没有识别的组块总数;

c = 识别错误的组块总数;

召回率(*recall*) = $\frac{a}{a+b} \times 100\%$;

精确率(*precision*) = $\frac{a}{a+c} \times 100\%$;

漏召率(*leakage*) = $\frac{b}{a+b} \times 100\%$;

错误率(*error*) = $\frac{c}{a+c} \times 100\%$;

$$recall + leakage = 1;$$

$$precision + error = 1.$$

对于一个测试文本,它包含的全部组块分为两部分,一部分是由组块标注器召回的正确组块(a),一部分是漏召的组块(b).组块标注器错误召回的组块数目为 c ,它对于系统性能的影响最大,因此要保证尽量降低错误率.

精确率和召回率是一对互相矛盾的指标.当基本组块规则数目 $c(r)$ 增多时,呈召回率上升、精确率下降的趋势;当规则数目减少时,精确率上升,召回率下降.当然规则数目的大小在某一个区间内才会呈现这种趋势.当过大或过小时,精确率和召回率都会受到极大的影响.建立规则集时,阈值的选择是一个重要的问题.我们的原则是在保证精确率和召回率基本平衡的基础上,尽量降低误召率和提高精确率.即使召回率不十分理想,尚可由后续手段予以进一步的处理.反过来,如果真正的组块划分错误,则会造成该句中其它组块划分的错误.表2为两种模型的实验结果比较,可以看出通过规则和统计相结合的方法,可以增大正确识别的组块数目 a ,不被识别的组块数目 b 就相应减少,而文本中组块的总数 $a+b$ 是一个恒量,从表中也可以看出召回率提高了约4.5个百分点.组块标注器识别的组块数目没有明显地增长,错误识别的组块数目已明显下降,从表3中的数据可以知道错误率降低了约7~8个百分点,这样就大大提高了系统的性能.

表3 两种模型的性能比较

测试类型	规则标注模型			规则和统计结合模型		
	精确率	召回率	错误率	精确率	召回率	错误率
封闭测试	89.9	87.5	10.1	96.2	92.1	3.8
开放测试	88.2	86.9	11.8	94.6	91.3	5.4

相对于精确率,召回率较小,因为和召回率相对的是漏召率,漏召对于系统的影响较小,而错误召回不仅错误识别了本组块,还会波及其它组块.因此我们宁可增大漏召率,而尽量减少错误率,这是系统组块标注的一个原则.所以识别结果中召回率相对较低不能作为系统性能评价的主要因素.

在使用统计和规则相结合的方法后,错误率仍为4%左右,下面我们分析一下错误存在的几方面原因:

(1) 由于测试评价标准是手工组块标注的结果,多少会有一些主观因素在其中,对于原始语料选择不同的人标注得到不同的标注结果.如果这几种标注结果作为一个并集,自动标注的结果只要属于这个并集就为正确,则精确率和召回率都大大地提高.定义为伪错误.

(2) 由于在组块划分过程遇到歧义问题,解决的方法是最大匹配原则,虽然对于出现错误多的规则引入统计方法,但也只能局限于小范围内集中降低错误率,还是不可避免地会产生一些划分错误.对于普遍的规则要解决这些错误,我们下一步考虑采用错误驱动策略对规则进行限制,当出现歧义问题时,根据限定条件进行组块划分.

(3) 使用统计方法进行组块划分使用了二元模型,只考虑到了前后一个词对当前词划分的作用,忽略了远距离词对组块划分的影响.可以通过提高模型的元数降低错误.同时因为训练统计模型的语料仍然不够大,校对过程中也有一些错误存在,这些对模型参数以至组块划分也有一定的副作用.

6 结 论

本文提出了一种有效的组块分析方法.它的基本原理是由语言学知识得到初步的组块划分语料,通过校正和学习不断对规则进行调整,完善规则模型,并在不断增大的标注语料基础上对统计模型进行训练,得到组块划分的统计模型.规则和统计方法并用,对文本进行组块标注,提高识别的精确率和召

回率,降低错误率.从封闭测试和开放测试的试验结果来看,两种方法结合进行标注的正确率分别达到了96.2%和94.6%,比单纯的规则方法约提高了7%.本文方法的显著特点是统计信息从语料库中得到,覆盖性和一致性好,同时规则的语法分析器确定性强,实现简单,效率高.实验结果表明,在兼顾识别效率、精确率、召回率的基础上,系统取得了令人满意的效果.

从对文本进行组块标注的实践中,要充分发挥规则和统计相结合方法的能力,关键在于找到规则标注和统计标注的最佳组合.因此,对于汉语组块自动标注,就要从大量语料的处理中,总结哪些语言现象更适用于用规则方法处理,哪些更适合用统计方法处理,合理运用不同处理方法,以获得最好的标注效果.

汉语中的组块分析是处于对语句进行分词标注和完整句法分析之间的一个中间步骤.它可以平缓一下分词标注中的错误,对于一些未登录词和常用语的识别也有很好的效果.同时语法分析是汉语处理的一大难题,组块分析减少了完整语法分析的复杂度.本文采用的组块分析技术,得到大规模的组块标注库,不仅为下一步的组块连接(attachment)问题打下了基础,而且为机器翻译、信息提取、文献检索等一系列应用提供了支持.

参 考 文 献

- 1 K W Church. A stochastic parts program and noun phrase parser for unrestricted text. In: Proc of 2nd Conf on Applied Natural Language Processing. Austin, Texas, 1988. 136~143
- 2 Abney, Steven. Parsing by chunks. In: Robert Berwick, Steven Abney, Carol Tenny eds. Principle-Based Parsing. Dordrecht: Kluwer Academic Publishers, 1991. 257~278
- 3 孙宏林. 从标注语料库中归纳语法规则:“V+N”序列试验分析. 见:语言工程. 北京:清华大学出版社,1997. 157~163
(Sun Honglin. Acquiring grammatical rules by induction from tagged corpus: A case study on “V+N” sequence. In: Language Engineering (in Chinese). Beijing: Tsinghua University Press, 1997. 157~163)

- 4 K H Chen, H H Chen. Extracting noun phrases from large-scale texts: A hybrid approach and its automatic evaluation. In: The 32nd Annual Meeting of the Association for Computational Linguistics. Las Cruces, New Mexico, 1994
- 5 A S Y Tse, Kam-Fai Wong *et al.* Effectiveness analysis of linguistics and corpus-based noun phrase partial parser. In: Proc of the Natural Language Processing Pacific Rim Symposium (NLPRS'95). Taljoi: Academic Press, 1995. 252~257
- 6 赵军, 黄昌宁. 基于转换的汉语基本名词短语识别模型. 中文信息学报, 1999, 13(2): 1~7
(Zhao Jun, Huang Changning. Recognition model of Chinese BaseNP based on transformation. Journal of Chinese Information Processing(in Chinese), 1999, 13(2): 1~7)
- 7 周强, 孙茂松, 黄昌宁. 汉语最长名词短语的自动识别. 软件学报, 2000, 11(2): 195~201
(Zhou Qiang, Sun Maosong, Huang Changning. Automatic identification of Chinese maximal noun phrases. Journal of Software(in Chinese), 2000, 11(2): 195~201)
- 8 Abney, Steven. Chunk stylebook. University Tubingen. Tech Rep; Sfs, 1996. <http://www.sfs.nphil.uni-tuebingen.de/~abney/Papers.html#96i>
- 9 周强. 规则和统计相结合的汉语词类标注方法. 中文信息学报. 1995, 9(3): 1~10

(Zhou Qiang. Chinese POS tagging method with rules and statistics combined. Journal of Chinese Information Processing (in Chinese), 1995, 9(3): 1~10)



李素建 女, 1975年生, 博士研究生, 主要研究方向为自然语言处理、机器翻译、知识挖掘。



刘群 男, 1966年生, 博士, 主要研究方向为自然语言处理、机器翻译、知识挖掘。



白硕 男, 1956年生, 研究员, 博士生导师, 主要研究方向为计算语言学、自然语言处理、网络安全。

Pacific Graphics 2002

征文通知

Pacific Graphics 2002 (PG2002)国际会议将于2002年10月9~11日在北京清华大学召开。Pacific Graphics 是国际上具有重要影响的图形学学术会议,会议论文集由IEEE Computer Society出版社出版,会议的优秀论文在Graphical Models和Computer Graphics Forum上以专辑形式发表。会议的语言为英语。有关论文投稿信息,请查阅会议网页。欢迎投稿和参加会议。

会议网页: <http://ncc.cs.tsinghua.edu.cn/pg2002>

征文截止时间: 2002年4月30日

联系方式: 清华大学计算机科学与技术系 胡事民 邮编:100084

电话:010-62782052 传真:010-62782052

E-mail: pg2002@tsinghua.edu.cn