

## 口语对话中的语句分组<sup>\*</sup>

徐为群<sup>+</sup>, 徐波, 黄泰翼

(模式识别国家重点实验室(中国科学院 自动化研究所),北京 100080)

### Grouping Utterances in Spoken Dialogues

XU Wei-Qun<sup>+</sup>, XU Bo, HUANG Tai-Yi

(National Laboratory of Pattern Recognition (Institute of Automation, The Chinese Academy of Sciences), Beijing 100080, China)

+ Corresponding author: E-mail: weiqun.Xu@gmail.com, http://www.ia.ac.cn

**Xu WQ, Xu B, Huang TY. Grouping utterances in spoken dialogues. Journal of Software, 2006,17(2):250-258.**  
<http://www.jos.org.cn/1000-9825/17/250.htm>

**Abstract:** In this paper, the interaction patterns and their automatic analysis in spontaneous spoken information-seeking dialogues are studied. First, based on previous work from discourse analysis (i.e., exchange as basic interaction unit in Birmingham School) and Systemic Functional Grammar (i.e., Halliday's speech function), a principled scheme is proposed to model interaction patterns with utterance groups. Then a dialogue corpus is annotated with this scheme and further analyzed. Some main factors affecting the structure of utterance group are distinguished. Based on these, an algorithm is established to analyze utterance groups and is evaluated in the corpus. The results achieve a correct rate of 55.4%~84.2% for overall utterance tags, depending on the different recognition performances of the extended sentence type and utterance topic.

**Key words:** interaction pattern; utterance group; dialogue; computational analysis

**摘要:** 研究了信息类自然口语对话中的交互模式及其自动分析.首先,基于话语分析中的 Birmingham 学派关于交互模式的工作和 Halliday 关于言语功能的分析,提出使用语句组来刻画交互模式,并建立原则性分类体系;然后,对语料中的交互模式进行标注分析;随后,根据影响语句组结构的主要因素建立交互模式分析算法,并在语料中进行实验.实验结果表明,语句组的整体分析正确率可达到 55.4%~84.2%——取决于不同来源的扩展句子类型和语句主题的分析结果.

**关键词:** 交互模式;语句组;对话;计算分析

**中图法分类号:** TP18 **文献标识码:** A

在文献[1]中,我们提出一个针对对话管理(对话系统的模块之一,起着中心控制作用)、面向任务的对话模型.该模型区分对话中的 5 个阶次的话语单元(即对话、对话相(包括开始相、主体相和结束相)、对话段(取决于主体相的任务如何分解为子任务)、语句组和语句)以及 3 个层次的动态过程(即宏观任务层、介观(或中间)交互层和微观语句层).其中,采用语句组来刻画介观层的交互模式.本文将对语句组或交互模式进行深入研究.这一工作至少具有以下两方面的意义:一方面,通过语句组建模和语料标注与分析,加深了对(信息查询类)对话

\* Supported by the National Natural Science Foundation of China under Grant No.60375018 (国家自然科学基金)

Received 2004-12-10; Accepted 2005-02-03

中的局部话语结构(即交互模式)的认识;另一方面,通过具体的计算分析,也为可能的实际应用(如对话系统中的对话管理)奠定了基础。

为了建立交互模式分类体系,本文第 1 节首先介绍相关的基础性工作,包括来自话语分析的关于交互基本单元(即 Birmingham 学派所提出的语换)的工作,以及 Halliday 关于言语功能的工作。第 2 节提出采用语句组来刻画交互模式,给出分类依据并建立原则性语句组分类体系。第 3 节对话料中的交互模式进行标注,并对语料中的语句组分布进行统计。为了实现交互模式的自动识别,第 4 节首先分析与语句组相关的若干因素,然后据此建立语句组分析算法,并对对话语料中的语句组进行自动分析。最后讨论相关工作。

## 1 前人的基础工作

Birmingham 话语分析学派的工作对本文关于交互模式的工作提供了直接的启发。但是,他们关于语换的分类缺少原则性分类准则。在寻找分类准则时,注意到 Halliday 在系统功能语法中关于言语功能的工作。下面,我们分别予以简略介绍(受篇幅限制,未能深入展开。请感兴趣者参阅相关文献,以了解进一步细节)。

### 1.1 作为交互基本单元的语换

由 Sinclair 与 Coulthard 所创立的 Birmingham 话语分析学派,最初为分析课堂交互结构建立了一套体系<sup>[2,3]</sup>。在这一体系中,区分了 5 个阶次的话语单元,按从大到小的顺序分别为:课或交互、事务(transaction)、语换(exchange)、语移(move)和行为(act)。该体系后来经过改进,用于其他类型的口头交互,如法庭话语<sup>[4]</sup>、日常会话<sup>[5]</sup>等。在该系统中,语换被定义为交互的基本单元。下面,根据文献[5]中的阐述加以介绍。

语换分为组织性的和会话性的。所谓组织性语换,是指关于问候、招呼和组织会话等的语换;而会话性语换,是指关于会话主旨的语换。此处关心的是会话性语换。会话性语换包括 3 个主要的,即起引(elicit)、告知(inform)和指示(direct)以及 3 个受限-起引的(bound-elicit,即受限于前一个语换,且以起引开始),即澄清(clarify)、重复(repeat)和重起始(re-initiate)。所有的会话性语换,除指示外,都有 I(R/I)R(F<sup>n</sup>)这样的结构。其中,I 和 R 是必须的;R/I 是可选的;F 是可选和不可预料的。表 1 给出了语换系统及其实现语移。

Table 1 Exchange

表 1 语换

	Elements of structure	Structures	Classes of move
Organizational: Boundary	Framing (Fr)	Fr	Fr: framing
Organizational: Structuring, Greet, Summon	Initiation (I) Response (R)	IR	I: Opening R: answering
Conversational exchange	Initiation (I) Response/Initiation (R/I) Response (R) Feedback (F)	I(R/I)R(F <sup>n</sup> )	I: eliciting, informing, directing R/I: eliciting, informing R: informing, acknowledging, behaving F: acknowledging

### 1.2 言语功能

在系统功能语法中,反映对话本质的言语功能(或人际功能,3 大元功能之一)受言语角色(或者是给予,或者是需求,即交换方向)以及所交换的内容(或者是信息,或者是实物和服务)两个因素的影响,可以分为 4 种基本类型(即提供、命令、陈述和提问)及其相应的积极或消极的响应,见表 2<sup>[6]</sup>。

Table 2 Speech function

表 2 言语功能

Commodity	Role	Initiation	Positive response	Negative response
Goods-and-Service	Giving	Offer	Acceptance	Rejection
Goods-and-Service	Demanding	Command	Undertaking	Refusal
Information	Giving	Statement	Acknowledgement	Contradiction
Information	Demanding	Question	Answer	Disclaimer

从上述基础工作的简单介绍中不难看出这两者间的互补性。Birmingham 话语分析学派对语换结构进行了

细致的分析,但缺少原则性的分类;而系统功能语言学关于言语功能的工作虽然具有一定的原则性,但是没有提供刻画实际交互模式结构的基础概念.因此,下面根据 Halliday 在系统功能语法中对言语功能的分析,对 Birmingham 学派的话语分析模型中的语换系统进行改进,建立新的语句组分类体系.首先给出分组依据,然后建立基本语句组和复杂语句组分类,并给出语料中的实例.

## 2 作为交互模式的语句组

### 2.1 分组依据

Halliday 划分言语功能的两个依据/变量,在应用到手头的信息查询类对话语料时,交换内容固定为信息.进一步分析发现,由于语料中对话者角色的不对称,单独使用言语角色来分析语料中的对话交互是不够的.因此,有必要将言语角色进一步分解为信息流向和交互起始/主动者.下面将分别讨论信息流向和交互起始者及其对交互模式的影响,并区分基本/复杂语句组.

信息流向.在信息查询类对话中,信息的流动方向主要分为两种,一种是从服务方流向客户方,另一种则是从客户方流向服务方.通常,客户方提供了基本信息以后,信息主要是从服务方流向客户方.在服务方提供信息时,若缺少关键信息,服务方会及时查询客户方.此时,通常会出嵌入式的前提性语句组.或者,若该信息为少数、可列举的,则服务方会渐进地提供信息.此时可能出现列举性语句组.我们的对话语料属于旅游信息领域,从客户方流向服务方的典型信息主要有:旅游线路、景点或大致区域(必需的)、旅游时间、游客人数等.而从服务方流向客户方的典型信息主要有:某条旅游线路的有无、可选线路、发团时间、周期、交通方式、价格、食宿条件、行程安排等.

交互起始者.在信息查询对话中,服务方和客户方的角色是不对称的,因而不同的说话者虽然采用相同的言语角色,但由于不同的习惯或规范,起始的交互模式也有所不同.若说话者角色是对称的,这一因素将不会影响交互模式.如果不考虑对话者角色的不对称或者对话者角色是对称的,那么,交互模式将可以仅用言语角色区别,即在信息查询对话中,将只需区分信息索求和信息提供两类交互.

### 2.2 基本语句组

根据以上两个分类依据,即信息流向和交互起始者(C:客户方;S:服务方),对话语料中反复出现的交互模式,可以用 4 个类型的基本语句组来描述,见表 3.

Table 3 Basic utterance groups

表 3 基本语句组

		Direction of information flow	
		S=>C	C=>S
Initiator	Customer	CISC	CICS
	Server	SISC	SICS

进一步的各类语句组实例如下(NLPR-TI 为本文所使用的语料,概况参见第 3.1 节):

- |  |  |
|--|--|
| <p>1. CISC 由客户方起始的,信息从服务方流向客户方<br/>【选自 NLPR-TI-26】<br/>C: 那个 它这个是 大概那个 坐什么交通工具呢<br/>S: 嗯 有双卧的 有卧往机返的<br/>C: 嗯 双卧的 嗯卧往机返<br/>S: 对</p> <p>2. SISC 由服务方起始的,信息从服务方流向客户方<br/>【选自 NLPR-TI-32】<br/>S: 乘火车去<br/>C: 呃 乘火车去<br/>S: 对</p> | <p>3. CICS 由客户方起始的,信息从客户方流向服务方<br/>【选自 NLPR-TI-47】<br/>C: 我想 那个 暑假 去 西安 那边 玩 去<br/>S: 暑假 啊<br/>C: 啊 对</p> <p>4. SICS 由服务方起始的,信息从客户方流向服务方<br/>【选自 NLPR-TI-31】<br/>S: 你几个人啊<br/>C: 我两个人 我和我爱人<br/>S: 两个人</p> |
|--|--|

### 2.3 复杂语句组

虽然大多数交互模式可以用上述 4 种基本语句组类型覆盖,但在真实对话语料中还是有一些例外.这些例外通常比上述基本语句组要复杂,本文称其为复杂语句组.经过分析,可以归纳出下面 5 种.

当对话中某一方对另一方所说的不能肯定、或者没有理解、或者没有听清甚至没有听见时,就会请求重复.这时,基本语句组就会插入两个或更多的语句,从而出现复杂语句组(CUG-A);当对话中某一方应手头信息的短缺时,会请求另一方稍候片刻,以便查找所需的信息.这时,在继续暂停了的交互之前,基本语句组中就会插入一对或两对额外的语句,从而出现复杂语句组(CUG-B);嵌入式复杂语句组除了上述两种情况以外,还有第 3 种情况,即前提式复杂语句组(CUG-C):当某一方在交互中缺少某项必需的信息时,会首先索求该信息.一旦对方提供了该信息,嵌入的语境组结束,外层的语句组继续.复杂语句组除了嵌入式外,还可能是线性的,有两种情况:一种是信息提供者渐进地提供信息(CUG-D);另一种是对话者跨语轮进行自我修正或者提供更精确的信息(CUG-E).

上述各类复杂语句组分别举例如下:

#### 1. CUG-A【选自 NLPR-TI-01】

C: 您是散客还是团体

S: 嗯(升调)

C: 散客还是团体

S: 我大有三四个人吧

C: 三个人或者四个人是吧

S: 对

#### 2. CUG-B【选自 NLPR-TI-28】

C: 它这个一般就是说这一次要多少天呢

S: 您稍等我给您看一下哈

C: 嗯

S: 6天

C: 噢是6天

#### 3. CUG-C【选自 NLPR-TI-27】

C: 那那个什么时候发团呢

S: 几个人哪

C: 几个人啊就那么几个人三四个人吧

S: 三四个人

C: 嗯

S: 嗯每星期五

C: 每一个礼就是每个礼拜五

S: 嗯对

#### 4. CUG-D【选自 NLPR-TI-44】

C: 好呢我想问一下就是这个费用

S: 费用如果您双卧的话两千三百八一个人

C: 一千两千三百八一个人

S: 嗯如果双飞是三千六百三

C: 三千六百三

S: 对

#### 5. CUG-E【选自 NLPR-TI-06】

C: 这一共是几天了一共

S: 七天

C: 七天

S: 哦九天了

C: 九天

根据导致基本语句组复杂的原因(是任务还是非任务或交际)和结构(是嵌入还是线性),上述 5 种复杂语句组可以进一步归类于表 4 中.

Table 4 Complex utterance groups

表 4 复杂语句组

	Task	Non-Task
Embedded	A, B	C
Linear	-	D, E

### 3 语句组标注

在文献[1]中,我们区分了语句的3个层次类型(utterance type,简称 UT),即:UT-1——语句的句子类型或语气(或扩展句子类型,XST);UT-2——语力或对话行为,指的是言语行为中施事行为的语力;UT-3——语句在语句组中的角色.这里关注的交互模式识别,也就是 UT-3 的分析.本节将对对话语料中的交互模式进行标注.下面,我们先给出语料概况和标注时所采用的准则,然后介绍具体标注及统计分析.

#### 3.1 实验语料

本文所使用的语料是实验室自行建立的关于旅游信息咨询的对话语料(NLPR-TI)<sup>[7]</sup>,由 58 个自然人、每人通过电话进行的对话组成,约 5.5 小时.工作中,将整个对话语料随机分成 3 部分:(1) 训练集,40 个对话;(2) 测试集,10 个对话;(3) 开发集,8 个对话.\*

#### 3.2 标注准则

语句组标注就是要依次标注对话中语句在语句组中的角色(即 UT-3),标注所用到的基本 UT-3 标记包括 4 个:I(起始)、R(回答)、F(反馈)、A(确认).另外,还用 X 来标记不能确定的,即不属于以上 4 种的.

在对每个语句标注 UT-3 时,采用以下准则:

- I—起始:如果说话者开始谈及新的主题(通常使用疑问句),标记该语句为 I.但起始的语句不必是疑问句,疑问句也不必起始一个语句组;
- R—回答:如果说话者给上一个起始的疑问句提供答复,则标记该语句为 R;
- F—反馈:如果说话者对对方所提供的信息表达某种程度的不确定(使用升调,或者显式地要求确认),或者有点自言自语的味道,通常部分或完全重复对方的语句,或者重述对方的意思,则标记该语句为 F;
- A—确认:如果说话者对对方的反馈进行响应,或者显式地确认,或者部分或完全地重复对方的语句;
- X—其他:当不能确定为上述 4 种的任意一种时,使用此标记.

#### 3.3 标注分析

根据上述准则,我们对 NLPR-TI 语料进行了标注.在从标注过的语料中将复杂语句组挑选出来之后,根据起始说话者角色和信息流向(人工确定),自动地将基本语句组分类.但由于某些信息流向的可能双向性(如时间,既可能为某条线路的预定发团时间,也可能为客户的打算出发时间.前者从服务方流向客户方,而后者却是从客户方流向服务方的),导致一些基本语句组的分类发生错误.因此,在统计不同类型和不同形状的语句组之前,先进行了人工检查并更正自动分类错误.最终的基本语句组和复杂语句组在语料中的分布分别在表 5 和表 6 中给出.下面进一步分析的是占据了语句组总数大多数的基本语句组,而对于只占少数(约占 7.8%)的复杂语句组,只给出其分布.

**Table 5** Distribution of basic utterance groups  
**表 5** 基本语句组分布

CISC		CICS		SISC		SICS	
Shape	No.	Shape	No.	Shape	No.	Shape	No.
IR	316	I	4	IF	72	IR	72
IRF	302	IFA	2	I	21	IRF	61
IRFA	113			IFA	16	IRFA	24
I	21					I	8
$\Sigma=752$		$\Sigma=6$		$\Sigma=109$		$\Sigma=165$	

**Table 6** Distribution of complex utterance groups  
**表 6** 复杂语句组分布

Type	A	B	C	D	E	$\Sigma$
No.	21	2	35	6	23	87

从表 5 可以看出:客户方主动、信息由服务方流向客户方的语句组(CISC)占绝大多数;而几乎没有客户方主动、信息由客户方流向服务方的语句组(CICS);服务方主动、服务方和客户方间的信息流动相差不多.进一步分

\* 这样一个分类对下面的语句组自动分析没有直接关系.但由于分析中使用了 XST 和语句主题(后者也采用了 XST 作为输入),而 XST 的识别采用了统计方法(NBC 和 HMM,参见文献[9]),其中必须区分语料的训练部分、测试部分和开发部分.

析,可以得出以下结论:

信息流向:分服务方到客户方( $S \Rightarrow C$ )和客户方到服务方( $C \Rightarrow S$ ).从上表可得: $S \Rightarrow C: C \Rightarrow S = (752+109):(6+165) = 861:171$ .因此,对话中信息主要由服务方流向客户方.

主动权:分客户方主动(CI)和服务方主动(SI).从上表可得: $CI:SI = (752+6):(109+165) = 758:274$ .因此,客户方比服务方更主动.

言语角色:分信息索求(information-seeking,简称 IS)(包括 CISC 和 SICS 语句组)和信息提供(information-providing,简称 IP)(包括 CICS 和 SISC 语句组).从表 5 可得: $IS:IP = (752+165):(6+109) = 917:115$ .因此,对话中更多的是信息索求.更重要的是,这两种交互模式在形状上也截然不同:IS 类的典型模式为 IR,可能的形式为  $I(R(F(A)))$ ;而 IP 类的典型模式为 IF,可能的形式为  $I(F(A))$ .回答(R)只在信息索求交互中出现,而在信息提供中出现.并且可以推断,如果有更多的 CICS 出现,其形状将会与 SISC 相似.

由以上定量分析所得出的结论与我们的常识是相吻合的.

## 4 交互模式识别

为了建立自动交互模式的分析算法,首先考察一下当前可用的可能有用的信息(即 UT-1 或 XST,UT-2 和语句主题)与语句组(或 UT-3)的关系(这里关注的是信息类对话主体部分的交互模式).

### 4.1 UT-3与UT-1和UT-2

总体上看,UT-1 是语句的浅层类型,大多有着明显的表层线索;UT-2 是语句的深层意图,表达具体的交际功能;而 UT-3 则是语句在语句组(局部话语结构)中的角色,表达更为普遍的交际功能.下面,我们来具体讨论 UT-3 与 UT-1 和 UT-2 之间的关系.

起始:在手头的面向信息查询的对话语料中,由于信息索求类交互占了大多数,因此,作为起始的 UT-3 相应地大部分由信息索求语力承担,并进一步地由特殊疑问句、一般疑问句和选择疑问句实现.另外,在少数信息提供类交互中,作为起始的 UT-3 由部分告知语力承担,以陈述句实现;

回应:作为回应的 UT-3 主要取决于处于起始位置的语力和 XST.如果是特殊疑问或选择疑问来索求信息,那么其回应就是告知,通常以(省略形式的)陈述句实现;但如果以一般疑问句来索求信息,那么其回应就是肯定或否定回答(绝大多数是肯定回答,因为提问者主观上通常是有偏的,即倾向于肯定回答),通常是短小的语句(功能性的肯定和否定短句).

反馈:此处的反馈与作为语力的反馈是有区别的.这里的反馈是指与语力中起始反馈相应的 UT-3.主要紧接着作为回应的语句(也有少部分跟在处于起始位置的告知之后),以表示理解、不理解或请求确认:当表示理解时,通常(部分或全部地)重复上一语句,或用自己的理解给出改述;当表示不理解时,会显式地请求重复;当请求确认时,通常采用附加问,或者零问(即采用疑问语调而字面是陈述句).

确认:确认是指与语力回应反馈相应的 UT-3.作为 UT-3 的确认,是对作为 UT-3 的反馈的反馈或回应.通常也是短小的功能性的肯定和否定短句.也有可能是上一语句的部分或全部的重复,但是极少出现.

### 4.2 UT-3与语句主题

通过语料分析发现,语句主题具有在语句组内保持连续或不变的特点(详见文献[8]),因而,当语句主题在两句之间发生改变时,这就意味着后一句很可能开始了一个新的语句组.但由于复杂语句组的存在,使得导致主题发生改变的语句不一定就是语句组的开始.为了说清这一问题,必须先区分语句主题是如何改变的.

由于采用属性-值对(attribute-value pair)来表示主题,因此当主题改变时,根据改变的是主题的属性还是值,可以分为强改变(主题属性改变)和弱改变(主题值改变).如果主题发生强改变,在排除是前提式复杂语句组中暂停的外部语句组的继续的情况下,就可以确定产生主题改变的这一语句的 UT-3 就是起始.一旦新的语句组开始,那么,就可以根据语句主题在语句组内保持不变以及上面所描述的 UT-3 与 XST 的关系,依次确定语句组内语句的 UT-3.如果主题强改变是前提式复杂语句组中暂停的外部语句组的继续所引起的,那么就必须续接外部

语句组进行处理.如果主题发生了弱改变,这就表明出现了线性的复杂语句组(导致主题发生改变的语句的 UT-3 为 R).

### 4.3 语句组分析算法

根据上述 UT-3 与其他语句类型以及与语句主题的关系,建立以下的语句组分析算法(因为对话开始相和结束相的交互模式较简单,这里从略,而把重点放在介绍主体相的语句组分析上):

首先,根据语句主题是否发生改变来判断是否出现新的语句组(主题为空的语句不会产生改变).根据主题变化的不同情况,依照第 4.2 节来决定如何处理当前语句.如果主题未发生改变,那就根据上一句的 UT-3 和 UT-1,以及当前语句的 UT-1,依照第 4.1 节中所给出的关系来确定当前语句的 UT-3.具体算法描述如下:

算法. 语句组分析算法.

Input: One dialogue, with utterances tagged with XST and topic.

Output: UT-3 for each utterance.

//Main

for  $i$  from 1 to  $n$  //  $n$  = no. of utterances in the dialogue

$UT-3[i] = X$ ;

if  $XST[i]$  is in  $XST-O$  //  $XST-O$ : XST set in the opening phase

$GroupingInOpening()$ ; //Details omitted

else if  $XST[i] \in XST-C$  //  $XST-C$ : XST set in the closing phase

$GroupingInClosing()$ ; // Details omitted

else

$GroupingInBody()$ ;

//Subfunction

$GroupingInBody()$

if  $UTopic[i]$ .Attribute changes

if  $UTopic[i] \neq GTopicLast$  // Not continue the utterance topic before last

$UT-3[i] = I$ ;

else

Continues to process the unfinished group before last

else if  $UTopic[i]$ .Value changes // when there appears linear complex group

$UT-3[i] = R$ ;

else //identify topic according to XSTs of last and current utterances

switch ( $UT-3[i-1]$ )

case I:

if ( $(XST[i-1]$  is WHQ) and ( $XST[i]$  is DCL))

or ( $(XST[i-1]$  is YNQ) and ( $XST[i]$  is PST or NEG))

$UT-3[i] = R$ ;

if ( $(XST[i-1]$  is DCL) and ( $XST[i]$  is DCL or TGQ))

$UT-3[i] = F$ ;

case R:

if ( $(XST[i-1]$  is DCL) and ( $XST[i]$  is DCL or TGQ))

$UT-3[i] = F$ ;

case F:

if ( $(XST[i-1]$  is DCL or TGQ) and ( $XST[i]$  is PST, NEG or DCL))

$UT-3[i] = A$ ;

case A:

case X:

$UT-3[i] = X$ .

#### 4.4 实验与结果

根据上述算法,采用不同来源的 XST 以及相应的语句主题\*\*,分别对语料 I 和语料 II 进行了分析,结果在表 7 中给出.从表中可以看出,语句组的分析结果受多方面因素的影响:首先是 XST 和语句主题标注的结果最好,启发式的和基于 NBC 的其次(两者区别甚小),而基于 HMM 的最差;从各个 UT-3 的识别结果来看, $I > R > F > A$ .这可以部分地归结于误差扩散,因为在语句组中,排在后面的 UT-3 的识别要受前面 UT-3 识别结果的影响.

Table 7 Result of utterance group analysis (correct rate (%))

表 7 语句组分析结果(正确率(%))

XST source	Utterance topic	Corpus	<i>I</i> (%)	<i>R</i> (%)	<i>F</i> (%)	<i>A</i> (%)	Overall (%)
Annotated	Annotated	I	89.0	88.5	83.5	72.8	84.2
		II	86.4	83.1	77.7	65.8	76.8
Heuristic	Algorithmic/Heuristic	I	86.0	83.3	72.3	57.1	77.7
		II	79.9	72.1	67.9	52.6	68.0
NBC	Algorithmic/NBC	I	87.4	84.5	74.3	59.9	79.2
		II	82.2	74.4	70.5	57.9	70.5
HMM	Algorithmic/HMM	I	69.5	62.5	53.6	51.6	60.8
		II	65.1	59.9	51.8	47.4	55.4

## 5 小结与讨论

本文首先建立了语句组分类体系,然后对话语料进行了标注分析.在进一步分析语句组与其他语句类型以及语句主题之间的关系之后,建立了语句组分析算法,并在语料中进行了实验.

与语句组标注相关的工作不是很多,其中最相关的是 HCRC Maptask 语料中的语弈(conversational game,与语句组类似)标注分析<sup>[9]</sup>.但这些工作没有对语句组作细致的分类,而是认为每一种起始语力开始一个语弈或语句组,并且也没有进行自动分析.DRI<sup>\*\*\*</sup>也有关于话语结构标注的研究,称作共识单元(common ground unit).但关于 CGU 的工作似乎不是非常令人满意,也没见到进一步的深入研究.

关于交互模式的自动分析,相关工作只有 Wright-Hastie 等人基于文献[9]的标注体系所进行的工作<sup>[10]</sup>.他们通过分析语移在语弈中的位置来辅助识别语移,但将交互模式的分析等同于分析语句在语弈中位置,过分地简化了交互模式分析.交互模式的分析就是要分析语句在语句组中的角色.另外,关于交互模式的计算分析,还有采用基于规划的方法,如文献[11].但是,由于基于规划方法有着知识获取、计算复杂性等一系列目前难以逾越的问题,很难成为当前的选择.

进一步的工作包括:如何将上述面向信息查询类对话的语句组分类体系,扩展到其他更为复杂的对话;如何采用机器学习的方法来实现语句组的自动识别;如何采用统一的框架来实现 3 种语句类型的识别,以及收集更多的语料来进行经验分析.更具挑战性的工作是,如何利用本文关于局部话语结构的工作来研究对话高层或全局话语结构.

\*\* XST 的来源有 4 种,即人工标注的和 3 种自动分析的(分别采用启发式、朴素 Bayes 分类器或 NBC 以及隐 Markov 模型或 HMM).语句主题的来源则包括标注的和 3 种对应于 XST 自动识别的算法分析的(因为语句主题的自动算法识别部分地依赖于 XST).关于 XST 和语句主题的分析及其结果细节,请参见文献[9].这里给出定性的结果:XST 识别的正确率,NBC>启发式>HMM;语句主题分析结果,XST 标注>{NBC,启发式}>HMM.NBC 与启发式相当.

\*\*\* DRI(discourse resource initiative),一个国际性话语资源行动组织,曾经于 1998 年在日本举行第 3 次研讨会.研讨会的主题之一就是话语结构标注.有关的研讨会报告参见 Core, Mark, Masato Ishizaki, Johanna Moore, Christine Nakatani, Nobert Reithinger, David Traum, and Syun Tutiya. 1999. The report of the third workshop of the discourse resource initiative. Technical Report, No.3 CC-TR-99-1, Chiba Univeristy and Kazusa Academia Hall. Chiba Corpus Project ([http://www.stanford.edu/~jurafsky/ Coreetal99.pdf](http://www.stanford.edu/~jurafsky/Coreetal99.pdf)),标注手册参见 Nakatani, Christine H., David R. Traum. 1999. Coding discourse structure in dialogue (Version 1.0). Technical Report, UMIACS-TR-99-03, University of Maryland (<http://hdl.handle.net/1903/991>).



致谢 非常感谢审稿老师所提出的宝贵、中肯、建设性的意见.同时,还要感谢胡晟和胡日勒所提供的帮助.

### References:

- [1] Xu WQ, Xu B, Huang TY, Xia HR. Bridging the gap between dialogue management and dialogue models. In: Jokinen K, McRoy S, eds. Proc. of the 3rd SIGdial Workshop on Discourse and Dialogue. Philadelphia: Association for Computational Linguistics, 2002. 201-210.
- [2] Sinclair JM, Coulthard M. Towards an Analysis of Discourse: The English Used by Teachers and Pupils. Oxford University Press, 1975.
- [3] Sinclair JM, Coulthard M. Towards an analysis of discourse. In: Coulthard M, ed. Advances in Spoken Discourse Analysis. London, New York: Routledge, 1992. 1-34.
- [4] Coulthard M. Forensic discourse analysis. In: Coulthard M, ed. Advances in Spoken Discourse Analysis. London, New York: Routledge, 1992. 242-258.
- [5] Francis G, Hunston S. Analyzing everyday conversation. In: Coulthard M, ed. Advances in Spoken Discourse Analysis. London, New York: Routledge, 1992. 183-196.
- [6] Halliday MAK. An Introduction to Functional Grammar. 2nd ed, London: Edward Arnold, 1994.
- [7] Xu B, Huang TY, Zhang X, Huang C. A Chinese spoken dialogue database and its application. In: Proc. of the 2nd Int'l Workshop on East-Asia Language Resources and Evaluation. Taipei, 1999.
- [8] Xu WQ, Xu B, Huang TY. Utterance Topic Identification in Spoken Dialogues. Chinese Information Processing, 2005, 19(4): 89-96.
- [9] Carletta J, Isard A, Isard S, Kowtko JC, Doherty-Sneddon G, Anderson AH. The reliability of a dialogue structure coding scheme. Computational Linguistics, 1997,23(1):13-31.
- [10] Hastie WH, Poesio M, Isard S. Automatically predicting dialogue structure using prosodic features. Speech Communication, 2002, 36(1):63-79.
- [11] Lochbaum KE. A collaborative planning model of intentional structure. Computational Linguistics, 1998,24(4):525-572.

### 附中文参考文献:

- [8] 徐为群,徐波,黄泰翼.口语对话中的语句主题分析.中文信息学报,2005,19(4):89-96.



徐为群(1973 - ),男,江苏泰兴人,博士,主要研究领域为对话和话语的计算.



黄泰翼(1936 - ),男,研究员,博士生导师,主要研究领域为语音识别,口语处理.



徐波(1966 - ),男,研究员,博士生导师,CCF高级会员,主要研究领域为语音识别,人机语音通信.