

一种基于语境的对话模型^{*}

刘红阁

(中国科学院数学与系统科学研究院数学研究所 北京100080)

A Context-Based Talker

LIU Hong-Ge

(Institute of Mathematics, Academy of Mathematics and System Sciences, The Chinese Academy of Sciences, Beijing 100080)

Abstract In order to make computer understand human languages much better and achieve certain tasks according to human intentions, it is required to carefully model the elementary concepts and the underlying theoretical principles for driving a conversation generating and developing. Under the given linguistic circumstances, to some extent, it might be hopeful to build up a computational model which can predict and control the sustaining development of conversational activity. This paper is going to make a formalized primary discussion on contextual features which influence conversational activity, design an algorithm processing ellipses, and build a Context-Based Talker (CBT). The empirical analysis shows, to some extent, CBT can control and guarantee the coherence and natural development of conversational activity.

Keywords Context-based talker, Context, CS-Net, Belief, Ontology

1. 引言

从1998年起,我们承担了一项国家自然科学基金重点课题:常识性知识的实用性研究。我们致力于建立一个大规模的常识知识库 Pangu,使得我们的系统接受用户关于常识问题的自然语言形式的询问,比如“牛有几条腿?”,并且产生基于常识知识库搜索和问题求解的回答。

在常识知识库搜索和问题求解的过程中,我们发现要想保证人机对话连贯自然的良性发展,各类语境信息尤其认知语境起到了至关重要的作用,而已有的对话模型很少考虑语境在对话中扮演的角色。认知科学告诉我们,人们认识事物时总是遵循一个基本规律,即以相关的认知环境为出发点来认识新事物,增加新信息,并进而改变已有的认知环境,为下一次认知活动做准备。人和机器之间的对话,也是一个认知的过程,只有在机器掌握了对方向和自身的认知语境情况下,才能对人的问题做出近似人的解释和响应。Sperber 和 Wilson 也认为,对话语理解起重要作用的是有听话者的一系列假设所构成的认知语境。认知语境是在话语理解过程中不断选择的结果,而不是在理解之前预先确定的。对话语的最佳解释就是花费最小的努力,从认知语境中选择最相关的假设,来获得最大的语境效果。

让人感觉和我们的机器交流就像和一个正常的人交流一样而没有障碍,虽然不是我们所有的目的,但是却是研究机器话语智能的一个主要手段。CBT 就是结合 Pangu 自身特点和需要而设计的,它从分析各类语境信息及其在对话过程中的动态改变着手做出话语选择。

2. 相关工作比较

首先来介绍一下语境的定义, Jef Verschueren 把语境划

分为交际语境和语言语境,并提出了语境的动态生成观。所谓交际语境包括语言使用者、心理世界、社交世界和物理世界。以上的因素都会不同程度地影响到对话者在使用和理解语言时做出的语言选择。Jef 对语境的划分是系统和全面的,但是建立一个综合这些因素的对话模型是非常复杂困难的,考虑到 Pangu 的特点,我们定义了在实际对话模型中最直接的影响因素。我们把语境划分成认知语境、语言语境和物理世界语境,认知语境主要涉及对话者的信念、意图、愿望和情感等心理因素。Power 1973 年提出的 Perceptual Control Theory (PCT)认为,所有的行为都是受认知状态的控制,先有认知再有行为,所有行为是靠执行行为的人的认知状态所决定的。语言语境,即我们通常所说的上下文,主要包括指代和省略两大方面。物理世界语境主要包括天气,时间和空间的指示关系:就时间而言,它包括事件时间,说话时间和指示时间;空间指示即地点指示关系可分成两种:绝对空间关系和参照指示对象的相对空间关系。

Robbert-Jan Beun^[2]设计了以认知为核心驱动的双机对话模型,他认为双方是为了完成一定任务而进行交流的,既然为了解决问题,就应该遵守合作和一致性等原则。他假设双方是完全合作的,A 不会去问 B 自己不相信的问题,对话者从来不误导对方,也从不提供缺少证据的断言,因此不去检查交流中存在的与事实违背的不一致。这其实是对现实对话的简化,也是设计者的美好愿望,但现实中却不是这样。交际中双方必须合作吗?说话人一定会做到表达真实、充分、相关和清晰吗?Grice 的合作原则及其各准则具有普遍性吗?Ayse Pinnar Saygin^[3]也试图通过分析 Grice 的合作原则及各会话准则的违反之间的关系来获取深层会话含意,但事实上,这种方法缺乏足够的理论依据,缺少严密性。Beun 的模型中仅分析了影响话语交流的认知语境,但是他的思想还是对我们有

^{*} 本文工作受到了中国科学院脑智科学基金以及973计划(编号为:2002CB312004)的资助。刘红阁 博士生,研究方向为自然语言理解和语用理论。

很大的启发,我们的模型中也采用了他定义的部分基本表示符号。我们是想处理真正符合现实的对话,使得我们的机器模拟某种意义上真正的人,所以意愿建立一个比较严谨和完备的对话模型。

3. 支持 CBT 的知识库结构

3.1 Pangu 的结构

Pangu 中有两种知识单元:agent 和 ontology,agent 用来表示知识之间纵向关系(继承),ontology 用来表示知识之间复杂的横向联系。CBS agent 是我们提出的一种新型的常识知识库的 agent,C,B 和 S 分别代表 Capability, Belief 和 Strategy。为了模仿人的行为,我们的 Belief 和 Strategy 分别表示静态和动态的知识,其实,Belief 类似一个人的头脑中存储的认识世界的知识,而 Strategy 是由一组规则组成,他模仿的是人利用知识(Belief)解决问题的能力。因为 Ontology 的结构和本文无关,这里不再介绍,详见文[1]。

3.2 CS-Net

CS-Net 是我们设计的用来表示 agent 的 Belief 和 Capabilities 的一种语义网络,每个 CS-Net 由一组关联的语义三元组构成。一般来说,每一个语义三元组由三个结点组成,其中第一和第三结点是概念结点,第二结点是第一结点和第三结点之间的关系结点,概念结点和关系结点又可以用别的语义三元组来修饰。例如:“猴子生活在什么地方”用 CS-Net 应该表示为:

(猴子,生活:L01,*);

L01:(position(生活),is,!x);

其中 position 是我们定义的关键字,L01引导的是修饰三元组。

3.3 ECWA

Reiter 的所谓封闭世界假设(Closed World Assumption,简称 CWA)的说法是:任何不能被知识库证明的命题都认为是错误的命题。比如如果知识库中没有说明“牛有两个尾巴”,那么就认为命题“牛有两个尾巴”是错误的。传统的 CWA 忽略了“不知道”出现的可能性,仅给出“是”和“否”两个答案。殊不知,“否”和“不知道”应该严格区别,而且把 CWA 直接应用到话语理解中是非常危险的。为此,我们提出了扩展的封闭世界假设(ECWA),思想是:如果人和牛是动物仅有的两个儿子结点,那么所有人和牛的共同属性都应该从动物这个 agent 获得。如果在动物里没有提到“跳舞”这个概念,但在人中提到“人会跳舞”,那么我们就可以肯定推断出“牛不会跳舞”。另外如果在三个结点中都没有提到“跳舞”,那么答案就是“不知道”(因为知识库不可能是完备的,回答“否”是很危险的)。通过合理组织知识库的结构,我们就把二值逻辑转换成了三值逻辑。

4. CBT 的设计

大致思想是:① 定义一组描述对话者在话语行为中的语境特征量,比如信念、意图、愿望、情感(通过多次正式的人机对话测试,我们发现:在人机对话中,为使机器模拟成某种意义上真正的人,让机器的响应具有一定的情绪和情感,是机器通过测试的一个关键。所以在 CBT 中,我们对情绪和情感给予了较为严格的划分,给出了测评的量化指标。并且考虑了情绪和情感的连续性变化,尽量保证机器和人不会出现情绪和情感的跳跃)等。② 结合历史语境和新产生的话语行为等信

息,制定一套较完备的语境更新规则,以完成在对话过程中动态更新对话者语境特征信息。③把对话策略(通常说的对话技巧)表示成规则,即根据三类语境信息,选择执行最符合对话者意图的行为。

CBT 是建立在人机对话基础上的,人的认知状态是隐藏在人的内心的,而且机器无法控制人所做出的语言选择。尽管如此,我们还是结合对话历史分析人的认知状态,以便机器作出具有满足人需求的最大语效的话语选择。

CBT 有黑板和规则库两大部分构成,在黑板模型中存放三类语境信息,规则库中存放语境更新规则和行为执行规则。假定所有的行为都是受三类语境因素的控制,在每一轮对话开始的时候,人需要把物理世界语境信息输入到黑板模型中,语言语境为空,机器和人的信念也就是各自所有的知识和常识,情感为正常态。此时需要首先由人主动开始对话,向机器提出问题,这时人就有了一个愿望或者意图,从此构成了认知语境的不平衡,对话开始伴随着解决问题的开始,直到认知语境中的所有状态达到了新的平衡,对话结束。在每一轮对话中,只有物理世界语境会保持恒定不变,其它的语境信息都可能在对话过程中动态改变,各类语境信息的动态改变过程利用堆栈来保存。

CBT 的详细流程如图1。

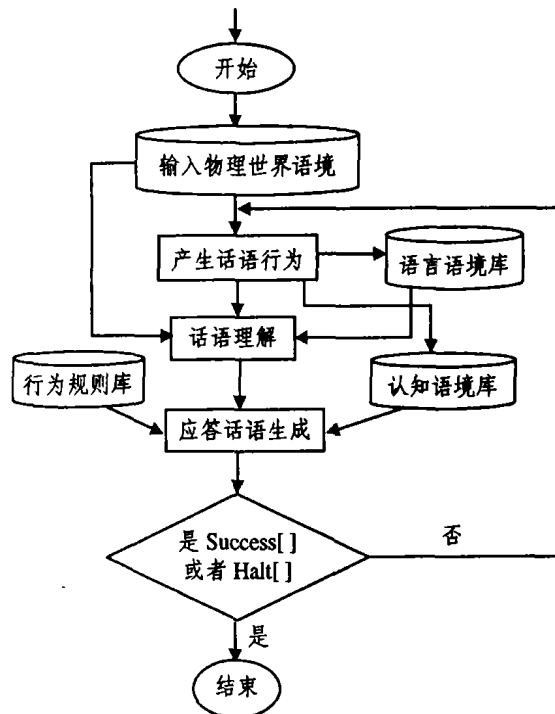


图1 CBT 的处理流程

4.1 物理世界语境在 CBT 中的应用

物理世界语境主要涉及时间、空间和天气这些内容,使用它们主要用来进行时空推理,另外有了这些信息,也能有助于话语的理解。

4.2 语言语境在 CBT 中的应用

语言语境是在交互过程中动态创建的,为此我们需要保留整轮谈话的历史记录在一个堆栈里。这里对人的语言有一点要求,就是尽量使用完整或近乎完整的语言,机器使用的语言不必担心,一定会是完整的。通过对人生成的话语的指代和省略分析就完成了语言语境在对话模型中的应用。指代问题是一个很复杂并且普遍存在的问题,也是话语分析中不可忽视的问题,当一句话中出现了一个“不明身份”的词和词组时,它往往是某人或某些已知身份的词或词组的代表。准确确定

不明身份的指代对象,是进行正确话语理解的前提。进行指代分析可以从三个层次来处理:①语法、语义分析;②句内常识分析;③句群常识分析,具体可参照文[4]。

在进行省略分析的时候,除了分析省略单句内的语法结构,句间结构和一致性还须从对话历史记录中搜寻省略的关键语法信息,根据所能获取的关键信息去常识知识库中搜索满足问题的解。利用语言语境进行省略分析的方法如下:

4.2.1 对话历史记录 H 的描述 H 是一个由对话历史记录构成的双向链表,每一个结点描述了每一次对话内容,以及对话序列的层次结构。每一个结点的后趋指针指向下一个对话结点,如果该结点是对前面某结点的解释,则前趋指针指向解释的结点,如果该结点不对任何其它结点的解释,则前趋指针指向自身。

例如: A_1 :老虎有几条腿?

B_1 :老虎跑得快吗?

A_2 :老虎跑得非常快。

B_2 :老虎有4条腿(在常识知识库中隐含有这样的常识:跑得快的动物有4条腿)。

表示成层次: $A_1(B_1 A_2)B_2$

历史记录 H 如图2:

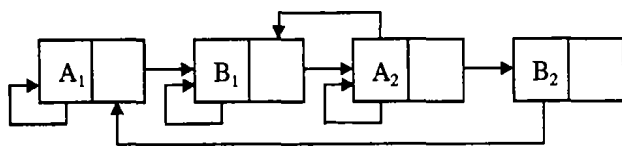


图2

上图中, A_2 是对 B_1 的解释, B_2 是对 A_1 的解释。

每个结点的对话内容用一个六元组表示:

(Counter, Identity, Content, Emotion, Feeling, AnsAgents, ListAgent)

其中 Counter 是回合标记,初始为0,从1开始记数;Identity 表示说话者的身份,取值为“人”和“机器”;Content 是话语的 CS-Net 表示;Emotion 表示说话者当前的情绪值;Feelings 表示说话者当前的情感值;AnsAgents 表示当前的应答是由哪些 Agents 给出的,只对机器有效;ListAgent 表示该句话语的 Agent 列表,排在最前面的是主 Agent。

在下面的算法中, $(i, \text{人}, \text{Content}, \text{Emotion}, \text{Feeling}, \text{AnsAgent}, \text{ListAgent})$ 简记为 A_i , $(i, \text{机器}, \text{Content}, \text{Emotion}, \text{Feeling}, \text{AnsAgent}, \text{ListAgent})$ 简记为 B_i 。

对话回合: (A_i, B_i)

n 个回合的对话历史: $A_1 B_1 A_2 B_2 \dots A_n B_n$ 称为 n 个回合的完全对话历史, $A_1 B_1 A_2 B_2 \dots A_n$ 称为 n 个回合的不完全对话历史。

主关键字:Content 中提问变量的范畴,一般是主三元组的头中的内容。例如:几条?对应的三元组表示: $(\text{number}(*), \text{is}, !x)$;主关键字为 number。

非主关键字:Content 中的非主关键字。

4.2.2 基本假设

① 由机器作出的回答都是完整的句子

② A_{n+1} 是对 n 个回合完全对话历史的深入而不是话题的转移

③ 省略的信息可以在本轮上文中的完整表达式中找到

4.2.3 算法

(1) if A_{n+1} Content 是信息完整的 return(交由其它策略

处理)

(2) 获取 A_{n+1} Content 的主关键字 PrimaryKey

(3) 获取 A_{n+1} Content 的非主关键字 InfoList

(4) if B_n ListAgent = “” 根据 H 的层次结构,获取前 $n-1$ 个不完全对话历史中最外层解释 B_k

(4.1) Agents = B_k ListAgent

(4.2) $i = k$

else

$i = n$

(5) if $i = 0$ return (没有可用的上文信息)

(6) if PrimaryKey = “*” 取出所有 B_i ListAgent 的 belief 构成集合 T

else 取出所有 B_i ListAgent 中所有包含 PrimaryKey 的 belief 构成集合 T

(7) if $T = \emptyset$ return (无相关信息)

(8) 获取 A_i Content 中的非主关键字信息 Vice

(9) 合并 Vice, InfoList, A_{n+1} ListAgent 成集合 K

(10) if PrimaryKey = “*” 对 T 中的每条 belief, 计算其对集合 K 的包含度, 取包含度最大且包含度大于 r (r 为阈值, 定为 0.68) 的 belief 作为应答。

else 对 T 中的每条 belief, 计算其对集合 K 的包含度, 取包含度最大且包含度大于 r (r 为阈值, 定为 0.70) 的 belief 作为应答

(11) 成功返回

(1) 中对 A_{n+1} Content 的完整性判断是非常困难的,因为省略形式是多种多样而又无处不在的。例如“毛多吗?”从句法上来看是完整的主谓句,但分析上文语境可能准确信息应为:“牛的毛多吗?”在这种情况下,我们只有采用折中的办法,不把这类问题作为省略来处理。

4.3 认知语境在 CBT 中的应用

在下面的具体应用中,我们假设机器和人只有在表达陈述的观点时,才会引起情绪和情感的连续性变化。按照 CBT 的需要,我们把情绪分成5种表现状态:愤怒、生气、正常、高兴和兴奋,分别取值为:-2, -1, 0, 1, 2;把情感也分成5种表现状态:绝望、哀伤、正常、欣慰和幸福,分别取值为:-2, -1, 0, 1, 2。当通过分析得知人的情绪值是负值时,机器向对话人表示适当的歉意,如果是正值,机器向对话人表示适当的谢意(我们认为情绪值的降低是因为说了伤害或侮辱对方的话而引起对方情绪值的反向变化,而情绪值的升高是因为说了赞赏或祝福对方的话而引起对方情绪值的正向变化);如果分析得知人当前情感值为负值时,机器应向对话人表示适当的安慰或同情,如果是正值,机器向对话人表示适当的祝贺或祝福(我们认为情感值的降低是因为自身发生了不愉快或者伤心的事而引起的,而情感值的升高是因为自身有了高兴或开心的事而引起的)。我们对情绪和情感的划分不是严格的,但这样的划分在进行对话分析的时候基本足够,也减轻了处理的复杂度。

4.3.1 基本符号和函数

p 代表简单命题, $p \rightarrow q$ 是复合命题,其中 p 是命题的前提, q 是命题的结论。不允许出现这样的命题: $p \rightarrow p$, 因为它不能够带来新的知识,只能造成重复。

$B_x p$: x 相信 p

$B_x \rightarrow B_y p$: x 相信 y 不相信 p

$SB_x p$: x 假设相信 P

$D_x p$: x 希望 p

$I_x(p_1, p_2, \dots, p_n)$: x 有知道 $p_1, p_2, \dots, p_n$ 的意图, p_n 是最近 x 想知道的意图。

E_x : x 当前的情绪值

F_x : x 当前的情感值

$Is(r, s)$: r 就是 s , 它们是等价关系。比如: Is (龙眼, 桂圆)

$Is_a_kind(r, s)$: r 是 s 的一种。比如: Is_a_kind (东北虎, 虎)

$Check(p)$: 进行 p 的常识检查, 如果冲突, 则返回 -1, 如果不冲突返回 1, 检查不出结果返回 0。

$State(p)$: 判断 p 是否是陈述句, 如果是返回 true, 否则返回 false。

$Fathers(p)$: 返回 P 的祖先命题集。比如 p : 牛有腿吗? $Fathers(p) = \{p_1, p_2, p_3, p_4\}$, p_1 : 家畜有腿吗? p_2 : 陆生哺乳动物有腿吗? p_3 : 哺乳动物有腿吗? p_4 : 动物有腿吗?

$Brothers(p)$: 返回 P 的兄弟命题集。比如 p : 黄牛有腿吗? $Brothers(p) = \{p_1, p_2\}$, p_1 : 奶牛有腿吗? p_2 : 水牛有腿吗?

$Contact(x, p) = \{q \mid B_x(q \rightarrow p)\}$: 如果 $q \rightarrow p$ 并且 $r \rightarrow p$, 则 $Contact(x, p) = \{q, r\}$

$Del(B_x I_x(p_n))$: 如果 $B_x I_x(p_1, p_2, \dots, p_n)$, 则经过此函数, 结果为: $B_x I_x(p_1, p_2, \dots, p_{n-1})$

$Add(B_x I_x(p_n))$: 同上一个函数正好相反

$Feeling(x)$: 获取 x 当前的情绪值

$Emotion(x)$: 获取 x 当前的情感值

4.3.2 基本推理规则

I1: $B_x p \& B_x(p \rightarrow q) \rightarrow B_x p$

I2: $B_x p \& B_x q \leftrightarrow B_x(p \& q)$

I3: $B_x(p \rightarrow q) \& B_x(r \rightarrow q) \leftrightarrow B_x(p \vee r \rightarrow q)$

4.3.3 基本话语行为

$State[x, p][E_x][F_x]$: x 使用陈述句表达 p , 其中 E_x 和 F_x 分别代表 x 当前的情绪值和情感值

$Question[x, p]$: x 使用疑问句表达 p

$Ignorance[x, p]$: x 没有 p 的见解, 放弃

$Supposition[x, p]$: x 假设 p 成立

$Learning[x, p]$: x 学习 P

$FSupport[x, y]$: x 响应 y 的不稳定情感, 包括安慰、同情, 支持, 鼓励和祝福等。

$Success[]$: 一轮对话成功结束

$Halt[]$: 一轮对话因为意外而被强制中段

4.3.4 对话策略

S0: $D_x p \& B_x p \& B_x I_x(\emptyset) \rightarrow Question[x, p]$: 每一轮都由此状态开始, 人首先提出问题。

S1: $F_y \neq 0 \rightarrow FSupport[x, y]$

S2: $B_x I_x(\dots, p) \& (B_x p \parallel SB_x p) \rightarrow State[x, p][E_x][F_x]$

S3: $B_x I_x(\dots, p) \& \neg B_x p \& \neg SB_x p \& q \in Contact(x, P) \& \neg B_x \rightarrow B_x q \rightarrow Question[x, q]$

如果当前 p 是 y 想要知道的命题, 但是 x 却不能够直接给出答案, 不过 x 知道和 P 关联的一些命题, 如果确定了其中的一个, 就能够确定 P , 所以 x 主动询问 y 是否能确定这些关联命题。一般 x 会拿出和 p 第一个关联的命题询问, 如果 y 给不出答案, x 再依次取出后面的关联命题询问。

S4: $B_x I_x(\dots, p) \& \neg B_x p \& \neg SB_x p \& q \in Contact(x, P) \& B_x \rightarrow B_x q \& E_x \geq -1 \rightarrow Supposition[x, q]$

如果 x 不能确定 y 想知道的 p , 并且 x 知道 y 也不能确定和 p 相关联的命题, 若此时 x 的情绪不是很坏的话, x 花费

代价假设 p 的第一关联 q 成立。

S5: $B_x I_x(\dots, p) \& \neg B_x p \& \neg SB_x p \& Contact(x, P) = \emptyset \& E_x \geq 0 \rightarrow Question(x, Fathers(P))$

如果 x 不能确定 y 想知道的 p , 并且没有和 p 关联的命题, x 此刻的情绪在正常以上, 问 P 的祖先命题, 因为属性有继承功能。

S6: $B_x I_x(\dots, p) \& \neg B_x p \& \neg SB_x p \& Contact(x, p) = \emptyset \& \neg B_x(Fathers(p)) \& E_x \geq 0 \rightarrow Question(x, Brothers(p))$

类似 S5, 由于查找兄弟节点的代价要比查找祖先节点的代价要高, 所以优先级必 S5 要低。

S7: $B_x I_x(\dots, p) \& \neg B_x p \& \neg SB_x p \& \neg B_x(Fathers(P)) \& \neg B_x(Brothers(P)) \& State(P) \& Check(x, p) = 0 \& E_x \geq -1 \rightarrow Learning(y, p)$

如果从 p 的祖先和兄弟结点都无法得出答案, p 是陈述命题, 且 x 对 p 的常识检查也无法得出答案, 而且情绪值不太坏, x 就把 p 学习成自己的 belief。

S8: $B_x I_x(\dots, p) \& \neg B_x p \& \neg SB_x p \& Contact(x, P) = \emptyset \& \neg B_x(Fathers(P)) \& \neg B_x(Brothers(P)) \& \neg State(P) \& E_x \geq -1 \rightarrow Ignorance(x, q)$

如果 p 的关联命题为空, 且从祖先和兄弟结点都无法给出答案, 而且情绪值不太坏, p 是 y 说出的非陈述句, x 忽略 p , 选择放弃。

S9: $B_x I_x(\dots, p) \& \neg B_x p \& \neg SB_x p \& E_x \leq -2 \rightarrow Halt[x, p]$

如果 x 不知道 p , 且 x 因为 y 心情极坏, 则 x 强制中断和 y 的对话。

S10: $B_x I_x(\emptyset) \& \neg (D_x p \& \neg B_x p) \& \neg (D_x p \& \neg SB_x p) \rightarrow Success[]$: x 和 y 的认知状态重新达到了新的平衡, 成功退出。

以上策略的优先级是逐渐降低的。

4.3.5 认知语境的更新规则

U1: $State[x, p][E_x] \& B_x I_x(p) \rightarrow B_y p \& Del(B_x \rightarrow B_y p) \& Del(B_x I_x(p)) \& (E_y = E_x + E_x)$

U2: $State[x, p][E_x] \& B_x \rightarrow I_y(p) \& Check(y, p) \rightarrow B_y(p) \& (E_y = E_x + E_x)$

如果 x 陈述了 p , 但 y 没有想知道 p 的意图, 此时进行常识检查, 如果没有冲突, y 就相信 p 是正确的。

U3: $Question[x, p] \rightarrow B_y I_x(p)$

U4: $Supposition[x, p] \rightarrow SB_x(p)$

U5: $Ignorance[x, p] \rightarrow B_y \neg B_x p \& Del(B_x I_x(p)) \& Del(D_x p)$

U6: $Learning[x, p] \rightarrow B_x p$

U7: $FSupport[x, y] \rightarrow |F_y| = |F_x| + 1$

U8: $SB_x p \& Check[x, p] = -1 \rightarrow Del(SB_x p) \& \neg B_x p$

如果原有 x 假设 P 成立, 但是由于知识的更新, x 得到 P 是不成立的, 此时应删除假设。

U9: $Success[] \rightarrow \odot$ 成功退出不会对认知语境带来改变。

U10: $Halt[] \rightarrow \odot$ 强行中断对话不会对认知语境带来改变。

4.3.6 实例分析

接下来让我们来看一个人机对话的实例, 它以上面介绍的认知语境的构成元素和规则为基础, 首先我们提供一个实例的抽象解释, 而后会给出具体的自然语言对话实例。

表1中, M (Machine)代表机器, H (Human)代表人。

假设 M 已有知识是: $q \in Brothters(p)$, $q \rightarrow \neg p$

表1所示的该轮对话中, M 和 H 试图去解决问题: p 是否

(下转第72页)

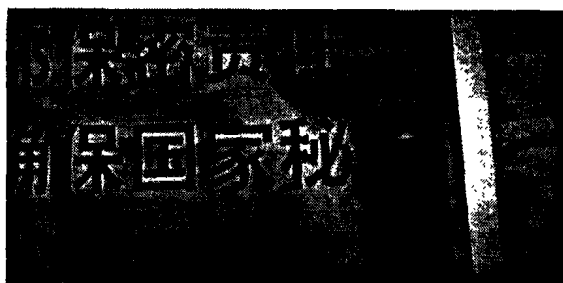


图2 全景图像

结束语 本文提出的仿射变换模型,能够在图像偏差较大的情况下实现图像的自动整合。如果只采用层次模型直接对齐图像,则两图像的偏差仅能在几个像素范围内。在实际实现中,工作量太大,算法无实际意义。该方法速度快,工作量

小,优化参数少,可避免陷入局部极小的问题。在实际实现过程中,一般图像偏差控制在几十个像素内,即可快速高质量地建立全景图像。

参考文献

- 1 Zelinik-Manor. 2, Irani. M. Multi-Trame Estimate of Planar Motion. IEEE Transaction on PAMI, 2000, 22(10)
- 2 Szeliski R. Video Mosaics for Virtual Environment. IEEE CGQA, March. 1996
- 3 Sawhney S H. True Multi-Trame Alignments and its Application to Mosaicing and Lens Distortion Connection. IEEE Transacton on PAMI, 1999, 21(3)
- 4 Shum H, Szeliski R. System and Experiments Paper : Construction of Panoramic Image Mosaics with Global and Local Alignment. International Journal of Computer Vision, 2000, 36 (2): 101~130
- 5 马颂德. 计算机视觉. 科学出版社, 1997
- 6 Anadan K R P. Representation of scene from collection of images. IEEE CVPR 1995

(上接第35页)

为真,初始状态有 H 提出问题,终止的时候问题得到了解决, H 和 M 都相信 p 是假的。

首先 H 询问 M 是否 p 是真的, M 接到问题后没有直接的 belief 给出答案。所以只有先去询问自身的 p 祖先命题集,但从祖先也无法得出答案,只好转向兄弟命题集去求解问题。由于 q 是 p 的兄弟命题集的一个元素, q 是成立的,根据封闭世界假设,得出 p 不成立,此时 H 提出的问题得到了解决,成功退出。

为了让这个例子更加具体实用,来看下面的实例:

机器具有知识:奶牛会产奶。

奶牛和黄牛都是牛的儿子结点。

H 询问 M: 黄牛会产奶吗?

经过上面的推理过程 M 能够得出答案: 黄牛不会产奶。满足了 H 的愿望,双方的愿望达到了新的平衡,问题得到了解决。

表1

	M			ACTIONS	H	
	B _M	B _M I _H	B _M →B _H		B _H I _M	B _M →B _H
初始 S ₀		∅	∅		∅	∅
				Question [H, p]		
S ₅		⟨p⟩	∅		∅	∅
				Question [M, Fathers(p)]		
S ₈		⟨p⟩	∅		Fathers(p)	∅
				Ignorance[M, Fathers(p)]		
S ₆		⟨p⟩	∅		∅	Fathers(p)
				Question [M, Brothers(p)]		
S ₂	→p	⟨p⟩	∅		∅	Fathers(p)
				State [M, →p][0][0]		
S ₁₀	→p	∅	∅		∅	Fathers(p)
				Success[]		

结束语 通过分析当前的认知语境,而获得对话语的适当解释,很多人做过这种方法的尝试,但都没有收到很好的效果。本文设计的 CBT 提供了一套话语理解过程中认知语境信息的形式化生成和更新规则,并且具备了常识检查和机器学

习的能力,虽然所能处理的问题还比较简单,但有理由相信,规则和知识库的逐渐丰富将会带来可喜的效果。

语境在话语理解中是非常重要的参考因素,并且它是个随着时间、空间、周围环境的变化以及交流的进行不断变化着的量,这正是我们话语分析中遇到的最大困难。难以做到对语境的精确把握,也就难以做到对话的连贯和自然。纵观科学发展的历史,但凡一门技术和科学,如果没有一个可靠的理论作保证,就像空中楼阁虚幻而不会长久。尽管我们见到了不少对话程序成功地通过了各类测试,但是这并不能说明我们找到了机器理解自然语言的有效办法。话语理解缺乏的不但是对各类信息的形式化描述,更缺乏扎实可靠有效的理论基础。我们期待不久的将来能够努力建立一套更加完整、实用且可计算的话语解释理论。

参考文献

- 1 Lu Ruqian, Zhang Songmao. PANGU - An agent-oriented knowledge base. In: Processing of Conference on Intelligent Information Processing (16th wcc2000), 486~493
- 2 Beun R-J, On the generation of coherent dialogue. Pragmatics & Cognition, 2001, 9: 1, 37~68. ISSN 0929-0907
- 3 Saygin A P, Cicekli I. Pragmatics in human-computer conversations. Journal of Pragmatics, 2002, 34: 227~258
- 4 周月娇. 基于常识的二义性排除; [硕士论文]. 中科院数学所, 1998
- 5 Brown G, Yule G. Discourse Analyse. Cambridge University Press, 1983
- 6 Peccei J S. Pragmatics. Routledge Press, 1999
- 7 Wang Hui, Zhan Weidong, Yu Shiwen. The Development of Semantic Knowledge-base of Contemporary Chinese and its Applications. The Second China-Japan Natural Language Processing Joint Research Promotion Conf. C/NLP, 2002. 211~225
- 8 杨帆. 图灵测试中对话处理研究; [硕士论文]. 中科院数学所, 2001
- 9 何自然, 冉永平. 语用和认知—关联理论研究. 外语教学和研究出版社, 2001
- 10 Zheng Hong, Lu Ruqian, et al. Ontology-Based Semantic Cache in AOKB. Journal of Computer Science & Technology, 2002, 17(5): 657~664
- 11 Bezuidenhout A, Cutting J C. Literal meaning, minimal propositions, and pragmatic processing. Journal of Pragmatics, 2002, 34: 433~456
- 12 Krahmer E, Piwek P. Presupposition projection as proof construction. In: H. Bunt, R. Muskens, eds. Computing Meaning, 1999, 1: 281~300
- 13 Halpern J, Moses Y. Knowledge and common knowledge in a distributed environment. Journal of the ACM, 1990, 37: 549~587
- 14 Bunt H. Context Representation for Dialogue Management. In: P. Bouquet, L. Serafini, eds. Modeling and Using Context. Lecture Notes in Artificial Intelligence Vol. 1668. Berlin, Springer
- 15 Lu Ruqian, et al. Agent oriented commonsense knowledge base. Acta Scientia, 2000, 43(6): 641~652