

自然语言理解的并行处理¹⁾*

雷晓军

中国科学院心理研究所, 北京

摘 要

传统的自然语言理解的范式是:先对句子进行句法分析,然后在此基础上再进行语义分析,得到句子的意义结构。显然,这与人的语言理解是不同的。一些心理学实验表明,句子理解时句法和语义是同时起作用的。

本文提出一个自然语言理解的新的范式,即句法和语义同时加工的模式。如果仅用句法规则对句子进行分析,其结果往往有几个歧义的句法结构,而其中只有一个是正确的。句法加工的同时,再加入语义的工作,则只得到一个句法结构。联结主义正在认知科学界兴起,它给句法和语义并行加工提供了方法。本文使用局部联结模式实现语言理解并行处理的范式,并在计算机上编制了这样的程序,它可以理解简单的汉语句子。

一、问题的导出

自然语言理解是认知科学的一个重要研究课题。经过三十年的研究,人们对自然语言有了一种共同的认识,即不再从句法的或语义的角度看待自然语言理解,而是将句法和语义合起来一同处理自然语言。有的研究者还把推理过程从语义因素中分离出来,作为一个单独在自然语言理解中起作用的重要因素。

以串行的符号加工思想为背景的认知科学在自然语言理解的问题上不可避免地要受到串行思想的影响。传统的自然语言处理的范式是:首先对输入的句子进行句法分析,得到一个句法树,或一个功能结构(f-structure),然后在语义规则和推理的作用下,将这个句法树或功能结构转化为一个深层的语义结构(或称为意义结构),这个过程可以用下图表示出来。

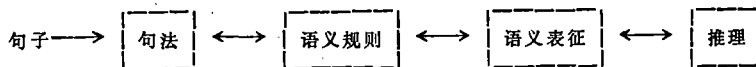


图1 传统的自然语言理解的范式

显然,上述的自然语言理解的范式是有缺陷的。句法和语义在自然语言处理中的作用往往是同时发生的,而不一定是先句法后语义。这种传统范式受到了串行加工思想的束缚。本世纪八十年代,认知科学家们提出了一种新的认知过程的加工范式:联结主义并

* 本文得到国家自然科学基金、科委863基金和北京大学信息科学中心的资助。

1) 本文于1990年6月2日收到。

行加工模式^[1-3]。这种强调认知的并行加工模式同时可以应用于自然语言的理解上,而克服传统的自然语言理解范式造成的困难。

本文谈的是使用并行加工模式的自然语言处理的工作。我们首先讨论句法和语义的工作,建立一个句法和语义并行的自然语言理解模型,最后讨论这个并行的自然语言理解模型的实现。

二、并行的自然语言理解加工模式

1. 句子理解中句法和语义的同时作用

首先,我们来看一下语言理解中句法和语义的并行加工。

在日常生活中,人们说的话并不一定全都是符合语法的,但是,这并没有造成理解上的困难。对于不符合语法的句子,如何利用句法形成一个功能结构而向语义结构转化呢?这是对“先句法后语义”范式的一个挑战。除非事先规定好不允许不符合句子出现,否则“先句法后语义”范式将陷入困境。而且,人们不仅可以理解不符合语法的句子,而且可以理解完全违背语法的句子。对于句(1),

(1) 烧火柴纵火犯仓库。

人们很快就会理解它的意思,尽管它是完全违背语法的,而且对它所形成的语义结构与句(2)的语义结构可能是一样的。

(2) 纵火犯用火柴烧仓库。

自然语言中许多词是歧义词,在不同的场合有不同的意思。存在适当的上下文时,人们很少意识到歧义词,很自然地选择了一种确切的意思。如果一个歧义词出现在一个孤立的句子中,则这句话变成了歧义句,几种理解都可以接受。例如,句(3)。

(3) 小王一天不吃饭也不行。(朱德熙^[4])

“饭”可以有两种意思:(1)米饭;(2)每天定时吃的饭食。由于词的歧义,这句话有两种意思:(1)小王一天也不能饿着;(2)小王天天都得吃米饭。在有适当的上下文的情境下,上下文选择一种词义而排除另一种词义,歧义消除。我们可以证明上下文对歧义词的一种词义的选择。例如,对英文句(4),

(4) The astronomer married a star. (Charniak^[5])

(那个天文学家同一个明星结了婚。)

人们在理解时起初会有一些困惑。“star”有两种意义:(1)星体;(2)明星。人们的困惑在于选择了星体这个意义,而使得整个句子难以理解。“astronomer”(天文学家)在这里是上文,它触动了“star”星体的意义。

上面的例子说明了一个问题,即,句法和语义在句子理解中的作用是同时发生的。

2. 句法和语义的作用及它们之间的交互作用

我们已经知道在语言理解中句法和语义的并行工作,那么它们的工作是彼此独立的,还是相互联系的,下面我们将讨论这个问题。

在句子理解过程中,句法和语义从不同的层次对句子进行加工,句法从表层结构对句子的句法信息进行加工,而语义从深层结构对句子的意义进行加工。这两个方面的加工是必不可少的。同时,这两种加工不是彼此独立的,而是相互联系的,它们相互给对方提

供信息。这种联系是并行加工的基础。

人们在理解一句话时很难意识到他们是怎样对它进行加工的，但是把这句话的某个词换上另一个词，他们就可能会觉得这句话有些不对。学过语法的人可能会指出这种不对是由于换上去的词与原来的词的词性不同，而这种差异造成整个句子的不通，没有学过语法的人则只能感到句子不对，却不能讲出什么道理。这说明理解句子时，确实存在着句法规则的作用。例如，句(4)，

(4)那个孩子很聪明。

若将“聪明”换作“智慧”，则有句(5)，

(5)那个孩子很智慧。

“聪明”是形容词，“智慧”是名词，词性的差异造成了病句。

理解句子时不仅存在着句法规则的作用，而且还存在着语义规则的作用。例如，句(6)和(7)，

(6)饭我吃了。(7)我饭吃了。

它们的句法树是完全一致的。

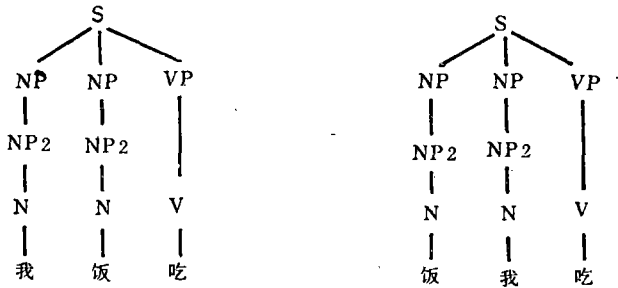
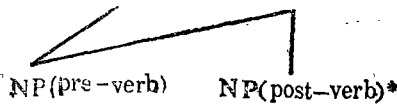


图2 句子(6)、(7)的句法分析树

决定“我”是“吃”的施事，“饭”是“吃”的受事的工作，是由语义规则(8)做的。这样，我们才能弄清谁吃和谁被吃。

(8) ANIMATE—AGENT—吃—OBJECT—FOOD



从上面所列举的例子可以看到，句法和语义规则是句子理解所不可缺少的两个因素。我们将从下面的分析看到这两个因素的相互作用。

语义规则对句子的操作是同句法分析一起进行的，而且彼此间相互传递信息，相互消除歧义。如果只有句法分析，那么往往会造成句法上的歧义。

我们举一个例子说明只用句法分析而出现的歧义以及句法和语义规则的并行工作对歧义的消除，句(9)，

(9)小王饭吃了。

对于图(3)中a、b两种句法结构，如果不使用语义规则是无法排除b的。句子(9)并

* pre-verb—指位于动词前，post-verb—指位于动词后。

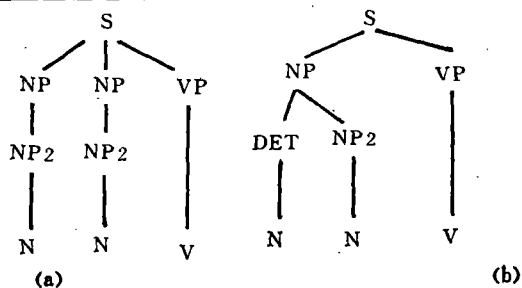


图 8 句子(9)的两种可能的句法结构a和b

不是一个歧义句,但是单纯的句法分析得到了两种句法结构,最终要通过语义规则来排除。如果在句法分析的同时,语义也在工作,而且与句法分析随时进行信息交换,那么句法歧义将不再存在。

3. 句法和语义并行的语言理解模式

为此,我们提出一个语言理解的并行加工模型,如图(4),

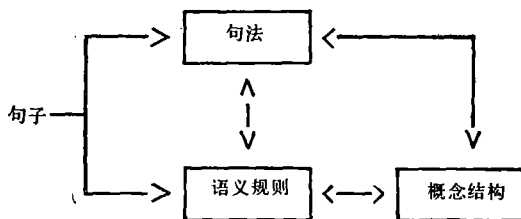


图 4 句法和语义同时加工的自然语言理解范式

在这个模式中,句子同时激发起句法和语义的工作,在加工过程中,句法和语义互相交换信息,使各自的工作得以继续进行。

三、自然语言理解的并行模式的实现

1. 并行加工模式的采用

传统的语言理解范式以串行加工思想为基础,因而它采用“先句法后语义”的加工方式。这种思想不适合句法和语义同时加工的语言理解模式。联结主义^[6,7]正在认知科学中兴起,它为并行加工模式提供了一种方法。

在认知加工的联结主义模式中,存在着两种模式:分布联结模式和局部联结模式。分布联结模式与局部联结模式相比,更接近人的认知加工模式,它完全摒弃了传统的符号加工的方式。但是,目前它的加工能力却是有限的。

虽然在自然语言理解中已经有一些工作运用了分布式联结模式,如英语动词的学习^[8]和对已分析好的句法树的句法成分赋予语义格^[9],但是这些工作只是自然语言处理的一些简单的局部工作,而且是从整个复杂认知过程中割裂出来的孤立的工作。这些工作更多地是说明分布式模式的功能,而对自然语言处理的研究的作用不大。

自然语言理解涉及到句法、语义以及知识的处理,这三个部分的工作不是简单的认知操作过程,它们超出了目前分布联结模式的所能处理的范围。

局部联结模式并没有完全摒弃符号加工系统思路下的自然语言理解的工作,而是从并行加工和大规模联结计算的角度上,重新看待以往的工作并加以新的使用。目前这种

局部联结式模式已被运用于知识表征和自然语言处理这些复杂的认知过程中,取得了一些令人瞩目的成绩^[10,11]。因此,我们采用的是局部联结模型。

2. 局部联结模式

局部联结模式是由若干个节点以及节点间的联结组成的并行网络。节点间的联结具有联结强度,用数值的大小表示,联结可以是正的也可以是负的,分别表示强化和抑制。在并行网络中,有两种节点,一种是表示句法规则和语义规则中的各个成分的,另一种是起着抑制作用的约束节点。如下列规则的网络表示为图5:

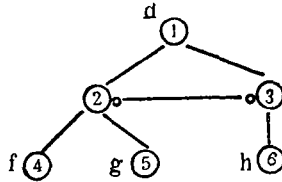


图5 规则(10)、(11)的网络结构

(10) $d \rightarrow fg$

(11) $d \rightarrow h$

规则(10)和(11)中的f、g和h都可能激活d,但最后只能存在一条路径: 1、2、4、5或1、3、6。节点2、3是约束节点,它们彼此间互相抑制竞争通向节点1的路。图中所有联结强度都为1,如果节点5处于激活状态,那么节点2的激活值大于节点3的,节点2将抑制节点3,节点3处于关闭状态。节点1—6是同时工作的,即它们各自同时接受和输出与它们相联结节点的信息。每个节点的加工能力是:

$$(a) \text{input}_i = \sum W_{ij} \times O_j$$

$$(b) \text{state}_i = \begin{cases} 0 & \text{input}_i < \text{threshold}_i \\ 1 & \text{input}_i > \text{threshold}_i \end{cases}$$

$$(c) \text{output}_i = \begin{cases} 0 & \text{state}_i = 0 \\ 1 & \text{state}_i = 1 \end{cases}$$

网络的工作是它的节点之间相互交换信息,在彼此的信息交换过程中,有的节点处于激活状态,有的则处于抑制状态,经过一段时间的信息交换,网络的工作结束,即信息交换停止。我们引入一个衡量信息交换结束的计算公式:

$$\sum \text{input}_i(t) = \sum \text{input}_i(t+1)$$

3. 句法和语义的局部联结模式的形成

句法的研究是自然语言研究中做得最多也是最好的,人们已经发现了许多句法系统^[12-16]和诸多的句法规则。为了方便起见,我们不采用复杂的句法系统,而只是用短语句法规则作为句法。这样做的原因是:(1)目前的工作尚处于尝试阶段,使用简单的句法规则就行了;(2)把句法和语义同时考虑,这样可能会产生一种新的句法,即考虑与语义交互作用的句法系统。

到目前为止,我们利用自然语言理解的并行模式处理的句子时比较简单的,每个句子

* input表示节点i接受的输入; W_{ij} 表示节点i与j之间的联结强度; O_j 表示节点j的输出; State表示节点的状态; threshold表示节点激活的阈限; Output表示节点的输出。

只有一个关系词,而且还尽量把句子限制于某个具体领域的知识中。这样做的目的是我们不可能一上来就处理非常复杂的语言,而且通过对简单的句子的理解,我们有可能看出复杂句子的理解过程,为以后的工作奠定好基础。

在我们的工作中,使用了下列的句法规则:

- * (12) $S \rightarrow NP VP$ (13) $S \rightarrow NP NP VP$
 (14) $VP \rightarrow V NP$ (15) $VP \rightarrow V$
 (16) $NP \rightarrow DET NP_2$ (17) $NP \rightarrow NP_2$
 (18) $DET \rightarrow N (De)$ (19) $DET \rightarrow V (De)$
 (20) $DET \rightarrow Det (De)$ (21) $NP_2 \rightarrow N$
 (22) $NP_2 \rightarrow Adj NP_2$

语义规则不象句法规则那样,可以被列举出来,它们存在于关系词的关系和部分名词的关系图式之中,随着词的不同而有着很大的区别。如何确定语义规则的问题已经超出了我们这里讨论的范围,在本文中我们只是将列举的例子把所需的语义规则写出来。

网络中表示语义规则成分(如AGENT,OBJECT等)的节点除了上述的加工规则外,还输入和输出激活传递能量。激活传递能量的作用控制网络节点被激活的数量,不至于使网络处于过多的冗余计算中去。一个节点接受外界的输入后,它便有了称为激活传递能量,如果这个能量能够使这个节点激活,那么它把这个能量再传递给与相联结的节点,激活传递能量的传递公式是:

$$(d) E_i = \frac{E_j}{N} \quad (N \text{ 是与节点 } j \text{ 联结的节点的数目其中包括节点 } i)$$

一个节点把它的能量传递之后,它本身就不再具有能量了,同时它也不接受这个能量的反传。因此,这些节点的加工规则要比上述的规则复杂一些,它们的加工规则为:

$$(a') \text{ input}_i = \sum W_{ij} \times O_j$$

$$(b') \text{ state}_i = \begin{cases} 0 & \text{input}_i < \text{threshold}_i \text{ and } E_i < E \\ 1 & \text{input}_i > \text{threshold}_i \text{ or } E_i > E \end{cases}$$

$$(c') \text{ output}_i = \begin{cases} 0 & \text{state}_i = 0 \\ 1 & \text{state}_i = 1 \end{cases}$$

再加上规则(d)。

表征句法规则的网络与表征语义规则的网络(其中也包含简单的知识表征)的工作是相对独立的,如激活传递能量并不在句法网络上传递,两种网络的节点大多各自联结。尽管如此,两种网络的联系还是十分密切的,它们各自有一部分节点互相联系,表示句法与语义的交互作用,而且这两种网络上的节点是同时工作的。

因为我们的并行网络的设计是在局部联结模式下进行的,它的节点之间的联结强度是由外界确定的,而不象分布式联结模式那样,通过学习^[17]自行建立节点间的联结。由外界确定的联结强度的原则是尽量快地理解句子、尽量减少网络的冗余计算。我们在这

* S—句子 NP—名词短语 VP—动词短语 DET—限定短语 N—名词 Det—限定词 V—动词 De—的 Adj—形容词

里不涉及这个技术问题。

最后,我们举两个用自然语言理解并行的模式处理的句子。句子(23)与句子(4)是一样的,

(23) The astronomer married a star.

对它的网络处理为见图 6。

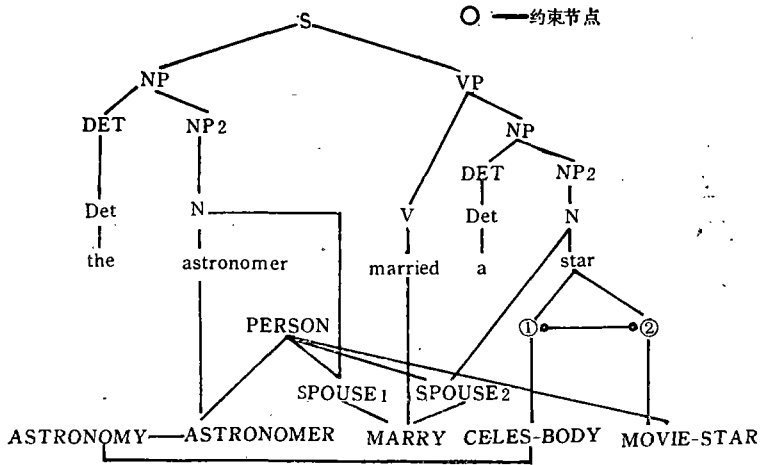


图 6 句子(23)理解的并行网络

如果节点①的激活值大于节点②的,那么节点②将被抑制,此时对句子(23)的理解会出现困惑。这个句子的最终的网络运算是节点②抑制节点①。

句法分析也帮助语义规则确定句子的意义结构。句子(24)中有两个可以充当施事的词,而哪一个是的问题需要借助句法的参与。

(24)小猫的水小狗喝了。

它的理解网络见图 7。

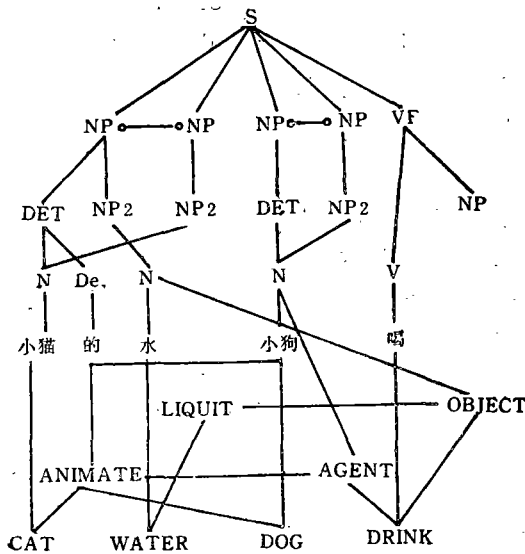


图 7 句子(24)理解的的网络

通过网络的计算,“小狗”成为施事,而“小猫”是限定“水”的成分,这种限定是由句法规则确定的。

四、总 结

通过对上面工作的讨论,可以得到下列的几点结论:

(1)句法和语义并行的语言理解模式可以排除句子分析中句法结构的歧义性,从而使得这个模式更接近人的句子理解过程。

(2)使用局部联结模式可以实现上述句法和语义并行的语言理解模式,得到一个计算机程序。

(3)在有本文第三部分列举的句法规则和一些语义规则的条件下,计算机程序可以理解简单的汉语句子。

(4)在目前的工作中,我们忽略了知识在语言理解中的作用。当句子变得较为复杂时,知识的作用将会很清晰地表现出来。同样,知识可以表征在局部联结模式的网络中,而在语言理解中起作用。关于知识表征的并行处理,我们已在另一篇文章中较为详细地讨论。

参 考 文 献

- [1] Hinton, G. E., & Anderson, J. A., *Parallel models of associative memory*. Hillsdale, NJ, Erlbaum, 1981.
- [2] McClelland, J. L., Rumelhart, D. E., & the PDP Research Group. *Parallel distributed processing: Explorations in the microstructure of cognition (Vol. II)*. Cambridge, MA, Bradford Books, 1986.
- [3] Rumelhart, D. E., & McClelland, J. L. & the PDP Research Group. *Parallel distributed processing: Explorations in the microstructure of cognition (Vol. I)*. Cambridge, MA, Bradford Books, 1986.
- [4] 朱德熙, 汉语句法中的歧义现象. *中国语文*, 1980, 2.
- [5] Charniak, E., *Passing markers: A theory of contextual influence in language comprehension*. *Cognitive Science*, 1983, 7, 171—190.
- [6] Feldman, J. A., *A connectionist model of visual memory*. In G. E. Hinton & J. A. Anderson Eds., *Parallel models of associative memory*. Hillsdale, NJ, Erlbaum, 1981.
- [7] Feldman, J. A., & Ballard, D. H., *Connectionist model and their properties*. *Cognitive Science*, 1982, 6, 205—54.
- [8] Rumelhart, D. E., & McClelland, J. L., *On learning the past tenses of English verbs*. In J. L. McClelland, D. E. Rumelhart, & the PDP Research Group Eds., *Parallel distributed processing: Explorations in the microstructure of cognition (Vol. II)*. Cambridge, MA, Bradford Books, 1986.
- [9] McClelland, J. L., & Kawamoto, A. H., *Mechanisms of sentence processing: Assigning roles to constituents*. In D. E. Rumelhart, J. L. McClelland, & the PDP Research Group Eds., *Parallel distributed processing: Explorations in the microstructure of cognition (Vol. II)*. Cambridge, MA, Bradford Books, 1986.
- [10] Fahlman, S. E., *NETL: A system for representing and using real-world knowledge*. Cambridge, MIT Press, 1979.
- [11] Cottrell, G. W., & Small, S. L., *A connectionist scheme for modelling word sense disambiguation*. *Cognition and Brain Theory*, 1983, 6, 89—120.
- [12] Perlmutter, D. M., & Postal, P. M., *Toward a universal characterization of passivization*. In Perlmutter, D. M., Ed., *Studies in relational grammar I*. Univ. of Chicago Press, Chicago, 1977.

- [13] Chomsky, N., *Lectures on government and binding*. Dordrecht, Foris, 1981.
- [14] Gazdar, G., Klein, E., Pullum, G. K., & Sag, I. A., *Generalized Phrase Structure Grammar*. Harvard Univ. Press, Cambridge, 1985.
- [15] Halliday, M. A. K., *Categories of the Theory of Grammar*. Word, 1981, 17, 241—292.
- [16] Bresnan, J. W., *A realistic transformational grammar*. In M. Halle, J. W. Bresnan, and G. A. Miller, eds., *Linguistic theory and psychological reality*. MIT Press, Cambridge, 1978.
- [17] Rumelhart, D. E., Hinton, G. E., Williams, R. J., *Learning internal representations by error propagation*. In D. E. Rumelhart, J. L. McClelland, & the PDP Research Group Eds., *Parallel distributed processing: Explorations in the microstructure of cognition (Vol. I)*. Cambridge, MA, Bradford Books, 1986.

THE PARALLEL PROCESSING IN UNDERSTANDING NATURAL LANGUAGE

Lei Xiaojun

Institute of Psychology, Academia Sinica

Abstract

The traditional paradigm of natural language understanding is that the syntactic structure of a sentence is analysed first, according to the structure, then the semantic structure is produced, which is apparently different from the process of human language understanding. There is some psychological evidence that syntax and semantics played a parallel role in understanding language.

This article proposes a new model of natural language comprehension, that is, syntax and semantics are parallel in comprehending sentences. When analysing the sentence by syntax, it might produce some ambiguous syntactic structures, in which only one is correct. If semantics is put in working with syntax at the same time, only the correct result could be obtained. Connectionism is becoming popular in the field of cognitive science, it provides a parallel method in processing sentences. In our work, a local connectionist model was used to realize the paradigm of natural language understanding. The model has been simulated by computer, which can process simple Chinese sentences.