

基于 Concept-Relation 模型的知识联通^{*})

Based-Concept-Relation-Model Knowledge Interconnection

雷玉霞¹ 陆跃飞²

(中国科学院软件研究所 北京100080)¹

(中国科学院计算技术研究所智能信息处理重点实验室 北京100080)²

Abstract It is necessary for computer to have a large knowledge base, while a lot of practice has proved that computer more needs broad knowledge interconnection in order to really realize intelligence. MindNet has dealt with the semantic path, which is kind of word-interconnection. In National Knowledge Infrastructure (NKI), we regard a closed-semantic statement as a whole knowledge unit, and discuss the interconnectivity between two statements. Knowledge interconnection is necessary in the view of knowledge engineering, and is feasible in the view of cognition and semantics. In this paper, we firstly introduce concept-relation-model, which is a suitable knowledge representation model for knowledge interconnection, and then we discuss knowledge interconnection based on the properties of relations in CR-model.

Keywords CR-model, Generalization, Specialization, Generalized path, Knowledge interconnection

1 引言

Johnson 认为知识不是孤立的而是互相联系的^[1],认知学指出思维过程体现为相关知识间的连接过程。那么如何在相关的知识之间建立联通?如何判定两条知识之间是否存在有意义的联通?又如何建立有意义的联通?这是知识处理中很重要的问题。形象地讲,知识联通就是在知识之间建立有意义的“由此及彼”的桥梁。总体上说,从知识工程角度看,要提高机器的智能程度需要知识间广泛的知识联通^[2];从认知和语义学角度讲,知识联通是可行的^[3]。另外,Johnson 的三阶段知识获取模型也说明知识不是孤立的而是互相联系的,这些说明研究知识联通是必要的、可行的。

知识库是计算机实现智能的重要基础,是计算机智能的来源。即使建立了大型知识库,而不把有关联的知识联系起来,也只能说为计算机堆积了大量的“知识”,所以还需要把有关的知识联通起来。在知识库方面,研究人员取得了许多成果,例如普林斯顿大学的 WordNet,它的最具有特色之处是根据词义而不是根据词形来组织词汇信息^[4];美国 D. Lenat 教授等研制的大型常识知识库系统 Cyc^[5];Richardson 等创建的 MindNet,它是从两种机器可读的词典(MRD)中自动构建起来的词汇知识库,代表了从自然语言文本中获取、组织、访问、开采语义信息的一般方法等,充分约束的、扩展的语义关系路径(Semantic Relation Paths)已经证明,词跟词之间的关系不被连接起来,就没有意义,MindNet 为在能支持常识推理的规模上抽取世界知识提供了可信的前景^[6,7];知网(HowNet),是个以揭示概念之间以及概念所具有的属性之间的关系为基本内容的常识知识库^[8]。MindNet, WordNet 和 HowNet 等都是把词汇作为研究对象,没有考虑把语句作为完整的知识单位加以研究,而这在基于知识的查询系统中是需要的。基于这个考虑,我们提出了基于 CR-模型(Concept-Relation-模型)的知识联通。

基于 CR-模型的知识联通有两层含义:一是建立广义路

径(语义封闭的陈述之间的路径);二是发现知识库中所蕴含的丰富的信息。在国内,有相关的研究,例如孟庆武和程虎等研究员研究过基于知识联接的自动推理方法,其方法是在程序本身建立知识间的广泛内在互连^[9]。虽然知识联接和知识联通有所区别,但是这也从另一个角度说明了研究知识联通在知识处理中的意义,即可以有助于知识推理。

本文首先介绍了 CR-模型(Concept-Relation 模型),因为合适的知识模型是研究知识联通的基础,然后对陈述的泛化和特化进行较深入的探讨,以便在知识之间建立正确的广义路径;接着主要讨论了 CR-模型中关系的性质,在此基础上探讨两条知识之间是否存在联通;最后总结全文并提出进一步讨论的问题。

2 CR-模型(Concept-Relation 模型)及广义路径

不同于 MindNet 的加权语义路径(可看作词汇联通的途径),我们的知识联通是陈述之间的联通。我们首先提出了一个合适的知识表示模型,即 CR-模型,这是知识联通的基础。为了表示两条知识之间联通路径,我们又引入了知识泛化、知识特化和广义路径等,广义路径就是通过知识泛化或特化在陈述之间建立的路径,确切地说,就是通过陈述之间的蕴含关系建立起来的路径。下面,我们逐一探讨这些问题。

2.1 CR-模型(Concept-Relation 模型)

从形式上讲,CR-模型是一个三元组 (C, R, S) ,其中 $S \subseteq C \times R \times C$, C 是 IS-A 概念主干,IS-A 记作 $\sqsubset$,即给定两个概念 C_1 和 C_2 ,如果 $C_1 \sqsubset C_2$,则 C_1 IS-A C_2 ,或者 C_1 是 C_2 的实例(Instance),或者 C_1 是 C_2 的子概念(Sub-concept),或 C_2 是 C_1 的超概念(super-concept)。 R 是所有关系的集合, S 是所有陈述的集合。为了清晰起见,我们下面假定 $R \in R$ 是个二元关系,每个陈述 $S \in S$ 的形式为: $S = CRC'$,其中 $C, C' \in C, R \in R$,并且 CR-模型中的陈述是语义封闭的。我们把这种形式的陈述看成一条知识。

^{*})本课题得到国家973项目(编号:G1999032701)资助。

CR-模型中的 $C=GC \cup IC$, 其中 GC 表示不是个体的概念的集合, IC 表示个体的集合。对于 IC 中的每一个元素 C , 都存在 GC 中的一个元素 C' , 使得 C 是 C' 的实例。假定 $\sqsubset$ 是一个偏序, 有最大元素 T , T 表示所有实体的概念, 所以给定两个概念 $C, C' \in C$, C 和 C' 有唯一的最小上界, 记作 $Lub(C, C')$ 。如果存在最大下界, 则记作 $Glb(C, C')$ 。

为了便于讨论陈述的泛化和特化, 我们需要把陈述进行分类, 又因为我们是根据关系本身的性质讨论陈述之间的联通性的, 所以先对关系进行分类。为了简单起见, 我们假定 $R = \{R\}$ 是由一个二元关系组成的。根据 R 的定义域 (Domain) 和值域 (Range), 我们将 R 区分为以下四类 (相应地, 陈述也有以下四类形式):

- (1) R 是1-类的, 如果 $Dom(R) \subseteq IC$, 并且 $Range(R) \subseteq IC$;
- (2) R 是2-类的, 如果 $Dom(R) \subseteq IC$, 并且 $Range(R) \subseteq GC$;
- (3) R 是3-类的, 如果 $Dom(R) \subseteq GC$, 并且 $Range(R) \subseteq IC$;
- (4) R 是4-类的, 如果 $Dom(R) \subseteq GC$, 并且 $Range(R) \subseteq GC$ 。

2.2 CR-模型中的广义路径

类似于 MindNet 中的加权语义路径, 我们在研究基于 CR-模型的知识联通时, 也需要一个广义路径, 这是研究知识联通的有效途径。这里涉及到三个问题: 一是如何建立广义路径; 二是如何建立正确的、有意义的广义路径; 三是如何判定两个陈述之间存在有意义的广义路径。这些都是知识联通中的关键问题。为了解决这些问题, 我们借用 Version Space 方法中的泛化或特化的思想^[10], 研究陈述之间的泛化和特化。

给定陈述 $S = C1RC2$, 称另一个陈述 $S' = C1'RC2'$ 是 S 的一个泛化 (相应地, 称 S 是 S' 的一个特化), 如果满足以下任何一条:

- (1) $C1 \sqsubset C1'$, 并且 $C2 = C2'$;
- (2) $C2 \sqsubset C2'$, 并且 $C1 = C1'$;
- (3) $C1 \sqsubset C1'$, 并且 $C2 \sqsubset C2'$ 。

很显然, 并不是陈述的每个泛化都是有意义的。例如, 设 $S = C1RC2$, 其中 $C1 = \text{人}$, $R = \text{是一个 (IS-A)}$, $C2 = \text{脊椎动物}$ 。如果 $C1' = \text{哺乳动物}$, 则 $S' = C1'RC2$ 是“哺乳动物是一种脊椎动物”, 这是有意义的。但是, 如果 $C1' = \text{动物}$, 则 $S' = C1'RC2$ 就是“动物是一种脊椎动物”, 是一个不正确的陈述。注意, 这里的“哺乳动物”意思是概念上的哺乳动物, 是对“哺乳动物”的所有实例都成立的。

如果 $C1' = \text{脊椎动物}$, 则 $S' = C1'RC2$ 是一个正确的陈述, 即“脊椎动物是一种脊椎动物”。如果我们假设 R 在上层本体中是自反的, 则 S' 是一个平凡的陈述。这说明有些泛化是不正确的, 而有些泛化即使是正确的, 也不一定是有意义的。我们需要排除无意义的泛化。

分析表明, 给定一个基于 IS-A 概念主干的陈述, 存在两类泛化和两类特化。准确地讲, 给定 $S = C1RC2$, 设 $Dom(RC2) = \{C \in C | CRC2 \in S\}$ 。设 α 是从 $C1$ 到 T 的路径上节点的集合。如果有 $\alpha \sqsubseteq Dom(RC2)$, 则对于 $C1$ 的任何上位概念 $C, S' = CRC2$ 是 S 的有意义泛化。如果 $\alpha \not\sqsubseteq Dom(RC2)$, 则对于 $C \in Dom(RC2) \cap \alpha, S' = CRC2$ 是 S 的有意义的泛化; 否则 $S' = CRC2$ 是 S 的无意义的泛化。特化的情况类似于泛化。

我们暂且把 S 的泛化和特化称为 S 的分类结论, 简称为

S 的 TCs。如果 S' 是 S 的泛化, 我们就称 S' 是 S 的一个 G-TC; 反之, S' 就是 S 的一个 S-TC。对于“是一个 G-TC, 和是一个 T-GC”这两个关系, 很显然有下面的命题:

命题2.1 “是一个 G-TC”和“是一个 T-GC”作为关系是传递的、非自反的”。

在 TCs 中, 有些是有意义的, 而有些是没有意义的。为了定义有意义的 TCs, 我们根据 R 所满足的性质将 R 作如下分类:

- (1) R 是左上正的, 如果对于 $C1 \sqsubset C1', (C1, C2) \in R$ 蕴含 $(C1', C2) \in R$;
- (2) R 是右上正的, 如果对于 $C2 \sqsubset C2', (C1, C2) \in R$ 蕴含 $(C1, C2') \in R$;
- (3) R 是左下正的, 如果对于 $C1' \sqsubset C1, (C1, C2) \in R$ 蕴含 $(C1', C2) \in R$;
- (4) R 是右下正的, 如果对于 $C2' \sqsubset C2, (C1, C2) \in R$ 蕴含 $(C1, C2') \in R$ 。

命题2.2 给定一个陈述 S , 设 S' 是 S 的一个 TCs。 S' 是有意义的, 如果 S' 满足下面的任一条:

- (1) S' 是 S 的一个泛化, 并且或者 $S' = C1'RC2$ 并且 R 是左上正的, 或者 $S' = C1RC2'$ 并且 R 是右上正的;
- (2) S' 是 S 的一个特化, 并且或者 $S' = C1'RC2$ 并且 R 是左下正的, 或者 $S' = C1RC2'$ 并且 R 是右下正的。

证明: 由关系 R 的四个分类和对有意义泛化的分析可证。

如果 R 是1-类或2-类的, 即 $Dom(R) \subseteq IC$, 则陈述 $S = C1RC2$ 应该被量词量化。通常地讲, 量词应该包括: 对每一个, 几乎全部, 有些, 不是每一个和没有。设 Q 表示一个量词, Q' 表示 Q 的对偶, 则有

Q	Q'
对每一个	有些
不是每一个	每一个都不
没有	不是每一个

设 $Q = \text{不是每一个}$, 则 $Qx\varphi(x)$ 意思是并不是每个 x 都满足 φ , 等价于 $Qx\varphi(x) = \neg \forall x\varphi(x)$ 。其对偶是 $\neg Q \rightarrow \equiv \neg \neg \forall \rightarrow \equiv \forall \neg$ 。如果 $Q = \text{没有}$, 则 $Q \equiv \forall \neg$, 其对偶有两种形式: 一是 $\neg Q \rightarrow \equiv \neg \forall \neg \rightarrow \equiv \forall$, 意思是不是每一个; 另一种是 $\neg Q \rightarrow \equiv \exists \rightarrow$, 意思是有些不是。这些形式是等价的。

有许多形如 $S = Qx \in C1RC2$ 的陈述, $S = Qx \in C1RC2$ 是指 $C1$ 的一个量化是和 $C2$ 是 R -相关的。例如, 假设 $Q = \text{有些}$, 则 $S = Qx \in C1RC2$ 意思是对于 $C1$ 的有些实例 x , 有 $S' = xRC2$, 其中 $x \in IC$ 。很显然, 对于种类陈述有下面的命题:

命题2.3 给定两个概念 C 和 D , 并且 $C \sqsubset D$, 并且给定一个陈述 $S = QxRC2$, 则有:

- (1) 如果 $Q = \text{对每一个}$, 则 $S = Qx \in DRC2$ 蕴含 $S' = Qx \in CRC2$;
- (2) 如果 $Q = \text{有些}$, 则 $S = Qx \in CRC2$ 蕴含 $S' = Qx \in DRC2$;
- (3) 如果 $Q = \text{不是每一个}$, 则 $S = Qx \in CRC2$ 蕴含 $S' = Qx \in DRC2$;
- (4) 如果 $Q = \text{没有}$, 则 $S = Qx \in DRC2$ 蕴含 $S' = Qx \in CRC2$ 。

3 基于 CR-模型的知识联通

通过陈述的泛化或特化, 可以在陈述之间建立广义路径,

有了前面的讨论作为基础,我们现在开始讨论基于 CR-模型的知识联通,为了清晰起见,我们仅限于基于 IS-A 概念主干上的知识联通。给定两个陈述 $S=C_1RC_2$, 和 $S'=C_1'RC_2'$, 我们试图找到一个有意义的陈述 T , T 是 S 和 S' 的一个 TC, 这里把 T 称为 S 和 S' 的一个联通, T 像 S 和 S' 之间的一个桥梁, 在知识工程中的意义体现在有利于知识推理。在这小节中, 我们首先分析知识联通, 然后给出一个实例。

3.1 基于 IS-A 概念主干上的知识联通

为了便于讨论, 我们假设 R 是左右区间封闭的。一个关系 R 是左区间封闭的, 是指对于任何概念 $C, D, E, F \in C$, 如果 $(C \subseteq D \subseteq E) \wedge (C, F) \in R \wedge (E, F) \in R$, 则 $(D, F) \in R$ 。类似地, 一个关系是右区间封闭的, 是指对于任何概念 $C, D, E, F \in C$, 如果 $(E \subseteq D \subseteq F) \wedge (C, E) \in R \wedge (C, F) \in R$, 则有 $(C, D) \in R$ 。

给定两个陈述 $S=C_1RC_2$, 和 $S'=C_1'RC_2'$, 下面讨论两个陈述之间的联通。首先, 我们考虑简单的情况, 即 $C_2=C_2'$ 。想找到这样的 T , 我们首先要找到 C_1 和 C_1' 的 $\text{Lub}(C_1, C_1')$, 记作 C 。如果 $(C, C_2) \in R$, 则令 $T=CR C_2$ 。其次, 我们找到 C_1 和 C_1' 的 $\text{Glb}(C_1, C_1')$, 记作 D 。如果 $(D, C_2) \in R$, 则令 $T=DR C_2$ 。如果 C 和 D 都不存在, 则 S 和 S' 之间不能进行知识联通。不难证明, 有下面的命题:

命题3.1 给定两个陈述 $S=C_1RC_2$ 和 $S'=C_1'RC_2$, 设 T 是 S 和 S' 的一个联通, 则如果 T 是存在的, 则一定有一结论成立: (1) $\text{Lub}(C_1, C_1') \in \text{Dom}(R)$; (2) $\text{Glb}(C_1, C_1') \in \text{Dom}(R)$ 。

现在, 我们考虑 $C_2 \neq C_2'$ 的情况。因为 $\text{Lub}(C_1, C_1')$ 是一定存在的, 所以我们先研究 $\text{Lub}(C_1, C_1')$ 的情况。我们定义一个概念: 关系 R 是和-封闭的, 如果对于任何 $C, D, E \in C$, 有

$$(3.1) (C, E) \in R \wedge (D, E) \in R \rightarrow (\text{Lub}(C, D), E) \in R,$$

或

$$(3.2) (C, D) \in R \wedge (C, E) \in R \rightarrow (C, \text{Lub}(D, E)) \in R$$

如果 C, D, E 满足 (3.1), 则我们记作 $C, D; E$ 。如果 C, D, E 满足 (3.2), 则我们记作 $C; D, E$ 。

命题3.2 设在 C 下 C 构成一个概念树。对于任意给定的两个陈述 $S=C_1RC_2$ 和 $S'=C_1'RC_2'$, 如果 R 是和-封闭的, $S_1=C_1'RC_2 \in S, S_2=C_1RC_2' \in S$, 则 S 和 S' 存在一个联通 T 。

证明: 由假设 R 是和-封闭的可知, S 和 $S_1=C_1'RC_2$ 有一个共同的泛化, 例如 $T_1 = \text{Lub}(C_1, C_1')RC_2$, 同理 $S_2=C_1RC_2'$ 和 S' 有一个共同的泛化, 例如 $T_2 = \text{Lub}(C_1, C_1')RC_2'$, 进而可知 T_1 和 T_2 存在一个共同的泛化, 例如 $\text{Lub}(C_1, C_1')R \text{Lub}(C_2, C_2')$ 。设 $T=CRD$, 则 T 是 S 和 S' 的一个联通。

这个命题表明给定两个陈述 $S=C_1RC_2$ 和 $S'=C_1'RC_2'$, 如果 R 是和-封闭的, 并且 $C_1, C_1'; C_2, C_1, C_1'; C_2'$ 和 $\text{Lub}(C_1, C_1'); C_2, C_2'$, 则 S 和 S' 一定存在联通; 否则不一定存在联通。

再考虑更特殊的情况, 即如果 $\text{Glb}(C_1, C_1')$ 和 $\text{Glb}(C_2, C_2')$ 是存在的。为了讨论这种情况, 我们定义: 一个关系 R 是交-封闭的, 即如果对于任何 $C, D, E \in C$, 有

$$(3.3) (C, E) \in R \wedge (D, E) \in R \rightarrow (\text{Glb}(C, D), E) \in R,$$

或

$$(3.4) (C, D) \in R \wedge (C, E) \in R \rightarrow (C, \text{Glb}(D, E)) \in R$$

如果 C, D, E 满足 (3.3), 则我们记作 $C, D; E$ 。如果 C, D, E 满

足 (3.4), 则我们记作 $C; D, E$ 。对于这种特殊的情况, 也有类似命题:

命题3.3 假设在 C 下 C 构成一个层状概念图。对于给定的两个陈述 $S=C_1RC_2$ 和 $S'=C_1'RC_2'$, 如果 R 是交-封闭的, $S_1=C_1'RC_2 \in S, S_2=C_1RC_2' \in S$, $\text{Glb}(C_1, C_1')$ 和 $\text{Glb}(C_2, C_2')$ 是存在的, 则 S 和 S' 存在一个联通 T 。

证明: 类似于命题3.2的证明。

这个命题表明给定两个陈述 $S=C_1RC_2$ 和 $S'=C_1'RC_2'$, 如果 R 是交-封闭的, 并且 $C_1, C_1'; C_2, C_1, C_1'; C_2', \text{Glb}(C_1, C_1')$ 和 $\text{Glb}(C_2, C_2')$ 是存在的, 另外 $\text{Glb}(C_1, C_1'); C_2, C_2'$, 则 S 和 S' 一定存在联通; 否则不一定存在联通。

命题3.4 给定两个陈述 $S=C_1RC_2$ 和 $S'=C_1'RC_2'$, 设 T 是 S 和 S' 的一个联通, 则如果 T 是存在的, 则一定有以下的一结论成立:

$$(1) \text{Lub}(C_1, C_1') \in \text{Dom}(R);$$

$$(2) \text{Glb}(C_1, C_1') \in \text{Dom}(R);$$

$$(3) \text{Lub}(C_2, C_2') \in \text{Range}(R);$$

$$(4) \text{Glb}(C_2, C_2') \in \text{Range}(R);$$

$$(5) \text{Lub}(C_1, C_1') \in \text{Dom}(R), \text{ 并且 } \text{Lub}(C_2, C_2') \in \text{Range}(R);$$

$$(6) \text{Lub}(C_1, C_1') \in \text{Dom}(R), \text{ 并且 } \text{Glb}(C_2, C_2') \in \text{Range}(R);$$

$$(7) \text{Glb}(C_1, C_1') \in \text{Dom}(R), \text{ 并且 } \text{Lub}(C_2, C_2') \in \text{Range}(R);$$

$$(8) \text{Glb}(C_1, C_1') \in \text{Dom}(R), \text{ 并且 } \text{Glb}(C_2, C_2') \in \text{Range}(R)。$$

证明: 由前面的分析易证。如果 $C_2=C_2'$, 则该命题和命题3.1等价。

3.2 实例

我们具体分析一个简单的知识联通。设 $S_1=\langle \text{饮食过量} \rangle \langle \text{引起} \rangle \langle \text{消化不良} \rangle \in S; S_2=\langle \text{暴饮暴食} \rangle \langle \text{引起} \rangle \langle \text{消化紊乱} \rangle \in S$, S_1 和 S_2 是否存在有意义的联通? 如何进行联通? 已知在 IS-A 概念主干中有“消化紊乱 $\subseteq$ 消化障碍, 消化不良 $\subseteq$ 消化障碍”; “饮食过量 $\subseteq$ 饮食不当”和“暴饮暴食 $\subseteq$ 饮食不当”。在这里“引起”作为一个二元关系, 即是左上正的, 又是右上正的。根据前面的讨论, 可以分别对 S_1 和 S_2 进行正确的、有意义的泛化, 然后找到 S_1 和 S_2 的联通 T (见图1)。

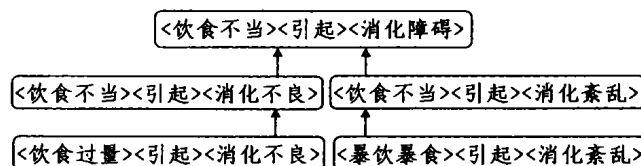


图1 基于 CR-模型的知识联通图

可见, 我们研究的知识联通不同于 MindNet 的词汇联通, 而是把语义封闭的陈述作为完整的知识单元加以讨论的, 其特点就是通过分析 CR-模型中关系的性质来探讨知识之间的联通性的。

总结和进一步工作 本文首先提出了一个合适的知识表示模型, 即 CR-模型, 这是陈述联通的基础。我们所讨论的陈述的形式是 $\langle C \rangle \langle R \rangle \langle C' \rangle$, 其中 C 和 C' 是概念, R 是个二元关系, 它是语义封闭的, 并且这种形式的陈述是一条知识。

(下转第58页)

示;(5)提供锚点,降低每次知识获取的难度(跳跃度)。

事物是普遍联系和不断发展的,我们需要智能的系统观。从长远来看,对人工智能的基础学科—认知科学的研究才是问题的出路。例如,从发展心理学来看,人的认知发展过程是连续的,任何一个智能行为的发生都不是孤立的,不是可分离的,不能割断发展的前后联系性,就某一单个智能行为就事论事地进行模拟是很片面的,在人工智能问题上也应该有历史观。通过对关于 Nature 和 Nurture 问题的讨论,我们可以获得许多关于智能系统的模型构造设计和学习样本设计的极有价值的启发,所以应该给智能系统设计学习子系统,用一种发展的、进化的、循序渐进的观点来实现系统功能的进步^[4]。

传统的符号化人工智能采用的是“再现智能”的方法,设计的模型往往是对一种以某种形式表现、应用于某一局部领域的智能现象的模拟,对模型的认知意义并不刻意追求。这种指导思想对智能就事论事,缺乏对其本质的一个深刻洞察,没有将此智能现象作为人所有智能的一部分放在人的心智体系中进行全面系统地考察、放在认知发展全过程的前因后果序列中考察、放在无法回避的认知约束下进行考察。我们已经不满足于其中的形式推理,而转向实现更朴素和本源的基于语义的推理。要想根本性地提高人工智能系统的性能,实现向“解释智能”的转变就必须将人工智能放在认知科学这个交叉学科的背景下进行修正,后者通过适当地使用还原和整合的策略说明高层次的智能现象如何突现。认知是包括知识获取、存储、回忆和使用的全部精神活动。认知的相关性就是认知结构在所加工信息的形态、加工流程、结构的组成、学习对结构的影响等方面的依赖关系。对这些相关性的研究对于知识获取、知识表示和知识应用的分阶段模型和联合模型的结构,以及构造过程研究具有重大意义。认知的相关性约束要求我们考虑智能的发生、发展和完善的过程,以不割裂的观点看待智

能行为的模拟,这就是智能模型构造的系统观点。通过对认知发展过程的研究使我们能更清楚、全面地看到智能的本质,而不仅仅是它的外表。认知相关性要求和系统的智能观是建立在统一的认知结构的基础之上,这个结构的基础和统一的体现是对信息的直接表达,这是一种对知觉信息的表达,是对语义的直接表达,它是心智的计算理论中心智表达和心智计算的核心,通过基于神经网络的直接表达的实现能够找到整体主义和还原主义、天赋论和建构论的结合点。语言是最具代表性的智能行为,对个人言语的计算结构研究能最大限度的体现认知心理学对于言语的信息加工算法、生理心理学对言语加工模型的体系结构、计算神经科学对于这个加工结构的实现、发展心理学对计算模型的进化作用的影响。

参考文献

- 1 Kirsh D. Foundations of AI: the big issues. Artificial Intelligence, 1991, 47: 4~20
- 2 Collins W A, Kuczaj S A II. Developmental Psychology: childhood and adolescence, New York: Macmillan Publishing Company, 1991. 167~199
- 3 林崇德. 发展心理学. 北京:人民教育出版社, 1995. 10~11, 50~63
- 4 危辉. 智能脑机制的认知结构核心之元态——表象式语义网及其构造: [北京航空航天大学研究生院博士学位论文]. 1998
- 5 张钹, 张铃. 问题求解理论及其应用. 北京:清华大学出版社, 1990. 4~5
- 6 (美)威廉·卡尔文. 大脑如何思维. 杨雄里、梁培基译. 上海:上海科学技术出版社, 1996. 32~33
- 7 朱智贤. 儿童心理学(1993年版). 北京:人民教育出版社, 1993. 125~129, 158~167, 211~229
- 8 朱平, 危辉, 何新贵. 动态模糊语义网及其并行执行. 系统工程与电子技术, 1998, 20(6): 53~56

(上接第55页)

文章着重讨论了三个问题:如何判定两条知识之间是否存在联通?如何建立正确的联通?如何保证联通是有意义的?针对这些问题,我们主要是通过分析 CR-模型中关系的性质逐一解决的。对于第一个问题,我们提出了两个概念的最小上界和最大下界等概念,然后由命题3.1~命题3.4判定知识之间是否存在联通。对于后两个问题,先是提出知识泛化和特化,两条知识正是通过泛化或特化建立联通的。为了保证是正确的、有意义的联通,文章采用了如下的策略:给定 $S = C1RC2$, 设 $Dom(RC2) = \{C \in C | CRC2 \in S\}$, α 是从 $C1$ 到 T 的路径上节点的集合。如果有 $\alpha \subseteq Dom(RC2)$, 则对于 $C1$ 的任何上位概念 C , $S' = CRC2$ 是 S 的有意义泛化。如果 $\alpha \not\subseteq Dom(RC2)$, 则对于 $C \in Dom(RC2) \cap \alpha$, $S' = CRC2$ 是 S 的有意义的泛化;否则 $S' = CRC2$ 是 S 的无意义的泛化。

在知识联通中,我们认为还存在很多问题值得讨论。例如如何利用本体在概念之间建立联通,如何在不同的本体之间建立联通?如何建立一个合适的联通测度?对于概念联通,我们可以借鉴 MindNet 的加权语义路径等技巧。如果这些问题得到较好的解决,我们认为将对 ITS 和知识共享方面都是有意义的。

致谢 中国科学院计算技术研究所大规模知识处理课题

组的曹存根研究员对本文的工作给予了指导,博士生张春霞同学对本文的完成提出了许多有益的建议,在此一并表示感谢。

参考文献

- 1 Johnson P E. What kind of expert should a system be? Medicine and Philosophy, 1983, 8: 77~97
- 2 Gallant S I. Connectionist expert systems. Communications of the ACM, 1988, 31: 152~169
- 3 伍谦光. 语义学导论. 湖南:湖南教育出版社, 1987. 6~241
- 4 Miller G A, et al. Introduction to WordNet: An On-line Lexical Database. Princeton University, 1993
- 5 Lenat D B. Cyc: A Large-Scale Investment in Knowledge Infrastructure. Communications of the ACM, 1995, 38(11): 1~38
- 6 Richardson S D, Dolan W B, Vanderwende L. MindNet: acquiring and structuring semantic information from text. Microsoft Research Technical Publications (MSR-TR-98-23), 1998. 网址: <ftp://ftp.research.microsoft.com/pub/tr/tr-98-23.doc>
- 7 Richardson S. Determining similarity and inferring relations in a lexical knowledge base: [PhD. dissertation]. City University of New York, 1997
- 8 董振东. 语言知识库和常识库的建设——《知网》简介. 见:钟义信. 中国人工智能进展. 北京:北京邮电大学出版社, 2001. 967~970
- 9 孟庆武, 程虎, 等. 基于知识联接的自动推理方法及应用. 见:史忠植, 怀进鹏, 田启家. 人工智能进展. 北京:清华大学出版社, 2001. 29~34
- 10 史忠植. 高级人工智能. 北京:科学出版社, 1998. 101~125