

计划语言和计算语言学

刘海涛

(青海铝厂计算机中心)

计算语言学是信息时代语言研究的产物，它的主要研究对象为人—机语言交流问题。计划语言是为了解决人—人语言交际困难，而提出的人造语言系统。由于二者之间关系的紧密性，某些学者认为，计算语言学仅仅是一种精确化的国际语学（计划语言学）。

从公元 2 世纪至今，已出现了近千种计划语言方案，它们几乎包括了人们可想到的所有人工语言或符号体系的类型。例如采用图像直接表意的方案，可以作为现代计算机技术中人机图形界面的元素来运用，由于一般的语言方案都系统地对人类的知识进行过较细致的分析，使用它们来作为 GUI 中的 ICO 想必会比现用的图符要易于识别和记忆。

就计划语言的语音而言，它一般没有难发和难以辨别的音，为了抽取出符合简单易发这一原则的音素，计划语言学者们已做了大量的工作，这些语音方面的研究成果，能对语音识别与合成工作者提供一定的启迪。

与自然语言相比，计划语言最大的优势当在于语法方面。民族语言是历史发展的产物，这使得语言在各个层次都充满了复杂和不尽合理之处，在语法规则里包含了许多例外，所有这一切给语言的学习和使用带来了巨大的困难。为了克服民族语的这许多弊病，简化国际交流，人们根据自身的需要创造出了计划语言。虽然计划语言的语法结构各异，但总的来说它比民族语更易学习和使用，特别是在规则性方面更是引人注目。为了表现语言结构的规则性，设计者们采用了各种不同的方式。设计良好的计划语言可以做到在结构上无歧义，而这种单义的结构是非常利于计算机分析处理的。如：世界语中有一套显式标记句法关系和结构的符号，这组符号具有表现句子结构的通用性。它的这种显式标注句法关系的特性，早在计算机产生之前就被有识之士所注意了，甚至与自动翻译思想的产生密切相关。现代机器翻译的实践也证明，在世界语中很容易做到没有句法结构歧义。研究表明，世界语独特的词法结构也非常适合于计算机处理。

近年来随着计算语言学领域的进展，人们的研究重点已从过去单纯的句法研究逐步转向语义、语用的综合性研究。这其中语义研究尤为活跃，由于语义较之句法而言是一种模糊的、不易界定的东西，所以其研究方法也大为不同。语义研究的热点已由过去的“义素”或“成分分析”转向“语义场”。“语义场”研究的本质就是对人类知识的划分和表示，而这种认识自然和表现世界的方法，在计划语言中屡见不鲜。人类对于世界的认识是不断变化的，这种变化来自于人类本身的进步和社会的发展。随着时间的推移，人类对知识的分类也是会变化的，换言之，世界上没有通用和固定的知识分类体系。计划语言的历史已经向我们证明了这一点。

狭义理解的计算语言学大概包括研究语言的语音、语言的分析和语义的计算机处理，而计划语言研究的本质是寻求更适宜的人类语言交流工具。二者都需要研究人类处理语言的机制和方法，都是为了克服人类的语言障碍，都要构建某种有别于自然语言的语言，以便于人或计算机理解所要表达或传递的意义。语言理解在很大的程度上是一种选择的技术，选择的依据是人类或机器关于世界的知识，选择的机制是建立于概率和模糊推理之上的。于是乎知识的表示又成为问题的核心，我们目前还没有全能的知识表示方法，但综合研究人类对于知识的表示方法的体系，可能会找到一条通向成功的路。在这一思想指导下，我们结合两门最关心人类知识表示和言语交际的学科——计划语言学和计算语言学，对此问题做了一些探索研究，希望有助于计算语言学的研究。