

HNC 知识处理研究院在北京中关村科技园区成立

HNC 研究院(工商注册名称为:北京大正语言知识处理研究院)由中国科学院声学研究所和北京麦纳科技发展有限公司合资组建,是一家高科技研发机构。研究院的中心目标是,以中国科学院雄厚的技术力量为后盾,以创新的 HNC 理论和技术为基础,致力于自然语言的计算机处理,研究和开发自然语言理解的核心技术,建立语言知识处理基地,开发语言理解的应用系统,提供智能化知识服务及其解决方案,其中包括机器翻译、语言信息搜索和咨询、情报发掘、网络信息过滤、电子客户服务、电子文本校对等。

HNC 是 Hierarchical Network of Concepts(概念层次网络)的缩写。HNC 理论是一个关于自然语言理解处理的创新理论,由中国科学院声学研究所资深研究员黄曾阳先生潜心钻研 8 年而创立。该理论在深入挖掘汉语特点的基础上,以意义表达和语言理解为主线,建立了一种模拟大脑语言感知过程的自然语言表述模式和计算机理解处理模式,适用于任何语种语言的描述和处理。该理论问世以来,引起了有关专家学者的广泛关注和认同,有些著名语言学家把它视为当前我国中文信息处理的三大流派之一,认为该理论体现了语言学、认知科学和计算机科学的有机结合。该理论也得到了多项重大项目的支持,其中有中国科学院“八五”重大项目、国家“九五”重点科技攻关项目、国家重点基础研究发展规划(973)项目等,现已经形成了称之为“句类分析的汉语语句理解处理”的核心技术,该技术已获得“模拟大脑语言感知过程的自然语言类分析方法”发明专利。由信息产业部主持的“九五”科技成果鉴定会,对此项专利技术评价为“在汉语语句理解方面达到国际领先水平。”今年 4 月,在武汉华中师范大学语言学系成功召开了 HNC 与语言学研究研讨会。

HNC 理论和句类分析技术是自然语言理解处理领域的突破性进展,是我国语言信息处理事业发展的一个结晶,也是中国人拥有自主知识产权的一项重要知识资产。HNC

研究院的成立将把这一高科技成果推向产品化和产业化。

自然语言处理是伴随着计算机的问世而诞生的一门交叉科学,其目标是实现语言信息的计算机自动处理。语言是人类表达和传递信息的最重要工具,实现语言信息处理的自动化是人类现代文明的巨大需求。随着信息和网络时代的到来,这一需求将日益迫切。对语言信息进行处理,不能只把语言作为符号来处理,而是要首先实现计算机对语言的理解,并在理解的基础上由计算机自动把语言中传递的信息提取出来,加以组织,形成有用的知识,这就是语言知识处理的使命。语言知识处理势将成为 21 世纪新经济的重要组成部分,它也正是当今发达国家意欲争先解决的重大技术难题之一。语言知识处理涉及四大应用领域,即机器翻译、自动文摘、智能检索、人机交互(友好界面)。这几方面的研究,虽都已有几十年的发展历程,但至今都还未形成市场认同的成熟技术,其根本原因在于缺少对语言进行理解处理的理论和技术。要在这方面获得突破,关键在于要把计算机对自然语言的理解推进到接近人类认知的水平。HNC 理论和技术为实现这一突破创造了可能条件,使我国能够在语言知识处理这一信息领域的国际竞争中取得领先地位。HNC 研究院的使命也在于此。

这次 HNC 研究院的成立,也是我国国家科研机构和企业界合作的结果,这是一种有创新意义和突破意义的发展模式,即把国家项目的科研成果先通过研究院进行二次研究和开发,然后再以组织项目公司的形式最终推向市场。这个新成立的 HNC 研究院,实质上是一种以基础研究为主的国家科研机构和以市场运作为主的项目公司之间的过渡。这样做的意义在于,一方面能使科研成果更顺利地向市场转化,另一方面也可为科研的深入提供反馈信息。

(林杏光供稿)

科技动态

微型宇宙探测器群

据英《新科学家》2001 年 1 月 6 日报道:美国航天科学家们从社会性昆虫(如蚂蚁遍地搜食物蜜蜂到处采蜜)的活动中得到启示,计划用成百上千个微型宇宙探测器组成集团军进行小行星探测。在太阳系,除了有九大行星外,还有数量多达几千个的小行星。这些小行星中可能有丰富的矿物和水资源。分布在火星和木星轨道之间的小行星数量占大多数,形成所谓小行星带。如果派遣单个的宇宙飞船去探测小行星带中某个小行星,其效率之低是可想而知的。

目前, NASA 的研究人员计划发射约 1 000 个微型宇宙探测器探测小行星带。这些微型探测器可以降落在直径大于 1 公里的小行星上,在那里寻找矿物和水资源。

另外,美国马里兰 Goddard 宇宙飞行中心的 Steven Curtis 说,他们准备先发射称为“自主纳米技术探测群”的

大约 100 个微型探测器到小行星带去收集岩石的数据。每个探测器都装有重量不大于 1 公斤的传感器。这个探测群体中有一个探测器是“领导”,其余的探测器则是“工人”,每个“工人”随时向“领导”报告它所探测到的数据。在充当“领导”的探测器内装有收集数据的计算机,一旦收集的数据“装满”了计算机,“领导”就把数据转给飞回地球的“信使”进行整理分析。

Curtis 说,计划发射 1 000 个微型探测器的“蚁海”战术还有一个目的,就是即使其中某个探测器失灵,也只是损失一个“单兵”,不致使整个飞行使命失败。这些探测器可以从地球引力之外的外层空间释放,当 100 平方米的太阳能帆板在外层空间展开时,其发出的电能足可以让成群的微型探测器坚持 3 年航行到达小行星带,进行宇宙探测。

(刘先曙)