

秦一男：一种英文句法结构解析的新方法（2015）

自然语言的句法结构解析，就是应用计算机科学与技术的原理和方法，对自然语言语句的句法成分和结构进行判断和划分的过程。这个过程在很大程度上影响和制约着计算机对自然语言的最终处理效果，并由此影响和制约着人们通过计算机而获得的对自然语言的间接理解效果。

本文是一篇跨学科、交叉性、综合性的学术论文。本文主要研究英文句法结构解析中的句法成分远距离相关问题、复合句式的深层递归嵌套问题、从句边界的划定问题，并给出了解决问题的方案。

本文提出了英语语言的三条确定性句法规律，并基于大量的实验数据指出了三种句法结构解析软件——Berkeley Parser、Stanford Parser、中国科学院计算技术研究所的改进型Shift-Reduce句法分析软件的重大技术漏洞，进而指出了当前的计算机句法结构解析的主流方法——PCFG方法的理论瓶颈所在。在此基础之上，本文给出了一种英文句法结构解析的新方法。

本文给出的新方法，依据数学原理和相应的计算机技术，运用复合函数的数学思想，通过建立矩阵模型、建立线性模型、构造递归函数进行对英文语句的句法结构解析；同时，综合运用数学归纳法等方法对重要的结论进行了证明。本文给出了单侧同向保序和单侧同向不保序两种整体插空方法，并运用集合族的生成处理并列句法成分。本文为英文语句建立了一套全新的数学模型，揭示了英文语句的深层数学规律。