

# 2025 年全国硕士研究生招生考试 国防科技大学自命题科目考试大纲

科目代码：F0808

科目名称：语言信息处理专业综合

## 一、考试要求

主要考查考生对计算语言学基础理论和研究方法的理解与掌握；熟练运用相关理论和方法研究自然语言语音、词法、句法、语义、语用知识的形式化表示与分析能力；熟悉并了解国内外多语言信息智能处理的最新学术动态及其在国防科技领域的应用前景。

## 二、考试内容

### 1. 计算语言学基础理论部分

计算语言学的学科属性、研究范型和发展简史；形式语法、依存语法、格语法、语义网络语法等理论和方法；计算音系学、计算语义学、计算话语学等分支学科研究内容；句法剖析的主要算法；语义的各种表示方法。

### 2. 语言信息处理技术部分

正则表达式；自动分词；词类标注；有限状态形态分析；N元语法；词语相似度计算；语音合成与语音识别、信息抽取、问答系统、自动文摘、对话与会话智能代理、机器翻译等自然语言处理应用技术的基本原理。

## 三、考试形式

考试形式为闭卷，时间为 90 分钟，满分 100 分。

题型：计算语言学基础理论部分（论述题，50 分），语言信息处理技术部分（问答题，50 分）。

#### 四、参考书目

1. 《计算语言学》（修订版），刘颖，清华大学出版社，2014 年。

2. 《自然语言处理基本理论和方法》，陈鄞，哈尔滨工业大学出版社，2017 第 2 版。

3. 《自然语言处理综论（第 2 版）》，Daniel Jurafsky, James H. Martin 著，冯志伟，孙乐译，电子工业出版社，2018 年。

4. 《自然语言处理与计算语言学》，Bhargav Srinivasa-Desikan, 何炜译，人民邮电出版社，2020 年。