

最短路径中文分词系统用户手册

罗毅凡

学号：2024202715

1 系统概述

本系统是一个基于 N-最短路径 (N-Shortest Path) 算法的高效中文分词工具。它通过构建有向无环图 (DAG)，结合大规模词典统计信息，利用动态规划思想寻找句子中概率最高的 N 种切分方式。

2 编译与运行

2.1 1. 准备数据

请确保目录下包含核心词典文件（通常命名为 `dict.txt`）。词典格式应为：

词语	词频	词性
----	----	----

例如：

做人	500	v
梦想	300	n
咸鱼	50	n

2.2 2. 运行程序

运行程序，即可启动命令行交互，实时输入中文文本实现分词

3 操作指南

3.1 交互模式

程序启动后进入交互模式，出现提示符：`>` 最短路径匹配分词，请输入中文文本。

步骤：

- 在提示符后直接输入待分词的中文句子。

2. 按下 Enter 键提交。
3. 系统将输出最优分词结果 (Top-1) 以及备选路径 (Top-N)。
4. 输入 `ctrl + C` 退出程序。

3.2 运行示例

Input:

› 做人如果没梦想，跟咸鱼有什么区别

Output:

N-最短路径分词结果 (Top 3)

Result 1: 做人/如果/没/梦想/， /跟/咸鱼/有/什么/区别/。

Result 2: 做/人/如果/没/梦想/， /跟/咸鱼/有/什么/区别/。

Result 3: 做人/如果/没/梦/想/， /跟/咸鱼/有/什么/区别/。