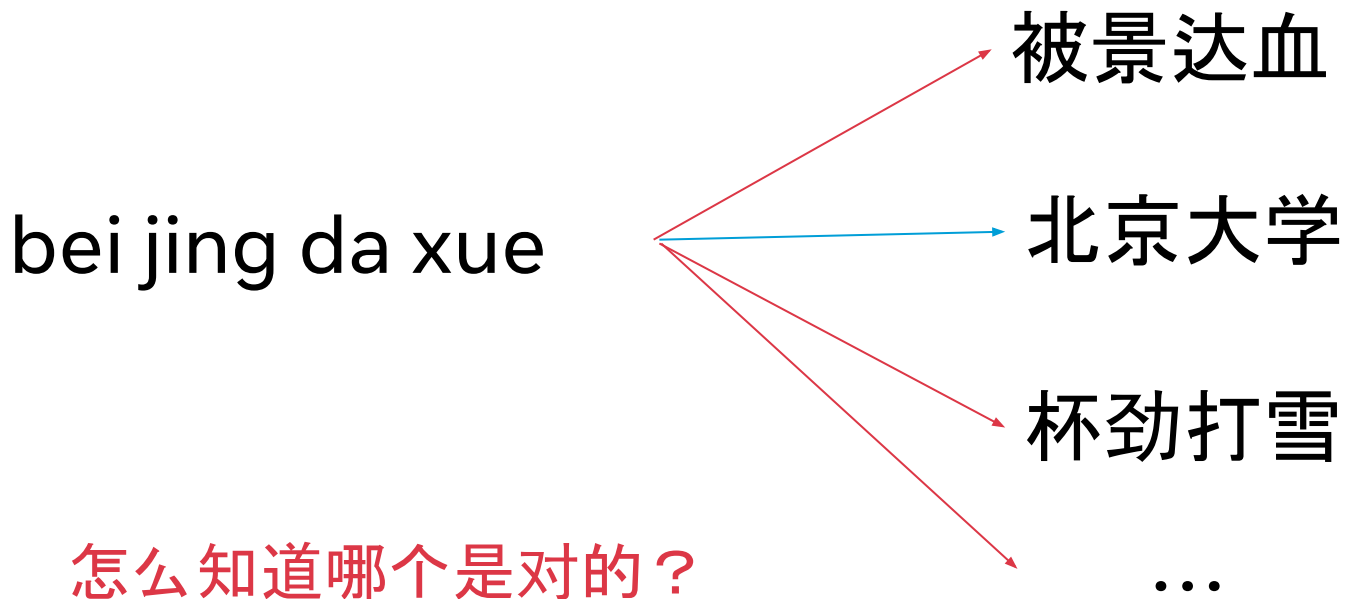


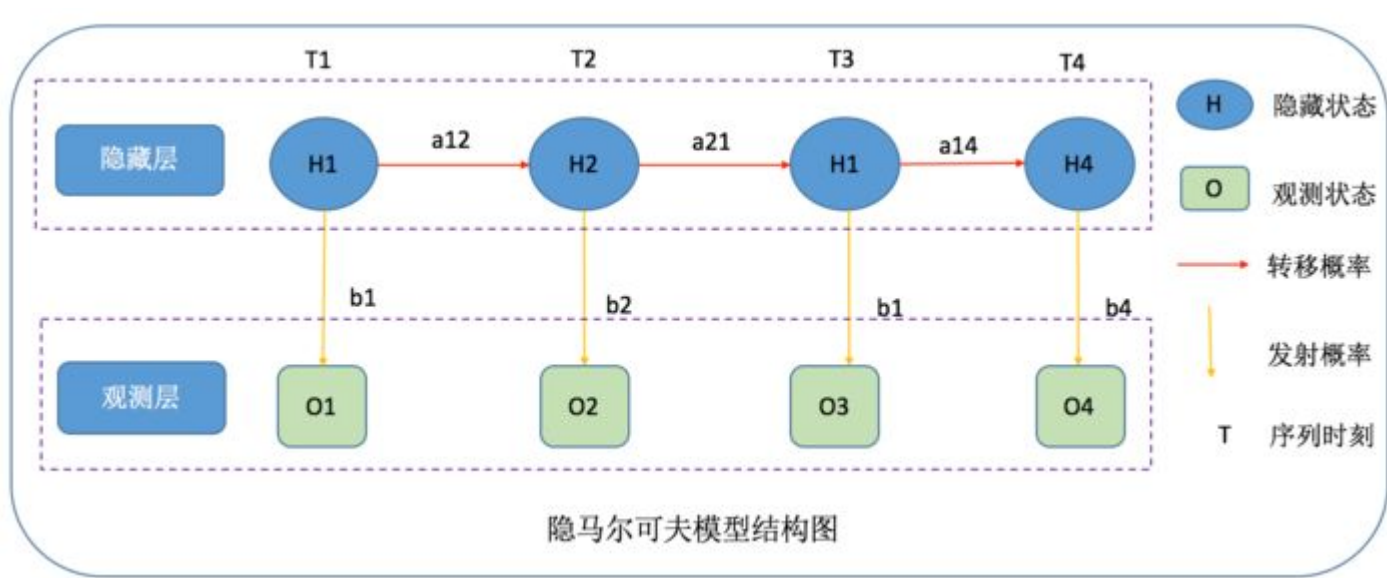
从拼音到汉字 转换方法 大作业展示



马尔可夫模型训练

汉字

拼音



Emission Probabilities / 发射概率: 获取从拼音(观察)到汉字(隐含状态)的可能性 `{ '我': { 'wo': 1.0 }, '了': { 'liao': 0.5, 'le': 0.5 } }`

Transition Probabilities / 转移概率: 获取一个汉字被另一个汉字跟踪的可能性 `{ '我': { '们': 0.3, '是': 0.8 }, '了': { } }`

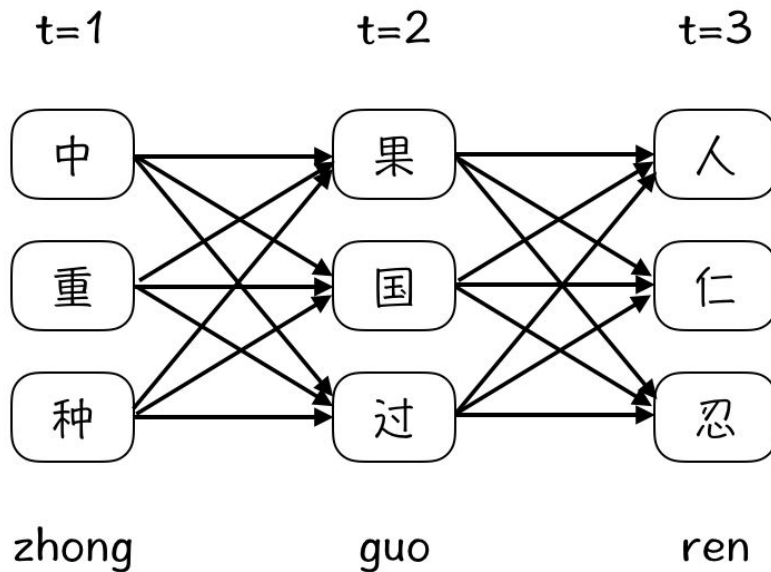
Start Probabilities / 开始概率: 获取以汉字开头的序列的可能性 `{ '我': 0.006, '了': 0.00001 }`

Viterbi 解码



北京大学
PEKING UNIVERSITY

根据一个拼音序列，用
计算效率高的方式找到
最可能的汉字序列。



Demo