

Cip基本概念

中文信息处理研究内容

- 字符层

-

- 内容层

-

现代汉语的特点

- 字
 - 象形
 - 发音少
- 字汇
- 词汇
- 字频、词频
 - 有何作用？

中文在计算机内的表示

- ASCII
- GB2312
 - 区位码、国标码、机内码
- GBK
- Big-5
- Unicode
- ISO10646
- GB13000

汉字简繁-繁简转换

- 如何做？
- 难度在哪里？

Base64编码

- 为何需要它？
- 如何编码？
- 如何解码？

汉字编码基本概念

- 码元
- 码长
- 编码空间和编码效率
- 单码和重码

汉字编码分类

- 流水码
 - 有何缺点？
- 音码
- 形码
- 音形码/形音码
 - 有何优点？
 - 有何缺点？

中文平台和中文操作系统

- 中文平台和中文操作系统区别？
- 中文操作系统设计方法？

汉字输入

- 汉字输入如何分类？
- 汉字输入码对照表有哪些常用结构？
 - 各有什么优缺点？

思考题

- P20 4,6,8
- P38 3,5,11
- P50 2,3
- P73 5,12
- P86 1,2
- P124 1,4,5