第二章　计算机信息的表示和存储

思维导图 · 中心向四支放射 · 对应知识点小结另见同名文件

信息的表示和存储

全章结构总纲

一、进位计数制

- **数制三要素**
  - 基数R：用R个数码，逢R进一
  - 数位：以小数点为基准
  - 位权 = R的整数次幂R^i / R^{-j}
- **非进位计数制**
  - 数码与位置无关：罗马数字
  - Ⅰ Ⅴ Ⅹ Ⅰ Ⅰ Ⅰ Ⅰ
- **常用进位制**
  - 十进制：逢十进一
  - 六进制：时/分/秒、度/分/秒
  - 二十四进制：时 | 十二进制：年月、12小时制
  - 二进制：计算机内部唯一使用
- **按权展开式**
  - $(666.66)_{10} = 6\times10^2+6\times10^1+6\times10^0$
  - $+6\times10^{-1}+6\times10^{-2}$

二、数制转换

- **十 → 二/八/十六**
  - 整数：除基数取余，逆序排列
  - 小数：乘基数取整，顺序排列
- **二/八/十六 → 十**
  - 按位权值展开求和
- **二 ↔ 八**
  - 每3位一组→一位八进制
  - 整数右→左、小数左→右
  - 不足在两端补0
- **二 ↔ 十六**
  - 每4位一组→一位十六进制
  - 补0规则同上
- **字母标识**
  - 二进制B、八进制Q
  - 十进制D、十六进制H
- **典型例题**
  - $(110.01)_2 = 6.25$
  - $(63.24)_8 = (110011.0101)_2$
  - $(1100101001.1101)_2 = (1451.64)_8$

三、信息编码

- **ASCII 码 ★**
  - 7位二进制表示1个字符，$2^7 = 128$个
  - 可打印95个：编码32～126
  - 控制字符33个：0～31和127
  - 单字节存储，最高位为0
  - 'A' = 1000001B = 41H = 65D
- **汉字编码四类 ★**
  - 交换码（国标码GB）：低7位，最高位0
  - 机内码：各字节最高位置1
  - 输入码（外码）：拼音、五笔、区位码
  - 字形码：点阵或矢量，用于显示打印
- **输入码四类**
  - 数字码：区位码、国标码、电报码
  - 拼音码：简单易学，重码率高
  - 字形码：等长、重码低、速度快
  - 音形码：自然码、太极码
- **ASCII 对比规律**
  - 数字< 大写字母 < 小写字母
  - 小写a = 61H，则d = 64H
  - 数字2 = 50，则数字5 = 53

四、存储与汉字标准

- **存储术语 ★**
  - 位bit：最小计量单位（一个0或1）
  - 字节Byte：8位位模式，容量基本单位
  - 字Word / 字长：CPU一次处理的位数
  - 内存地址：无符号二进制整数
- **容量换算（1024）**
  - 1B = 8bit | 1KB = 1024B | 1MB = 1024KB
  - 1GB = 1024MB | 1TB = 1024GB | 1PB = 1024TB
  - 1EB = 1024PB | 1ZB = 1024EB | 1YB = 1024ZB
- **GB2312-80**
  - 汉字6763个：一级3755（拼音）
  - 二级3008（部首）
  - 其他682个，合计7445
- **区位码**
  - 区位码 = 区号×100 + 位号
  - '国'在25区90位→2590
  - 每区可容纳94个汉字
- **GBK / GB18030**
  - GBK与GB2312完全兼容
  - GB18030采用单、双、四字节
  - 收录繁简汉字2万7千余字
  - 编码空间150万个码位以上
- **字形点阵**
  - 有笔画格记1，无笔画记0
  - 16×16点阵：每行2字节×16行 = 32字节