

# 编程入门基础

—— 编码方式

**主讲教师：耿宇航**

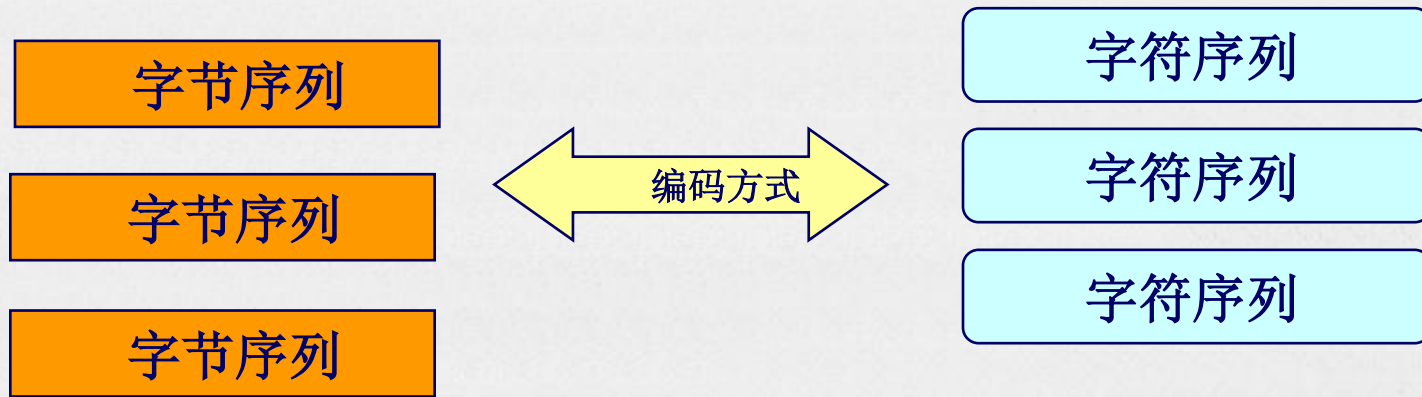


# 数据如何转化为byte流？

- 需要某种规则----编码方式
- 最为常用的编码方式：字符与字节序列的关系
- 因为历史原因，编码方式不是唯一一种。
  - 字符的转化问题，就存在多种
  - unicode 之 utf-8 国际标准
  - gb2312 gbk 国家标准
  - big5 台湾等

# 字节编码与字符的转换

- **String**类提供了转化手段
  - `String(byte[] b, int start, int len, String codeName)`
- **InputStreamReader**提供了直接读字符的手段
  - `int read()` 返回-1表示流结束, 否则是字符(需要强制转换)

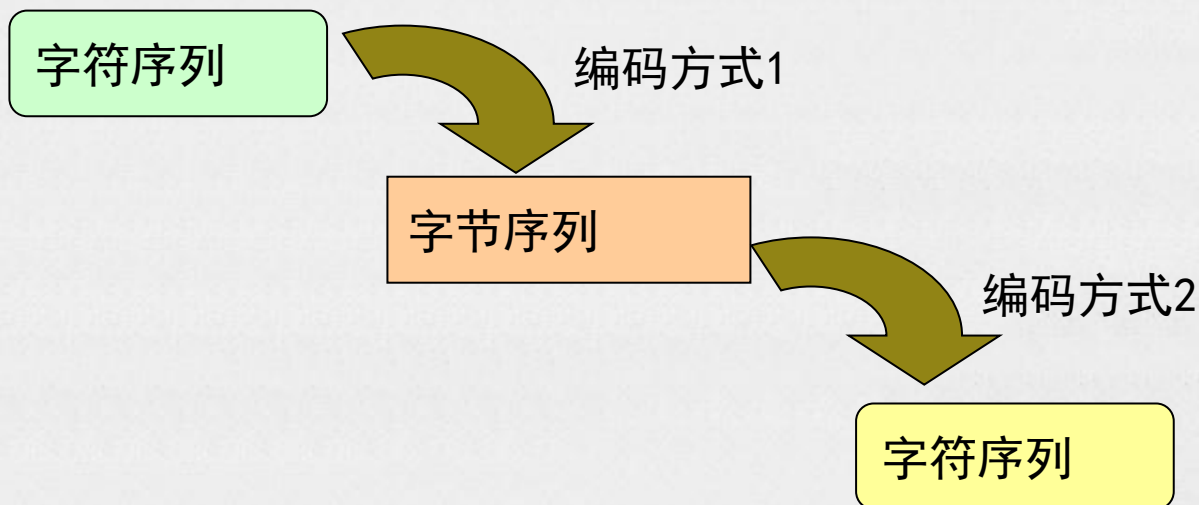


# 编码方式的作用

- 把某个字符转换为一个字节序列
  - 编码方式不同，则转换结果不同
  - 有的是固定长度，有的是可变长度
- 把一个字节序列解释为若干字符
  - 编码方式不同，则解释结果不同
  - 甚至可能会无法解释

# 乱码问题

- 因为编码方式的多样性，会导致乱码问题！
- 字节序列中并不会注明是用哪种编码方式，因而我们在进行双向转换的时候如果选择了不当的编码方式，就会出现怪异结果





# 谢 谢 ！