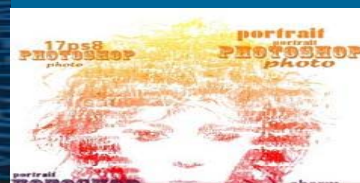




北京理工大学
BEIJING INSTITUTE OF TECHNOLOGY

第二讲

计算机信息数字化基础



□ 通用字符编码集UCS

■ 字符信息编码与标准交换

□ 西文字符的编码与标准交换

□ 汉字信息的编码与交换

□ 通用字符编码集UCS

国际标准ISO 10646定义了**通用字符集(UCS)**，它包含了世界上大多数可书写的字符系统。

UCS定义了**两种编码**格式：

UCS-4：对每一个字符采用4字节编码

UCS-2：对每一个字符采用2字节编码

UCS与其他字符集都是**双向兼容**的，即相互转换不丢失信息。

□ 通用字符编码集UCS

ISO 10646定义了一个31位的字符集，迄今为止只分配了前65534个码位。

这个UCS分为两个16位子集：

- 16位BMP：基本多语言面（ Basic Multilingual Plane ）；
- 被编码在16位BMP以外的字符都属于非常特殊的字符（ 比如象形文字 ）。