

ASCII and Unicode Basics SEC (Semester-1)

By

Santosh Kumar Lal

Dept. of Commerce

Sariya College, Suriya

ASCII Basics

ASCII (American Standard Code for Information Interchange) is a character encoding standard developed in the early 1960s to represent text in computers and communication devices.

- ◆ **Key Points:**

- **Developed by:** American National Standards Institute (ANSI)
- **Year:** 1963
- **Bit Size:** 7 bits
- **Total Characters:** 128 (0–127)
- **Includes:**
 - English alphabets (A–Z, a–z)
 - Digits (0–9)
 - Punctuation symbols (?, !, &, etc.)
 - Control characters (e.g., NULL, ENTER, TAB)

◆ **ASCII Structure:**

- **0–31:** Control characters
- **32–126:** Printable characters
- **127:** Delete (DEL)

◆ **Limitations:**

- Supports only English language characters.
- Cannot represent characters from other languages like Hindi, Chinese, Arabic, etc.
- Extended ASCII (8-bit, 256 characters) was introduced later but still limited.

◆ **Example:**

- A → 65
- a → 97
- 0 → 48

यूनिकोड (Unicode) की मूल बातें

Unicode एक अंतरराष्ट्रीय कैरेक्टर एन्कोडिंग मानक है जो दुनिया की लगभग सभी भाषाओं के अक्षरों और चिन्हों को कंप्यूटर में दर्शाने के लिए बनाया गया है।

- ◆ मुख्य बिंदु:

विकसित किया गया द्वारा: Unicode Consortium

- शुरुआत: 1991
- कैरेक्टर संख्या: 1,49,000+ (लगातार बढ़ती हुई)
- बिट साइज़: 8-बिट, 16-बिट, 32-बिट (UTF-8, UTF-16, UTF-32)

◆ Unicode की विशेषताएँ:

- विश्व की लगभग सभी भाषाओं का समर्थन
- हिंदी, संस्कृत, तमिल, चीनी, अरबी आदि भाषाएँ शामिल
- Emojis 😊 और विशेष चिन्ह भी शामिल

◆ सामान्य Unicode एन्कोडिंग प्रकार:

- UTF-8: सबसे लोकप्रिय, वेब में व्यापक उपयोग
- UTF-16: Windows और Java में उपयोग
- UTF-32: फिक्स्ड-लेंथ एन्कोडिंग

◆ उदाहरण:

- A → U+0041
- ₹ (भारतीय रुपया चिन्ह) → U+20B9
- क → U+0915

ASCII और Unicode में अंतर

आधार	ASCII	Unicode
भाषाएँ	केवल अंग्रेज़ी	सभी प्रमुख भाषाएँ
कैरेक्टर सीमा	128	1 लाख से अधिक
बिट साइज़	7-बिट	8/16/32-बिट
उपयोग	प्रारंभिक कंप्यूटर	आधुनिक सिस्टम और इंटरनेट

निष्कर्ष (Conclusion)

ASCII कंप्यूटर के शुरुआती दौर का एक महत्वपूर्ण कैरेक्टर एन्कोडिंग मानक था, लेकिन इसकी सीमाओं के कारण Unicode विकसित किया गया। आज के समय में Unicode वैश्विक संचार और इंटरनेट के लिए अत्यंत आवश्यक है।