ASCII & Unicode

IGCSE Computer Science Revision Sheet

■ **Big Idea:** ASCII and Unicode are character encoding systems — they allow computers to represent letters, numbers, symbols, and emojis in binary.

What is ASCII? & What is Unicode?

ASCII	Unicode
American Standard Code for Information Interchange Uses **7 bits** to represent each character. Can represent **128** characters. Includes uppercase and lowercase letters, numbers, and punctuation. English language only. Example: Capital A = denary **65** = binary **01000001**	**Universal Character Encoding Standard** Uses **variable length** (e.g. UTF-8 uses 8–32 bits). Can represent **millions** of characters. Supports worldwide languages and writing systems. Includes emojis and special symbols. Used in all modern websites, apps, and operating systems.

KEY FACT: ASCII = 7 bits, 128 characters, English only. Unicode = variable bits, millions of characters, all languages.

Teacher note: The most common exam error is confusing ASCII bit size (7) with byte size (8). ASCII uses 7 bits — though often stored in 8. Be precise.

ASCII — Sample Character Values

Character	Denary Value	Binary Value	Character	Denary Value	Binary Value

A	65	01000001	a	97	01100001
B	66	01000010	b	98	01100010
C	67	01000011	c	99	01100011
Z	90	01011010	z	122	01111010
0	48	00110000	Space	32	00100000

EXAM TIP: Uppercase A = 65. Lowercase a = 97. The difference between uppercase and lowercase is always 32 — a useful pattern to remember.

Teacher note: Ask students: if A = 65, what is B? What is C? This quick exercise builds confidence with ASCII values and shows the sequential pattern.

ASCII vs Unicode — Full Comparison

Feature	ASCII	Unicode
Bit size	7 bits	Variable (UTF-8: 8–32 bits)
Characters	128	Over 143,000 (millions possible)
Languages	English only	All world languages and scripts
Includes emojis	No	Yes
File size	Smaller	Larger (more bits per character)
Modern use	Legacy systems	Standard for all modern software
Backwards compatible	—	Yes — first 128 Unicode values match ASCII

EXAM TIP: ASCII is smaller and older. Unicode supports far more characters and is the modern standard. Use both terms precisely in compare questions.

www.futurelogiceducation.com
hello@futurelogiceducation.com

Teacher note: The backwards compatibility point often surprises students — Unicode was designed to include ASCII so existing systems could transition without breaking.

Why Unicode Replaced ASCII

✔ **128 characters was not enough.**
As computing spread globally, 128 characters could not cover non-English languages, scripts, or symbols.

✔ **Global communication required it.**
The internet needed a single standard that worked for Arabic, Chinese, Hindi, and thousands of other languages.

✔ **Emojis and symbols needed support.**
Modern communication includes emoji, mathematical symbols, and currency signs that ASCII cannot represent.

✔ **UTF-8 kept file sizes manageable.**
UTF-8 (a common Unicode format) uses 1–4 bytes per character — English text stays the same size as ASCII.

Teacher note: Ask students: "Why can't we just stick with ASCII?" This discussion naturally covers globalisation, modern apps, and the scale of the internet.

Quick Check Questions

1.	How many bits does ASCII use to represent each character?
2.	How many characters can ASCII represent?
3.	Give two reasons why Unicode was needed to replace ASCII.
4.	What is the ASCII denary value of capital A?
5.	Explain why Unicode files are often larger than ASCII files.
6.	What does UTF-8 stand for and what is it used for?

Answers on the next page →

Answer Guide

1.	ASCII uses 7 bits to represent each character.
2.	ASCII can represent 128 different characters.
3.	Any two of: ASCII only supports English; Unicode supports all world languages and scripts; ASCII cannot represent emojis or modern symbols; the internet required a global standard.
4.	The ASCII denary value of capital A is 65.
5.	Unicode uses more bits per character than ASCII because it needs to represent a far greater number of characters, therefore each character requires more storage space, increasing file size.
6.	UTF-8 stands for Unicode Transformation Format – 8-bit. It is a common Unicode encoding format that uses between 1 and 4 bytes per character, and is the standard format for web pages and modern software.

■ **FutureLogic Summary:** ASCII = 7 bits, 128 characters, English only. Unicode = variable bits, millions of characters, all world languages. Unicode replaced ASCII because the world needed a global standard.