ICS 2019 Problem Sheet #6

Mahiem Agrawal
CS '22

6.1

a)

53 - S	69 - i	6d - m	61 - a
69 - i	65 - e	65 - e	6e - n
6d - m	64 - d	2c - ,	63 - c
70 - p	20 - _	20 - _	65 - e
6c - l	74 - t	62 - b	2e - .
69 - i	6f - o	65 - e	0a - newline
63 - c	20 - _	63 - c	2d - -
69 - i	74 - t	6f - o	20 - _
74 - t	68 - h	6d - m	4a - J
79 - y	65 - e	65 - e	6f - o
2c - ,	20 - _	73 - s	6e - n
20 - _	65 (~~space~~) - e	20 - _	20 - _
63 - c	78 - x	65 - e	46 - F
61 - a	74 - t	6c - l	72 - r
72 - r	72 - r	65 - e	61 - a
72 - r	65 - e	67 - g	6e - n

~~6d~~ 6b - k ~~Final word~~.
6c - l
~~6d~~ - i
6e - n
0a - (newline)

∴ Simplicity carried to the extreme, becomes elegance
-Jon Franklin

b)

UTF-8	Unicode	Name	Symbol
7c	U+007C	Vertical line	|
7b	U+007B	left-curly-Bracket	{
2d	U+002D	Hyphen	-
7d	U+007D	Right curly Bracket	}
e2 88 ~~8d~~	U+222A	Union	∪
7b	U+007b	left curly Bracket	{
e2 8892	U+2212	Minus	−
7d	U+007d	Right curly Bracket	}
e2 88 aa	U+222A	Union	∪
7b	U+007B	left curly Bracke	{
c2 ad	U+00AD	Soft HYPHEN	-
7d	U+007D	Right curly bracket	}
e2 88 aa	U+222A	Union	∪
7D	U+007B	left curly bracket	{
e2 80 93	U+2013	En-dash	-
7d	U+007B	Right curly Bracket	}
e2 88 aa	U+222A	Union	∪
7b	U+007B	Left Curly Bracket	{
e2 80 91	U+2011	Non-Breaking Hyphen	-
7d	U+007D	Right curly Bracket	}
e2 88 aa	U+22A	Union	∪
7b	U+007B	left curly bracket	{
e2 80 94	U+2014	Em Dash	—
7d	U+007D	Right Curly Bracket	}

e2 88 aa	U+222A	Union	∪
7b	U+007B	left coily Bracket	{
e2 80 92	U+2012	Figole Dash	‒
7d	U+007D	Right Coily Bracket	}
7c	U+007C	Veitical line	|
20	U+0020	Spare	
3d	U+003D	Equal	=
20	U+0020	Spare.	
37	U+0037	Digit Seven	7
0a	U+000A	End of line.	

|{-}∪{-}∪{}∪{-}∪{—{∪{-}|=7

c) UTF-32 uses 5 bytes to store single character
UTF-8 uses 3 bytes to store single charact

In UTF-32 Chinese text uses
$5 \times 800\,000$
$= 4000\,000$ bytes.

In UTF-8 chinese text uses.
$3 \times 800\,000$
$=$
$2\,400\,000$ bytes

6.2) Converting them

a) 2019-10-15T13:15:00
2019-10-13T17:15:00
2019-10-13T13:1500.
2019-10-13T17:15:06
2019-10-13T7 13:15:00
2019-10-13T17:15:00

∴ Dates 3, 5 are same and Dates 2, 4 and 6 are same.

b) As this helps us represent a unique case when the the time in UTC is known but the offset to the local time is unknown.

c) As all Unix's signed 32-bit integer started from 1970 the time format

03:14:07 on 19 January 2038 ~~cannot~~ is be ~~represented as all bits are 1~~ the latest time that can be shown. other dates are not possible, therefore we call it Year 2038 problem.

The solution to this is using more bits such as a 64-bit system to store the date.