

Programming Assignment #3

Due: 12월 12일 23:59

목표: 문자열 인코딩/디코딩 하기

(1) 개요

1바이트짜리 문자는 총 256가지의 경우가 있으며, 이들 중 플랫폼에 상관 없이 언제 어디서나 눈으로 볼 수 있는 문자는 일부에 불과하다. 예를 들어, 아스키코드 값이 65 이상 90 이하이면 A부터 Z 중 하나이므로 어느 기기를 쓰든지, 어느 언어권에서 어떤 운영체제를 쓰든지 항상 보일 것이지만, 아스키코드 값이 128 이상 255 이하이면 눈에 보이지 않을 것이다. 또한, 2바이트짜리 문자 (예: 한글, 히라가나, 가타가나, 한자 등)는 그 언어권이 아닌 운영체제에서 보면 깨져서 보일 것이다. 하지만, 그런 문자들을 문자 자체로는 읽을 수 없어도, 어느 플랫폼에서나 눈으로 확인할 수 있어야 하는 경우가 존재한다. 이를 위하여 다양한 인코딩/디코딩 방법이 개발되었는데, 본 프로젝트에서는 그 중 한 방법을 C로 구현해 보고자 한다.

참고로, 원본 정보를 다른 것으로 변환하는 것을 인코딩이라 하고, 변환된 것을 원본으로 되돌리는 것을 디코딩이라 한다.

(2) 인코딩 하기

- 인코딩은 다음과 같은 과정을 거친다.
- A) 원본 문자열을 맨 앞에서부터 문자 세 개씩 나눈다.
 - B) 각 세 개짜리 문자는 총 24 ($=8 \times 3$)비트로 구성되어 있다. 이를 앞에서부터 차례로 6비트씩 네 개로 나눈다.
 - C) 각 6비트짜리들은 다음과 같은 규칙에 의하여 문자로 바뀐다.

6비트값	문자	6비트값	문자	6비트값	문자	6비트값	문자
000000	A	010000	Q	100000	g	110000	w
000001	B	010001	R	100001	h	110001	x
000010	C	010010	S	100010	i	110010	y
000011	D	010011	T	100011	j	110011	z
000100	E	010100	U	100100	k	110100	0
000101	F	010101	V	100101	l	110101	1
000110	G	010110	W	100110	m	110110	2
000111	H	010111	X	100111	n	110111	3
001000	I	011000	Y	101000	o	111000	4
001001	J	011001	Z	101001	p	111001	5
001010	K	011010	a	101010	q	111010	6
001011	L	011011	b	101011	r	111011	7
001100	M	011100	c	101100	s	111100	8
001101	N	011101	d	101101	t	111101	9
001110	O	011110	e	101110	u	111110	+
001111	P	011111	f	101111	v	111111	/

예를 들어, 세 개짜리 문자가 Sm5이면, 다음과 같이 변환된다.

원본	S								m								5							
ASCII Code	0	1	0	1	0	0	1	1	0	1	1	0	1	0	0	1	1	0	1	0	1	0	1	
6비트	0	1	0	1	0	0	1	1	0	1	1	0	1	0	0	1	1	0	1	0	1	0	1	
인코딩결과	U				2				0				1											

- 따라서 인코딩 결과는 U201이 된다.
- D) 만일 원본 문자열의 길이가 3의 배수가 아니면 마지막에는 문자가 1개 또는 2개가 남는데, 그런 경우는 인코딩 결과의 빈 곳에 =을 채워준다.
- 예를 들어, 마지막에 남은 문자가 S이면, 다음과 같이 변환된다.

원본	S																								
ASCII Code	0	1	0	1	0	0	1	1																	
6비트	0	1	0	1	0	0	1	1	0	0	0	0													
인코딩결과	U								w								=								

- 따라서 인코딩 결과는 Uw==이 된다.
- 다른 예로, 마지막에 남은 문자가 Sm이면, 다음과 같이 변환된다.

원본	S								m															
ASCII Code	0	1	0	1	0	0	1	1	0	1	1	0	1	1	0	1								
6비트	0	1	0	1	0	0	1	1	0	1	1	0	1	1	0	1	0	0						
인코딩결과	U								2								0							
																	=							

따라서 인코딩 결과는 U20= 이 된다.

참고로, 인코딩 결과의 길이는 무조건 4의 배수이다.

(3) 디코딩 하기

위 (2)의 역순으로 하면 된다. 이 때, 입력되는 문자열의 길이는 무조건 4의 배수가 되어야 한다.

(4) 프로그램 실행 방법

프로그램을 실행하면 인코딩, 디코딩 여부를 결정하게 한다. 즉, 다음과 같은 화면이 뜬다.

```
C:\Windows\system32\cmd.exe
Enter a number (1: Encoding / 2: Decoding):
```

만일 1, 2 이외의 다른 수를 입력하면 다시 입력하라는 메시지가 뜬다.

```
C:\Windows\system32\cmd.exe
Enter a number (1: Encoding / 2: Decoding): 3
Wrong number! (1/2): 10
Wrong number! (1/2): -1
Wrong number! (1/2):
```

1을 입력하면 1024바이트 이하의 문자열을 입력하라는 메시지가 뜬다.

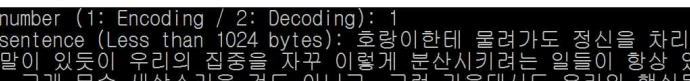
```
C:\Windows\system32\cmd.exe
Enter a number (1: Encoding / 2: Decoding): 1
Enter a sentence (Less than 1024 bytes):
```

이 때, 입력하면 다음과 같은 결과가 나온다.

```
C:\Windows\system32\cmd.exe
Enter a number (1: Encoding / 2: Decoding): 1
Enter a sentence (Less than 1024 bytes): My grades are down from A's to D's. I'm
way behind in history. I lost myself in fantasies of you and me together. I don
't know why--I--I but dreaming's all I do. I won't get by--I--I on mere imagination.
Upside down. Bouncing off the ceiling. Inside out. Stranger to this feeling. Go
t no clue what I should do. But I'll go crazy if I can't get next to you.

Encoded Message: TXkgZ3JhZGVzIGFYZSBkb3duIGZyb20gQSDzIHRvIEQncy4gSSdtIHdheSBiZWWh
pbm0gaW4gaGlzZG9yeS4gSSBsb3N0IG15c2VzSiBpbIBmYW50YXNpZXMgb2YgeW91IGFuZCkgZSB0b2d
lIGh1ci4gSSBkb24ndCBrbm93IHdoeS1JLUkgybnV0IGRyZWFiaw5nJ3MgYXNpIEkgZG8uIEkgd29uJ3R0
gZ2V0IGJ5LlUkTSB5bnIbZTJJIIGltYWdpbmF0aW9uLlBVCm9pZGUGZG93b1I4gQm91bmNpbmN0b2ZmIHR0
eSBZbJWlIaW5nLlBJbnNpZGUGb3V0LlBTdHJhbmdlciB0bYB0aGlzIGZlZWxpbmculEdvdCBubyBjbHV
lIHdoYXQgSSBZag91bGQgZG8uIEU1dCBJJ2xslGdvIGNyYXp5IGlmIEkgY2FuZ3QgZ2V0IG5leHQgdG8
geW91lg==
계속하려면 아무 키나 누르십시오 . . .
```

영어 뿐 아니라 한글이나 한자도 될 것이다.



C:\Windows\system32\cmd.exe

```

Enter a number (1: Encoding / 2: Decoding): 1
Enter a sentence (Less than 1024 bytes): 후랑이한테 물려가도 정신을 차리면 된다
        그런 말이 있듯이 우리의 집중을 자꾸 이렇게 분산시키려는 일들이 항상 있을 거
        다, 으레. 그게 무슨 새삼스러운 것도 아니고. 그런 가운데서도 우리의 핵심목표는
        올해 달성해야 될 것은 이것이다 하는 것으로 정신을 차리고 나아가면 우리의 에너지
        를 분산시키는 걸 해낼 수 있다는 마음을 가지셔야 될 거라고 생각합니다.

Encoded Message: yk02+8DMx9HF1yC5sLfBskGt1SDBpL3FwLsgwve4rrjp1LXltNmOwiCx17ex1L
1iWmWgwNa17cDM1L/suk7AxyDB/chfwLsgwNqy2SDAZLe4sN0gutC76r3DxbC3wbTC1MDptenAzCDH1
7vz1MDWwLsgsMW02SgwglI3uS4gsdew1CC05q7281Lv1u++9urevv+4gsm21tSC+xrTPs00u1LHXt7Eg
sKG/7rWl vk21tSC/7L1uWmcgx9m9yb jxx6W0wiC/w8fY1lTevLRH2L7f1LXJ1LDNwLogwMywzcDMtNk
gx8+0wiCwzc2k4184gwaS9xc71ML3uK6w7SCzqr7GSGK46SC/7L1uWmcgv6GzysH2uKYgutC76r3Dxb
C0wiCwySDH2L0+1Lz21MDW1NmOwiC4tsC9wLsgsKHB9rzFvt8gtckgsMW287Dt1Lv9sKLH1bTPtNku
계속하려면 아무 키나 누르시오 . . .
  
```

마찬가지로, 2번을 선택하면 다음과 같이 1365 미만의 바이트만큼 입력하라는 메시지가 뜨고, 인코딩된 문자열을 넣으면 디코딩이 될 것이다.

```
C:\Windows\system32\cmd.exe
```

```
Enter a number (1: Encoding / 2: Decoding): 2  
Enter a ciphertext (Less than 1365 bytes): TXkgZ3JhZGVzIGFySSBkb3duIGZyb20gQSDzIHVrIEQncy4gSSdtIHdheSBiZWlhbWQgaW4gaGlzdG9yeS4gSSBsb3N0IG15c2VsZiBpbmYw50YXNoZXMgb2YgeW91IGFuZCBtZSB0b2dldGhlci4gSSBkb24ndCBrbm93IHdoeS1JLUKgyNV0IGRyZWftalw5nJ3MgYWxslEkqZG8uIEkqd29uJ3QgZ2V0IGJ5LUktSSBvb3IbtXJJIGltYWdpbmF0aW9uLlBVcHNpZGUgZG93bi4gQm91bmNpbmcgb2ZmIHRobzSiZWlsaW5nLiBJbnNpZGUgb3V0LiBTdHJhbmdldiB0byB0aGlzIGZlZWxpbnmcuLEdvdcUbyb3IjbnVlIHdoeXQgSSBzaG91bGQgZG8uIEJ1dCBBJ2xsIGdvIGNyYXp5IGlmIEkgY2FuJ3QgZ2V0IG5leHQgdG8geW91Lg==  
  
Decoded Message: My grades are down from A's to D's. I'm way behind in history. I lost myself in fantasies of you and me together. I don't know why—I-I but dreaming's all I do. I won't get by—I-I on mere imagination. Upside down. Bouncing off the ceiling. Inside out. Stranger to this feeling. Go to no clue what I should do. But I'll go crazy if I can't get next to you.  
계속하려면 아무 키나 누르십시오 . . .
```

```
C:\Windows\system32\cmd.exe
Enter a number (1: Encoding / 2: Decoding): 2
Enter a ciphertext (Less than 1365 bytes): yK02+8DMx9HF1yC5sLfBsKG1tSDBpL3FwLsg
wve4rrjpILXltNm0wiCx17exILi7wMwgnA17cDMIL/suK7AxyDB/cHfwLsgwNqy2SDAzLe4sNQgutC
76r3DxbC3wbTCIMDPtenAzCDH17vzIMDWwLsgsMW02Swgwi3uS4gsdew1CC5q728ILv1u++9urevv+
4gsM21tSC+xrTPs00uILHXt7EgskG/7rWlvK21tSC/7LiuwMcgx9m9ybjxx6W0wiC/w8fYILTeVlrH2
L7fILXJILDNwLogwMywzcDMtNkgx8+0wiCwzcC4t84gwaS9xcC7IML3uk6w7SCzqr7GskG46SC/7Liu
wMcgv6GzysH2uKYgutC76r3DxbC0wiCwySDH2L0+ILz2IMDWtNm0wiC4tsC9wLsgskHB9rzFvt8gtck
gsMW287DtILv9sKLH1bTPtNku

Decoded Message: 호랑이한테 물려가도 정신을 차리면 된다는 그런 말이 있듯이 우리
의 집중을 자꾸 이렇게 분산시키려는 일들이 항상 있을 거다, 으레. 그게 무슨 새삼
스러운 것도 아니고. 그런 가운데서도 우리의 핵심목표는 올해 달성해야 될 것은 이
것이다 하는 것으로 정신을 차리고 나아가면 우리의 에너지를 분산시키는 걸 해낼 수
있다는 마음을 가지셔야 될 거라고 생각합니다.
계속하려면 아무 키나 누르십시오 . . .
```

하지만, 디코딩의 경우, 입력한 메시지의 길이가 4의 배수가 아니면 다음과 같은 오류메시지가 뜬다.

```
C:\Windows\system32\cmd.exe
Enter a number (1: Encoding / 2: Decoding): 2
Enter a ciphertext (Less than 1365 bytes): yK02+8DMx9HF1yC5sLfBsKG1tSDBpL3FwLsg
wve4rrjpILXltNm0wiCx17exILi7wMwgnA17cDMIL/suK7AxyDB/cHfwLsgwNqy2SDAzLe4sNQgutC
76r3DxbC3wbTCIMDPtenAzCDH17vzIMDWwLsgsMW02Swgwi3uS4gsdew1CC5q728ILv1u++9urevv+
4gsM21tSC+xrTPs00uILHXt7EgskG/7rWlvK21tSC/7LiuwMcgx9m9ybjxx6W0wiC/w8fYILTeVlrH2
L7fILXJILDNwLogwMywzcDMtNkgx8+0wiCwzcC4t84gwaS9xcC7IML3uk6w7SCzqr7GskG46SC/7Liu
wMcgv6GzysH2uKYgutC76r3DxbC0wiCwySDH2L0+ILz2IMDWtNm0wiC4tsC9wLsgskHB9rzFvt8gtck
gsMW287DtILv9sKLH1bTPtNk

The input text cannot be decoded!
계속하려면 아무 키나 누르십시오 . . .
```

(5) 주의 사항

- 컴파일이 안 되면 0점임. 제출 직전에 컴파일 되는지 반드시 확인할 것!
- FTP에 PA3 폴더를 만들고 그 안에 c 파일만(헤더파일도 있으면 h파일도) 제출할 것
- 하루씩 Delay 될 때마다 20%씩 감점
- 이 Programming Assignment는 본인의 정치 성향과는 무관합니다.