

컴퓨터 구조

데이터표현

고려대학교 세종캠퍼스 인공지능사이버보안학과
구 자 훈

목차

❖ 학습 목표

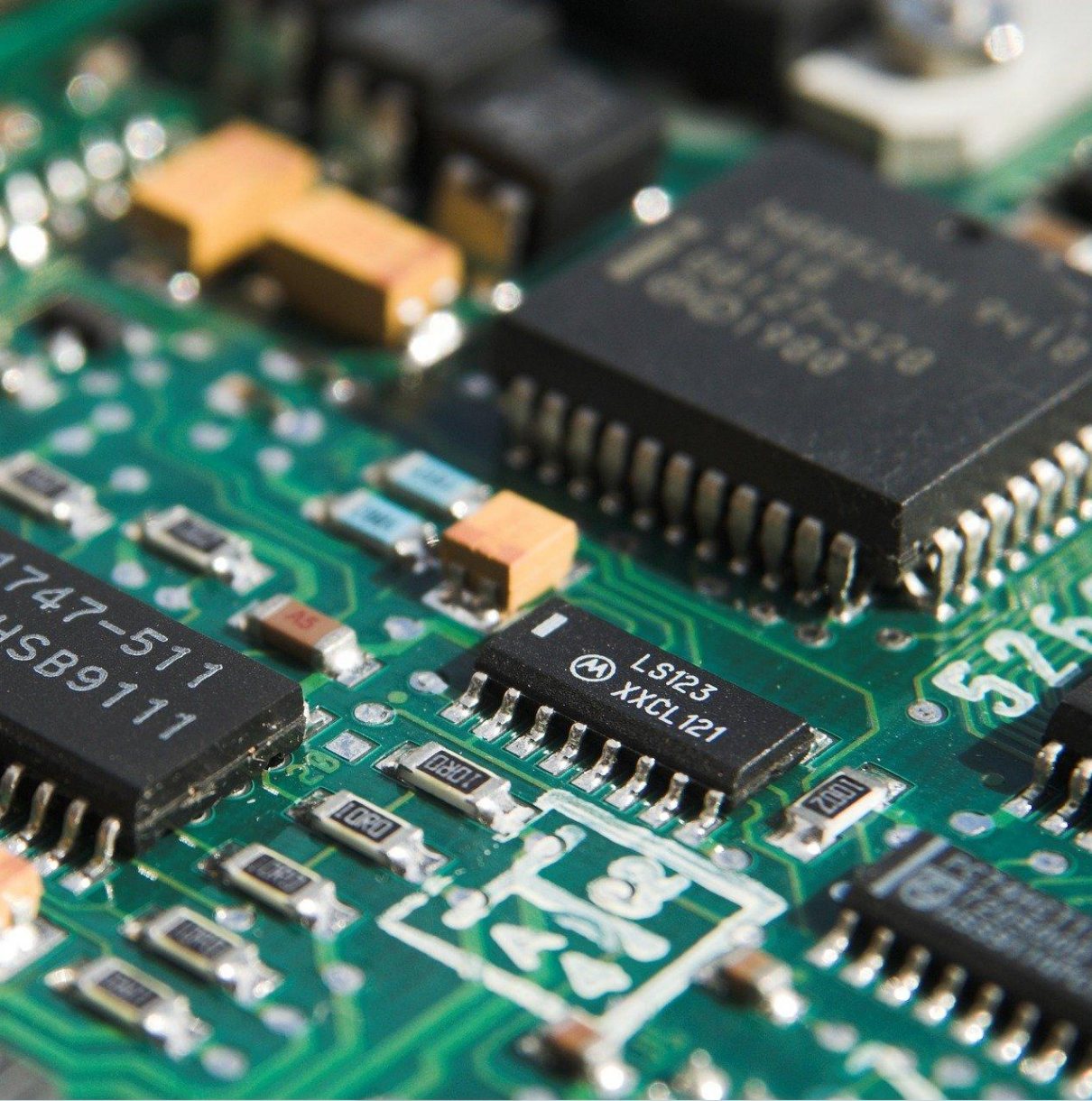
- 수를 표현하는 원리를 이해한다.
- 10진수와 임의의 R진수 간 표현을 변환할 수 있다.
- 문자를 코드(code)로 표현하는 방법을 이해한다.

❖ Part 1.

- 디지털 시스템
- 수의 체계
- 진법 변환
- 코드

❖ Part 2.

- 퀴즈
- 요약



Part 1

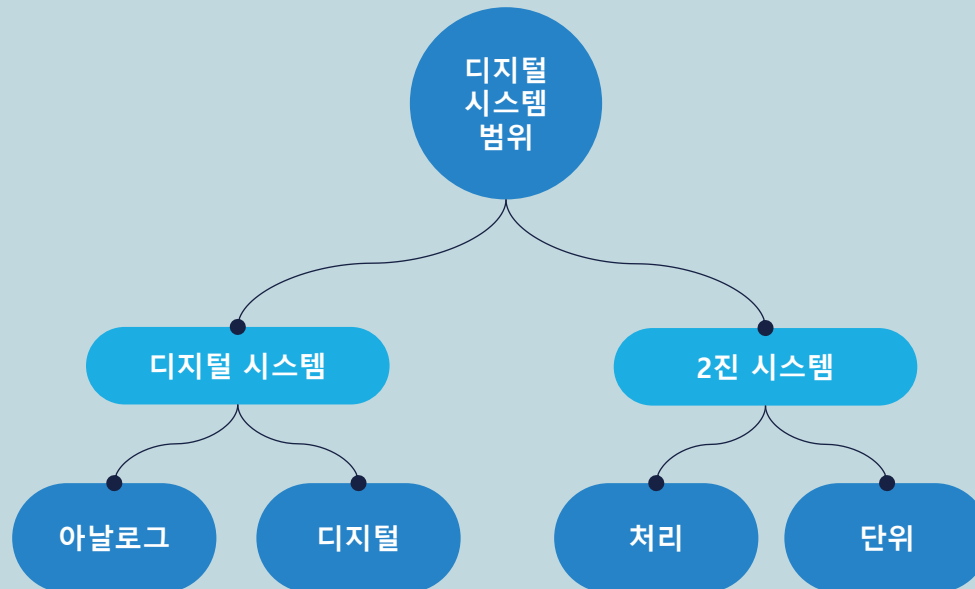
2.1 디지털 시스템

2.2 수의 체계

2.3 진법 변환

2.4 코드

2.1 디지털 시스템



01 디지털 시스템

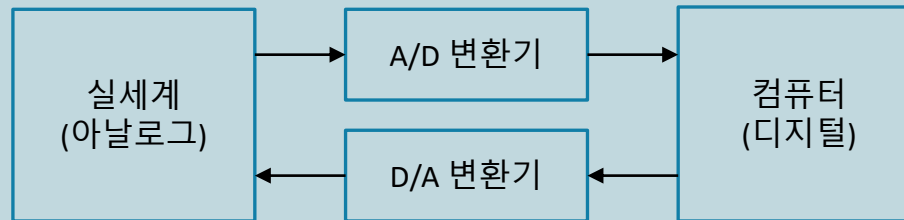
디지털 시스템

- digit + al
- 불연속 값(discrete value)을 취급(처리, 계산)하는 시스템
- 컴퓨터는 디지털 시스템의 일종

아날로그

- 실세계(read-world)는 아날로그(analog)
- 연속적인 값(continuous value)
- 무게, 길이, 압력, 거리, 높이 등등

신호 변환



01 디지털 시스템 : 아날로그 & 디지털

디지털

Digital signal

- 불연속(discrete)
- 숫자로 표현
- 실세계의 값을 숫자로 변환

Digital System

- 온라인 상태로 저장 가능
- 정확하다

아날로그

Analog signal

- 연속(continuous)
- 실세계에 존재한다

Analog system

- 온라인 상태로 저장 불가
- 부정확하다

왜 디지털 신호로?

“디지털 신호는 보다 안정적(stable)이다.”
(= 값이 딱 나누어 떨어진다)

- 숫자를 아날로그(사람의 필기) / 디지털로 처리한다면?

01 ^{디지털 시스템} 디지털 시스템 : 아날로그 & 디지털

10진수

사람은 10진수를 사용한다.

- 디짓(digit): 자릿수, 손가락
- 아라비아 숫자
: 0, 1, 2, 3, 4, 5, 6, 7, 8, 9

2진수

기계는 2진수를 사용한다 : 안정된 상태가 2개

- 스위치: 열림(off)/닫힘(on).
- OCR 카드: 표시가 없음/까만 색.
- 펀치 카드: 구멍을 뚫지 않음/구멍을 뚫음.
- 전기 신호: 전류가 흐르지 않음/흐름.
- 자기(자석): N극/S극.

02 디지털 시스템 : 2진 시스템

2진 시스템이란?

- 2진수로 표현된 데이터를 처리하는 시스템
 - 비트 (bit) = binary + digit: 2진수 한 자리, 0 또는 1
 - MSB (Most Significant Bit): 가장 왼쪽 자리의 비트
 - LSB (Least Significant Bit): 가장 오른쪽 자리의 비트

quiz

- $(12345678)_{10}$
 - MSD = _____
 - LSD = _____
- $(00001111)_2$
 - MSB = _____
 - LSB = _____

2진 데이터의 단위



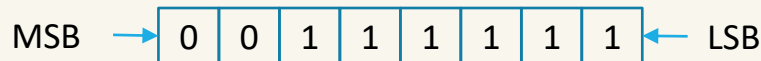
(a) 비트(bit)



(b) 니블(nibble)



(c) 바이트(byte)

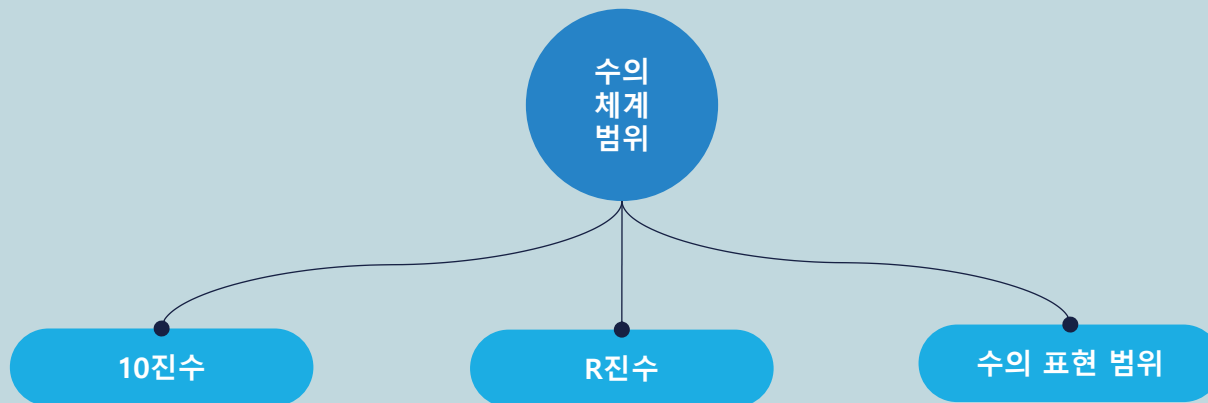


(d) MSB와 LSB

2.2 수의 체계

학습 내용

- 수를 표현하는 원리 (weighted number)
- 무게? 수에서의 무게는 무엇인가?
- R진법의 수(R진수)의 표현과 크기



01 수의 체계 : 10진수

무게 수 (weight number)

- 숫자가 가지고 있는 무게 (클수록 무거움)
- 로마자 10^n / 별도의 기호 一, 十, 百, 千, 萬, 億, 兆

1234.56

- 기호(symbol): {0, 1, 2, 3, 4, 5, 6, 7, 8, 9}
- 자리에 따라 무게가 다름

수의 표현	...	1	2	3	4	.	5	6	...
자릿수	...	3	2	1	0		-1	-2	...
무게	...	10^3	10^2	10^1	10^0		10^{-1}	10^{-2}	...

- EX) $(1234.56)_{10} = 1 \times 10^3 + 2 \times 10^2 + 3 \times 10^1 + 4 \times 10^0 + 5 \times 10^{-1} + 6 \times 10^{-2}$
 $= 1000 + 200 + 30 + 4 + 0.5 + 0.06$
 $= 1234.56$

- 값 (value) : 수가 나타내는 값은 수식으로 표현 가능

$$\begin{aligned} V(N) &= A_{n-1} \times 10^{n-1} + A_{n-2} \times 10^{n-2} + \cdots + A_{-m} \times 10^{-m} \\ &= \sum_{k=-m}^{n-1} A_k \times 10^k \end{aligned}$$

02 수의 체계 : R진수

기수 (radix)

- 수를 나타내는데 기초가 되는 수
(10진수에서는 0,1,2,3,4,5,6,7,8,9)

R진법의 수

- 기호(symbol): {0, 1, 2, ..., R-1}
- 값 (value) :

기호	...	A_3	A_2	A_1	A_0	.	A_{-1}	A_{-2}	...
자릿수	...	3	2	1	0		-1	-2	...
무게	...	R^3	R^2	R^1	R^0		R^{-1}	R^{-2}	...

$$\begin{aligned} V(N) &= A_{n-1} \times R^{n-1} + A_{n-2} \times R^{n-2} + \dots + A_{-m} \times R^{-m} \\ &= \sum_{k=-m}^{n-1} A_k \times R^k \end{aligned}$$

- EX) $(1A4C)_{16}$
 - 기호: {0, 1, 2, 3, 4, 5, 6, 7, 8, 9, A, B, C, D, E, F}
 - 값: $1 \times 16^3 + 10 \times 16^2 + 4 \times 16^1 + 12 \times 16^0 = 4,096 + 2,560 + 64 + 12 = 6,732$

03 수의 체계

수의 표현 범위

수의 종류

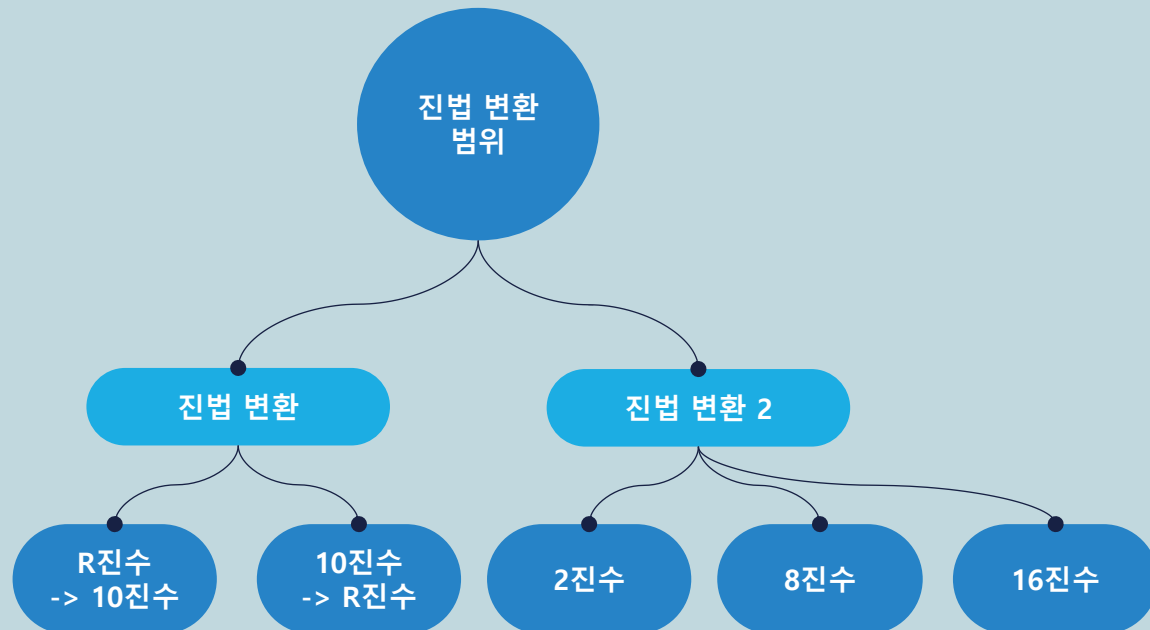
- 부호 없는 수(unsigned number): 0을 포함한 양수
- 정수(signed number): ..., -1, 0, 1, ... (음수, 0, 양수)
- 실수(real number): 소수점을 포함하는 수

수의 표현 범위

- 10진수 n자리: $0 \sim 10^n - 1 = 0 \sim 9...9$ (9가 n개)
- R진수 n자리: $0 \sim R^n - 1$
 - 예) 4자리 십진수 $(10^4 - 1) = 9999$ 표현 범위 0~9999
 - 8자리 2진수 $(2^8 - 1) = 255$ 표현 범위 0~255

사람은 필요에 따라 숫자를 무한정 늘려서 쓸 수 있음
하지만 컴퓨터는 주어진 비트 수에 따라 제한을 받게 됨
따라서 한정된 자리수로 표현할 수 밖에 없는 수의 표현을 배움

2.3 진법 변환



01 진법 변환 진법 변환

진법 변환

- R진수 -> 10진수

$$\begin{aligned} V(N) &= A_{n-1} \times R^{n-1} + A_{n-2} \times R^{n-2} + \cdots A_{-m} \times R^{-m} \\ &= \sum_{k=-m}^{n-1} A_k \times R^k \end{aligned}$$

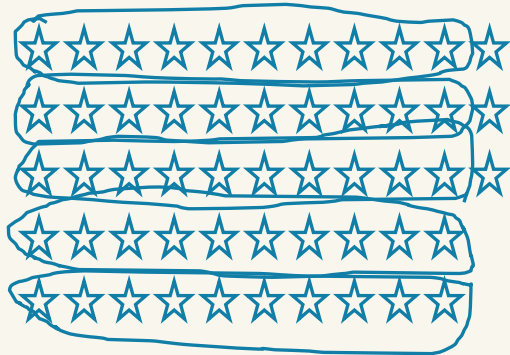
- 10진수 -> R진수

구성

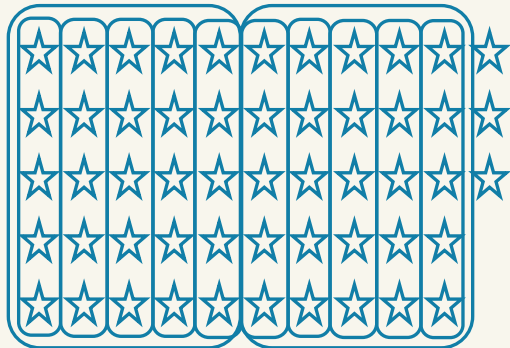
- 10진수를 R진수로 변환
- 2진수, 8진수, 16진수

01 진법 변환 진법 변환 : 10진수를 R진수로 변환

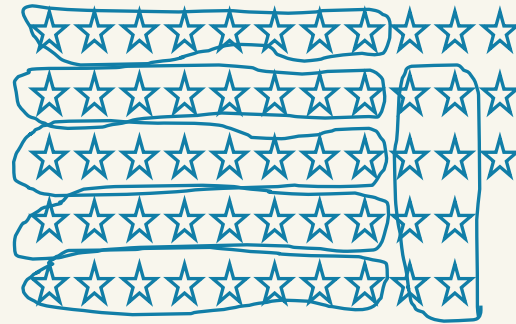
절대 수는 불변, 진법에 따라 표현이 다르다



()₁₀



()₅



()₈



()₂

01 ^{진법 변환}진법 변환의 원리

$$\begin{aligned} V(N) &= A_{n-1} \times 10^{n-1} + A_{n-2} \times 10^{n-2} + \cdots A_0 \times 10^0 \\ &= X_{k-1} \times R^{n-1} + X_{k-2} \times R^{n-2} + \cdots X_0 \times R^0 \end{aligned}$$

$$V(N)/R = X_{k-1} \times R^{n-2} + X_{k-2} \times R^{n-3} + \cdots X_1 \times R^0 \quad \text{remains} \quad X_0$$

R개씩 묶으면

- 묶음의 수가 몫이고, 나머지가 1의 자리 숫자이다.

10진수를 R로 나누고

- 몫은 몫의 수이고, 나머지가 R^0 자리의 숫자이다.

01 진법 변환 R진수로 변환

2진수로 몇 개?

_____개



$527_{10} \rightarrow 8$ 진수

02 ^{진법 변환} 2진수, 8진수, 16진수

컴퓨터는 2진수를 사용하지만,
사람이 보기 불편하기 때문에 8, 16진수로 표현한다.

예시

- 1100010110010001
1_100_010_110_010_001 \leftrightarrow (142621)₈
- 1100010110010001
1100_0101_1001_0001 \leftrightarrow (C591)₁₆ = C591h

02 ^{진법 변환} 2진수, 8진수, 16진수

2진수를 배우는 이유

컴퓨터는 2진수를
사용하기 때문에
사람도 2진수에
익숙해져야 한다.

10 진수	2 진수	8 진수	16 진수
0	0000	00	0
1	0001	01	1
2	0010	02	2
3	0011	03	3
4	0100	04	4
5	0101	05	5
6	0110	06	6
7	0111	07	7
8	1000	10	8
9	1001	11	9
10	1010	12	A
11	1011	13	B
12	1100	14	C
13	1101	15	D
14	1110	16	E
15	1111	17	F

02 ^{진법 변환} 2진수 ↔ 8진수, 16진수

2진수 ↔ 8진수: $2^3 = 8$

0000110110000010 ↔ _____ ↔ (142621)₈
0100111101111011 ↔ _____ ↔ (067072)₈

2진수 ↔ 16진수: $2^4 = 16$

0000110110000010 ↔ _____ ↔ (C591)₁₆
0100111101111011 ↔ _____ ↔ (6E3A)₁₆

02 ^{진법 변환} 2진수 ↔ 8진수, 16진수

2진수 ↔ 8진수: $2^3 = 8$

1100010110010001 ↔ _____	↔ (142621) ₈
0000110110000010 ↔ _____	↔ (006602) ₈
0110111000111010 ↔ _____	↔ (067072) ₈
0100111101111011 ↔ _____	↔ (047573) ₈

2진수 ↔ 16진수: $2^4 = 16$

1100010110010001 ↔ _____	↔ (C591) ₁₆
0000110110000010 ↔ _____	↔ (0D82) ₁₆
0110111000111010 ↔ _____	↔ (6E3A) ₁₆
0100111101111011 ↔ _____	↔ (4F7B) ₁₆

02 진법 변환 진법 변환 : 2^k

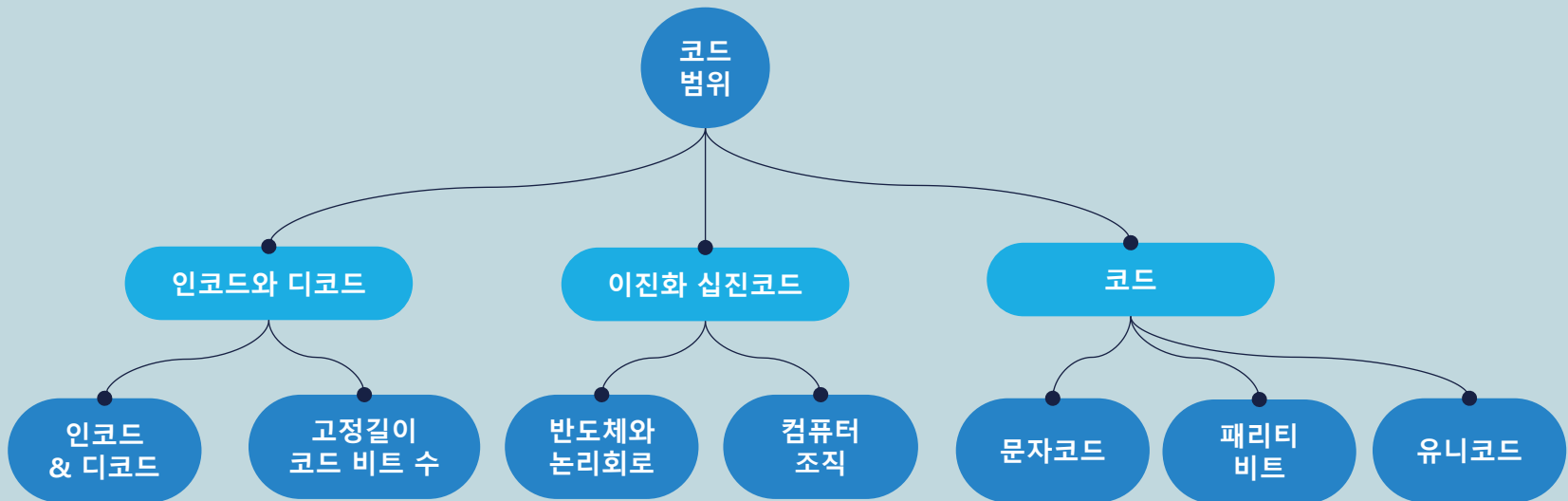
2^0	1	2^4	16	2^8	256	2^{30}	1Giga
2^1	2	2^5	32	2^9	512	2^{40}	1 Tera
2^2	4	2^6	64	2^{10}	1024 =1K		
2^3	8	2^7	128	2^{20}	1Mega		

2.4 코드

코드

숫자 이외의 데이터를 2진수로 표현하는 방법

- 문자 코드
- 신호, 음성, 영상
 - Analog-to-Digital (A/D): 영상/음성같은 아날로그데이터는 샘플링(sampling)과 양자화(quantization) 과정을 거쳐 숫자로 변환되는 과정



01 ^{코드}인코드와 디코드

코드

- 유한개의 원소로 구성된 집합에 대하여
- 각 원소를 서로 구별할 수 있도록 각 원소에 부여하는 숫자

코드의 종류

- 고정 길이 코드: 원소에 부여된 2진수의 길이가 같다.
- 가변 길이 코드: 원소에 부여된 2진수의 길이가 다르다.

동작

인코드 (encode)

- 원소 기호 → 코드

디코드 (decode)

- 코드 → 원소 기호

01 코드 예제

{♠, ◇, ♥, ♣}에 대한 코드

	가변코드		고정코드	
집합의 원소	코드 1	코드 2	코드 3	코드 4
♠	0	1	00	101
◇	10	11	01	111
♥	110	111	10	110
♣	1110	1111	11	000

- 코드 1: ♣♥♠◇♥ → 1110_110_0_10_110 → ♣♥♠◇♥
- 코드 2: ♣♥♠◇♥ → 1111_111_1_11_111 → ???
- 코드 3: ♣♥♠◇♥ → 11_10_00_01_10 → ♣♥♠◇♥
- 코드 4: ♣♥♠◇♥ → 000_110_101_111_110 → ♣♥♠◇♥

01 ^{코드} 고정 길이 코드 비트 수

- 집합의 원소 수가 N 일 때, 고정 길이 코드의 비트 수는?

N 을 2^k 로 표현하려면 2^k 표현할 수 있는 수가 N 보다는 커야 함

- 2^k 으로 표현할 수 있는 수의 개수 $\geq N$

$$\log_2 2^k \geq \log_2 N$$

$$k \geq \log_2 N$$

최종적으로 k 는 $\log_2 N$ 보다 커야 한다.

$$\lceil \log_2 N \rceil$$

02^{코드} 이진화 십진 코드

BCD

- Binary Coded Decimal
- BCD코드는 2진수로 10진수를 표현하는 코드
- 8421 BCD코드가 가장 보편적으로 사용됨
- 각 자리에 무게를 부여하고 만든 코드

3초과 BCD 코드

- 코드 값에 $3_{10} / (0011)_2$ 을 더해서 만든 코드
- 각 자리의 무게는 의미가 없음
- 임의의 수에 1의 보수를 취하면 결과로 10진수로 해석한 9의 보수 반환
 - 보수: 각 자리의 숫자의 합이 어느 일정한 수가 되게 하는 수.
예를 들어 10의 7에 대한 **보수**는 3
ex) $X + y = 9$ 이면 x, y 는 서로 9에 대한 보수
- 3초과 코드에서 3을 해당하는 0110의 보수인 1001은 6에 해당하는 초과 코드

02^{코드} 이진화 십진 코드

자보수 특성

- 자보수 : self-complementary
- 3초과 코드처럼 보수를 취했을 때 / 보수가 나오는 경우

10 진수 기호	0	1	2	3	4	5	6	7	8	9
8421 BCD 코드	0000	0001	0010	0011	0100	0101	0110	0111	1000	1001
3초과 BCD 코드	0011	0100	0101	0110	0111	1000	1001	1010	1011	1100

03 ^{코드} 문자 코드

ASCII(American Standard Code for Information Interchange)

ANSI (American National Standards Institute) 제정

가장 널리 쓰이는
문자 code의 표준

영어 권 이외의
언어는 표현 불가능

원래 7 bit 체계.
지금은 8 bit로 확장

- byte 크기의 memory cell에 맞춤
- 한 문자의 시작은 0으로, 나머지는 코드 값으로 구성

03^{코드} 문자 코드 : 예제

b ₃ b ₂ b ₁ b ₀	b ₆ b ₅ b ₄							
	000	001	010	011	100	101	110	111
0000	NUL	DLE	SP	0	@	P	`	p
0001	SOH	DC1	!	1	A	Q	a	q
0010	STX	DC2	"	2	B	R	b	r
0011	ETX	DC3	#	3	C	S	c	s
0100	EOT	DC4	\$	4	D	T	d	t
0101	ENQ	NAK	%	5	E	U	e	u
0110	ACK	SYN	&	6	F	V	f	v
0111	BEL	ETB	'	7	G	W	g	w
1000	BS	CAN	(8	H	X	h	x
1001	HT	EM)	9	I	Y	i	y
1010	LF	SUB	*	:	J	Z	j	z
1011	VT	ESC	+	;	K	[k	{
1100	FF	FS	,	<	L	\	l	
1101	CR	GS	-	=	M]	m	}
1110	SO	RS	.	>	N	^	n	~
1111	SI	US	/	?	O	_	o	DEL

"Good Morning!"

03^{코드} 패리티 비트(parity bit)

(parity bit) : parity (특히 보수·지위의) 동등함

패리티 비트

- 데이터가 올바른지 검사하기 위하여 추가하는 비트
- 비트의 수가 잘 맞으면 제대로 데이터가 입력되었다는 증빙

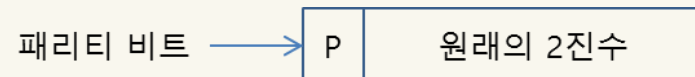
짝수 패리티

- even parity
- 1이 짝수가 되도록 추가하는 비트

홀수 패리티

- odd parity
- 1이 홀수가 되도록 추가하는 비트

짝수와 홀수 패리티 구하기



- A = (아스키 코드) 100_0001 (짝수) 0100_0001 (홀수) 1100_0001
- T = (아스키 코드) 101_0100 (짝수) 1101_0100 (홀수) 0101_0100

03^{코드} 유니코드(Unicode)

유니코드

- 세계 각국의 언어 표현
- 국제적으로 통용되는 16비트 문자 체계
- 1991년에 버전 1.0
현재 2025년 17.0

배치

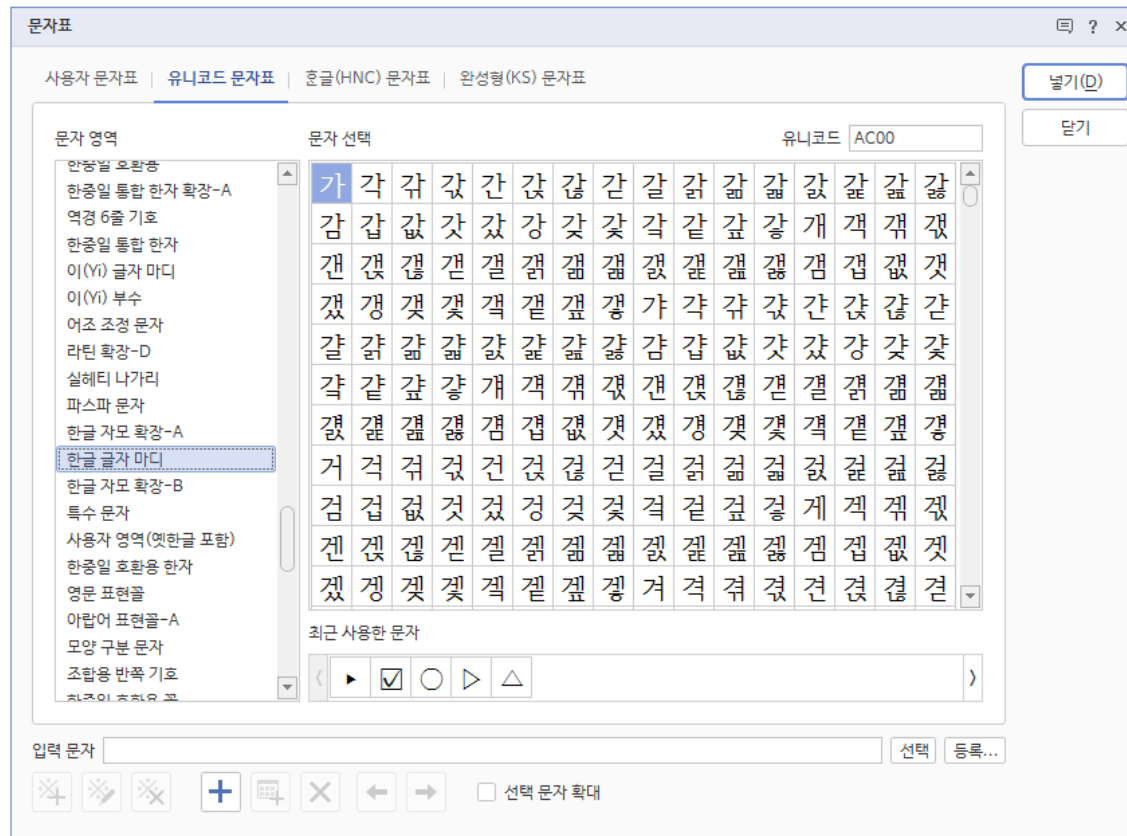
- U+0000 ~ U+007F 영역에
영문자 배치
- 이후 여러 나라 문자
- 한글은 U+AC00 부터 U+D7A3까지
11,172 글자 정의
- 문서편집기의 문자표에서
유니코드확인 가능

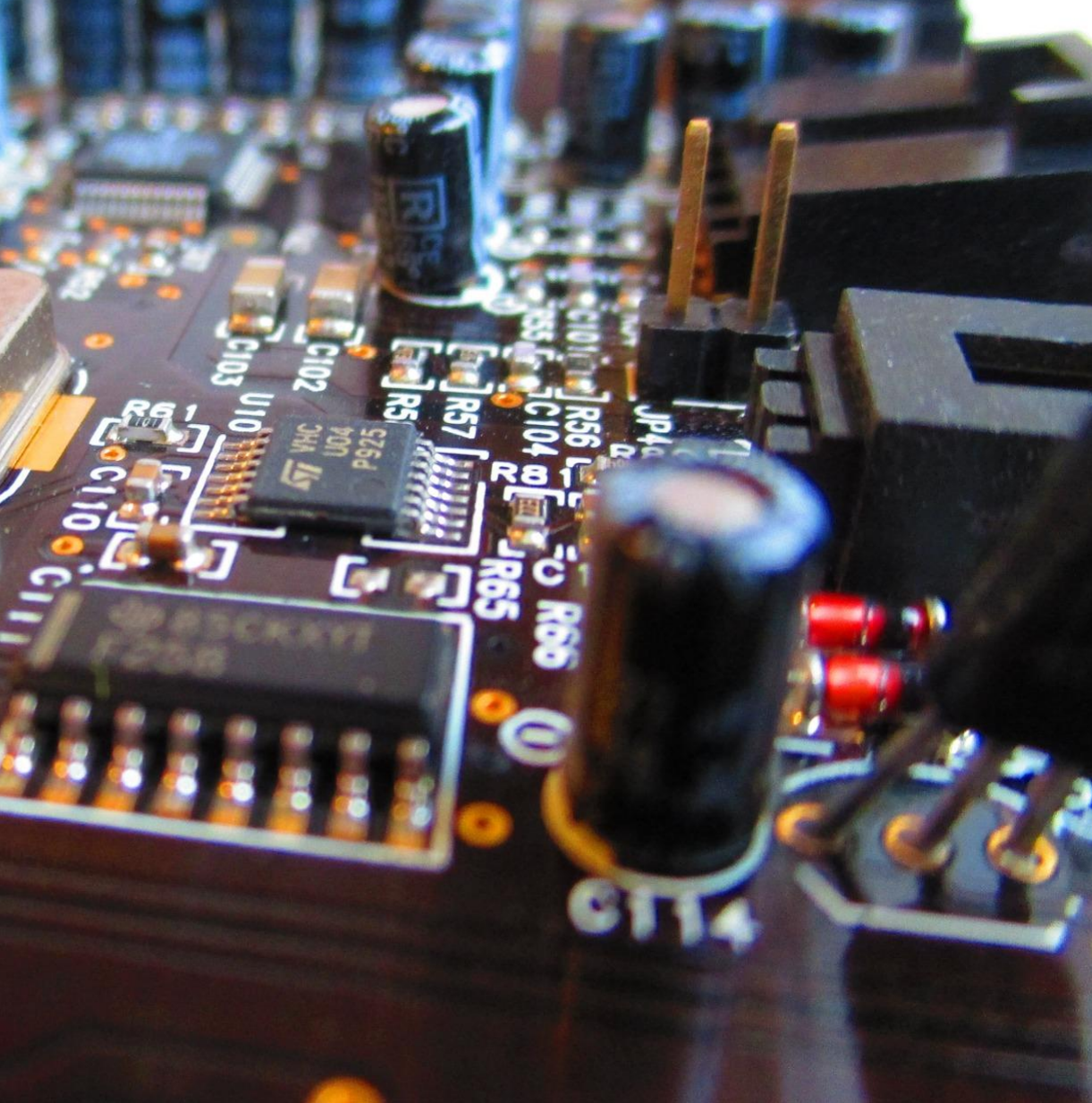
예제

- 홍길동: D64D AE38 B3D9

03^{코드} 유니코드(Unicode)

유니코드 (예제 / 한글)





Part 2

2.4 퀴즈

2.5 요약

01 데이터 표현 문제

진법 변환

- (1) $(10110001101011)_2$ 를 8진수로
- (2) $(10110001101011)_2$ 를 16진수로
- (3) $(523)_8$ 을 2진수로
- (4) $(D1AF)_{16}$ 을 8진수로

10진수 구하기

- (5) $01001100 =$ _____
- (6) $10000011 =$ _____

01 데이터 표현 문제

진법 변환

- (1) $(10110001101011)_2$ 를 8진수로
- (2) $(10110001101011)_2$ 를 16진수로
- (3) $(523)_8$ 을 2진수로
- (4) $(D1AF)_{16}$ 을 8진수로

10진수 구하기

- (5) $01001100 =$ _____
- (6) $10000011 =$ _____

문단
(1) 26 153 (2) 2C6B (3) 0001 0101 0011
(4) 150 657 (5) 76 (6) 131

02 디지털 시스템 요약

디지털 시스템

- 숫자를 처리하는 시스템
- A/D 변환기: 아날로그 신호를 숫자로 변환

2진 시스템

- 2진수 데이터를 처리하는 시스템
- 안정된 상태가 2개인 소자.
- 비트/니블/바이트
- MSB/LSB

03 수의 체계 요약

수의 표현 원리 = 무게 수 (weighted number)

기호

◦ R개 (0, 1, 2, ..., R-1)

기호	...	A_3	A_2	A_1	A_0	.	A_{-1}	A_{-2}	...
자릿수	...	3	2	1	0		-1	-2	...
무게	...	R^3	R^2	R^1	R^0		R^{-1}	R^{-2}	...

수의 크기

$$\begin{aligned} V(N) &= A_{n-1} \times R^{n-1} + A_{n-2} \times R^{n-2} + \cdots A_{-m} \times R^{-m} \\ &= \sum_{k=-m}^{n-1} A_k \times R^k \end{aligned}$$

04 진법 변환 요약

10진수를 R진수로 변환

- 나누기 R: 몫과 나머지(R^k 자리의 수)

2^k 계산

- $2^{10} = 1K$, $2^{20} = 1M$, $2^{30} = 1G$

2진수, 8진수, 16진수

- $23 = 8$: 2진수를 3자리씩 읽으면 8진수
- $24 = 16$: 2진수를 4자리씩 읽으면 16진수

05^{코드} 요약

코드

- 유한 개의 원소로 구성된 집합의 원소에
- 각 원소를 구별할 수 있도록 부여한 숫자

동작

- 인코드(encode): 원소에 대하여 숫자를 부여하는 과정
- 디코드(decode): 숫자를 보고 원래의 원소를 찾는 과정

자주 사용되는 코드

- BCD (Binary-Coded Decimal)
- 문자 코드: 아스키 코드, 유니코드

06 데이터 표현 요약

디지털 시스템

- A/D 변환기, D/A 변환기
- 이진 시스템: 비트(bit), 바이트(byte)

수의 체계

- 무게 수(weighted code)
- 진법 변환, 2진수, 8진수, 16진수

코드

- 인코드 / 디코드
- BCD, ASCII

다음 시간에 배울 내용: 논리회로 기초

- 논리 연산 규칙 (부울 대수)
- 논리회로 동작 표현(논리식, 논리 회로도, 진가표)
- 여러 가지 논리 게이트 소개