

# 데이터 표현

Heeseung Jo

## 2.1 디지털 시스템

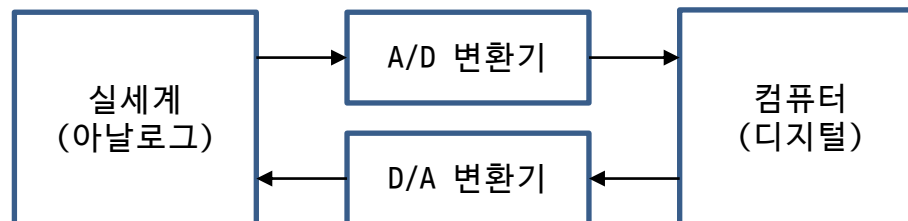
### 디지털 시스템이란?

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

### 실세계(real-world)는 아날로그(analog)

- 연속적인 값(continuous value)

### 신호 변환



# 아날로그 & 디지털

## Analog signal

- 연속(continuous)
- 실세계에 존재

## Analog system

- 온라인 상태로 저장 불가
- 부정확

## Digital signal

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

## Digital system

- 온라인 상태로 저장 가능
- 정확

# 숫자 표현

사람은 10진수 사용

- Digit: 자릿수, 손가락
- 아라비아 숫자: 0, 1, 2, 3, 4, 5, 6, 7, 8, 9

기계는 2진수 사용: 안정된 상태가 2개

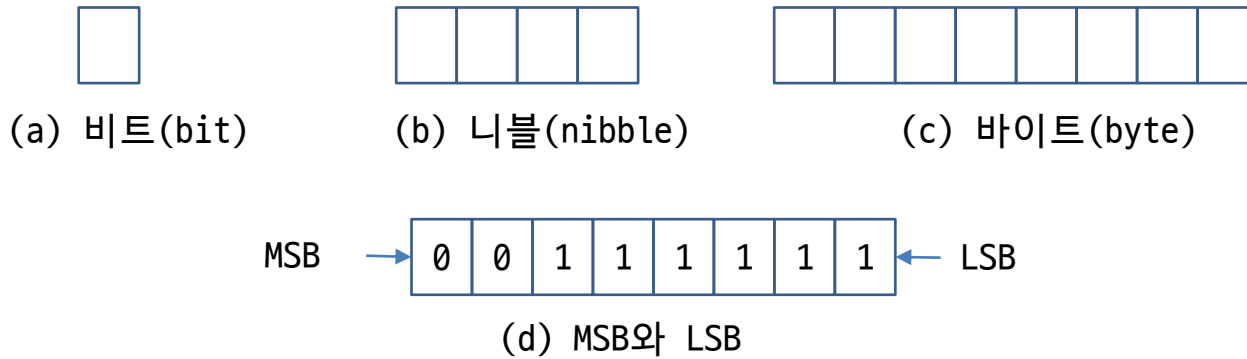
- 스위치: 열림(off)/닫힘(on)
- OCR 카드: 표시가 없음/까만 색
- 펀치 카드: 구멍을 뚫지 않음/구멍을 뚫음
- 전기 신호: 전류가 흐르지 않음/흐름
- 자기(자석): N극/S극

# 2진 시스템(binary system)

2진수로 표현된 데이터를 처리하는 시스템

비트 (bit) = binary + digit: 2진수 한 자리, 0 또는 1

2진 데이터의 단위



MSB (Most significant bit) vs. LSB (Least significant bit)

- 예제 2-1

-  $(12345678)_{10}$ : MSD = \_\_\_\_\_, LSD = \_\_\_\_\_

-  $(00001111)_2$ : MSB = \_\_\_\_\_, LSB = \_\_\_\_\_

## 2.2.1 10진수

무게 수(weighted number)

1234.56

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

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

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

값

$$V(N) = A_{n-1} \times 10^{n-1} + A_{n-2} \times 10^{n-2} + \dots + A_{-m} \times 10^{-m}$$
$$= \sum_{k=-m}^{n-1} A_k \times 10^k$$

## 2.2.2 R진수

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

### R진법의 수

- 기호(symbol):  $\{0, 1, 2, \dots, R-1\}$

- 값

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

예:  $1A4C_{16}$

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

## [예제 2-2] 크기 계산

수의 크기를 10진수로 적으세요

표현이 잘못된 수는 그 이유를 설명하세요

1.  $519.87_{10}$

2.  $127.4_8$

3.  $127.4_5$

4.  $1BE_{16}$

5.  $1101.1_2$

## 2.2.3 수의 표현 범위

### 수의 종류

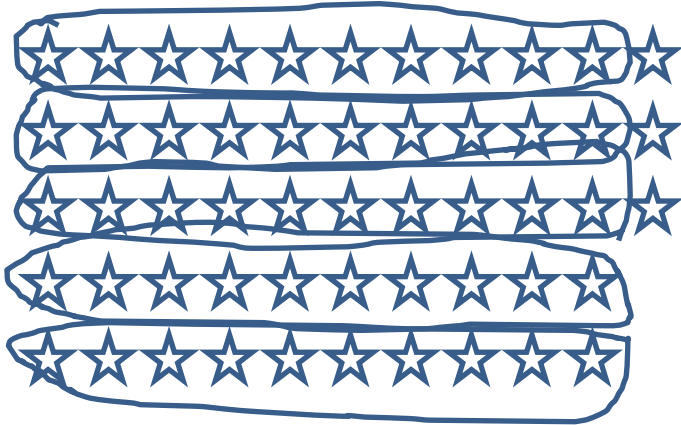
- 부호 없는 수(unsigned number): 0을 포함한 양수
- 정수(signed number): ..., -1, 0, 1, ... (음수, 0, 양수)
- 실수(real number): 소수점을 포함하는 수

※  $R^n = 10\dots0$  (0이 n개)

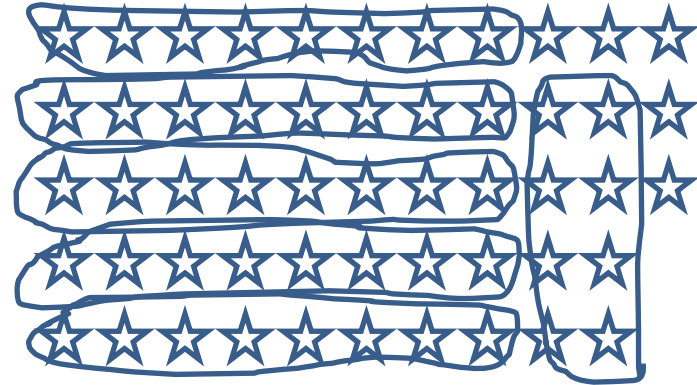
### 수의 표현 범위

- 10진수 n자리:  $0 \sim 10^n - 1 = 0 \sim 9\dots9$  (9가 n개)
- R진수 n자리:  $0 \sim R^n - 1$

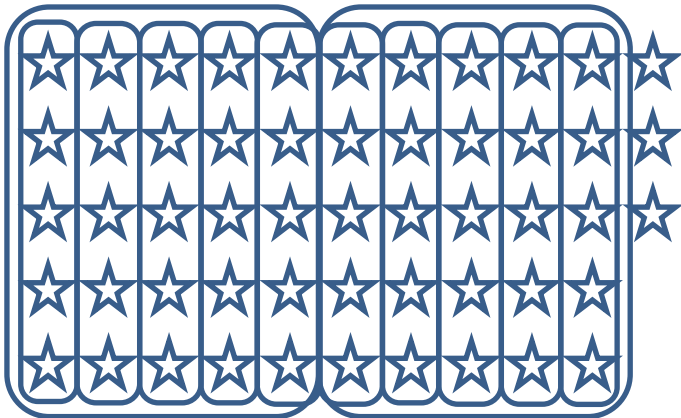
## 2.3.1 10진수를 R진수로 변환



( )<sub>10</sub>



( )<sub>8</sub>



( )<sub>5</sub>



( )<sub>2</sub>

절대 수는 불변, 진법에 따라 표현이 다름

# 진법 변환 원리

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

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

R개씩 묶으면

- 묶음의 수가 몫이고, 나머지가 1의 자리 숫자이다

10진수를 R로 나누고

- 몫은 몫의 수이고, 나머지가  $R^0$  자리의 숫자이다

## 2.3.2 2진수, 8진수, 16진수

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

# 진법 변환

2진수  $\leftrightarrow$  8진수:  $2^3 = 8$

1100010110010001	$\leftrightarrow$ _____	$\leftrightarrow$ (142621) <sub>8</sub>
0000110110000010	$\leftrightarrow$ _____	$\leftrightarrow$ (006602) <sub>8</sub>
0110111000111010	$\leftrightarrow$ _____	$\leftrightarrow$ (067072) <sub>8</sub>
0100111101111011	$\leftrightarrow$ _____	$\leftrightarrow$ (047573) <sub>8</sub>

2진수  $\leftrightarrow$  16진수:  $2^4 = 16$

1100010110010001	$\leftrightarrow$ _____	$\leftrightarrow$ (C591) <sub>16</sub>
0000110110000010	$\leftrightarrow$ _____	$\leftrightarrow$ (0D82) <sub>16</sub>
0110111000111010	$\leftrightarrow$ _____	$\leftrightarrow$ (6E3A) <sub>16</sub>
0100111101111011	$\leftrightarrow$ _____	$\leftrightarrow$ (4F7B) <sub>16</sub>

# 예제

---

## [예제 2-6] 진법 변환

(1)  $10110001101011_2$ 를 8진수로

(2)  $10110001101011_2$ 를 16진수로

(3)  $523_8$ 을 2진수로

(4)  $D1AF_{16}$ 을 8진수로

# $2^k$

$2^k$

$$2^0 = 1$$

$$2^1 = 2$$

$$2^2 = 4$$

$$2^3 = 8$$

$$2^4 = 16$$

$$2^5 = 32$$

$$2^6 = 64$$

$$2^7 = 128$$

$$2^8 = 256$$

$$2^9 = 512$$

$$2^{10} = 1024 = 1K$$

$$2^{16} = 64K$$

$$2^{19} = \text{-----}$$

$$2^{20} = 1\text{Mega}$$

$$2^{24} = \text{-----}$$

$$2^{27} = \text{-----}$$

$$2^{30} = 1\text{Giga}$$

$$2^{32} = \text{-----}$$

$$2^{34} = \text{-----}$$

$$2^{40} = 1\text{Tera}$$

$$2^{43} = \text{-----}$$

$$2^{47} = \text{-----}$$

## 2.4.1 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  $\rightarrow$  코드
- 디코드(decode): 코드  $\rightarrow$  원소 기호

# 코드 예제

원소 {♠, ♢, ♡, ♣}에 대한 코드

집합의 원소	코드체계 1	코드체계 2	코드체계 3	코드체계 4
♠	0	1	00	101
♢	10	11	01	111
♡	110	111	10	110
♣	1110	1111	11	000

코드 예:

- 코드 1: ♣♡♠♢♡ → 1110\_110\_0\_10\_110 → 1110110010110 → ♣♡♠♢♡
- 코드 2: ♣♡♠♢♡ → 1111\_111\_1\_11\_111 → 1111111111111 → ???
- 코드 3: ♣♡♠♢♡ → 11\_10\_00\_01\_10 → 1110000110 → ♣♡♠♢♡
- 코드 4: ♣♡♠♢♡ → 000\_110\_101\_111\_110 → 000110101111110 → ♣♡♠♢♡

# 고정길이 코드 비트 수

집합의 원소 수가  $N$ 일 때,  
고정길이 코드의 비트 수는?  $\lceil \log_2 N \rceil$

[예제 2-9] 집합의 원소 수에 대한 고정길이 코드의 최소 비트 수는?

- (1) 원소 8개
- (2) 원소 10개
- (3) 원소 16개
- (4) 원소 25개

## 2.4.2 이진화 십진 코드

### BCD (Binary Coded Decimal)

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

### 자보수(self-complementary) 특성

- 보수: 0을 1로, 1을 0으로 바꾼 수
- $[X + (X\text{의 보수}) = 9\text{에 대한 코드}]$  인 특성 (3초과 BCD)
  - 3 (0110)의 보수 1001 (6)
  - $3 + 6 = 9$

### [예제 2-10] $(1225)_{10}$

- 8421 BCD 코드: \_\_\_\_\_
- 3초과 BCD 코드: \_\_\_\_\_

## 2.4.3 문자 코드

ASCII(American Standard Code for Information Interchange)

b <sub>3</sub> b <sub>2</sub> b <sub>1</sub> b <sub>0</sub>	b <sub>6</sub> b <sub>5</sub> b <sub>4</sub>							
	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

[예제 2-12] Good Morning!

-> \_\_\_\_\_

# 패리티 비트 (parity bit)



## 패리티 비트

- 데이터가 올바른지 검사하기 위하여 추가하는 비트
- 짝수 패리티(even parity): 1의 개수가 되도록 추가하는 비트
- 홀수 패리티(odd parity): 1의 개수가 되도록 추가하는 비트

## [예제 2-13] 짝수와 홀수 패리티 구하기

- A = (아스키 코드) \_\_\_\_\_ (짝수) \_\_\_\_\_ (홀수) \_\_\_\_\_
- T = (아스키 코드) \_\_\_\_\_ (짝수) \_\_\_\_\_ (홀수) \_\_\_\_\_

b <sub>3</sub> b <sub>2</sub> b <sub>1</sub> b <sub>0</sub>	b <sub>6</sub> b <sub>5</sub> b <sub>4</sub>						
	000	001	010	011	100	101	110
0000	NUL	DLE	SP	0	@	P	`
0001	SOH	DC1	!	1	A	Q	a
0010	STX	DC2	"	2	B	R	b
0011	ETX	DC3	#	3	C	S	c
0100	EOT	DC4	\$	4	D	T	d
0101	ENO	NAK	%	5	E	U	e

# 유니코드 (Unicode)

---

## 유니코드

- 세계 각국의 언어 표현
- 국제적으로 통용되는 16비트 문자 체계
- 1991년에 버전 1.0, 현재 2018년 11.0

## 배치

- U+0000 ~ U+007F 영역에 영문자 배치
- 이후 여러 나라 문자
- 한글은 U+AC00부터 U+D7A3까지 11,172 글자 정의
- 문서편집기의 문자표에서 유니코드확인 가능