

計算機方式論

第6章 データ形式 - 文字の表現

1

文字の表現

- ◆ 数字、カナ文字、アルファベットの表現:
文字データを用いる
6～8ビットで表現
- ◆ 各文字とビットパターンの対応付け
コード化
文字コード表で表す
- ◆ コード化の様々な方式
ISOコード、JISコード、ASCIIコード、
EBCDICコード

2

情報交換用7単位符号(7bits)

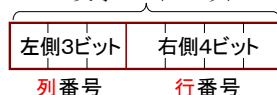
◆ISOコード

国際標準化機構(International Standardization Organization, ISO)で定めた国際基準の文字コード

◆この規格を基に、各国の基準が規格化されている

◆表は国際基準版IRVの文字コード(1967年の規格)

◆現在の国際基準版
ISO/IEC646はASCIIと完全一致
文字コード7ビット



ISO-国際標準化機構の国際基準版
IRV(International Reference Version)

		左側3ビット						
		0	1	2	3	4	5	6 7
0	NUL	DLE	SP	0	@	P	ˆ	p
右1	SOH	DC1	!	1	A	Q	a	q
2	STX	DC2	"	2	B	R	b	r
側3	ETX	DC3	#	3	C	S	c	s
4	EOT	DC4	□	4	D	T	d	t
4 5	ENQ	NAK	%	5	E	U	e	u
6	ACK	SYN	&	6	F	V	f	v
ビ7	BEL	ETB	ˆ	7	G	W	g	w
8	BS	CAN	(8	H	X	h	x
ツ9	HT	EM)	9	I	Y	i	y
A	LF	SUB	*	:	J	Z	j	z
トB	VT	ESC	+	;	K	[k	{
C	FF	FS	,	<	L	\	l	
D	CR	GS	-	=	M]	m	}
E	SO	RS	.	>	N	ˆ	n	˘
F	SI	US	/	? O	—	o	DEL	

3

情報交換用7単位符号(7bits)

◆ISOコード

0,1列及びSP,DELは制御文

字である。SP(空白文字)、
NUL(空白、媒体の空きを埋める)、BEL(ベル、ベルを鳴らす)、
LF(改行)、FF(書式送り)、
CR(復帰)、SO(シフトアウト)、
SI(シフトイン)、ESC(エスケープ)、DEL(抹消)等。

◆コード21～7Eは図形文字である。SPは図形文字でもある。

文字SP 01000000
20₁₆

文字K 10010111
4B₁₆

ISO-国際標準化機構の国際基準版
IRV(International Reference Version)

		左側3ビット						
		0	1	2	3	4	5	6 7
0	NUL	DLE	SP	0	@	P	ˆ	p
右1	SOH	DC1	!	1	A	Q	a	q
2	STX	DC2	"	2	B	R	b	r
側3	ETX	DC3	#	3	C	S	c	s
4	EOT	DC4	□	4	D	T	d	t
4 5	ENQ	NAK	%	5	E	U	e	u
6	ACK	SYN	&	6	F	V	f	v
ビ7	BEL	ETB	ˆ	7	G	W	g	w
8	BS	CAN	(8	H	X	h	x
ツ9	HT	EM)	9	I	Y	i	y
A	LF	SUB	*	:	J	Z	j	z
トB	VT	ESC	+	;	K	[k	{
C	FF	FS	,	<	L	\	l	
D	CR	GS	-	=	M]	m	}
E	SO	RS	.	>	N	ˆ	n	˘
F	SI	US	/	? O	—	o	DEL	

4

情報交換用7単位符号(7bits)

◆JISコード(JIS X 0201)

ローマ文字用7単位符号

		左側3ビット						
		0	1	2	3	4	5	6 7
0	NUL DLE SP	0	@	P	ˆ	p		
右 1	SOH DC1	!	1	A	Q	a	q	
2	STX DC2	"	2	B	R	b	r	
側 3	ETX DC3	#	3	C	S	c	s	
4	EOT DC4	\$	4	D	T	d	t	
4 5	ENQ NAK	%	5	E	U	e	u	
6	ACK SYN	&	6	F	V	f	v	
ビ 7	BEL ETB	ˆ	7	G	W	g	w	
8	BS CAN	(8	H	X	h	x	
ツ 9	HT EM)	9	I	Y	i	y	
A	LF SUB	*	:	J	Z	j	z	
ト B	VT ESC	+	;	K	[k	{	
C	FF FS	,	<	L	¥	l		
D	CR GS	-	=	M]	m	~	
E	SO RS	.	>	N	ˆ	n	ˆ	
F	SI US	/	? O	—	o	DEL		

片仮名用7単位符号

		左側3ビット						
		0	1	2	3	4	5	6 7
0	NUL DLE SP	ー	タ	ミ				
右 1	SOH DC1	。ア	チ	ム				
2	STX DC2	「	イ	ツ	メ			
側 3	ETX DC3	」	ウ	テ	モ			
4	EOT DC4	、	エ	ト	ヤ			
4 5	ENQ NAK	・	オ	ナ	ユ			
6	ACK SYN	ヲ	カ	ニ	ヨ	定	定	
ビ 7	BEL ETB	ア	キ	ヌ	ラ			
8	BS CAN	イ	ク	ネ	リ			
ツ 9	HT EM	ウ	ケ	ノ	ル	義	義	
A	LF SUB	エ	コ	ハ	レ			
ト B	VT ESC	オ	サ	ヒ	ロ			
C	FF FS	ヤ	シ	フ	ワ			
D	CR GS	ユ	ス	ヘ	ン			
E	SO RS	ヨ	セ	ホ	ˆ			
F	SI US	ッ	ソ	マ	ˆ	DEL		

5

情報交換用7単位符号(7bits)

◆ASCIIコード

(American Standard Code for Information Interchange)

		左側3ビット						
		0	1	2	3	4	5	6 7
0	NUL DLE SP	0	@	P	ˆ	p		
右 1	SOH DC1	!	1	A	Q	a	q	
2	STX DC2	"	2	B	R	b	r	
側 3	ETX DC3	#	3	C	S	c	s	
4	EOT DC4	\$	4	D	T	d	t	
4 5	ENQ NAK	%	5	E	U	e	u	
6	ACK SYN	&	6	F	V	f	v	
ビ 7	BEL ETB	ˆ	7	G	W	g	w	
8	BS CAN	(8	H	X	h	x	
ツ 9	HT EM)	9	I	Y	i	y	
A	LF SUB	*	:	J	Z	j	z	
ト B	VT ESC	+	;	K	[k	{	
C	FF FS	,	<	L	1			
D	CR GS	-	=	M]	m	~	
E	SO RS	.	>	N	ˆ	n	ˆ	
F	SI US	/	? O	—	o	DEL		

6

情報交換用8単位符号(8bits)

◆JISコード(JIS X 0201)

ローマ文字・片仮名用8単位符号

		左側4ビット										
		0	1	2	3	4	5	6	7	8	9	A B C D E F
0	NUL DLE SP	0	@	P	ˆ	p				ー	タ	ミ
1	SOH DC1	!	1	A	Q	a	q			。ア	チ	ム
右 2	STX DC2	"	2	B	R	b	r			「	イ	ツ
側 3	ETX DC3	#	3	C	S	c	s			」	ウ	テ
4	EOT DC4	\$	4	D	T	d	t			、	エ	ト
4 5	ENQ NAK	%	5	E	U	e	u			・	オ	ナ
6	ACK SYN	&	6	F	V	f	v			定	定	定
ビ 7	BEL ETB	ˆ	7	G	W	g	w			ア	キ	ヌ
8	BS CAN	(8	H	X	h	x			イ	ク	ネ
ツ 9	HT EM)	9	I	Y	i	y			ウ	ケ	ノ
A	LF SUB	*	:	J	Z	j	z			エ	コ	ハ
ト B	VT ESC	+	;	K	[k	{			オ	サ	ヒ
C	FF FS	,	<	L	¥	l				ヤ	シ	フ
D	CR GS	-	=	M]	m	~			ユ	ス	ヘ
E	SO RS	.	>	N	ˆ	n	ˆ			ヨ	セ	ホ
F	SI US	/	? O	—	o	DEL				ッ	ソ	マ

7

情報交換用8単位符号(8bits)

◆拡張2進化10進コード—EBCDICカタカナコード

(Extended Binary Coded Decimal Interchange Code)

		左側4ビット														
		0	1	2	3	4	5	6	7	8	9	A	B	C	D	E F
0	NUL DLE DS	SP	&	—	ソ					{	}	\				0
1	SOH DC1 SOS	。エ	/	ア	タ	〜				A	J					1
右 2	STX DC2 FS SYN	「	オ	イ	チ	ヘ				B	K	S				2
3	ETX TM	」	ヤ	ウ	ツ	ホ				C	L	T				3
側 4	PF RES BYP PN	、	ユ	エ	テ	マ				D	M	U				4
5	HT NL LF RS	・	ヨ	オ	ト	ミ				E	N	V				5
4 6	LC BS ETB UC	ヲ	ッ	カ	ナ	ム				F	O	W				6
7	DEL IL ESC EOT	ア		キ	ニ	メ				G	P	X				7
ビ 8	GE CAN	イ	ー	ク	ヌ	モ				H	Q	Y				8
9	RLF EM	ウ		ケ	ネ	ヤ				I	R	Z				9
ツ A	SMM CC SM	e	!		:	コ	ノ	ユ	レ							LVM
B	VT CU1 CU2 CU3	。¥	,	#						ロ						
ト C	FF IFS	DC4	<	*	%	@	サ	ヨ	ワ							
D	CR IGS ENQ NAK	()	—	'	シ	ハ	ラ	ン							
E	SO IRS ACK	+	;	>	=	ス	ヒ	リ	ˆ							
F	SI IUS BEL SUB		ˆ	?	"	セ	フ	ル	ˆ							E0

8

情報交換用8単位符号(8bits)

- ◆ **拡張2進化10進コードーEBCDIC英小文字コード**
(Extended Binary Coded Decimal Interchange Code)

		左側4ビット															
		0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
0	NUL	DLE	DS		SP	&	-							{	}	\	0
1	SOH	DC1	SOS		。	/				a	j	~		A	J		1
2	STX	DC2	FS	SYN	「					b	k	s		B	K	S	2
3	ETX	TM			」					c	l	t		C	L	T	3
4	PF	RES	BYP	PN	,					d	m	u		D	M	U	4
5	HT	NL	LF	RS	.					e	n	v		E	N	V	5
6	LC	BS	ETB	UC						f	o	w		F	O	W	6
7	DEL	IL	ESC	EOT						g	p	x		G	P	X	7
8	GE	CAN			—					h	q	y		H	Q	Y	8
9	RLF	EM								i	r	z		I	R	Z	9
A	SMM	CC	SM		e	!		:									LVM
B	VT	CU1	CU2	CU3	.	\$,	#									
C	FF	IFS		DC4	<	*	%	@									
D	CR	IGS	ENQ	NAK	()	_	'									
E	SO	IRS	ACK		+	;	>	=									
F	SI	IUS	BEL	SUB		?	"										E0

9

日本語文字コード

- ◆ **2バイト**から成り、漢字等を表現
- ◆ **JIS漢字コード**(JIS X 0208) : 情報交換用符号化漢字集合
- ◆ **シフトJISコード**: パソコン中心のコード
- ◆ **Unicode**(JIS X 0221): 国際符号化文字集合、ISO-10464-1のUCS-2、Windows NT, JABA, Macintoshで採用
- ◆ **EUCコード**: AT&Tの定めたコード、Unixで使われる

10

データの属性

- ◆ **命令語とデータ語の区別**:
計算機の処理サイクルで決まる。
フェッチサイクル ⇒ 命令語
実行サイクル ⇒ データ語

11

データの種別の指定方式

- ◆ **データの種別**:
2進表現・10進表現
固定小数点・浮動小数点等の区別をいかに行うか?
- ① 命令コード方式
- ② ディスクリプタ方式
- ③ タグ方式

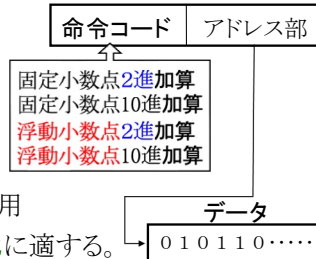
12

①命令コード方式

- ◆ **命令コード**で、データの種別を指定。アドレス部で直接データを指す。

命令の数:

2進表現・10進表現
固定小数点・浮動小数点
のデータ種別分必要!



- ◆ 現在の大多数のコンピュータで使用
 - ◆ 演算制御が**簡単**で、**高速化**に適する。
 - ◆ データの種別が増えると**命令数も増える**。
 - ◆ 対象とするデータの種別が異っても、一般にそれを**検出**できない。

13

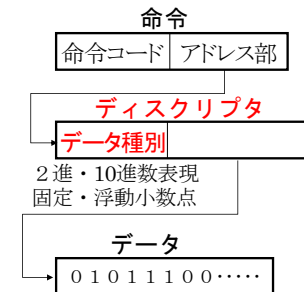
②ディスクリプタ方式

- ◆ 命令のアドレス部…データの**ディスクリプタ** (descriptor)を指す。

ディスクリプタが**データの種別**とデータの番地を記述する。

- ◆ 命令の種類が**少なく済む**。
異なる種別のデータの演算も**同じ命令**で可能。

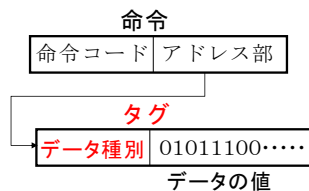
- ◆ ディスクリプタのアクセスのため**1回余分な主記憶アクセス**が必要。



14

③タグ方式

- ◆ 命令…**タグ**を指す。
タグがデータの種別とデータの値を記述。
- ◆ ディスクリプタ方式と同じで命令の種類が**少なく済む**。
異なる種別のデータの演算も**同じ命令**で可能。
- ◆ 主記憶アクセスは**少ない** (ディスクリプタ方式と比べて)。
- ◆ タグ中にデータ種別とデータ値をもつためタグが**長くなる**。



15

演習1ーデータ形式

- (1) 整数形コンピュータにおいて、2進数5桁の符号を考慮した数値データを考える。
 - (a) 正の整数15の2進数表示を書け。
 - (b) 負の整数-15の「1の補数による負数表現」の2進数表示を書け。
 - (c) 負の整数-15の「2の補数による負数表現」の2進数表示を書け。

16

演習2—データ形式

(2) 2進数5ビットのXとYは、先頭が符号ビットの正または負の整数を表す。

- ① XとYとの2進数加算 $(X)_2 + (Y)_2$ を行い、結果を書け。
- ② このとき、循環桁上げが必要なら、手順と結果を書け。
- ③ この結果を吟味せよ。この際、符号ビットに関する桁上げを考慮せよ。2進数加算 $(X)_2 + (Y)_2$ がオーバーフローのときは、オーバーフローと書け。

ただし、符号を考慮した値は、「1の補数による負数表現」とする。例えば、 $X=11001$ のとき、値-6を表す。

- (a) $X=00101$ 、 $Y=01001$
- (b) $X=00110$ 、 $Y=10110$
- (c) $X=11101$ 、 $Y=10110$
- (d) $X=10111$ 、 $Y=10110$

17

演習3—データ形式

(3) 整数形コンピュータで2進数6桁のデータにおいて、

- (a) 整数31の2進数表示を書け。
- (b) 負の整数-31の「符号と絶対値」による2進数表示を書け。
- (c) 負の整数-31の「1の補数」による2進数表示を書け。
- (d) 負の整数-31の「2の補数」による2進数表示を書け。
- (e) 「2の補数」を使ったとき表せる整数値の範囲を書け。

18

演習4—データ形式

(4) 11ビットを使った浮動小数点表現を考える。

下の図の浮動小数点表現の表す値の2進数表現と10進数表現を書け。

1	1001	111000
---	------	--------

ただし、左端の1ビットが仮数部の符号、その右隣の4ビットが指数部、残りの6ビットが仮数部を表す。

指数部、仮数部ともに2進数で表す。基数は、2進数とする。負の仮数は、符号と絶対値表現で表す。

仮数部は、正規化したうえでけち表現がとられている。

指数部は、バイアス値7とする。

19

演習5—データ形式

(5) データ形式に関し、次の問いに答えよ。

- (1) 数値表現には固定小数点方式と浮動小数点方式とがある。これらを概説せよ。
- (2) 浮動小数点方式の正規表現とけち表現を述べよ。
- (3) 浮動小数点方式のバイアス表現を述べよ。

20