

iEVO 2025

「①」や英小文字は使えるの？ 基礎的な IBM i の文字コードの話

2025/11/27

IBM Power エバンジェリスト 安井 賢克
ベル・データ株式会社



- EBCDIC は好きですか？
 - IBM i 独自である
 - 英小文字を処理できない
 - 理解しようとするといくつもの概念・用語が出てきて複雑だ
 - CCSID (Coded Character Set Identifier) , CS (Character Set) , CP (Code Page) , ES (Encoding Scheme) , JIS X 0208:1997, JIS X 0213:2004, ISO10646 . . .
 - 扱える文字数に制限がある
 - 「①」や「宥」、「m̐」などを扱えない
 - 外字に頼らざるを得なくなり、管理が面倒
 - シフトアウト・シフトインは意味不明なスペース？



EBCDIC に対する「モヤモヤ」を少しでも解消することをめざします
概要理解を優先し、必ずしも正確性にこだわらずに説明します

- EBCDIC (Extended Binary Coded Decimal Interchange Code) とは？

- Extended : 拡張された
- Binary Coded Decimal : 二進化十進数
- Interchange Code : データ交換用コード

商用計算、特に小数を含む場合は、
誤差が生じないので
現在も利用されることがあります

- Binary Coded Decimal (BCD) とは？

- 1964 年の System/360 登場前に主力だったデータ表現方法
- 0~9 のみが見使用され、十進数 1 桁と二進数 4 桁が対応する表記法
 - $0 = (0000\ 0000)_2$, $1 = (0000\ 0001)_2$, $2 = (0000\ 0010)_2$, \dots $9 = (0000\ 1001)_2$
 - $10 = (0001\ 0000)_2$, $11 = (0001\ 0001)_2$, \dots $15 = (0001\ 0101)_2$, $16 = (0001\ 0110)_2$ \dots
- 0~F が使用される、本来の十六進数・二進数表記
 - 0~9 は上記と同じ、 $10 = (A)_{16}$ 、 $11 = (B)_{16}$ 、 \dots $15 = (F)_{16}$ 、 $16 = (10)_{16}$ \dots
 - $10 = (0000\ 1010)_2$, $11 = (0000\ 1011)_2$, \dots $15 = (0000\ 1111)_2$, $16 = (0001\ 0000)_2$ \dots

- すなわち \dots

- 旧来の BCD に対して上位互換を維持し、文字を含むコード体系
- System/360 と共に登場

- 英文字だけでなく多くのヨーロッパ言語文字を含む
- 縦列における文字の割り当てが 0~9 で折り返されている
 - BCD の名残り
- 上位桁 0~F + 下位桁 0~F で表現され、256 文字を含む
 - 上位 4 ビット + 下位 4 ビット = 8 ビット = 1 (シングル) バイト
 - 1 バイト文字コード表
- 日本語文字を含むには 2 バイト (ダブル・バイト) が必要
 - システムのパワー不足のため当初はサポートできなかった

Code Page 00037 [ホスト・コード・ページ解説書](#)より

HEX DIGITS 1ST → 2ND ↓	4-	5-	6-	7-	8-	9-	A-	B-	C-	D-	E-	F-
-0	(SP) SP010000	& SM030000	- SP100000	ø LO610000	Ø LO620000	° SM190000	μ SM170000	^ SD150000	{ SM110000	}	\	0 ND100000
-1	(RSP) SP300000	é LE110000	/ SP120000	É LE120000	a LA010000	j LJ010000	~ SD190000	£ SC020000	A LA020000	J LJ020000	÷ SA060000	1 ND010000
-2	â LA150000	ê LE150000	Â LA160000	Ê LE160000	b LB010000	k LK010000	s LS010000	¥ SC050000	B LB020000	K LK020000	S LS020000	2 ND020000
-3	ä LA170000	ë LE170000	Ä LA180000	Ë LE180000	c LC010000	l LL010000	t LT010000	· SD630000	C LC020000	L LL020000	T LT020000	3 ND030000
-4	à LA130000	è LE130000	À LA140000	È LE140000	d LD010000	m LM010000	u LU010000	© SM520000	D LD020000	M LM020000	U LU020000	4 ND040000
-5	á LA110000	í LI110000	Á LA120000	Í LI120000	e LE010000	n LN010000	v LV010000	§ SM240000	E LE020000	N LN020000	V LV020000	5 ND050000
-6	ã LA190000	î LI150000	Ã LA200000	Î LI160000	f LF010000	o LO010000	w LW010000	¶ SM250000	F LF020000	O LO020000	W LW020000	6 ND060000
-7	ä LA270000	ï LI170000	Å LA280000	Ï LI180000	g LG010000	p LP010000	x LX010000	¼ NF040000	G LG020000	P LP020000	X LX020000	7 ND070000
-8	ç LC410000	ì LI130000	Ç LC420000	Ì LI140000	h LH010000	q LQ010000	y LY010000	½ NF010000	H LH020000	Q LQ020000	Y LY020000	8 ND080000
-9	ñ LN190000	β LS610000	Ñ LN200000	´ SD130000	i LI010000	r LR010000	z LZ010000	¾ NF050000	I LI020000	R LR020000	Z LZ020000	9 ND090000
-A	¢ SC040000	! SP020000	¡ SM650000	:	“ SM700000	” SM710000	“ SM720000	“ SM730000	(SHY) SP320000	1 ND011000	2 ND021000	3 ND031000
-B	· SP110000	\$ SC030000	, SP080000	# SM010000	» SP180000	“ SM200000	“ SP160000	“ SM080000	ô LO150000	û LU150000	Ô LO160000	Û LU160000
-C	< SA030000	* SM040000	% SM020000	@ SM050000	ð LD630000	æ LA510000	Ð LD620000	- SM150000	ö LO170000	ü LU170000	Ö LO180000	Ü LU180000
-D	(SP060000) SP070000	_ SP090000	' SP050000	ý LY110000	¸ SD410000	Ý LY120000	“ SD170000	ò LO130000	ù LU130000	Ò LO140000	Ù LU140000
-E	+ SA010000	; SP140000	> SA050000	= SA040000	þ LT630000	Æ LA520000	Þ LT640000	' SD110000	ó LO110000	ú LU110000	Ó LO120000	Ú LU120000
-F	 SM130000	⌋ SM660000	? SP150000	" SP040000	± SA020000	⌘ SC010000	® SM530000	× SA070000	õ LO190000	ÿ LY170000	Õ LO200000	(EO) ND000000

ここで折り返し

- 二通りの手法が実装された
 - 数字・英大文字とカタカナを中心に据え、英小文字を空き領域に割り当てる
 - コードページ **290**
 - 英小文字は当初は含まれていなかったが、現在は拡張されて英小文字を含むようになっている
→ システム上は拡張前後のコードページ **290** は識別されない
 - 英語環境との互換性に制約が生じ易い
 - 英小文字が不変文字セット（コード・ページ間でコード・ポイント割り当てが変わらない文字セット）の例外になっているため
 - 数字・英大小文字を中心に据え、カタカナを空き領域に割り当てる
 - コードページ **1027**
 - 英語環境との互換性に優れる
 - 英語圏にて開発された製品の実行環境として推奨されることが多い
 - 上記いずれにおいても、カタカナの表記サイズは英文字と同じ（半角カタカナ）
 - 日本語と同じサイズのカタカナはダブル・バイト文字コード表内で定義される（全角カタカナ）
- シングル・バイト文字コードのみでも日本語（カタカナ）は何とか取り扱えた

- カタカナ・英大文字・数字を中心とする文字コード表

- 旧 290* は英小文字を含まない
(現在は存在しない)

- 290* (拡張後・現行) は英小文字を含む

- 英小文字は空き領域に割り当てられている

- 割り当てられた値 (コードポイント) は 290 独自であり、英語圏製品・環境と相性の問題が生じ易い

- コマンド入力時になどにおいて正しく認識されない

- 英小文字利用時は推奨されない

* 文字セット拡張前後のコードページ 290 を便宜的に表現したものであり、システム上は両者は識別されません

Code Page 00290 [ホスト・コード・ページ解説書](#)より

HEX DIGITS 1ST → 2ND ↓	4-	5-	6-	7-	8-	9-	A-	B-	C-	D-	E-	F-
-0	(SP) SP010000	& SM030000	- SP100000	[SM060000] SM080000	ソ JS500000	~ SD190000	^ SD150000	{ SM110000	}	\$ SC030000	0 ND100000
-1	。 JQ700000	エ JE010000	/ SP120000	i LI010000	ア JA000000	タ JT100000	- SM150000	¢ SC040000	A LA020000	J LJ020000	€ SC200000	1 ND010000
-2	「 JQ710000	オ JO010000	a LA010000	j LJ010000	イ JI000000	チ JT200000	へ JH400000	\ SM070000	B LB020000	K LK020000	S LS020000	2 ND020000
-3	」 JQ720000	ヤ JY110000	b LB010000	k LK010000	ウ JU000000	ツ JT300000	ホ JH500000	t LT010000	C LC020000	L LL020000	T LT020000	3 ND030000
-4	、 JQ730000	ユ JY310000	c LC010000	l LL010000	エ JE000000	テ JT400000	マ JM100000	u LU010000	D LD020000	M LM020000	U LU020000	4 ND040000
-5	・ JQ740000	ヨ JY510000	d LD010000	m LM010000	オ JO000000	ト JT500000	ミ JM200000	v LV010000	E LE020000	N LN020000	V LV020000	5 ND050000
-6	ヲ JW500000	ツ JT310000	e LE010000	n LN010000	カ JK100000	ナ JN100000	ム JM300000	w LW010000	F LF020000	O LO020000	W LW020000	6 ND060000
-7	ア JA010000		f LF010000	o LO010000	キ JK200000	ニ JN200000	メ JM400000	x LX010000	G LG020000	P LP020000	X LX020000	7 ND070000
-8	イ JI010000	一 JX700000	g LG010000	p LP010000	ク JK300000	ヌ JN300000	モ JM500000	y LY010000	H LH020000	Q LQ020000	Y LY020000	8 ND080000
-9	ウ JU010000		h LH010000	` SD130000	ケ JK400000	ネ JN400000	ヤ JY100000	z LZ010000	I LI020000	R LR020000	Z LZ020000	9 ND090000
-A	£ SC020000	! SP020000		: SP130000	コ JK500000	ノ JN500000	ユ JY300000	レ JR400000				
-B	・ SP110000	¥ SC050000	,	# SM010000	q LQ010000	r LR010000	s LS010000	ロ JR500000				
-C	< SA030000	* SM040000	% SM020000	@ SM050000	サ JS100000		ヨ JY500000	ワ JW100000				
-D	(SP060000) SP070000	_ SP090000	' SP050000	シ JS200000	ハ JH100000	ラ JR100000	ン JN000000				
-E	+ SA010000	; SP140000	> SA050000	= SA040000	ス JS300000	ヒ JH200000	リ JR200000	・ JX710000				
-F	 SM130000	⌋ SM660000	? SP150000	" SP040000	セ JS400000	フ JH300000	ル JR300000	。 JX720000				(EO)

- 英大文字・数字を中心とする文字コード表
 - 英語圏との互換性が重視されたコードページ
 - 英大文字に割り当てられた値（コードポイント）は英語用コードページ **037** と共通
 - 英大文字・小文字はそれぞれに連続領域に割り当てられている
 - カタカナは空き領域に割り当てられている

HEX DIGITS 1ST → 2ND ↓	4-	5-	6-	7-	8-	9-	A-	B-	C-	D-	E-	F-
-0	(SP) SP010000	& SM030000	- SP100000	コ JK500000			- SM150000	^ SD150000	{ SM110000	}	\ SM070000	0 ND100000
-1		ウ JU010000	/ SP120000	サ JS100000	a LA010000	j LJ010000	~ SD190000	£ SC020000	A LA020000	J LJ020000	€ SC200000	1 ND010000
-2	。 JQ700000	エ JE010000	イ JI000000	シ JS200000	b LB010000	k LK010000	s LS010000	¥ SC050000	B LB020000	K LK020000	S LS020000	2 ND020000
-3	「 JQ710000	オ JO010000	ウ JU000000	ス JS300000	c LC010000	l LL010000	t LT010000	ヤ JY100000	C LC020000	L LL020000	T LT020000	3 ND030000
-4	」 JQ720000	ヤ JY110000	エ JE000000	セ JS400000	d LD010000	m LM010000	u LU010000	ユ JY300000	D LD020000	M LM020000	U LU020000	4 ND040000
-5	、 JQ730000	ユ JY310000	オ JO000000	ソ JS500000	e LE010000	n LN010000	v LV010000	ヨ JY500000	E LE020000	N LN020000	V LV020000	5 ND050000
-6	・ JQ740000	ヨ JY510000	カ JK100000	タ JT100000	f LF010000	o LO010000	w LW010000	ラ JR100000	F LF020000	O LO020000	W LW020000	6 ND060000
-7	ヲ JW500000	ツ JT310000	キ JK200000	チ JT200000	g LG010000	p LP010000	x LX010000	リ JR200000	G LG020000	P LP020000	X LX020000	7 ND070000
-8	ア JA010000	ー JX700000	ク JK300000	ッ JT300000	h LH010000	q LQ010000	y LY010000	ル JR300000	H LH020000	Q LQ020000	Y LY020000	8 ND080000
-9	イ JI010000	ア JA000000	ケ JK400000	` SD130000	i LI010000	r LR010000	z LZ010000	レ JR400000	I LI020000	R LR020000	Z LZ020000	9 ND090000
-A	¢ SC040000	! SP020000		: SP130000	テ JT400000	ノ JN500000	マ JM100000	ロ JR500000				
-B	・ SP110000	\$ SC030000	, SP080000	# SM010000	ト JT500000	ハ JH100000	ミ JM200000	ワ JW100000				
-C	< SA030000	* SM040000	% SM020000	@ SM050000	ナ JN100000	ヒ JH200000	ム JM300000	ン JN000000				
-D	(SP060000) SP070000	_ SP090000	' SP050000	ニ JN200000	フ JH300000	[SM060000]				
-E	+ SA010000	; SP140000	> SA050000	= SA040000	ヌ JN300000	ヘ JH400000	メ JM400000	・ JX710000				
-F	 SM130000	⌋ SM660000	? SP150000	" SP040000	ネ JN400000	ホ JH500000	モ JM500000	° JX720000				(EO)

Code Page 01027 [ホスト・コード・ページ解説書](#)より

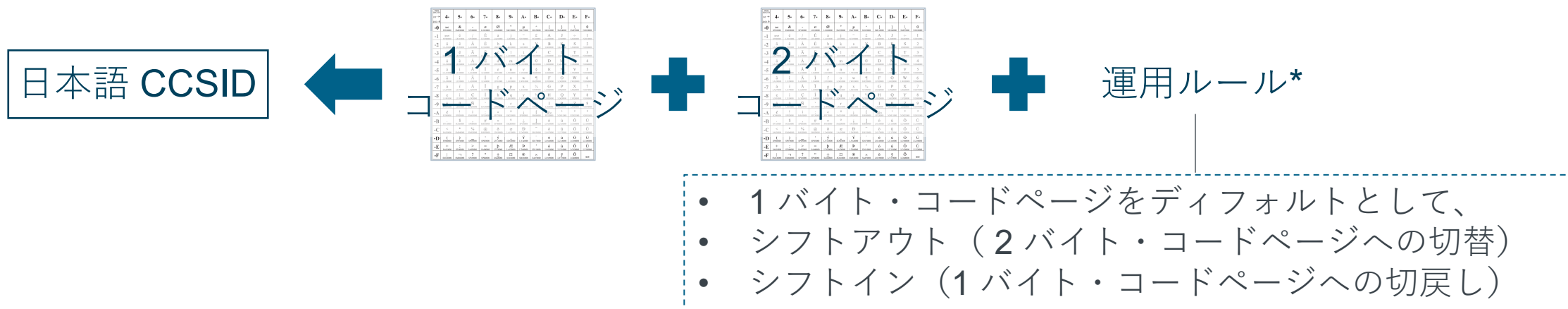
- ひらがな、カタカナ、漢字、特殊記号などを含む
 - コードページ 300 にあるカタカナは全角（ひらがなや漢字と同じサイズで表記）
 - コードページ 290・1027 にあるカタカナは半角（英大小文字と同じサイズで表記）
- 新旧二種類のコードページ
 - 旧 300 *
 - 第一・第二水準漢字、非漢字（計 6,879）を含む
 - 新 300 *
 - 第一・第二・第三・第四水準漢字、非漢字（計約 16,000 文字）を含む
 - 外字（ユーザー定義文字）に頼らなくても済むケースがある
 - 追加されている非漢字の例
 - ①、②、(有)、kg、𐤎、𐤏 など

* 実際には同一コードページ 300 における CS（Character Set）の違いによって両者は識別されます

- CCSID: Coded Character Set Identifier
 - EBCDIC、Unicode、Shift JIS などあらゆる文字コードを包含する IBM の文字セット識別情報
- EBCDIC 英語環境の CCSID
 - 1 バイト・コードページのみで構成される



- EBCDIC 日本語環境の CCSID
 - 1 バイト・コードページと 2 バイト・コードページの組み合わせ



- EBCDIC 日本語 CCSID の違い
 - 1 バイト・コードページは英語環境との互換性が高いか？
 - 2 バイト・コードページに含まれる文字数は十分に大きいか？

	EBCDIC CCSID	1 バイト・ コードページ	2 バイト・ コードページ
英語	037	037	-
日本語	5026	290	旧 300
	5035	1027	旧 300
	1399	1027	新 300
定義無し	65535	-	-

日本ではおそらく最も多くのユーザー

第三・第四水準漢字を含み文字数が多い

英語環境との互換性が高い

- EBCDIC と UNICODE ・ Shift JIS の共存環境における考慮点

- エミュレータ・ソフトによる翻訳が必要
- ホスト・コード・ページを正しく設定する

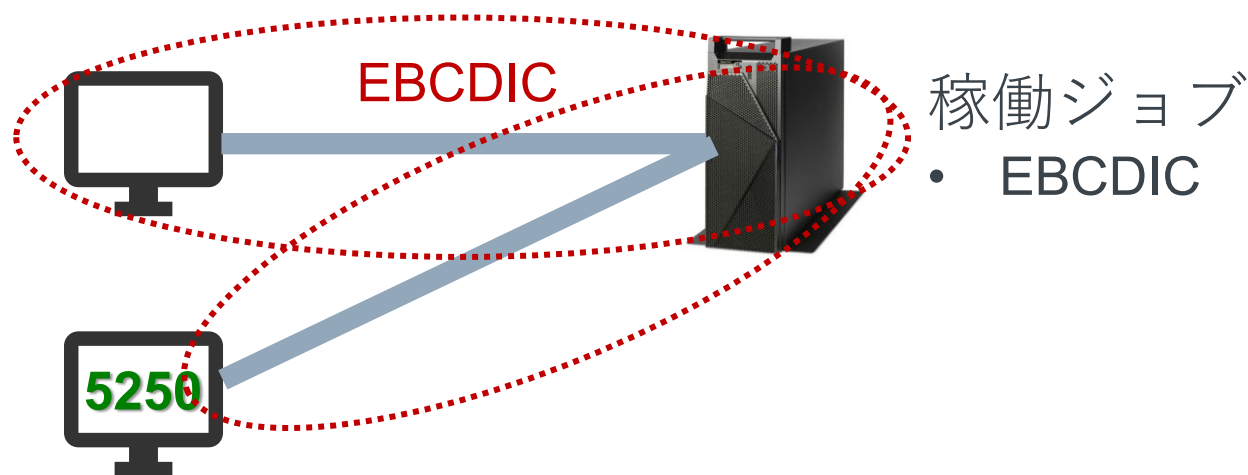
旧来の専用端末

- EBCDIC



現在の汎用端末

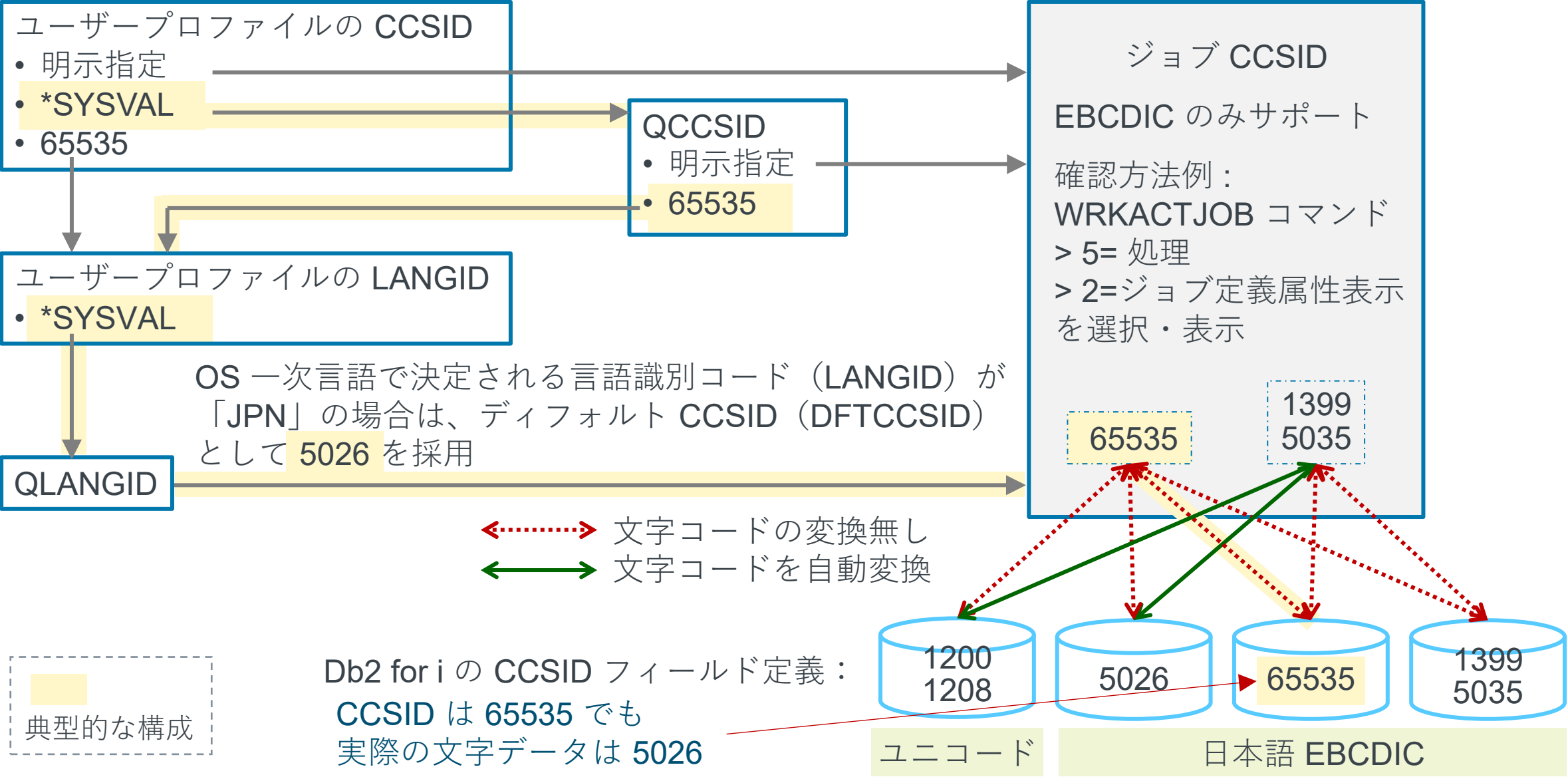
- OS
 - UNICODE ・ Shift JIS で稼働
- 5250 エミュレータ
 - 専用端末の「ふりをさせる」ためのソフトウェア
 - EBCDIC と UNICODE ・ Shift JIS の双方向同時翻訳を行う
 - 翻訳対象 EBCDIC を「ホスト・コード・ページ」で設定する
 - 「ホスト・コード・ページ」はサーバーの稼働ジョブ EBCDIC CCSID に対応させる




稼働ジョブ
• EBCDIC


Emulate: まねる、匹敵する


IBM i 稼働ジョブの CCSID はどうやって決定されるのか



- 英小文字利用可否は設定次第
 - サーバー上のジョブ **CCSID** に合わせてエミュレータのホスト・コード・ページを設定する







	ホスト・コード・ページ設定値	想定される ジョブ CCSID	1 バイト・ コードページ	2 バイト・ コードページ
英語	037 米国	037	037	-
日本語	930 日本語（カタカナ）	旧 5026	旧 290	旧 300
	930 日本（拡張カタカナ）	5026	290	旧 300
	939 日本（拡張ローマ字）	5035	1027	旧 300
	1399 日本語（Latin Unicode 拡張; JIS2004）	1399	1027	新 300
		65535	-	-

- デフォルト設定値
- 英小文字を入力・表示*できない
- 「930 日本（拡張カタカナ）」のサブセット

- 英小文字をサポートしない
- 旧 5026（旧 290 + 旧 300）は 5026（290 + 旧 300）に拡張されている

* ACS のエミュレータでは英小文字の入力は不可だが表示は可能

- 英小文字利用可否は設定次第
 - サーバー上のジョブ **CCSID** に合わせてエミュレータのホスト・コード・ページを設定する



	ホスト・コード・ページ	想定される ジョブ CCSID	1 バイト・ コードページ	2 バイト・ コードページ
英語	037 米国	037	037	-
日本語	930 日本語（カタカナ）	旧 5026	旧 290	旧 300
	930 日本（拡張カタカナ）	5026	290	旧 300
	939 日本（拡張ローマ字）	5035	1027	旧 300
	1399 日本語（Latin Unicode 拡張; JIS2004）	1399	1027	新 300
		65535	-	-

- 英小文字を入力・表示できる
- 「930 日本（カタカナ）」のスーパーセット

- 英小文字をサポートする
- 英小文字のコードポイントは独自であり、コマンド入力などにおいて支障が出る
- エミュレータのホスト・コード・ページが「930 日本語（カタカナ）」だと英小文字は入力されない

- 英小文字利用可否は設定次第
 - サーバー上のジョブ **CCSID** に合わせてエミュレータのホスト・コード・ページを設定する

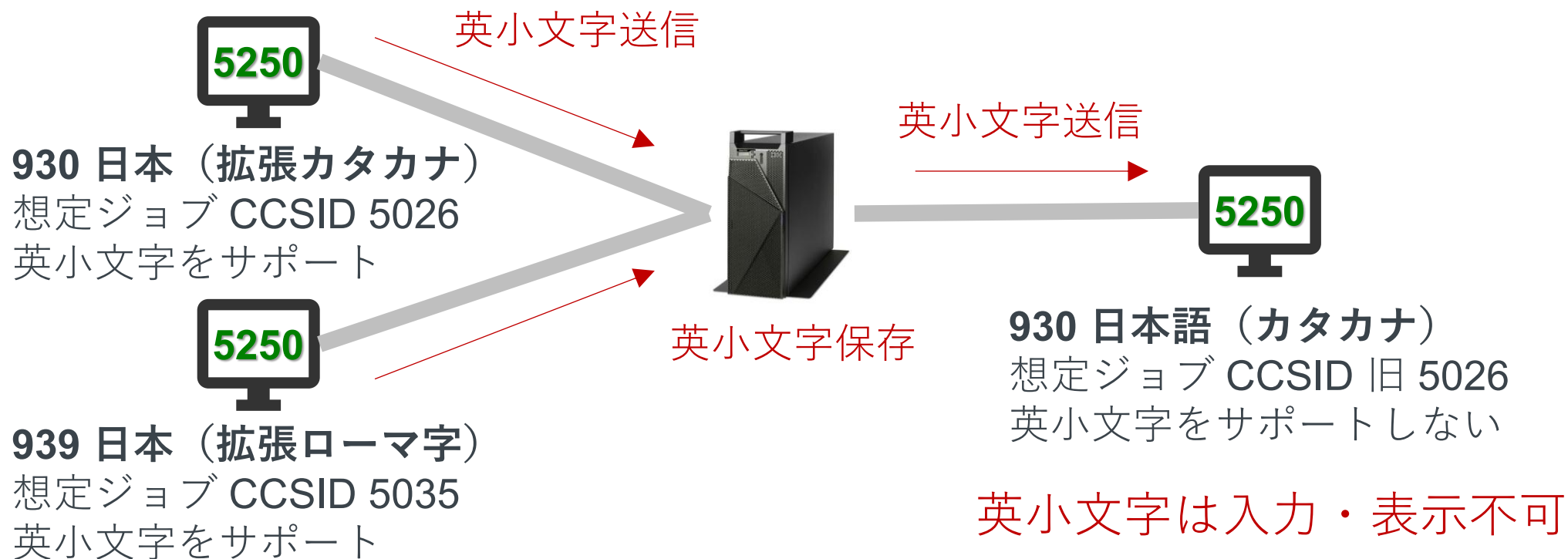


	ホスト・コード・ページ	想定される ジョブ CCSID	1 バイト・ コードページ	2 バイト・ コードページ
英語	037 米国	037	037	-
日本語	930 日本語（カタカナ）	旧 5026	旧 290	旧 300
	930 日本（拡張カタカナ）	5026	290	旧 300
	939 日本（拡張ローマ字）	5035	1027	旧 300
	1399 日本語（Latin Unicode 拡張; JIS2004）	1399	1027	新 300
		65535	-	-

• 英小文字を入力・表示できる

• 英小文字をサポートする
• 英小文字のコードポイントは英語圏と共通であり、設定値として推奨される

- エミュレータのホスト・コード・ページを単一システム上に混在させると・・・



- 日本語 **EBCDIC** でも英小文字は処理できる
 - エミュレータのホスト・コード・ページを正しく設定する
- 長期的には **CCSID 1399** が望ましい
 - 英語環境との互換性が高く、オープン系の各種ソフトウェアを取込み易い
 - 第三・第四水準漢字に加えて、数多くの非漢字をサポートできる
- 文字コード設定の変更は単純作業ではない
 - 全体の整合性を維持する
 - エミュレータのホスト・コード・ページ
 - システムのジョブ **CCSID**
 - データベース文字フィールドの **CCSID**: 実データと設定値
 - 設定を誤ると文字化けが生じる可能性がある
- 正しく構成すれば、文字コード自動変換機能を活かすことができる

- 文字コード概説動画
 - 文字コード概説 - 一般編 - : <https://www.e-bellnet.com/category/knowledges/doc/2004-03.php>
 - 文字コード概説 (IBM i 編) : <https://www.e-bellnet.com/category/knowledges/doc/mojicode-video-ibmi.php>
- 「IBM i の文字コードを考える」
 - アイマガジン記事 : <https://www.imagazine.co.jp/ibm-i/e3%81%ae%e6%96%87%e5%ad%97%e3%82%b3%e3%83%bc%e3%83%89%e3%82%92%e8%80%83%e3%81%88%e3%82%8b%e3%80%80%ef%bd%9eebcdic%e3%81%a8unicode%ef%bd%9c%e6%97%a2%e5%ad%98%e8%b3%87%e7%94%a3%e3%82%92%e6%b4%bb/>
- 「CCSID 1399 への移行のご紹介（メリットと移行の考慮点）」
 - IBM 提供資料 : <https://ibm.ent.box.com/s/kwwpgirry0vj933j2q0ccsqh4cz3u677>
- IBM の EBCDIC 日本語 2 バイト文字セット（コードページ 300）仕様
 - IBM Corporate Specification CH 3-3220-024 : <https://public.dhe.ibm.com/software/globalization/gcoc/attachments/CP00300.pdf>

ありがとうございました。

