

照合順序(COLLATE)について

作成日: 2017/12/15

作成者: 西村

更新履歴

更新日	更新概要	作業者
2017/12/15	・ 新規作成	西村
	・	
	・	
	・	
	・	

はじめに

この資料では、DB(RDBMS, SQL)で使用する照合順序(COLLATE)について簡単に説明します。

照合順序？

「照合順序」(COLLATE ※「これーと」, Collation)は、文字の順番(大きい、小さい、同じ)を示したものです。例えば

- ・ 「a」は「a」と同じ
- ・ 「a」は「b」より小さい
- ・ 「a」は「A」より小さい / 「a」は「A」と同じ

のような、「文字の並び順」「どの文字をどの文字と一緒にみなすか」を判断するための材料として照合順序が使われます。

照合順序は、主に下記の部分に影響します。

- ・ WHERE での比較 (「=」「<=」「>=」など)
- ・ ORDER BY での並び順

文字の並び順？

「文字をどのような並びで整頓するか(並べるか)」はいろいろなルールがありますが、だいたい基本的には

- ・ 使用している文字コード(文字エンコーディング。UTF-8、Shift_JIS など)のコード順で文字が並ぶ
 - ※ 半角の記号や英数字が前、全角文字(ひらがな、かたかな、漢字など)が後の方、という形です。文字コードによって文字に対するコード番号が異なるので、例えば Shift_JIS だと半角カタカナは最初の方に並んでくるのに、UTF-8 だと後ろの方に並んでくる、といったことがあります。
- ・ 「どの文字をどの文字と一緒にみなすか」によって、さらに順序が調整される

という形になります。

どの文字をどの文字と一緒にみなす？

「どの文字をどの文字と一緒にみなすか」の設定は DB 製品によっていろいろな設定がありますが、主に下記のようなものがあります。

- ・ 「a」「A」のような英大文字と英小文字
- ・ 「A」「A」、 「ア」「ア」のような半角と全角
- ・ 「は」「ば」「ぱ」のような濁点ありとなし
- ・ 「あ」「ア」のようなひらがなとカタカナ

※ 「a」「A」のような英大文字と英小文字を一緒にみなす / みなさないについては、特に「Case insensitive (CI。大文字小文字を区別せず、一緒にみなす)」「Case sensitive (CS。大文字小文字を区別し、一緒にみなさない)」と呼ばれることがあります。(「sensitive」は「敏感」という意味です)
例えば SQL Server の「Japanese_CI_AI」、MySQL の「utf8_general_ci」などにこの用語が出てきます。

照合順序の例

照合順序は、DB 製品によって名前や動きが異なります。

例

DB 製品	照合順序名	意味
MySQL	utf8_general_ci	<ul style="list-style-type: none">・UTF-8 の文字コード順 (utf8)・半角と全角や濁点を区別する (general)・英大文字と英小文字を区別しない (ci)
	utf8_unicode_ci	<ul style="list-style-type: none">・UTF-8 の文字コード順 (utf8)・半角と全角や濁点を区別しない (unicode)・英大文字と英小文字を区別しない (ci)
	utf8mb4_bin	<ul style="list-style-type: none">・UTF-8(4 バイト対応)の文字コード順 (utf8mb4)・バイナリコード順ですべて区別する (bin)
SQL Server	Japanese_CI_AS	<ul style="list-style-type: none">・日本語順 (Japanese)・英大文字と英小文字を区別しない (CI)・アクセント(濁点等)を区別する (AS)
	Japanese_CI_AI	<ul style="list-style-type: none">・日本語順 (Japanese)・英大文字と英小文字を区別しない (CI)・アクセント(濁点等)を区別しない (AI)
	Japanese_CS_AS	<ul style="list-style-type: none">・日本語順 (Japanese)・英大文字と英小文字を区別する (CS)・アクセント(濁点等)を区別する (AS)

SQL での指定の仕方

SQL では、主に下記で照合順序の指定ができます。

- ・ データベースやテーブル、列の作成時
- ・ WHERE
- ・ ORDER BY

基本的には、データベースやテーブル、列の作成時に照合順序を指定したのちは WHERE や ORDER BY などの SQL 文内で照合順序を指定することはありません。

WHERE で指定する例

WHERE 列名 = 値 COLLATE 照合順序名

WHERE name = 'あ' COLLATE utf8_general_ci
--

ORDER BY で指定する例

ORDER BY 列名 COLLATE 照合順序名

ORDER BY name COLLATE utf8_general_ci
