A.11.1 What CJK character sets are available in MySQL?

The list of CJK character sets may vary depending on your MySQL version. For example, the gb18030 character set is not supported prior to MySQL 5.7.4. However, since the name of the applicable language appears in the DESCRIPTION column for every entry in the INFORMATION_SCHEMA.CHARACTER_SETS table, you can obtain a current list of all the non-Unicode CJK character sets using this query:

```
mysql> SELECT CHARACTER_SET_NAME, DESCRIPTION
       FROM INFORMATION_SCHEMA.CHARACTER_SETS
       WHERE DESCRIPTION LIKE '%Chin%'
       OR DESCRIPTION LIKE '%Japanese%'
       OR DESCRIPTION LIKE '%Korean%'
       ORDER BY CHARACTER_SET_NAME;
+--------------------+-------------------------------+
| CHARACTER_SET_NAME | DESCRIPTION                   |
+--------------------+-------------------------------+
| big5               | Big5 Traditional Chinese      |
| cp932              | SJIS for Windows Japanese     |
| eucjpms            | UJIS for Windows Japanese     |
| euckr              | EUC-KR Korean                 |
| gb18030            | China National Standard GB18030 |
| gb2312             | GB2312 Simplified Chinese     |
| gbk                | GBK Simplified Chinese        |
| sjis               | Shift-JIS Japanese            |
| ujis               | EUC-JP Japanese               |
+--------------------+-------------------------------+
```

(For more information, see Section 26.3.4, "The INFORMATION_SCHEMA CHARACTER_SETS Table".)

MySQL supports three variants of the *GB* (*Guojia Biaozhun*, or *National Standard*, or *Simplified Chinese*) character sets which are official in the People's Republic of China: gb2312, gbk, and (as of MySQL 5.7.4) gb18030.

Sometimes people try to insert gbk characters into gb2312, and it works most of the time because gbk is a superset of gb2312. But eventually they try to insert a rarer Chinese character and it does not work. (For an example, see Bug #16072).

Here, we try to clarify exactly what characters are legitimate in gb2312 or gbk, with reference to the official documents. Please check these references before reporting gb2312 or gbk bugs:

• The MySQL gbk character set is in reality "Microsoft code page 936". This differs from the official gbk for characters A1A4 (middle dot), A1AA (em dash), A6E0-A6F5, and A8BB-A8C0.