

oracle 中如何判断一个字符串是否含有汉字

1.1 BLOG 文档结构图

oracle 中如何判断一个字符串是否含有汉字
1.1 BLOG 文档结构图
1.2 前言部分
1.2.1 导读
1.2.2 实验环境介绍
1.2.3 本文简介
1.3 实验部分
1.3.1 lengthb 和 length 函数结合 to_single_byte 函数
1.3.2 regexp_replace --替换其它字符为"
1.3.3 to_multi_byte 函数--全是全角字符
1.3.4 网友问题解答：某个字段数据是不是包含了全角字 ...
1.4 总结
1.5 About Me

1.2 前言部分

1.2.1 导读

各位技术爱好者，看完本文后，你可以掌握如下的技能，也可以学到一些其它你所不知道的知识，~O(∩_∩)O~：

①全角字符的判断，或者是含有汉字的字符串的判断

本文如有错误或不完善的地方请大家多多指正，ITPUB 留言或 QQ 皆可，您的批评指正是我写作的最大动力。

1.2.2 实验环境介绍

11.2.0.3 RHEL6.5

1.2.3 本文简介

看到网友问，怎么查询表中某个字段数据是不是包含了全角字符啊？这个问题涉及到几个函数：to_single_byte、length 和 lengthb，我之前做开发的时候研究的是如何判断一个字符串中是否包含中文，其实和这个本质是一样的，且看实验部分。



1.3 实验部分

1.3.1 lengthb 和 length 函数结合 to_single_byte 函数

---含有汉字，严格的说是含有全角字符

```
SELECT l.name,
       length(l.name),
       lengthb(l.name)
FROM   xb_link l
WHERE  length(l.name) != lengthb(l.name)
AND    length(l.name) < 20;
```

NAME	LENGTH(L.NAME)	LENGTHB(L.NAME)
(0000)-(220VAC输入01)	19	21
(-24VDC输出01)-(串联)	17	21
(接地排)-(ZXJ10)	13	16
(接地排)-(总配线架/接地条)	16	26
(接地排)-(总配线架)	12	19
(接地排)-(MA5103)	14	17
(接地排)-(综合柜)	11	17
(接地排)-(综合配线柜/接地排)	17	28
(03)-(13楼计算机中心)	15	21
(09)-(13楼计算机中心)	15	21

以下数据也满足条件：

```
SELECT l.id,
       l.name
FROM   xb_link l
WHERE  length(l.name) != lengthb(l.name)
AND    l.metacategory IN
      ('com.gxlu.ngrm.network.DDNCircuit',
       'com.gxlu.ngrm.network.FRCircuit',
       'com.gxlu.ngrm.network.ATMCircuit',
       'com.gxlu.ngrm.network.DDNOCircuit',
       'com.gxlu.ngrm.network.FROCircuit')
AND    l.id IN ('301898331', '301898335', '301908187', '301929403');
```

ID	NAME
301898331	103080877 (101001051)
301898335	101009051 (103067711)
301908187	101017112 (101017111)
301929403	103080531, 103080531

所以可以借助 to_single_byte 函数来解决。

```
SELECT l.id,
       l.name,
       to_single_byte(l.name),
       length(l.name) l1,
       lengthb(l.name) l2,
       length(to_single_byte(l.name)) l
FROM   xb_link l
WHERE  length(l.name) != lengthb(l.name)
AND    l.metacategory IN
      ('com.gxlu.ngrm.network.DDNCircuit',
       'com.gxlu.ngrm.network.FRCircuit',
       'com.gxlu.ngrm.network.ATMCircuit',
       'com.gxlu.ngrm.network.DDNOCircuit',
```

```
'com.gxlu.ngrm.network.FROCircuit')
```

```
AND l.id IN ('301898331', '301898335', '301908187', '301929403');
```

ID	NAME	TO_SINGLE_BYTE(LNAME)	L1	L2	L
301898331	103080877 (101001051) ...	103080877(101001051) ...	20	22	20
301898335	101009051 (103067711) ...	101009051(103067711) ...	20	22	20
301908187	101017112 (101017111) ...	101017112(101017111) ...	20	22	20
301929403	103080531, 103080531 ...	103080531,103080531 ...	19	20	19

1.3.2 regexp_replace --替换其它字符为”

```
SELECT
  l.id,
  l.name
FROM   xb_link l
WHERE  regexp_replace(TRIM(l.name),
                      '([A-Za-z0-9]|[:punct:][:space:])',
                      '') IS NOT NULL;
```

ID	NAME
300092460	主城区华为OSN10G电信移动互联互通环2/人和移动7500-人和移动3500 ...
300092461	主城区华为OSN10G电信移动互联互通环2/人和移动7500-人和移动3500 ...
300092462	主城区中兴10G电信联通互联互通环3/101-回龙-100-五里店/双向/10G ...
300092463	主城区中兴10G电信联通互联互通环3/100-五里店-102-红旗河沟(联通)/ ...
300092465	主城区中兴10G电信联通互联互通环3/104-照母山-106-南方花园(联通)/ ...

1.3.3 to_multi_byte 函数--全是全角字符

---全是汉字

```
SELECT l.name,
       to_multi_byte(l.name)
FROM   xb_link l
WHERE  l.name = to_multi_byte(l.name);
```

NAME	TO_MULTI_BYTE(
清山茶楼互联网 ...	清山茶楼互联网 ...
清升司法所 ...	清升司法所 ...
清升信用社 ...	清升信用社 ...
李明荣 ...	李明荣 ...

```
SELECT l.name,
       to_multi_byte(l.name)
FROM   xb_link l;
```

NAME	TO_MULTI_BYTE(
102	1 0 2
103	1 0 3
104	1 0 4
105	1 0 5
106	1 0 6

1.3.4 网友问题解答：某个字段数据是不是包含了全角字符？

```
drop table aa;
create table aa (col varchar2(255));
SELECT * FROM aa for update;
```

	COL
1	123ABC34
2	123ABa s f a s f C346 7
3	1 2 1 3
4	1 2 3我的家里1 2 3 15
5	asfsafda孩子

```
SELECT a.col,
       to_single_byte(a.col),
       length(a.col),
       lengthb(a.col),
       length(to_single_byte(a.col)),
       lengthb(to_single_byte(a.col))
FROM   aa a
WHERE  (lengthb(a.col) - length(a.col)) <>
       (lengthb(to_single_byte(a.col)) - length(to_single_byte(a.col)));
```

	COL	TO_SINGLE_BYTE(A.COL)	LENGTH(A.COL)	LENGTHB(A.COL)	LENGTH(TO_SINGLE_BYTE(A.COL))	LENGTHB(TO_SINGLE_BYTE(A.COL))
1	123ABa s f a s f C346 7	123ABasfasf C3467	17	26	17	17
2	1 2 1 3	1213	4	8	4	4
3	1 2 3我的家里1 2 3 15	123我的家里12315	12	22	12	16

1.4 总结

方法很简单，网友可能还有其他的办法，欢迎留言，对于不同的场景处理方式有很多种，我们应该学会灵活变通。

1.5 About Me

.....

本文作者：小麦苗，只专注于数据库的技术，更注重技术的运用

ITPUB BLOG：<http://blog.itpub.net/26736162>

本文地址：<http://blog.itpub.net/26736162/viewspace-1688209/>

本文pdf版：<http://yunpan.cn/QCwUAI9bn7g7w> 提取码：af2d

QQ：642808185 若加 QQ 请注明你所正在读的文章标题

创作时间地点：2015-06-05 10:00~ 2015-06-05 13:00 于外汇交易中心

<版权所有，文章允许转载，但须以链接方式注明源地址，否则追究法律责任!>

.....