

QGIS Application - Bug report #8332

QGIS uses comma as decimal separator for shapefile attribute data

2013-07-20 09:41 AM - marisn -

Status: Closed	
Priority: Normal	
Assignee:	
Category: Data Provider/OGR	
Affected QGIS version: master	Regression?: No
Operating System:	Easy fix?: No
Pull Request or Patch supplied:	Resolution: up/downstream
Crashes QGIS or corrupts data:	Copied to github as #: 17120
Description	
<p>I'm using QGIS with Latvian locale. It's decimal separator is comma and thousands separator - point.</p> <p>When creating new object, QGIS allows to use only point as a decimal separator (OK. So be it). Still in DBF file it's written with comma as a decimal separator and thus only QGIS sees a correct value. Other programs, like ogr2ogr and dbfdump see only whole part of number and drop anything after decimal separator (floor to integer with all 0's after comma). cat'ing DBF file is displaying correct comma separated values.</p> <p>Attachment contains few points from a larger shapefile that is correctly readable only by QGIS.</p> <p>QGIS versija 1.9.0-Master QGIS code revision e0a0a3a Compiled against Qt 4.8.4 Running against Qt 4.8.4 Compiled against GDAL/OGR 1.10.0 Running against GDAL/OGR 1.10.0</p> <p>cat comma_separated.dbf: 68,50010k 66,30010k</p> <p>ogrinfo: AUGSTUMS (Real) = 68.000 PIEZIMES (String) = 10k AUGSTUMS (Real) = 66.000 PIEZIMES (String) = 10k</p>	
Related issues:	
Related to QGIS Application - Bug report # 6110: KML export is generating dec...	Closed 2012-07-27

Associated revisions

Revision c64a051b - 2014-02-12 04:33 PM - Jürgen Fischer

ogr provider: write numeric data in C locale (fixes #8332)

History

#1 - 2013-07-20 10:32 AM - Jürgen Fischer

- Status changed from Open to Feedback

Platform? On Linux I couldn't reproduce the problem - even with LANG=lv_LV.UTF-8 (although I figure de_DE should have the same problem).

#2 - 2013-07-21 06:14 AM - marisn -

Gentoo Linux ~AMD64

```
$ locale
LANG=lv_LV
LC_CTYPE="lv_LV.utf8"
LC_NUMERIC="lv_LV.utf8"
LC_TIME="lv_LV.utf8"
LC_COLLATE="lv_LV.utf8"
LC_MONETARY="lv_LV.utf8"
LC_MESSAGES="lv_LV.utf8"
LC_PAPER="lv_LV.utf8"
LC_NAME="lv_LV.utf8"
LC_ADDRESS="lv_LV.utf8"
LC_TELEPHONE="lv_LV.utf8"
LC_MEASUREMENT="lv_LV.utf8"
LC_IDENTIFICATION="lv_LV.utf8"
LC_ALL=lv_LV.utf8
```

Here's output of a new shapefile created with QGIS:

```
$ cat rm_comma_test.dbf
   aidN
skaitlisN
*****  3,400 *****  6,500
$ dbfdump rm_comma_test.dbf
   id  skaitlis
(NULL)  3.000
(NULL)  6.000
```

As it's visible, 3,4 and 6,5 are turning into 3.0 and 6.0 in dbfdump and ogrinfo output. Interesting that QGIS accepts only 3.5 not 3,5 in it's attribute data form, still stores with comma.

#3 - 2013-07-21 06:54 AM - Jürgen Fischer

marisn - wrote:

As it's visible, 3,4 and 6,5 are turning into 3.0 and 6.0 in dbfdump and ogrinfo output. Interesting that QGIS accepts only 3.5 not 3,5 in it's attribute data form, still stores with comma.

strange. Doesn't happen here on debian. What type does the field have? Which OGR version is in play?

#4 - 2013-07-21 07:18 AM - marisn -

```
ogrinfo rm_comma_test.shp rm_comma_test
INFO: Open of `rm_comma_test.shp'
      using driver `ESRI Shapefile' successful.
```

Layer name: rm_comma_test
Geometry: Point
Feature Count: 2
Extent: (632458.729239, 382529.825215) - (646246.187935, 389491.361737)

Layer SRS WKT:

```
PROJCS["LKS92_Latvia_TM",  
  GEOGCS["GCS_LKS92",  
    DATUM["Latvia_1992",  
      SPHEROID["GRS_1980",6378137,298.257222101]],  
    PRIMEM["Greenwich",0],  
    UNIT["Degree",0.017453292519943295]],  
  PROJECTION["Transverse_Mercator"],  
  PARAMETER["latitude_of_origin",0],  
  PARAMETER["central_meridian",24],  
  PARAMETER["scale_factor",0.9996],  
  PARAMETER["false_easting",500000],  
  PARAMETER["false_northing",-6000000],  
  UNIT["Meter",1]]
```

id: Integer (10.0)

skaitlis: Real (10.3)

OGRFeature(rm_comma_test):0

id (Integer) = (null)

skaitlis (Real) = 3.000

POINT (646246.187935094581917 382529.825215096992906)

OGRFeature(rm_comma_test):1

id (Integer) = (null)

skaitlis (Real) = 6.000

POINT (632458.729239132953808 389491.361736992374063)

ogrinfo --version

GDAL 1.10.0, released 2013/04/24

#5 - 2013-09-29 04:58 AM - marisn -

- Target version changed from Version 2.0.0 to Future Release - High Priority

Still an issue with 2.x

QGIS versija 2.1.0-Master QGIS code revision b2396f6

Compiled against Qt 4.8.5 Running against Qt 4.8.5

Compiled against GDAL/OGR 1.10.0 Running against GDAL/OGR 1.10.0

#6 - 2014-01-26 10:56 AM - Giovanni Manghi

- Status changed from Feedback to Open

#7 - 2014-02-12 07:33 AM - Jürgen Fischer

- Status changed from Open to Closed

Fixed in changeset commit:"c64a051b1c716e5d34fc385b92ec7971ae8fec26".

#8 - 2014-09-20 01:07 PM - Andrey Isakov

- Status changed from Closed to Reopened

Same issue for me under Debian testing.

QGIS 2.4.0-Chugiak

Compiled against GDAL/OGR 1.10.1 Running against GDAL/OGR 1.10.1

```
andrey@crt-s1:~$ locale
LANG=en_US.UTF-8
LANGUAGE=
LC_CTYPE="en_US.UTF-8"
LC_NUMERIC=ru_RU.utf8
LC_TIME=ru_RU.utf8
LC_COLLATE="en_US.UTF-8"
LC_MONETARY=ru_RU.utf8
LC_MESSAGES="en_US.UTF-8"
LC_PAPER=ru_RU.utf8
LC_NAME="en_US.UTF-8"
LC_ADDRESS="en_US.UTF-8"
LC_TELEPHONE="en_US.UTF-8"
LC_MEASUREMENT=ru_RU.utf8
LC_IDENTIFICATION="en_US.UTF-8"
LC_ALL=
```

QGIS using the system locale uses comma as decimal separator instead of point.

If QGIS started using the LC_NUMERIC="C" qgis then decimal separator is point.

Related issue #6110

#9 - 2014-10-23 01:07 PM - Jürgen Fischer

- Category set to Data Provider/OGR

#10 - 2016-02-25 05:00 AM - Jürgen Fischer

- Status changed from Reopened to Feedback

The commit should fix the production of shape files. A newly created shape file shouldn't have commas. It doesn't fix the use of the broken comma_separated.shp. And I can't get QGIS to produce those broken files anymore. Does that still happen for others?

#11 - 2016-04-09 11:45 AM - Giovanni Manghi

- Status changed from Feedback to Closed

- Resolution set to not reproducible

closing for lack of feedback. Please reopen if necessary.

#12 - 2017-02-06 02:21 PM - Marco Bernasocchi

- Status changed from Closed to Reopened

- File data_with_locale.zip added

I just had a client producing the attached file in QGIS 2.18 which contains commas as separator in the dbf file. I also found a potential problem in gdal and reported it <https://trac.osgeo.org/gdal/ticket/6804>

#13 - 2017-02-07 08:39 AM - Marco Bernasocchi

- Resolution deleted (not reproducible)

- File locale_bug.zip added

I did more investigation and the issue seem to be in QgsVectorFileWriter.writeAsVectorFormat.

I wrote a test script

```
# use any locale with comma decimal separator
test_locale = 'it_IT.UTF-8'
orig_path = './test_data.shp'

# this file will have numbers with comma in the dbf
dest_path = './result.shp'

# end config

import sys, locale
from qgis.core import QgsVectorFileWriter, QgsVectorLayer, QgsApplication

# init QGIS
qgs = QgsApplication(sys.argv, False)
qgs.initQgis()

l = QgsVectorLayer(orig_path, 'test layer', 'ogr')
print 'feature Count %s' % l.featureCount()

old_locale = locale.getlocale(locale.LC_NUMERIC)
locale.setlocale(locale.LC_NUMERIC, test_locale)
QgsVectorFileWriter.writeAsVectorFormat(l, dest_path, 'UTF-8', l.crs(), 'ESRI Shapefile')
locale.setlocale(locale.LC_NUMERIC, old_locale)

msg = 'open %s with a text editor and you should find numbers separated by comma'
print msg % dest_path.replace('shp', 'dbf')
```

looks to me like this #20ea3e2f5a13857857f014938d33f84cd17ed785#diff-2572d1079b3ce82d94d0d5fef972b795R1727 is not working

Attached are some test data and the minimal script to reproduce the issue. You just need a locale that uses a comma as decimal separator

#14 - 2017-03-07 08:50 AM - Giovanni Manghi

- Status changed from Reopened to Feedback

- Priority changed from High to Normal

- Crashes QGIS or corrupts data changed from Yes to No

| Attached are some test data and the minimal script to reproduce the issue. You just need a locale that uses a comma as decimal separator

what versions are still affected? 2.18.4 and as well master/qgis3?

#15 - 2017-04-29 07:39 AM - Giovanni Manghi

- Target version deleted (Future Release - High Priority)
- Status changed from Feedback to Open

#16 - 2017-05-01 01:09 AM - Giovanni Manghi

- Regression? set to No
- Easy fix? set to No

#17 - 2017-11-13 02:39 PM - Even Rouault

- Description updated
- Status changed from Open to Feedback

@marco Do you reproduce with recent GDAL versions ? Normally this issue has been fixed in GDAL 2.0

#18 - 2018-02-22 12:54 PM - Giovanni Manghi

- Resolution set to up/downstream
- Status changed from Feedback to Closed

Even Rouault wrote:

| @marco Do you reproduce with recent GDAL versions ? Normally this issue has been fixed in GDAL 2.0

assuming this is fixed in GDAL, please reopen if necessary.

Files

comma_separated.zip	1.72 KB	2013-07-20	marisn -
data_with_locale.zip	148 KB	2017-02-06	Marco Bernasocchi
locale_bug.zip	151 KB	2017-02-07	Marco Bernasocchi