QGIS Application - Bug report #2496
Error trying to make distance matrix for "ID fields" with non ascii characters

2010-03-06 04:17 PM - Maxim Dubinin

Status:	Closed			
Priority:	Low			
Assignee:	cfarmer -			
Category:	Python plugins			
Affected QGIS version:		**Regression?:**	No	
Operating System:	All	**Easy fix?:**	No	
Pull Request or Patch supplied:		**Resolution:**	fixed	
Crashes QGIS or corrupts data:		**Copied to github as #:**	12556	

Description

1. Make point layer from admin.shp with Mean Coordinates tool

2. Using created point layer and poi-osm, try to use Distance Matrix tool using CATEGORY as a target field.

I'm getting this error, probably because CATEGORY contains values in cyrrilics.

See shape-files attached.

```
An error has occured while executing Python code:

Traceback (most recent call last):
  File "c:\\OSGeo4w\\apps\\qgis-dev\\python\\plugins\\fTools\\tools\\doPointDistance.py", line 105, in accept
    self.compute(point1, point2, field1, field2, outPath, matType, nearest, self.progressBar)
  File "c:\\OSGeo4w\\apps\\qgis-dev\\python\\plugins\\fTools\\tools\\doPointDistance.py", line 150, in compute
    self.linearMatrix(writer, provider1, provider2, index1, index2, nearest, distArea, matType, sindex, progressBar)
  File "c:\\OSGeo4w\\apps\\qgis-dev\\python\\plugins\\fTools\\tools\\doPointDistance.py", line 204, in linearMatrix
    if matType == "Linear": writer.writerow([unicode(inID), unicode(outID), float(dist)])
[[UnicodeEncodeError]]: 'ascii' codec can't encode characters in position 0-3: ordinal not in range(128)
```

Associated revisions

Revision da14012c - 2010-03-08 03:01 AM - cfarmer -

Updates distance matrix tool to allow non ascii characters in id fields. This is based on a workaround (from http://docs.python.org/library/csv.html), as the default Python csv function(s) do not support nonascii characters. Fixes #2496

git-svn-id: http://svn.osgeo.org/qgis/trunk/qgis@13019 c8812cc2-4d05-0410-92ff-de0c093fc19c

Revision ebc2e6f7 - 2010-03-08 03:01 AM - cfarmer -

Updates distance matrix tool to allow non ascii characters in id fields. This is based on a workaround (from http://docs.python.org/library/csv.html), as the default Python csv function(s) do not support nonascii characters. Fixes #2496

git-svn-id: http://svn.osgeo.org/qgis/trunk@13019 c8812cc2-4d05-0410-92ff-de0c093fc19c

Revision fb724f5e - 2010-05-14 12:44 AM - cfarmer -

now supports fields containing non-ascii chars. Fixes #2496. Patch from alexbruy

git-svn-id: http://svn.osgeo.org/qgis/trunk/qgis@13478 c8812cc2-4d05-0410-92ff-de0c093fc19c

Revision b2dc3549 - 2010-05-14 12:44 AM - cfarmer -

now supports fields containing non-ascii chars. Fixes #2496. Patch from alexbruy

git-svn-id: http://svn.osgeo.org/qgis/trunk@13478 c8812cc2-4d05-0410-92ff-de0c093fc19c

History

#1 - 2010-03-06 05:05 PM - Giovanni Manghi

Hi,

I'm trying to replicate here under linux but I don't see any Cyrillic char in the category column of the poi-osm layer you posted.

So... with the attached data making the matrix using just poi-osm works fine, and also making the matrix between poi-osm and the "mean points".

The real problem seems that the distance matrix GUI shows a wrong list of fields for the "input unique ID field" and/or for the "target unique ID field" when the input/target point layer is the "mean" one:

in fact it shows his "UID" field (the "real" type fields are not considered by this tool) but also the fields of the poi-osm are available. See screenshot.

If you choose the "wrong" unique ID field for the "mean" layer (belonging to the other layer) then a python error is thrown (see 2nd screenshot).

#2 - 2010-03-06 05:19 PM - Maxim Dubinin

I'm trying to replicate here under linux but I don't see any Cyrillic char in the category column of the poi-osm layer you posted.

Hmm, please check this screenshot, it does have cyrrilics
http://gis-lab.info/images/screenshots/20100306-bu0-101kb.jpg

Please note my setup for this tool also:
http://gis-lab.info/images/screenshots/20100306-3mo-37kb.jpg

I guess there are multiple issues here.

#3 - 2010-03-06 05:23 PM - Giovanni Manghi

> *Hmm, please check this screenshot, it does have cyrrilics*
> *http://gis-lab.info/images/screenshots/20100306-bu0-101kb.jpg*

Hmmm... then probably depends about the encondig I used to open the vector.

> *Please note my setup for this tool also:*
> *http://gis-lab.info/images/screenshots/20100306-3mo-37kb.jpg*

Ok, so the issues I described are confirmed. We can open a new ticket or change the title of this one.

#4 - 2010-03-07 01:24 PM - Giovanni Manghi

The same happens with non ascii chars, like "ç" or "õ". The error is the same

An error has occured while executing Python code:

Traceback (most recent call last):
  File "/usr/share/qgis/python/plugins/fTools/tools/doPointDistance.py", line 105, in accept
    self.compute(point1, point2, field1, field2, outPath, matType, nearest, self.progressBar)
  File "/usr/share/qgis/python/plugins/fTools/tools/doPointDistance.py", line 152, in compute
    self.regularMatrix(writer, provider1, provider2, index1, index2, nearest, distArea, sindex, progressBar)
  File "/usr/share/qgis/python/plugins/fTools/tools/doPointDistance.py", line 173, in regularMatrix
    writer.writerow(data)
[[UnicodeEncodeError]]: 'ascii' codec can't encode characters in position 3-4: ordinal not in range(128)

#5 - 2010-03-07 06:03 PM - cfarmer -
- Resolution set to fixed
- Status changed from Open to Closed

A fix has been made in 13019. This fix is based on a workaround (from http://docs.python.org/library/csv.html), as the default Python csv function(s) do not support non-ascii characters. As such, please test and confirm that the revision has indeed fixed the problem.

-Carson

#6 - 2010-03-08 03:00 AM - Giovanni Manghi

Replying to [comment:5 cfarmer]:

> *A fix has been made in 13019. This fix is based on a workaround (from http://docs.python.org/library/csv.html), as the default Python csv function(s)*
> *do not support non-ascii characters. As such, please test and confirm that the revision has indeed fixed the problem.*
>
> *-Carson*

the problem seem solved, thanks!

The problem regarding the wrong columns being shown is still there, but I believe that a different ticket has been filed meanwhile.

#7 - 2010-05-05 10:40 AM - Alexander Bruy
- Status changed from Closed to Feedback
- Resolution deleted (fixed)

Problem with non ASCII characters in fields values still exists. Tested under Linux with UTF-8 locale and QGIS 13404. Here is error log

Traceback (most recent call last):
  File "/usr/local/qgis-svn/share/qgis/python/plugins/fTools/tools/doPointDistance.py", line 143, in accept
    self.compute(point1, point2, field1, field2, outPath, matType, nearest, self.progressBar)
  File "/usr/local/qgis-svn/share/qgis/python/plugins/fTools/tools/doPointDistance.py", line 188, in compute
    self.linearMatrix(writer, provider1, provider2, index1, index2, nearest, distArea, matType, sindex, progressBar)
  File "/usr/local/qgis-svn/share/qgis/python/plugins/fTools/tools/doPointDistance.py", line 242, in linearMatrix
    if matType == "Linear": writer.writerow([unicode(inID), unicode(outID), float(dist)])
  File "/usr/local/qgis-svn/share/qgis/python/plugins/fTools/tools/doPointDistance.py", line 64, in writerow
    self.writer.writerow(row)
[[UnicodeEncodeError]]: 'ascii' codec can't encode characters in position 0-10: ordinal not in range(128)

Patch added

#8 - 2010-05-13 03:45 PM - cfarmer -

- Status changed from Feedback to Closed

- Resolution set to fixed

Patch applied in commit:b2dc3549 (SVN r13479).

Carson

Files

shape.zip	74.1 KB	2010-03-06	Maxim Dubinin
Screenshot.png	52 KB	2010-03-06	Giovanni Manghi
Screenshot-1.png	72.8 KB	2010-03-06	Giovanni Manghi
fix_2496.diff	1.24 KB	2010-05-05	Alexander Bruy