

QGIS Application - Bug report #14846

QGIS Union: wrong results

2016-05-20 12:30 AM - Enrico Fiore

Status:	Closed	
Priority:	Low	
Assignee:		
Category:	Processing/QGIS	
Affected QGIS version:	master	Regression?: No
Operating System:		Easy fix?: No
Pull Request or Patch supplied:		Resolution: duplicate
Crashes QGIS or corrupts data:		Copied to github as #: 22799
Description I see this bugs report #8456, but I think that the issues are different. The issues that I have noticed are: 1. Union returns records without any geometries and they are repeated more times. I suppose that they should be linked to the new geometries derived from the union output; 2. some records without geometry should have geometries, because the records have values that derive from the union of the two input tables records; 3. if you interrogate geometries, whit "identify" tool, in some cases it returns two results, but only one is linked to a geometry. You can find the relative record in the table, but if you select it in the table no geometry is selected; The same Union performed with GRASS (v.overlay in Processing tools) is correct. In attached the input file(circoscrizioni84 e quartieri84)and the fTools and GRASS results (circ84Quartieri84 / (circ84Quartieri84_grass)).		
Related issues:		
Duplicates QGIS Application - Bug report # 8456: Union tool produces wrong re...		Closed 2013-08-12
Duplicates QGIS Application - Bug report # 17131: Processing: Union and Inter...		Closed 2017-09-13

History

#1 - 2016-05-22 11:43 AM - Giovanni Manghi

- Priority changed from High to Severe/Regression
- Target version set to Version 2.16
- Operating System deleted (windows 7)

As discussed in Las Palmas this ftools must return the right results, warn the users in case there are invalid geometries, or be removed entirely.

#2 - 2016-06-06 02:35 AM - Matthias Kuhn

Is the processing version of the algorithm ok?

I think now that fTools is disabled by default we should mainly invest into making processing bulletproof and leave fTools as deprecated and "resolve" this issue with a warning in the plugin manager.

#3 - 2016-06-22 12:09 AM - Matthias Kuhn

- Status changed from Open to Feedback

#4 - 2016-07-03 12:59 AM - Giovanni Manghi

- Affected QGIS version changed from 2.14.2 to master
- Category changed from 44 to Processing/QGIS
- Status changed from Feedback to Open

Matthias Kuhn wrote:

| *Is the processing version of the algorithm ok?*

unfortunately no.

#5 - 2016-07-03 01:21 AM - Giovanni Manghi

- Subject changed from fTools Union: wrong results to QGIS Union: wrong results

#6 - 2016-07-03 11:40 AM - Matthias Kuhn

- Status changed from Open to Feedback

Apparently the output is a mix of lines and polygons which shapefile isn't able to handle.

When selecting the output format as GML a mixed geometry output is created. Does that look correct to you?

When the output isn't able to handle the geometry type (e.g. lines here), should it just dismiss it? We could implement this option in QgsVectorFileWriter (and expose it to the user or not in the algorithm configuration dialog).

Or is the mixed geometry output actually broken? Or are there better ideas?

#7 - 2016-07-04 12:45 AM - Giovanni Manghi

- Status changed from Feedback to Open

Matthias Kuhn wrote:

| *Apparently the output is a mix of lines and polygons which shapefile isn't able to handle.*

| *When selecting the output format as GML a mixed geometry output is created. Does that look correct to you?*

| *When the output isn't able to handle the geometry type (e.g. lines here), should it just dismiss it? We could implement this option in QgsVectorFileWriter (and expose it to the user or not in the algorithm configuration dialog).*

| *Or is the mixed geometry output actually broken? Or are there better ideas?*

Hi Matthias!

the input and output of the Union operation is expected to be always of polygon type

<http://desktop.arcgis.com/en/arcmap/10.3/tools/analysis-toolbox/how-union-analysis-works.htm>

this is also true for the Difference operations

[http://webhelp.esri.com/arcgisdesktop/9.2/index.cfm?TopicName=Symmetrical_Difference_\(Analysis\)](http://webhelp.esri.com/arcgisdesktop/9.2/index.cfm?TopicName=Symmetrical_Difference_(Analysis))

For the Intersection (that is kind of a Clip, except that in the output you have attributes from both input layers)

[http://webhelp.esri.com/arcgisdesktop/9.2/index.cfm?TopicName=Intersect_\(Analysis\)](http://webhelp.esri.com/arcgisdesktop/9.2/index.cfm?TopicName=Intersect_(Analysis))

#8 - 2016-07-04 12:55 AM - Giovanni Manghi

Matthias Kuhn wrote:

| *Apparently the output is a mix of lines and polygons which shapefile isn't able to handle.*

please notice that in the wrong output produced by qgis there are also duplicated polygons that are not supposed to be there (compare with grass and saga results).

#9 - 2016-07-04 05:12 AM - Matthias Kuhn

Ok, so it's just plain buggy.

#10 - 2016-07-04 08:41 AM - Giovanni Manghi

Matthias Kuhn wrote:

| *Ok, so it's just plain buggy.*

so, we ship again broken (f)tools?

#11 - 2016-07-05 10:53 AM - Matthias Kuhn

Again or still... Apparently we do.

#12 - 2016-07-06 06:12 AM - Giovanni Manghi

Matthias Kuhn wrote:

| *Again or still... Apparently we do.*

I will try to put all the (f)tools issues (regarding wrong results) in one ticket.

#13 - 2016-10-27 01:24 AM - Graeme Seggie

To aid fix, I am attaching some files (in response to Matthias Kuhn feedback on email group) showing the results from simple UNION operations against

shapefiles union'd with self and on each other.

Files attached:

Shapefile1 has three features - one of which (id=3) overlaps the other two (id=1 and id=2).

Shapefile2 has a single feature (id=10), which overlaps all of the three features in Shapefile1.

Results of the processing from Vector --> Geoprocessing Tools --> Union also attached

Shapefile1_union_Shapefile1 (15 features result)

Shapefile1_union_Shapefile2 (7 features result)

The first resulting shapefile shows nulls and features with seemingly no geometry, the latter looks more sensible.

Graeme

#14 - 2016-10-27 01:38 AM - Graeme Seggie

- File *QGIS_Union_Test_Shapefiles.zip* added

File attached 'QGIS Union Test Shapefiles.zip'

#15 - 2016-11-08 03:53 AM - Etienne Trimaille

- File *Screen_Shot_2016-11-08_at_18.43.02.png* added

- File *layer_A.geojson* added

- File *layer_B.geojson* added

I tried on 2.16 and 2.18 too. I got some overlapping polygons too.

With some transparency, we can see I got too many features.

These overlapping polygons are different if we do A union B or B union A.

I put my test data in two geojson.

On 2.14, the algorithm is not performing a union at all unfortunately. It's a kind of "difference".

#16 - 2016-11-12 11:26 PM - Etienne Trimaille

- File *layer_python_error_A.geojson* added

- File *layer_python_error_B.geojson* added

I'm working with the union algorithm again.

In my previous message, I haven't got any python error, just wrong results.

Today, I got a python exception :

```
2016-11-13T14:19:23 2 Uncaught error while executing algorithm
```

```
Traceback (most recent call last):
```

```
File "/Applications/QGIS.app/Contents/MacOS/../Resources/python/plugins/processing/core/GeoAlgorithm.py", line 203, in execute
self.processAlgorithm(progress)
```

```
File "/Applications/QGIS.app/Contents/MacOS/../Resources/python/plugins/processing/algs/qgis/Union.py", line 159, in processAlgorithm
    if diff_geom.isGeosEmpty() or not diff_geom.isGeosValid():
AttributeError: 'NoneType' object has no attribute 'isGeosEmpty'
```

I will try to have a look. Just in case, I'm going to attach again two others layers which are producing this error, whatever the order layer A/B in the parameters. Sorry, it's not small layers, I will try to reduce them and see which polygon is failing.

#17 - 2016-11-14 09:11 PM - Tim Sutton

- File *union_check_hazard.geojson* added
- File *union_check_aggregation.geojson* added

I took the sample dataset from @gustry and reduced it to a minimal dataset that can be used to replicate the issue. Please see attached files.

#18 - 2016-11-14 09:26 PM - Tim Sutton

- File *union_check_aggregation.geojson* added
- File *union_check_hazard.geojson* added

And an even simpler test dataset with one polygon each

#19 - 2016-11-15 04:01 AM - Enrico Fiore

I have others remarks concerning the union function and about the results that produces (I tested shapefile format and temporary layer file)
Now it is allowed to perform union between layer with different geometry type (point vs polygon; line vs polygon, line vs line etc..) the tool returns a result that could be not correct (I think so) because it is generated a geometry collection and not a well specified type (example: lines union returns mix of points and lines).

So at the end we have:

- *shapefile format returns a new layer with the information related to the two old layer but without combined parts. Because the format doesn't allow geometry collection;*
- *temporary file returns a new layer with the information related to the two old layer and the combined parts, but the geometries that can be visualized are only that concerning the input layer1. So if the input layer1 is a point layer then you can see only point, also if you have combined for example point vs polygon. But if you use identify tools it is possible to interrogate also the polygon.*

Isn't better allow union only between point layer or polygon layer? Or alert the user that the result can be wrong or not complete? (the same issue, if it is a issue, can be showed with intersection)

#20 - 2016-12-12 04:20 AM - Giovanni Manghi

see also:

#15962

#21 - 2017-01-08 09:17 PM - Martin Dobias

The problem with data from Etienne/Tim is that the input feature in hazard layer has invalid geometry (self-intersecting outer ring). The algorithm is correct, just the input data are not correct. Applying ST_Buffer(geom, 0) to the hazard geometry fixes the invalid geometry and union works as expected afterwards.

Currently there is a discussion among QGIS devs regarding how to best deal with invalid input geometries in processing algorithms... see <https://github.com/qgis/QGIS/pull/3865>

#22 - 2017-01-08 11:15 PM - Giovanni Manghi

Martin Dobias wrote:

Applying ST_Buffer(geom, 0) to the hazard geometry fixes the invalid geometry and union works as expected afterwards.

Currently there is a discussion among QGIS devs regarding how to best deal with invalid input geometries in processing algorithms... see <https://github.com/qgis/QGIS/pull/3865>

Using liblwgeom (ST_MakeValid) seems a very good option: it works better then 0 distance buffer (that frequently changes the input layer by removing nodes and/or entire area) and is available (also on Windows) without needing to install PostGIS.

#23 - 2017-02-08 02:45 AM - Alexander Bruy

- Status changed from Open to Closed
- Resolution set to duplicate

I close this, as there is another ticket (#8456) about union issues.

#24 - 2017-09-14 10:52 PM - Giovanni Manghi

- Priority changed from Severe/Regression to Low
- Description updated

new reference ticket for this matters is #17131

#25 - 2017-09-28 09:34 AM - Jürgen Fischer

- Duplicates Bug report #17131: Processing: Union and Intersection wrong results added

Files

unioni.zip	751 KB	2016-05-19	Enrico Fiore
QGIS_Union_Test_Shapefiles.zip	6.41 KB	2016-10-26	Graeme Seggie
layer_A.geojson	30 KB	2016-11-08	Etienne Trimaille
layer_B.geojson	55.4 KB	2016-11-08	Etienne Trimaille
Screen_Shot_2016-11-08_at_18.43.02.png	180 KB	2016-11-08	Etienne Trimaille
layer_python_error_A.geojson	568 KB	2016-11-12	Etienne Trimaille
layer_python_error_B.geojson	2.23 MB	2016-11-12	Etienne Trimaille
union_check_aggregation.geojson	53.4 KB	2016-11-14	Tim Sutton
union_check_hazard.geojson	41.4 KB	2016-11-14	Tim Sutton
union_check_aggregation.geojson	24.3 KB	2016-11-14	Tim Sutton
union_check_hazard.geojson	26.3 KB	2016-11-14	Tim Sutton