QGIS Application - Bug report #8456
Union tool produces wrong result

2013-08-12 04:23 AM - Alexander Bruy

Status:	Closed		
Priority:	High		
Assignee:			
Category:	Processing/QGIS		
Affected QGIS version:	2.18.4	**Regression?:**	Yes
Operating System:		**Easy fix?:**	No
Pull Request or Patch supplied:	No	**Resolution:**	duplicate
Crashes QGIS or corrupts data:	No	**Copied to github as #:**	17224

Description

Union tool from fTools (and also SEXTANTE) produces wrong result. If we look at an accepted definition of "union" (http://resources.arcgis.com/en/help/main/10.1/index.html#//000800000010000000) and we will see that where features are overlapping, then in the union it is expected to have duplicated feature. In the attached result "union18", produced by QGIS 1.8, there areas where it is supposed to have overlapped/repeated features originating from overlapped parts in the input layers.

There is also some disscussion in mailing list about this
http://osgeo-org.1560.x6.nabble.com/Union-Tool-Ftools-Sextante-How-should-the-real-result-be-tp5051653.html

Related issues:

Related to QGIS Application - Bug report # 7708: Union produces the wrong out...	**Closed**	**2013-04-25**
Related to QGIS Application - Bug report # 15962: Union & Intersection result...	**Closed**	**2016-12-08**
Duplicated by QGIS Application - Bug report # 14846: QGIS Union: wrong results	**Closed**	**2016-05-20**

History

#1 - 2013-08-12 04:30 AM - Giovanni Manghi

it is also necessary to review the resulting attributes -> it would necessary to give a look to all the discussions in particular #4567

If this is not going to be tagged as blocker (and #7708 closed because this ticket is replacing it) then I would say that the union tool has to be removed from qgis because it misleads the users with wrong results, this is very bad.

But honestly missing a union tool in a major release it would be bad too (I'm not asking anything, just making a point).

cheers!

#2 - 2013-08-12 05:04 AM - Daniel Vaz

I will work on this bug. (just for avoid double work)

I will take into account your considerations Giovanni.

Thanks in advance

#3 - 2013-08-12 05:07 AM - Giovanni Manghi

Daniel Vaz wrote:

> *I will work on this bug. (just for avoid double work)*
>
> *I will take into account your considerations Giovanni.*
>
> *Thanks in advance*

Thanks Daniel! very good news indeed!

#4 - 2013-08-14 06:04 PM - Daniel Vaz

- File 5.png added

Hi Giovanni, follow my first results in union features on same layer.

These results show 10 features instead of 11 expected.

I can't find yet a elegant solution to behaviour described in arcgis help, but I will keep trying.

Thanks

#5 - 2013-08-15 05:50 PM - Daniel Vaz

- File two_layers.tar.gz added

- File same_layer.tar.gz added

Please, I need some feedback on first results.

Thanks in advance.

#6 - 2013-08-15 06:24 PM - Giovanni Manghi

- Operating System deleted (all)

Hi Daniel, thanks for the effort.

Are you writing code from scratch or trying to use something already available? For example I see that shapely does this kind of operations, but I'm not sure it is a qgis dependency.

http://toblerity.org/shapely/manual.html

About the results you attached:

one of the issues I'm seeing is that in the expected results each inner boundary should delimit a separate feature. This way the result would be as the "output" figure here:

http://resources.arcgis.com/en/help/main/10.1/0008/GUID-F81C8E85-1DA1-4857-B872-6CE30DDE3984-web.png

- where the two circles and the rectangle overlap there are 3 features
- where the circles (1) overlap with the rectangle there are 2 features

- in the other places there is 1 feature

#7 - 2013-08-15 07:07 PM - Daniel Vaz

Please Giovanni, download the .tar.gz files and give it a try.

Turn edit mode on and move the features of results layer.

Thanks

#8 - 2013-08-16 02:36 AM - Giovanni Manghi
- File union.jpg added

Daniel Vaz wrote:

> *Please Giovanni, download the .tar.gz files and give it a try.*
>
> *Turn edit mode on and move the features of results layer.*
>
> *Thanks*

Hi Daniel, look at the (very bad) image I attached:

- the green must be a geometry, repeated/overlapped 3 times
- the red and blue must geometries, each repeated/overlapped 2 times
- the yellow are geometries not repeated

cheers!

#9 - 2013-08-16 05:33 AM - Daniel Vaz

Ok, I already have been understood it.

Is it the expected behaviour if the three geometries are inside the same layer, right?

What is expected if the two circles resides in one layer and the rectangle resides in another layer?

My attached results almost reach it, but the green don't repeat three times, but only two. I didn't find an elegant solution to get the green part be repeated three times.

But the current solution is worse than mine, although both are "wrong".

Thanks in advance

#10 - 2013-08-16 06:19 AM - Giovanni Manghi

Hi Daniel,

> *Is it the expected behavior if the three geometries are inside the same layer, right?*

I don't know if I understand correctly, but as far as I know, the union operation is supposed to have 2 (or more***) different inputs. I guess that eventually the same layer can be used two times, but I don't think it makes much sense.

> *What is expected if the two circles resides in one layer and the rectangle resides in another layer?*

This is the common situation, the union tool needs two inputs. In the example the circles would be in one layer and the rectangle in another.

The result of a union operation is just one layer, so in the example you are using all the resulting 11 features are in the same layer, a few of them repeated/overlapped.

(***)In QGIS union and intersect just allow to select 2 input layers. There is already a feature request to allow choose >2 input layers (as for example Arcgis does). I believe that in this case the operations are made by unioning/intesecting the layers 2 by 2 and then the result unioned/intersected with the following one.

> *But the current solution is worse than mine, although both are "wrong".*

I really appreciate your effort. I'm still convinced that we cannot ship a tool that we know produce a wrong result. I would be in favor to eventually remove (temporarily) the tool from qgis 2 and reinstate it later when a solution is found. But this obviously must be discussed in the dev mailing list.

#11 - 2013-08-16 06:04 PM - Daniel Vaz

Hi Giovanni,

I was thinking, could somebody download the two .tar.gz attached files and using arcgis do the union process and attach the results here?

The shapefiles are:
File: arcgis_dataset.shp (all features in same layer)
File: rectangle_only.shp and two_circles.shp

Maybe we need to discuss the requisite of a union tool instead of try to copy a arcgis tool.
I am not convinced that the requisite of the QGis union tool was completely clear. If you see the dev-list discussion you will face with opposite opinions.

See for example #4567 (a related bug):

> *You can see this easily applying transparency to the resultant layer. In that way you can detect overlapping polygons in a single layer. That must not exist as a result of a union function.*

The user expects that a union tool just union the overlapped area not create a new polygon.

Again, I think that we need to discuss the requisite of a union tool, not get a arcgis feature and try to reproduce.

Thanks in advance

#12 - 2013-08-18 04:44 AM - Giovanni Manghi
- File union_tests.zip added

Hi Daniel,

you may be right. If fact if you look for general union/intersection/etc definitions you will see that there isn't a standard one.

This is easily confirmed by testing other software. I have no access to Arcgis (but I'm almost sure that the results are the described one) but I tested gvSIG (that has an internal union tool) and Sextante (using Kosmo, as QGIS/Sextante does not have yet this tool).

I attach here the results, and as you will see are different (also including qgis) both on geometry side and attributes one.

So I agree that this would be good to be discussed... or not :)

To be in time for qgis 2.0 we can leave the qgis tool as it is now **but** the truth is that I'm personally that if compared with most of the union definitions the result provided by arcgis is the correct one.

I will try to provide you it as soon as possible.

cheers!

#13 - 2013-10-12 01:09 AM - Raymond Nijssen

- File arcgis_results.zip added

Hi Daniel and Giovanni,

Good to see you are working on this and taking the analyses results so seriously!

As you requested I tested this in arcgis. The 3 resulting shapefiles are in the zip:

arcgis_dataset_Union.shp
  Union of the shapefile with all 3 features.

two_circles_Union.shp
  Union of the 2 separate shapefiles, the one with 2 circles listed first.

two_circles_Union1.shp
  Union of just the shapefile with 2 circles, because I was curious...

It seems that unions within a layer are done first, and later the unions between layers. So arcgis will only create 2 polygons for the most central polygon, that should be 3 as described in the example.

Also notice that the upper and lower part of the rectangle results in a multipolygon.

For me, the name "union" does not make much sense, cause the result feels more like "split" to me. The OGC union operator would turns two overlapping polygons into one, not 3 or 4 (or even more). Maybe the arcgis way is not the best way for our users to do this.

In what cases do users need this functionality? And what would be the best or most logical result for them?

Regards,
Raymond

#14 - 2013-10-12 03:43 AM - Jukka Rahkonen

Hi,

What is "union" for rest of the world is "dissolve" for ESRI. This speed comparison deals loosely with that
http://lin-ear-th-inking.blogspot.fi/2013/02/jts-union-vs-arcgis-dissolve.html

I believe that union in QGIS should yield the same results as union in PostGIS http://postgis.refractions.net/documentation/manual-svn/ST_Union.html. Rather than using the ESRI terminology we could have a ESRI-FOSS4G lexicon about geographic operations with sample images.

I use union ofter very often for converting a bunch of polygons into a smallest possible number of polygons which cover the same total area.

#15 - 2013-10-12 12:23 PM - Giovanni Manghi

Hi Raymond and jratike80,

thanks for you comments.

For me in good whatever direction we think is more appropriate. I personally suggested to try do as esri does just because it is what most people would expect. But using another approach is good to. The main issue here is having a tool that works as advertised.

cheers!

#16 - 2016-12-12 05:35 AM - Giovanni Manghi
- Priority changed from Normal to Severe/Regression
- Category changed from 44 to Processing/QGIS

#17 - 2017-02-28 12:01 AM - Giovanni Manghi
- Affected QGIS version changed from master to 2.18.4
- Target version changed from Future Release - High Priority to Version 2.18

#18 - 2017-04-30 05:08 PM - Giovanni Manghi
- Regression? set to Yes

#19 - 2017-04-30 05:08 PM - Giovanni Manghi
- Priority changed from Severe/Regression to High

#20 - 2017-05-01 01:10 AM - Giovanni Manghi
- Easy fix? set to No

#21 - 2017-05-12 08:11 PM - Giovanni Manghi
- Status changed from Open to Closed
- Description updated
- Resolution set to duplicate

Merged with #15962

Files

union.zip	231 KB	2013-08-12	Alexander Bruy

5.png	65.6 KB	2013-08-14	Daniel Vaz
same_layer.tar.gz	3.5 KB	2013-08-15	Daniel Vaz
two_layers.tar.gz	3.71 KB	2013-08-15	Daniel Vaz
union.jpg	29.5 KB	2013-08-16	Giovanni Manghi
union_tests.zip	3.52 KB	2013-08-18	Giovanni Manghi
arcgis_results.zip	16.3 KB	2013-10-11	Raymond Nijssen