QGIS Application - Bug report #7514
big regression: column in the provided shapefile fails to load on QGIS master, shifts data to following columns

2013-04-05 12:33 AM - Mathieu Pellerin - nIRV

Status:	Closed		
Priority:	Severe/Regression		
Assignee:			
Category:	Data Provider		
Affected QGIS version:	master	**Regression?:**	No
Operating System:		**Easy fix?:**	No
Pull Request or Patch supplied:	No	**Resolution:**	upstream
Crashes QGIS or corrupts data:	No	**Copied to github as #:**	16463

Description

Spotted a pretty bad regression: one column this [shapefile](#) (60mb) fails to load on QGIS master, and result in data shifts to following columns. The data in the last column is lost, and any column-based style saved through .qml is effectively broken.

The shapefile .dbf has these columns:
PA_NAME_EN,C,50
PA_TYPE,C,5
PROTECTION,C,40
THREAT_STA,C,30
PROTECTION1,C,40
ELEVATION_,C,10
VEG_RECLAS,C,50
VEG_GAP_FI,C,70
AREA,N,16,3
PERIMETER,N,16,3
HECTARES,N,16,3
VEG_SIMPLE,C,60

When loaded under QGIS master, the PROTECTION1 column doesn't appear, and its data is shifted to the next column (ELEVATION_C), and the data of the next column is shifted to the next, etc. etc. The data of the last column VEG_SIMPLE is simply gone.

History

#1 - 2013-04-05 12:49 AM - Mathieu Pellerin - nIRV

Oups, fixed link http://www.licadho-cambodia.org/qgis/big_reg.zip

#2 - 2013-04-05 12:57 AM - Larry Shaffer

I can confirm on Mac 10.7.5 with latest master, in QBrowser.

However, running ogrinfo -so 'gap_landcover.shp' 'gap_landcover' returns:

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

    Layer name: gap_landcover

Geometry: 3D Polygon

Feature Count: 86668

Extent: (211823.109375, 1144262.940806) - (785034.000000, 1625281.000000)

Layer SRS WKT:

(unknown)

PA_NAME_EN: String (50.0)

PA_TYPE: String (5.0)

PROTECTION: String (40.0)

THREAT_STA: String (30.0)

PROTECTION: String (40.0)

ELEVATION_: String (10.0)

VEG_RECLAS: String (50.0)

VEG_GAP_FI: String (70.0)

AREA: Real (16.3)

PERIMETER: Real (16.3)

HECTARES: Real (16.3)

VEG_SIMPLE: String (60.0)

PROTECTION1 doesn't show up. OGR sees the second field as having the same name, i.e. it seems to only read the first 10 characters of the field name. Looks like it is a [set function of the shapefile driver - see Creation Issues](). I can only assume the same holds true for reading.

#3 - 2013-04-05 01:06 AM - Mathieu Pellerin - nIRV

Larry, the strange thing is, with osgeo4w's QGIS 1.8, which is built against GDAL 1.9.2 just like QGIS master, it works fine.

#4 - 2013-04-12 11:37 PM - maning sambale

I confirm this issue in the latest nightly build.

#5 - 2013-04-12 11:55 PM - Even Rouault

If you look at the header of gap_landcover.dbf you'll see that there are 2 fields called PROTECTION. There can **not** be a PROTECTION1 field, since it would be 11 character wide, and DBF field names are limited to 10 characters. There's no way OGR can read PROTECTION1, since it is not in the DBF file... I'd say that the DBF is prone to cause great confusion to shapefile readers.

I cannot confirm nirvn comment : I have opened it with QGIS 1.8.0 against GDAL 1.9.2 on Linux, and it reports 2 PROTECTION fields, not PROTECTION1. And the attributes for each of the PROTECTION field show up correctly in the attribute table. But this can cause various issues when a field is retrieved by its name, and not its index...

#6 - 2013-04-13 02:03 AM - Even Rouault

Note: After the change I've just done in http://trac.osgeo.org/gdal/changeset/25903, the ogr2ogr of GDAL 1.10 will be able to convert the above bogus shapefile in a shapefile with unique field names.

Before :
$ ogrinfo gap_landcover.shp -al -so
INFO: Open of `gap_landcover.shp'
    using driver `ESRI Shapefile' successful.

Layer name: gap_landcover
Geometry: Polygon
Feature Count: 86668
Extent: (211823.109375, 1144262.940806) - (785034.000000, 1625281.000000)
Layer SRS WKT:
(unknown)
PA_NAME_EN: String (50.0)
PA_TYPE: String (5.0)
PROTECTION: String (40.0)
THREAT_STA: String (30.0)
PROTECTION: String (40.0)
ELEVATION_: String (10.0)
VEG_RECLAS: String (50.0)
VEG_GAP_FI: String (70.0)
AREA: Real (16.3)
PERIMETER: Real (16.3)
HECTARES: Real (16.3)
VEG_SIMPLE: String (60.0)

Conversion:
$ ogr2ogr gap_landcover_unique.shp gap_landcover.shp
Warning 6: Normalized/laundered field name: 'PROTECTION2' to 'PROTECTI_1'
Warning 1: Field 'PROTECTION' already exists. Renaming it as 'PROTECTI_1'

After:
$ ogrinfo gap_landcover_unique.shp -al -so
INFO: Open of `gap_landcover_unique.shp'
    using driver `ESRI Shapefile' successful.

Layer name: gap_landcover_unique
Geometry: Polygon
Feature Count: 86668
Extent: (211823.109375, 1144262.940806) - (785034.000000, 1625281.000000)
Layer SRS WKT:
(unknown)
PA_NAME_EN: String (50.0)
PA_TYPE: String (5.0)
PROTECTION: String (40.0)
THREAT_STA: String (30.0)
PROTECTI_1: String (40.0)
ELEVATION_: String (10.0)
VEG_RECLAS: String (50.0)
VEG_GAP_FI: String (70.0)
AREA: Real (16.3)
PERIMETER: Real (16.3)
HECTARES: Real (16.3)
VEG_SIMPLE: String (60.0)

#7 - 2013-04-17 06:59 PM - Mathieu Pellerin - nIRV

- Resolution set to upstream

- Status changed from Open to Closed

Since this seems to have to be handled through gdal, closing with upstream resolution.