# pyDelphin and gTest Updates

Mi(ke|chael) Wayne Goodman

2016-06-17

# pyDelphin

## Releases

2 major (and 1 minor) releases since last year:

- [v0.4.0](#) (2016 Mar 03)
  - Better TDL list support
  - LaTeX export of DMRS
  - Semantic versioning and release on PyPI
- [v0.4.1](#) (2016 Apr 20)
  - Bugfixes, improved testing, error checking on MRS parsing
- [v0.5.0](#) (2016 Jun 18?)
  - Basic EDS support
  - `to_dict()` and `from_dict()` on Mrs, Dmrs, EDS
  - "Proper" modeling of UDF (derivation) terminals
  - REST client
  - More thorough and up-to-date documentation

# Project Changes

- Project moved from https://github.com/goodmami to https://github.com/delph-in

- Project released on PyPI

  ```
  $ pip install pydelphin
  ```

- More extensive unit tests (from ~50% (v0.3) to ~60% (v0.5.0) code coverage)

- Dropped Python 2.6 support

# UDF derivation tree support

```
>>> from delphin import derivation
>>> d = derivation.Derivation.from_string(
...     '(root'
...     ' (1 rule -1 -1 -1'
...     '  (2 word -1 -1 -1'
...     '   ("word" 3'
...     '    "token [ +FORM \\"word\\" ]"))))'
... )

>>> d.entity
'root'

>>> d.daughters
[<UdfNode object (rule) at 140540560397368>]

>>> d.daughters[0].daughters[0].daughters[0]
<UdfTerminal object (word) at 140540560319336>

>>> d.daughters[0].daughters[0].daughters[0].tokens[0]
<UdfToken object (3 'token [ +FORM \\"word\\" ]') at 140540560400456>

>>> d.daughters[0].daughters[0].daughters[0].tokens[0].tfs
'token [ +FORM \\"word\\" ]'
```

# Better support for TDL, particularly lists
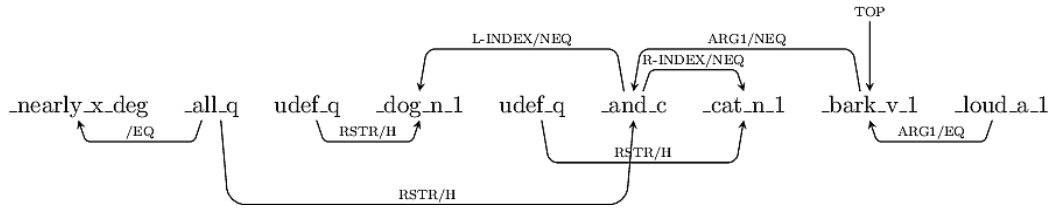
```
>>> from io import StringIO
>>> from delphin import tdl

>>> tdlfile = StringIO(
...     'type := supertype & [ ATTR < "string1", "string2" > ].'
... )
>>> t = next(tdl.parse(tdlfile))

>>> t['ATTR'].values()
['"string1"', '"string2"']

>>> t['ATTR.FIRST']
'"string1"'
>>> t['ATTR.REST.FIRST']
'"string2"'
```

# LaTeX Export for DMRS



```latex
\begin{dependency}[text only label,label style={above,font=\footnotesize},edge unit distance=2ex]
  \begin{deptext}[column sep=10pt]
    \_nearly\_x\_deg \& \_all\_q \& udef\_q \& \_dog\_n\_1 \& udef\_q \&
    \_and\_c \& \_cat\_n\_1 \& \_bark\_v\_1 \& \_loud\_a\_1 \\
  \end{deptext}
  \deproot[edge unit distance=2ex]{8}{TOP}
  \depedge[edge below]{2}{1}{/EQ}
  \depedge[edge below]{2}{6}{RSTR/H}
  \depedge[edge below]{3}{4}{RSTR/H}
  \depedge[edge below]{5}{7}{RSTR/H}
  \depedge[edge above]{6}{4}{L-INDEX/NEQ}
  \depedge[edge above]{6}{7}{R-INDEX/NEQ}
  \depedge[edge above]{8}{6}{ARG1/NEQ}
  \depedge[edge below]{9}{8}{ARG1/EQ}
\end{dependency}
```

# EDS output

```
>>> from delphin.mrs import eds, simplemrs
>>> m = simplemrs.load_one('nearly-all-dogs-bark-loudly.mrs')
>>> print(eds.dumps_one(m))
{e2: (fragmented)
|e5:_nearly_x_deg<0:6>[]
 _1:_all_q<7:10>[BV x3]
 _2:udef_q<11:15>[BV x10]
 x10:_dog_n_1<11:15>[]
 _3:udef_q<16:24>[BV x15]
 x3:_and_c<16:19>[L-INDEX x10, R-INDEX x15]
 x15:_cat_n_1<20:24>[]
 e2:_bark_v_1<25:29>[ARG1 x3]
 e20:_loud_a_1<30:37>[ARG1 e2]
}
```

# REST client

```
>>> from delphin.interfaces import rest

>>> response = rest.parse('All dogs chase some cat.',
...     params={'mrs':'json'})
>>> response.result(0)['mrs']
{'variables': {'h13': {'type': 'h'}, 'h6': {'type': 'h'}, 'h1': ...
>>> response.result(0).mrs()
<Xmrs object (all dog chase some cat) at 139620954273216>

>>> response = rest.parse('All dogs chase some cat.',
...     params={'derivation': 'json'})
>>> response.result(0).derivation()
<UdfNode object (sb-hd_mc_c) at 139620919391736>
```

# Documentation

See https://github.com/delph-in/pydelphin/wiki/API-Reference

# gTest

- Recall: gTest is a commandline tool for running basic tests on a grammar

- Already had: regression tests, coverage tests

- New semantic tests

  - Well-formedness

    - ARG0 on all EPs
    - No shared intrinsic variables
    - EPs have no more than one quantifier
    - All EPs have a label
    - Connectivity
    - Lo handle of HCONS is the label of some EP(s)

  - Connectedness

  - Semantic Headedness

# gTest Semantic Tests

```
$ ./gTest \
    -G ~/logon/lingo/erg \
    M \
    --skel-dir ~/logon/lingo/lkb/src/tsdb/skeletons/english/ \
    :mrs
:mrs (107 items; 0 ignored):
  results     :   445/107   (4.1589 per item)
  No MRS      :     0/445   ( 0.00% of results)
  Bad MRS     :     0/445   ( 0.00%)
  Ill-formed  :    11/445   ( 2.47%)
  Disconnected:     0/445   ( 0.00%)
  Non-headed  :    20/445   ( 4.49%)
```