Elan Python course
Session 1

Mart Lubbers: mart.lubbers@mpi.nl
Fransisco Torreira: francisco.torreira@mpi.nl

2014-10-02

Table of contents

Introduction importing

Introduction to the Classes

Basic functions

- Create or load an object
- Create tiers
- Adding and removing annotations
- Query information
- Writing to files

Advanced functions

- Converting and extracting
- Gaps and overlaps
- Filter, glue, shift and merge
- Excercise

How to import modules

How to import modules

```
1 from pympi import Eaf, TextGrid
2 import glob
3 import os
```

Use imported modules

Use imported modules

Indirect importing

```
1 import glob
2
3 files = glob.glob('./*.eaf')
```

Use imported modules

Indirect importing

```
1 import glob
2
3 files = glob.glob('./*.eaf')
```

Direct importing

```
1 from glob import glob
2
3 files = glob('./*.eaf')
```

Use imported modules

Indirect importing

```
1 import glob
2
3 files = glob.glob('./*.eaf')
```

Direct importing

```
1 from glob import glob
2
3 files = glob('./*.eaf')
```

Import the pympi module

```
1 from pympi import Eaf, TextGrid
```

Table of contents

Introduction importing

Introduction to the Classes

Basic functions

- Create or load an object
- Create tiers
- Adding and removing annotations
- Query information
- Writing to files

Advanced functions

- Converting and extracting
- Gaps and overlaps
- Filter, glue, shift and merge
- Excercise

Eaf, word with elan files

Eaf, word with elan files

Eaf: Documentation

`http://dopefishh.github.io/pympi/Elan.html`

Eaf, word with elan files

Eaf: Documentation

`http://dopefishh.github.io/pympi/Elan.html`

TextGrid: Documentation

`http://dopefishh.github.io/pympi/TextGrid.html`

Table of contents

Introduction importing

Introduction to the Classes

Basic functions

- Create or load an object
- Create tiers
- Adding and removing annotations
- Query information
- Writing to files

Advanced functions

- Converting and extracting
- Gaps and overlaps
- Filter, glue, shift and merge
- Excercise

Eaf object

Eaf object

Create empty object

```
1 from pympi import Eaf
2 a = Eaf()
```

Eaf object

Create empty object

```
1 from pympi import Eaf
2 a = Eaf()
```

Load from file

```
1 from pympi import Eaf
2 a = Eaf('/path/to/file/file.eaf')
3
4 # Full function
5 __init__(file_path=None, author='pympi')
```

TextGrid object

TextGrid object

Create empty object

```
1 from pympi import TextGrid
2 a = TextGrid(xmax=60)
```

TextGrid object

Create empty object

```
1 from pympi import TextGrid
2 a = TextGrid(xmax=60)
```

Load from file

```
1 from pympi import TextGrid
2 a = TextGrid('/path/to/file/file.TextGrid')
3
4 # Full function
5 __init__(file_path=None, xmin=0, xmax=None, codec='ascii', stream=False)
```

Create tiers

Create tiers

Eaf

```
1 | from pympi import Eaf
2 | a = Eaf()
3 | a.add_tier('tiername')
4 |
5 | # Full function
6 | a.add_tier(tier_id, ling='default-lt', parent=None, locale=None, part=None,
      ann=None, tier_dict=None)
```

Create tiers

Eaf

```
1 | from pympi import Eaf
2 | a = Eaf()
3 | a.add_tier('tiername')
4 |
5 | # Full function
6 | a.add_tier(tier_id, ling='default-lt', parent=None, locale=None, part=None,
      ann=None, tier_dict=None)
```

TextGrid

```
1 | from pympi import TextGrid
2 | a = TextGrid()
3 | tier1 = a.add_tier('tiername')
4 | tier2 = a.add_tier('tiername2', tier_type='TextTier')
5 |
6 | # Full function
7 | add_tier(name, tier_type='IntervalTier', number=None)
```

Adding and removing annotations

Adding and removing annotations

Eaf

```
1  from pympi import Eaf
2  a = Eaf()
3  a.add_tier('tier1')
4  a.add_tier('tier2')
5  a.insert_annotation('tier1', 5000, 10000, 'my first annotation')
6  a.insert_annotation('tier2', 10001, 20000) # This annotation will be empty
7
8  # Full function
9  insert_annotation(id_tier, start, end, value='', svg_ref=None)
10
11 a.remove_annotation('tier1', 7500)
12 a.remove_all_annotations_from_tier('tier2')
13
14 # Full functions
15 remove_annotation(id_tier, time, clean=True)
16 remove_all_annotations_from_tier(id_tier)
```

Adding and removing annotations, continued

Adding and removing annotations, continued

TextGrid

```
1  from pympi import TextGrid
2  a = TextGrid()
3  tier1 = a.add_tier('words')
4  a.add_tier('phones')
5
6  tier1.add_interval(5, 10, 'my first annotation')
7  tier2 = a.get_tier('phones') # Is equivalent to: tier2 = a.get_tier(2)
8  tier2.add_interval(10, 20) # This annotation will be empty
9
10 tier1.remove_interval(7.5)
11 tier2.clear_intervals()
12
13 # Full functions
14 remove_interval(time)
15 clear_intervals()
```

Querying information about annotations

Querying information about annotations

Eaf

```
1  a = Eaf()
2  # Add stuff or load a file
3  ann50ms = a.get_annotation_data_at_time(50)
4  start = ann50ms[0]
5  end = ann50ms[1]
6  value = ann50ms[2]
7  all_anns = a.get_annotation_data_for_tier('words')
8
9  # Full functions
10 get_annotation_data_at_time(id_tier, time)
11 get_annotation_data_for_tier(id_tier)
12 get_annotation_datas_between_times(id_tier, start, end)
```

Writing to file

Writing to file

Eaf

```
1 from pympi import Eaf
2 a = Eaf()
3 # Do stuff, add tiers, edit etc...
4 a.to_file('/some/file/path/file.eaf')
5
6 # Full function
7 to_file(file_path, pretty=True)
```

Writing to file

Eaf

```
1 | from pympi import Eaf
2 | a = Eaf()
3 | # Do stuff, add tiers, edit etc...
4 | a.to_file('/some/file/path/file.eaf')
5 |
6 | # Full function
7 | to_file(file_path, pretty=True)
```

TextGrid

```
1 | from pympi import TextGrid
2 | a = TextGrid()
3 | # Do stuff, add tiers, edit etc...
4 | a.to_file('/some/file/path/file.eaf')
5 |
6 | # Full function
7 | to_file(filepath, codec='utf-8')
```

Table of contents

Introduction importing

Introduction to the Classes

Basic functions

- Create or load an object
- Create tiers
- Adding and removing annotations
- Query information
- Writing to files

Advanced functions

- Converting and extracting
- Gaps and overlaps
- Filter, glue, shift and merge
- Excercise

Converting and extracting

Converting and extracting

```
1  from pympi import Eaf, TextGrid
2  a = Eaf()
3  a.add_tier('Phones')
4  a.add_tier('Words')
5  a.add_tier('Utterances')
6  # Do stuff with it
7  a_subset = a.extract(5000, 50000)
8  a_tg = a.to_textgrid(excluded_tiers=['Phones', 'Words'])
9
10 # Full functions
11 extract(start, end)
12 to_textgrid(excluded_tiers=[], included_tiers=[])
```

Calculating gaps and overlaps

Calculating gaps and overlaps

Some examples

```
1  from pympi import Eaf, TextGrid
2  a = Eaf()
3  a.add_tier('A_utterances')
4  a.add_tier('B_utterances')
5  # Do stuff with it
6  a.create_gaps_and_overlaps_tier('B_utterances', 'A_utterances', 'A_B_FTOs',
       5000)
7  for ann_data in a.get_annotation_data_for_tier('A_B_FTOs'):
8      begin = ann_data[0]
9      end = ann_data[1]
10     value = ann_data[2]
11     # Do something with the overlaps and gaps, for example process them into
12     # a data file
13
14 # Full functions
15 get_annotation_data_for_tier(id_tier)
16 create_gaps_and_overlaps_tier(tier1, tier2, tier_name=None, maxlen=-1)
```

Filter, glue, shift and merge

Filter, glue, shift and merge

Some examples

```
1  from pympi import Eaf, TextGrid
2  a = Eaf()
3  a.add_tier('A_utterances')
4  a.add_tier('B_utterances')
5  a.filter_annotations('A_utterances', 'A_utterances_filtered', filtex=['uh', 'uhu'])
6  a.glue_annotatations_in_tier('Phones', 'PhonWords', 85)
7  a_corrected = a.shift_annotations(-10) # this will shift 10ms to the right
8  merge_tiers(['A_utterances', 'B_utterances'], 'A_and_B_utterances')
9
10 # Full functions
11 filterAnnotations(tier, tier_name=None, filtin=None, filtex=None)
12 glue_annotations_in_tier(tier, tier_name=None, treshhold=85, filtin=None,
       filtex=None)
13 shift_annotations(time)
14 merge_tiers(tiers, tiernew=None, gaptresh=1)
```

Tips

Tips

- Use glob
- Check documentation, check documentation, check documentation...
- First write an outline(use the comments)