ScanLat

A Grammatical Scansion Tool for Classical Latin Poetry

Matthew Penza '19 & Christopher Ferri '18

TRA 301 / COS 401: Introduction to Machine Translation

Presented Friday 11 May 2018

Latin Poetic Meter (an over-simplification)

– – –

- Meter determined by syllable quantities, not stress accents
  - Short syllable = 1 mora
  - Long syllable = 2 morae

- Verses broken into feet (1 foot = 3 or 4 morae)
  - Iamb: 1 short + 1 long syllable (cărō)
  - Trochee: 1 long + 1 short (mēnsă)
  - Tribrach: 3 shorts (tĕmĕrĕ)

  - Dactyl: 1 long + 2 shorts (līt̆ŏră)
  - Anapaest: 2 shorts + 1 long (păt̆ŭlaē)
  - Spondee: 2 longs (fātō)

Scansion

A markup of the meter. Excerpt from Virgil's *Aeneid*:

```
-  u  u| -   u  u| -||  -| -    -|  - u  u |- -
Ārmă vĭ-rŭmquĕ că-nō, Trō-iae quī prīmŭs ăb ōrīs
- u u|-   -| - ||  u u| -   -| -  u  u| - -
Ītălĭ-ăm fā-tō   prŏfŭ-gŭs Lā-vīniăquĕ vēnĭt
 - u u | -       - |     -   - | - || - | - u  u |-  -
lītŏră, mŭlt(um) ĭl-l(e) ĕt tĕr-rīs  iăc-tātŭs ĕt ăltō
 -  u u| - ||  - | -   u u| -   -| - u  u |- -
vī sŭpĕ-rŭm,  sae-vae mĕmŏ-rĕm Iū-nōnĭs ŏb īrăm;
```

Existing solutions

- Manual scansion analysis:
  - The old standby
  - Tedious and time-consuming
  - Prone to human error
  - Easily able to accommodate known irregularities

- Classical Languages Toolkit's (CLTK) implementation:
  - Probabilistic approach to syllable quantities via ngrams
  - Already exists (!)
  - Does not leverage grammar → potentially inaccurate quantities

Our Solution: ScanLat

- Pros:
  - Determines vowel quantities grammatically
  - Avoids probabilistic errors

- Cons:
  - Dependent on performance of other tools
  - Some irregularities accounted for, but not all (yet)

Methods and Tools

- UDPipe: dependency parser
  - In TRA 301 we've worked mostly with constituency parsing
  - Trained on Universal Dependencies 2.0 treebanks (PROIEL)
  - Ideal for morphologically rich, free-order languages like Latin
- LatMor: morphological analyzer
  - Determines vowels long *by nature* from UDPipe's grammatical info
- Custom Python modules
  - Bridge gap between UDPipe and LatMor
  - Word syllabification: determine syllables long *by position*

Partial Demo

Questions?