LIN570: HW5 – morphological paradigms (100pts)

YOUR NAME (`UW NetID`)

Due date: 11pm on Nov 5, 2018 (Tuesday)

Since 2016, a shared task on morphological reinflection has been organized and extended to universal morphological reinflection (Cotterell et al., 2016, 2017).[1] The task is to generate a target inflected form based on given a lemma with morphological features. For this homework, you are going to implement the task 1 of CoNLL–SIGMORPHON 2018 Shared Task: Universal Morphological Reinflection. All the required files are under `~/dropbox/19-20/570/hw5/task1`.

Rubric:

2pts `hw.tar.gz` submitted, it should contain following files:

- `baseline_improved.sh`

2pts `readme.txt` or `readme.pdf` submitted

6pts All files and folders are present in expected locations

10pts Programs run to completion

5pts The output of programs on patas match submitted output

(75pts) Task 1: Type-level inflection

- Source form and target features:
  - `release V;V.PTCP;PRS`
- Target form:
  - `releasing`

Example of `train`:

```
reimburse       reimburses      V;3;SG;PRS
transliterate   transliterate   V;NFIN
tend    tended  V;PST
disallow disallowing V;V.PTCP;PRS
reconfirm       reconfirmed     V;V.PTCP;PST
```

The type of labels and their frequencies in `train`:

[1] `https://sigmorphon.github.io/sharedtasks/2018/`

```
2008 V;3;SG;PRS
2008 V;NFIN
2014 V;PST
1984 V;V.PTCP;PRS
1986 V;V.PTCP;PST
```

There are 77 in `all` and 9 languages in `surprise`.

- `wc -l ../all/english-*`
  - `1000 ../all/english-dev`
  - `1000 ../all/english-test`
  - `100 ../all/english-train-low`
  - `1000 ../all/english-train-medium`
  - `10000 ../all/english-train-high`
- `python2 baseline.py`
  - `english[task 1/low]:  0.772`
  - `english[task 1/medium]:  0.908`
  - `english[task 1/high]:  0.949`

More detailed information available at
`https://github.com/sigmorphon/conll2018/tree/master/task1/baseline`

1. use English and at least one other language (specify in your `readme`
2. use only `*-train-high` for extracting rules
3. working as a group up to two students is *possible* (let us know before Oct. 30) — HW submission will be individual
4. `./baseline_improved.sh` which will learn and evaluate your model (as in `baseline.py`)
5. your output will be similar to the original `baseline.py`
6. (50pts) your result will be hopefully better than the baseline result
7. (25pts) in `readme`, describe in detail what you *change/improve* alignment, span merging and/or rule extraction
8. in `readme`, describe how to apply transformation rules

References

Cotterell, R., Kirov, C., Sylak-Glassman, J., Walther, G., Vylomova, E., Xia, P., Faruqui, M., Kübler, S., Yarowsky, D., Eisner, J., and Hulden, M. (2017). CoNLL-SIGMORPHON 2017 Shared Task: Universal Morphological Reinflection in 52 Languages. In *Proceedings of the CoNLL SIGMORPHON 2017 Shared Task: Universal Morphological Reinflection*, pages 1–30, Vancouver. Association for Computational Linguistics.

Cotterell, R., Kirov, C., Sylak-Glassman, J., Yarowsky, D., Eisner, J., and Hulden, M. (2016). The SIGMORPHON 2016 Shared Task—Morphological Reinflection. In *Proceedings of the 2016 Meeting of SIGMORPHON*, pages 10–22, Berlin, Germany. Association for Computational Linguistics.