

Clarisse Taboy
Methods II – HW4 Write Up
April 5, 2021

1. At some point, some of my tests kept failing because I had mistakenly used two different symbols for the phoneme /ʃ/: ʃ (integral symbol) and ʃ (esh). The latter was obtained from copy/pasting from the hw instructions, the former from the keyboard shortcut Opt+B. I don't recall how I noticed this mistake, but from now on I feel the need to be more mindful about the exact symbols I use to avoid this mistake in the future.

2. At the start, I used Jupyter Notebook to write my rules, plus the snippet of code below (with 'rápido' as an example) to test whether each individual input string matched its expected output:

```
a = pynini.compose("rápido", G2P)
print(a.string())
```

I was still getting used to running code without directly inputting the string myself and seeing its output, but over time I got more comfortable with the abstractness of unit tests and relying mostly on them. However, I also had to make sure that I wrote good rules and tests.

3. Before asking Kyle about writing an example string under multiple tests when many rules are applied to the one word (like "árvore"), I had written tests for each individual rule. For example: `test_c_to_s_before_i_e`, `test_final_e_to_i_and_before_s`, etc. But if one example string didn't output the expected string, multiple tests would fail for that one word, so it seemed a bit messy. This is what made me wonder about whether it was okay put example strings under all the relevant tests, even if it meant that some examples appear in multiple tests. To test "entire grammars", or overall end result as I understood it, I changed my tests so that I tested words and not rules instead (which I assume is what is expected, based on the given example test "test_vez"). Though for words like "braço/s" and "case/s", I put both the singular and plural forms under one test rather than two separate ones. My reasoning was that the only difference between singular and plural forms was the addition of "-s", and writing separate tests seemed too much for a small difference. Thoughts?

4. In the stretch goals, I had a bit of trouble with the rule for "r" → /ɾ/ especially when it came to "árvore"/"arvuri". Initially I included "á" in the union of all vowels, but doing so gave me "arvuri". Since it seemed like "r" doesn't change only when it follows "á", I treated "á" like an exception to the general rule and so I put the union of "á" in its own variable "acute_a" instead of including it in the regular vowels. Not sure if this is okay to do, or if there are other/better conventions for this.

5. Running black and flake8 on my code would initially give me conflicting results, because black would format my lines to the default 88-character length, which flake8 complained about. So Kyle suggested I specify the length for black to 79, which solved the problem.