

Natural Language 2023/2024

MP1 Group 16 - Maria Campos (95629) and Tomás Pereira (95682)

Name of the group member	Contribution Estimation	Justification
Maria Campos	50%	Sketched and developed half of the transducers, collectively worked on the last three FSTs, wrote report
Tomás Pereira	50%	Sketched and developed half the transducers, collectively worked on the last three FSTs, wrote scripts

Tests

- 1) **t-studentX-en-mix**: Test with Student X's 18th bday in EN mixed condensed format
- 2) **t-studentX-pt-mix**: Test with Student X's 18th bday in PT mixed condensed format
- 3) **t-studentX-num**: Test with Student X's 18th bday in the numeric format

Delivery FSTs

- **mmm2mm.fst** was written from scratch.
- **mix2numerical** was obtained by **concatenating** mmm2mm.fst with the transducer copy_day_year2out.fst, which was written from scratch. Tested with tests **1)** and **2)**.
- **pt2en.fst** was obtained by **concatenating** month_pt2en.fst, written from scratch, with the transducer copy_day_year2out.fst.
- **en2pt.fst** was obtained by **inverting** pt2en.fst. Tested with test **2)**.
- **day.fst** was written from scratch.
- **month.fst** was written from scratch.
- **year.fst** was written from scratch.
- **datenum2text.fst** was obtained by **concatenating** month.fst with separator.fst, then the result was concatenated with day.fst and again concatenated with comma.fst and finally with year.fst. Tested with test **3)**.
- **mix2text.fst** was obtained by **composing** mix2numerical.fst with datenum2text.fst. Tested with tests **1)** and **2)**.
- **date2text.fst** was obtained by performing the **union** between mix2text.fst with datenum2text.fst. Tested with **all** tests.

Note: All FSTs were tested with intermediate tests adapted to each of them. These can be found and executed from the folder **/scripts**.

Auxiliary FSTs: The auxiliary developed FSTs were **copy_day_year2out.fst**, which copies the segment `"/dd/yyyy"` of the input to the output, **month_pt2en.fst**, which converts the month symbols from PT to EN, **separator.fst**, which reads a `" / "` and outputs an *eps* and **comma.fst**, which reads a `" / "` and outputs a `" , "` symbol.