TÉCNICO LISBOA

Report Mini-Project Nº1 (MP1) Natural Language

Frederico Silva (99222), Pedro Cruz (99297) - October 2023

Work distribution:

Frederico Lucas Marques da Silva (50%) | Pedro Agostinho da Cruz (50%)
Both members of the team worked equally and distributed the work in a proportional way.

Sources:

- **slashdate.txt:** Allows for the representation of the slash symbol. Used in transducers that represent the date in a numerical format.
- **slash.txt:** Allows for the recognition of the slash symbol.
- **comma.txt:** Allows for the representation of the comma symbol.

Both *slash.txt* and *comma.txt* are used in transducers that represent the date in a text format.

- **monthPtToEn.txt:** Converts the abbreviated names of the months from Portuguese to English. By reversing this transducer, it can make the translation from English to Portuguese. Therefore, this transducer was used to create both pt2en and en2pt transducers.
- **month.txt:** Allows for the conversion of the months from a numerical format to their textual representation in English. It accepts both the single and double-digit format when converting the first 9 months.
- **month2name.txt:** Converts both the English and Portuguese abbreviated month names to their corresponding non-abbreviated form always in English, this is used in transducers that convert date to text.
- **0-9.txt:** Transducer that mirrors the input symbol if its a number 0-9. When two of these tranducers are concatenated, they allow us to represent any day within a 30-day month from a numerical standpoint.
- **0sdays.txt:** Allows for the textual representation of the days that start on 0's digit, which includes all of those from 01-09. It has two different paths so that it can be used both with the first 9 days as well as the days 10 through 31.
- **10sdays.txt / 20days.txt / 30days.txt:** All of these allow for the representation of the days that start with the digits 1, 2, or 3 respectively allowing for the textual representation of the day in a numerical format. It is also important to note that: 10sday.txt allows for the representation of the 1st, 20sday.txt for the representation of the 2nd and the 20th, and 30sday.txt for the representation of the 30th and 31st.
- **mmm2mm.txt:** Allows for the conversion of a month's numerical representation into its abbreviated textual form in English. It's used in mmm2mm and mix2numerical transducers.
- **year.txt** Allows for the textual representation of the date's year from 2001 to 2099. For this it has one path for the first 9 years (2001-2009) that can be combined with the paths that represent every "decade" from the *20s* to the *90s* allowing for the conversion of every year in between. It also has a different unique path for the *10s* decade (2011-2019).

All the other transducers were obtained with transformations based on these transducers through operations such as concatenation, union, and reversal. The order to which this operations were applied can be seen on the compilation part of the file run.sh.