GF - Portuguese Extension

Rami Galil
Overview

- Portuguese language
- Starting with a new language (personal experience)
- Future work
Portuguese

- Romance language
- mutually intelligible with standard Spanish
- most of the lexicon is derived from Latin
New Language

- started without the functor for Romance languages
- using a small set of files: GrammarPor, LangPor, LexiconPor, ParadigmsPor, ResPor
- majority of the work is in GrammarPor and ResPor
New Language (cont.)

• morphological paradigms easier than syntax rules
• implemented: morphological paradigms for verbs, nouns and adjectives
• try to have a good smart paradigm in order to use the worst-case function as little as possible
Verbs

• verbs are the most inflectionable words in Portuguese (number, person, mood, tense, aspect and voice)

VForm =
\[\text{VInf} \mid \text{VInf TMood Number Person} \mid \text{VInf Verb NumPers} \mid \text{VPart Gender Number} \mid \text{VGer} \]

-- very simple for now
TMood = VPres Mood ;
Mood = Indic | Conjunct ;
NumPersI = SgP2 | PlP2 ; -- | PlP1 ;
Tense = Pres ;
Adjectives

• adjectives in Portuguese: some can be placed before the noun; some can be placed after
• how to deal with adjectives that can be placed before and after?
• in some cases the same adjective can mean different things depending on the position (homem rico vs. rico homem)

Adj : Type = {s : Gender => Number => Str ; isPre : Bool} ;
Results

- small extracts from test.trees and test.treebank

AdjCN (PositA brown_A) (UseN door_N)

AdjCN (PositA heavy_A) (AdjCN (PositA small_A) (UseN door_N))

Lang: AdjCN (PositA brown_A) (UseN door_N)
LangEng: brown door
LangPor: porta castanha

Lang: AdjCN (PositA heavy_A) (AdjCN (PositA small_A) (UseN door_N))
LangEng: heavy small door
LangPor: pequena porta pesada
Future Work

- complete the verbs param
- implement more syntax rules (predication and determination)
- check if the use of the functor for Romance languages is convenient in this case
- refactoring (not sure if all the definitions are in the files they should be)
Questions?