

Tem como objetivo reduzir uma palavra à sua forma base e agrupar diferentes formas da mesma palavra. Por exemplo, os verbos no tempo passado são alterados para presente (por exemplo, "foi" é alterado para "vai") e os sinônimos são unificados (por exemplo, "melhor" é alterado para "bom"), padronizando palavras com significado semelhante à sua raiz.

A lematização resolve as palavras em sua forma de dicionário (conhecida como lema), para a qual requer dicionários detalhados nos quais o algoritmo pode pesquisar e vincular palavras aos lemas correspondentes.

Por exemplo, as palavras "correr", "corre" e "correu" são todas formas da palavra "correr", portanto "correr" é o lema de todas as palavras anteriores.

A lematização também leva em consideração o contexto da palavra para **resolver outros problemas, como a desambiguação**, o que significa que ela pode discriminar entre palavras idênticas que têm significados diferentes, dependendo do contexto específico. Pense em palavras como "morcego" (que pode corresponder ao animal ou ao clube de "banco". Ao fornecer um parâmetro de parte do discurso para uma palavra (seja um substantivo, um verbo etc.), é possível definir um papel para essa palavra na frase e remover a desambiguação.

A implementação deste processo pode ser realizado com a lib stanza. O módulo de lematização retorna a forma do lema para cada palavra de entrada. Por exemplo, a sequência de entrada "Eu comi uma maçã" será lematizada em "Eu como uma maçã". Esse tipo de normalização de palavras é útil em muitos aplicativos do mundo real. Na Stanza, a lematização é realizada pelo LemmaProcessor.

Exemplo de utilização

```
import stanza
nlp = stanza.Pipeline(lang='en', processors='tokenize,mwt,pos,lemma')
doc = nlp('Barack Obama was born in Hawaii.')
print(*[f'word: {word.text+" "} \tlemma: {word.lemma}' for sent in doc.sentences for word in sent.words], sep='\n')
```

saída do código acima:

```
word: Barack      lemma: Barack
word: Obama       lemma: Obama
word: was         lemma: be
word: born        lemma: bear
word: in          lemma: in
word: Hawaii      lemma:
Hawaii word: .    lemma: .
```

O tutorial para obter mais detalhes pode ser acessado em : <https://stanfordnlp.github.io/stanza/lemma.html>