

CYK Probabiliste

Arthur Lapraye & Korantin Lévêque & Mélanie Viegas

Paris VII

28 juin 2016

L'Algorithme CYK

- Un algorithme de parsing ascendant

L'Algorithme CYK

- ▶ Un algorithme de parsing ascendant
- ▶ Complexité $\mathcal{O}(|G|n^3)$

L'Algorithme CYK

- ▶ Un algorithme de parsing ascendant
- ▶ Complexité $\mathcal{O}(|G|n^3)$
- ▶ Parsing tabulaire

L'Algorithme CYK

- ▶ Un algorithme de parsing ascendant
- ▶ Complexité $\mathcal{O}(|G|n^3)$
- ▶ Parsing tabulaire
- ▶ Extension aux grammaire hors-contexte probabilistes (PCFG)

L'Algorithme CYK

Les PCFG

- ▶ Les CFG : un quadruplet (Σ, V, S, P)

Les PCFG

- ▶ Les CFG : un quadruplet (Σ, V, S, P)
- ▶ Les CFG pondérées : ajout d'une fonction
 $f : p \mapsto \alpha, p \in P, \alpha \in \mathbb{R}$

Les PCFG

- ▶ Les CFG : un quadruplet (Σ, V, S, P)
- ▶ Les CFG pondérées : ajout d'une fonction $f : p \mapsto \alpha, p \in P, \alpha \in \mathbb{R}$
- ▶ Les CFG probabilistes : toutes les réécritures d'un non-terminal doivent sommer à 1.

Les PCFG

- ▶ Les CFG : un quadruplet (Σ, V, S, P)
- ▶ Les CFG pondérées : ajout d'une fonction
 $f : p \mapsto \alpha, p \in P, \alpha \in \mathbb{R}$
- ▶ Les CFG probabilistes : toutes les réécritures d'un non-terminal doivent sommer à 1.
- ▶

Transformer la grammaire en CNF



Le corpus Sequoia

- Un corpus diversifié



Notre implémentation du CYK

Evaluation

References I



Brian Roark, Richard Sproat.

Computational Approaches to Morphology and Syntax.
Oxford University Press, 2007.



Mariana Romanyshyn, Vsevolod Dyomkin.

The Dirty Little Secret of Constituency Parser Evaluation,
2014.

[http ://tech.grammarly.com/blog/posts/The-Dirty-Little-Secret-of-Constituency-Parser-Evaluation.html](http://tech.grammarly.com/blog/posts/The-Dirty-Little-Secret-of-Constituency-Parser-Evaluation.html)