

Communication

Campagne d'évaluation PARSEME

Nicolas ZAMPIERI - M2 IAAA - 2018/2019

Sommaire

- ▶ Introduction
- ▶ Campagne d'évaluation PARSEME
- ▶ Systèmes soumis en 2018
- ▶ Analyses
- ▶ Conclusion

Introduction

- ▶ Qu'est ce que PARSEME ?
 - ▶ Un réseau scientifique interdisciplinaire consacré aux rôles des expressions (MWE) dans l'analyse syntaxique.

Campagne d'évaluation PARSEME

► Tâche:

- Reconnaître automatiquement les expressions dans le texte.
 - *Se faire des idées*
 - *Prendre de la poudre d'escampette*
 - *Il vide son sac*
 - *Prendre le taureau par les cornes*

Campagne d'évaluation PARSEME

- ▶ Pourquoi sont-elles compliquées à reconnaître automatiquement ?
 - David fait une présentation.
 - Une présentation est faite par David.
 - Plaidez-vous coupable ou non coupable ?

Campagne d'évaluation PARSEME

- ▶ Pourquoi sont-elles compliquées à reconnaître automatiquement ?
 - David **fait** une **présentation**.
 - Une **présentation** est **faite** par David.
 - Plaidez-vous **coupable** ou **non coupable** ?

Campagne d'évaluation PARSEME

▶ Comment est évalué un système ?

▶ Deux variantes Complète / Par token :

▶ Bijection optimale

▶ Précision :

$$\bullet P = \frac{|AnS|}{|S|}$$

▶ Rappel :

$$\bullet R = \frac{|AnS|}{|A|}$$

▶ Score :

$$\bullet F = \frac{2 \cdot P \cdot R}{P + R}$$

Token	Attendu (A)	Système (S)
t1	1	1
t2	1	2
t3	2	2

▶ Complète :

$$\bullet P = R = F = 0$$

▶ Par token :

$$\bullet P = R = F = \frac{2}{3}$$

Systèmes

- ▶ 17 systèmes ont participé à la campagne d'évaluation (2018) :
 - 13 systèmes à la “*closed-track*”
 - 4 systèmes à la “*open-track*”
- ▶ Deux types de systèmes :
 - À base de réseaux de neurones (NN)
 - À base de Conditional Random Fields (CRF)

Systemes à base de NN

TRAPACC / TRAPACCs

- ▶ Features
 - Lemmes
 - Parties du discours (UPOS)
 - Arbre de dépendances
- ▶ Modèle
 - Modèle de transition (dérivé de MAST)
 - Réseau convolutionnel (CNN)
- ▶ Sortie
 - Format en Configuration

GBD-NER

- ▶ Features
 - Formes, Lemmes
 - Parties du discours (UPOS et XPOS), Attributs
 - Arbre de dépendances
- ▶ Modèle
 - Graphe de dépendances (GBD)
 - Réseau de neurones récurrents bidirectionnels (biRNN)
- ▶ Sortie
 - Format BIOES

Systemes à base de CRF

TRAVERSAL

- ▶ Features
 - Lemmes
 - UPOS
 - Arbre de dépendances
- ▶ Modèle
 - Hypergraphe, hyper-arc, hyper-nœud
 - Extension de CRF
- ▶ Sortie
 - Hyperpath

CRF-DepTree

- ▶ Features
 - Lemmes
 - UPOS
 - Arbre de dépendances
- ▶ Modèle
 - CRF au format XML (XCRF)
- ▶ Sortie
 - XML
 - Format IO

Systèmes

SHOMA

- ▶ Features
 - Formes, lemmes
 - UPOS
- ▶ Modèle
 - CNN
 - biRNN
 - CRF
- ▶ Sortie
 - Format BIO

Autres Systèmes

- ▶ Veyn
 - ▶ À base de RNN
- ▶ Mumpitz
 - ▶ À base de RNN
- ▶ Deep-BGT
 - ▶ À base de RNN-CRF

Analyses

Résultats généraux

Systèmes	MWE-Complète (F-mesure en %)
SHOMA	58.09
Traversal	54
TRAPACCs	49.74
TRAPACC	49.57
CRF-DepTree	43.91
GBD-NER	41.62
Veyn	36.94

Analyses

Résultats généraux

Systèmes	MWE-Complète (F-mesure en %)
SHOMA	58.09
Traversal	54
TRAPACCs	49.74
TRAPACC	49.57
CRF-DepTree	43.91
GBD-NER	41.62
Veyn	36.94

Analyses

- ▶ Quels types de systèmes sont le plus impactés par les MWEs non vues en entraînement ?

Systèmes	Corrélation de Spearman	P-value
SHOMA	-0.7357	0.00033
Traversal	-0.8841	5.133e-07
TRAPACC	-0.7120	0.00062
CRF-DepTree	-0.6374	0.0033
GBD-NER	-0.4556	0.0499
Veyn	-0.6201	0.0046

Analyses

- ▶ Quels types de systèmes sont le plus impactés par les MWEs discontinues ?

Systèmes	Corrélation de Spearman	P-value
SHOMA	-0.3887	0.0999
Traversal	-0.3448	0.1481
TRAPACC	-0.5484	0.0150
CRF-DepTree	-0.3238	0.1762
GBD-NER	-0.2343	0.3342
Veyn	-0.0096	0.9686

Analyses

- ▶ Quels types de systèmes dépendent le plus de la richesse en MWEs des corpus d'entraînement ?

Systèmes	Corrélation de Spearman	P-value
SHOMA	0.3684	0.1206
Traversal	0.3263	0.1727
TRAPACC	0.3684	0.1206
CRF-DepTree	0.4105	0.0808
GBD-NER	0.3912	0.0976
Veyn	0.1948	0.4241

Conclusion

Systèmes	Corrélation de Spearman (non vues)	Corrélation de Spearman (discontinues)	Corrélation de Spearman (richesse corpus)
SHOMA	-0.7357	-0.3887	0.3684
Traversal	-0.8841	-0.3448	0.3263
TRAPACC	-0.7120	-0.5484	0.3684
CRF-DepTree	-0.6374	-0.3238	0.4105
GBD-NER	-0.4556	-0.2343	0.3912
Veyn	-0.6201	-0.0096	0.1948

Merci pour votre attention !

Annexe 1

- ▶ Quels types de systèmes dépendent le plus de la taille des corpus d'entraînement ?

Systèmes	Corrélation de Spearman	P-value
SHOMA	0.0965	0.6942
Traversal	-0.0745	0.7615
TRAPACC	-0.2135	0.5886
CRF-DepTree	0.0535	0.8276
GBD-NER	0.1904	0.4348
Veyn	-0.0992	0.6861

Annexe 2

TRAPACC et TRAPACCs

- ❑ Configuration :
 - (σ, β, P)
- ❑ Shift :
 - $(\sigma, x|\beta, P) \Rightarrow (\sigma|x, \beta, P)$
- ❑ Merge :
 - $(\sigma|x|y, \beta, P) \Rightarrow (\sigma|(x, y), \beta, P)$
 - Si x et y appartiennent à une MWE.
- ❑ Reduce :
 - $(\sigma|x, \beta, P) \Rightarrow (\sigma, \beta, P)$
- ❑ Complete-MWE* :
 - $(\sigma|x, \beta, P) \Rightarrow (\sigma, \beta, P \cup \{x\})$
 - Si x est une MWE complète.

Annexe 3

TRAVERSAL

► Représentation des données

