

Description du sujet : Le groupe de travail "Prétopologie, TAL et temporalité" réunit des chercheurs des laboratoires LIFO et LLL spécialisés en Linguistique, Traitement Automatique des Langues et Apprentissage Automatique. Le sujet d'étude de ce groupe est l'analyse du discours et plus particulièrement la tâche consistant à extraire d'un texte annoté une structure temporelle prenant la forme d'un DAG (Graphe Orienté sans Cycle) d'événements verbaux pré-identifiés [Ning et al., 2017]. L'approche envisagée consiste à apprendre un espace prétopologique structurant (algorithme LPS [Caillaud&Cleuziou, 2018]) à partir d'un ensemble de prédicats censés capturer l'information linguistique, lexicale, syntaxique et sémantique portée par chaque événement temporel.

Description des tâches : Les stagiaires auront pour mission la réalisation d'une preuve de concept par le développement en collaboration d'une chaîne de traitement complète permettant d'extraire les relations prédictives à partir d'un corpus d'énoncés pré-annotés. Ils participeront à l'élaboration de ces prédicats en concertation avec l'équipe au regard des travaux récents sur cette problématique et évalueront quantitativement et qualitativement les structures temporelles issues de l'algorithme LPS relativement aux prédicats construits et aux approches existantes.

Le profil linguiste s'intéressera à la caractérisation de l'ensemble des traits essentiels à l'apprentissage pour chaque paire d'événements verbaux, tandis que le profil informaticien s'intéressera à la faisabilité et la mise en œuvre des méthodes proposées. Le binôme devra évaluer et interpréter l'évaluation des résultats de cette chaîne de traitement et préconiser un travail d'adaptation de la chaîne en vue d'une amélioration des résultats.

Encadrants :

- Encadrants principaux stage informatique : Anaïs Lefeuvre-Halftermeyer <anaïs.halftermeyer@univ-orleans.fr> et Gaëtan Caillaud <gaetan.caillaud@univ-orleans.fr>
- Encadrants principaux stage linguistique : Anne-Lyse Minard <anne-lyse.minard@univ-orleans.fr> et Anaïs Lefeuvre-Halftermeyer <anaïs.halftermeyer@univ-orleans.fr>
- Autres membres du groupe de travail : Sylvie Billot et Guillaume Cleuziou

Références :

[Caillaud&Cleuziou, 2018] Gaëtan Caillaud, Guillaume Cleuziou: Learning Pretopological Spaces to Model Complex Propagation Phenomena: A Multiple Instance Learning Approach Based on a Logical Modeling. CoRR abs/1805.01278 (2018)

[Ning et al., 2017] Qiang Ning, Zhili Feng, Dan Roth: A Structured Learning Approach to Temporal Relation Extraction. EMNLP 2017: 1027-1037

Paramita Mirza and Anne-Lyse Minard. HLT-FBK: a complete Temporal Processing system for QA TempEval. In Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015).