

Entreprise : AINUW

Contact : yann.gainche@ainuw.com, laurent.baudet@ainuw.com

Tuteur académique :

Contexte :

La société Ainüw développe une plate-forme WhoAml, fondée sur l'approche systémique, vouée à modéliser des systèmes et leurs interactions, à travers un modèle de langage des relations binaires entre les catégories de ces systèmes. Ces relations sémantiques sont nommées et la force de l'interaction que chacune induit entre deux systèmes est quantifiée.

Un objectif de cette plate-forme est, à partir d'une description minimale d'un système, de générer un comportement de ce nouveau système en interaction avec ceux existants. Les applications sont nombreuses, citons celle dans le cadre des personnages non-joueurs (PNJ) dans les jeux vidéo. Un tel personnage PNJ possède son caractère, et son mode de comportement dans le jeu, qui ne sont pas centraux au jeu, mais il reste important pour agrémenter les scènes et scénarios secondaires. La génération manuelle de la modélisation de ce type de personnage est fastidieuse, alors qu'en considérant ce personnage comme un système en interaction avec les autres, sa modélisation pourrait être générée à partir d'une description minimale des concepts qui le caractérisent, à l'aide de la plate-forme WhoAml.

Sujet :

La plate-forme WhoAml est en cours de construction. Elle comporte actuellement un dictionnaire comprenant un millier de catégories, elles-mêmes définies par des ensembles de mots. Le travail consiste à inférer les relations entre ces catégories, sur la base (1) des mots les décrivant, et (2) des relations lexicales entre ces mots (synonymie, proximité des champs lexicaux, etc.), et (3) d'autres ressources susceptibles d'être importantes dans ce cadre. La langue considérée est le français. Les ressources publiques peuvent être exploitées (corpus, outils lexicaux et linguistiques, outils génératifs, etc.). Le travail consiste ainsi à inférer, à partir de données lexicales à agréger, les arêtes d'un graphe d'un millier de nœuds, et de les étiqueter selon le type d'interaction (un type parmi 12), et selon la force de l'interaction.

Afin d'évaluer la qualité de l'inférence, la vérité terrain sera constituée d'un sous-ensemble complet de WhoAml qui sera rendu disponible, comprenant au moins une quinzaine de catégories décrites et dont les interactions seront elles aussi décrites. Deux à trois mesures de performances des relations inférées seront à définir, et permettront de comparer quantitativement les différentes approches d'inférence choisies.

Lors de la présentation orale du sujet, cette vérité terrain sera commentée.

Outils : les participants devront trouver des outils pré-existants pour agréger/exploiter les informations lexicales nécessaires pour inférer les arêtes du graphe et les étiqueter. Dans le groupe de participants, une connaissance de la programmation est nécessaire.