

ANR COMPO & ANR HEBBIAN

Alexis Nasr

COMPO

- ▶ Compositionality-capable Deep Learning Models of Natural Language
- ▶ Consortium
 - ▶ LIS
 - ▶ Laboratoire de Linguistique Formelle (Paris)
 - ▶ Laboratoire d'Informatique de Grenoble
 - ▶ Laboratoire de Psychologie Cognitive (Marseille)
- ▶ Durée : 48 mois
- ▶ Début : janvier 2024

Questions scientifiques

- ▶ Peut-on améliorer les capacités de **généralisation** des modèles de langages actuels en y introduisant des **biais cognitifs** ?
- ▶ Hypothèse : les modèles actuels généralisent mal
→ performances sur le jeu de données COGS
- ▶ Hypothèse : Le langage possède une **structure**, cette structure provient des **limites cognitives humaines**, en particulier **la mémoire**
→ l'humain peut difficilement mémoriser de longues séquences arbitraires, mais réussit mieux lorsque ces séquences possèdent une structure sous jacente.

Méthodologie

- ▶ Deux biais :
 - ▶ biais de mémoire
 - ▶ biais de structure
- ▶ Deux tâches :
 - ▶ segmentation en mots
 - ▶ prédiction de la structure sémantique
- ▶ Etudier l'impact des deux biais sur les deux tâches
 - ▶ Comment mesurer cet impact ?
 - ▶ On ne vise pas à améliorer l'état de l'art
 - ▶ Evaluation sur des données contrôlées ?
 - ▶ Comparer les performances des modèles aux performances humaines ?

Workpackages

- ▶ WP1 Identifying compositionality-related biases (LLF-Wisniewski)
- ▶ WP2 Structural biases for semantic prediction (LIG - Gaussier)
- ▶ WP3 Memory biases for segmentation (LIS - Nasr)
- ▶ WP4 Memory biases for semantic prediction (LLF - Crabbé)

WP1 Identifying compositionality-related biases

- ▶ Diagnostic de modèles existant
 - ▶ sondes (probes)
 - ▶ jeux de données contrôlés (COGS - natural stories)
- ▶ Comportement sur les phénomènes non compositionnels
 - ajouter des MWE dans COGS
 - expériences en traduction
- ▶ Etude des limites sur la mémoire de travail des modèles de langages autorégressifs
 - $a = 1; b = -a; c = b; d = -c; d = ?$
 - adaptation à des phénomènes linguistiques

WP2 Structural biases for semantic prediction

- ▶ Ajouter de la structure syntaxique explicite (ou pas) de manière non supervisée
 - parsing non supervisé
 - prise en compte de la structure dans un transformer
 - prediction jointe de la syntaxe et du modèle de langage
- ▶ Etudier les effets sur une tâche d'analyse sémantique explicite (production d'une structure)
- ▶ Evaluation sur :
 - ▶ COGS
 - ▶ Geoquery (construction de requêtes à partir de texte)
 - ▶ CFQ (Compositional freebase questions)

WP3 Memory biases for segmentation

- ▶ Tâche : segmentation d'un flux de phonèmes (acoustique ?) en unités lexicales
- ▶ Différente de la tâche de segmentation en TAL car on ne dispose pas de lexique, c'est en même temps de la segmentation et de l'acquisition lexicale
- ▶ Etude de modèles issus de la psycho-linguistique (PL) (PARSER, CBL)
- ▶ Etude du processus d'oubli dans les modèles PL
- ▶ Etude de l'oubli dans des réseaux de neurones
- ▶ Etude des modèles neuronaux hebbiens (*Cells that fire together wire together*)

Sous tâches

- ▶ Performances linguistiques des modèles PL
→ arrivent-ils à inférer le lexique ?
- ▶ Performances comportementales des modèles PL
→ corrélation des scores des mots avec des données issues de self-paced reading
- ▶ Etude de l'oubli dans des réseaux de neurones
→ oubli simulé (on cherche à produire les scores des modèles PL ou des données comportementales)
→ oubli réel (vecteurs d'état des RNN, matrices d'attention, ...)
→ quels moyens a-t-on de contrôler le processus d'oubli dans un RN
- ▶ Réseaux de neurones hebiens

WP4 Memory biases for semantic prediction

- ▶ Tâche de prédiction sémantique sur les données COGS, Geoquery et CFQ
- ▶ La tâche de prédiction est vue comme la construction d'un graphe pondéré par un score d'activation (soumi à l'oubli)
- ▶ Contraintes d'incrémentalité
- ▶ Contraintes de mémoire

Moyens

- ▶ une bourse de thèse
- ▶ des financements de stages

HEBBIAN

- ▶ Hebbian learning of sequences
- ▶ Consortium
 - ▶ Laboratoire de Psychologie Cognitive
 - ▶ LIS
 - ▶ INRIA Côte d'Azur
 - ▶ Bases Corpus Langage (Université Côte d'Azur)
- ▶ 48 mois

Apprentissage hebien

- ▶ Aussi appelé apprentissage associatif
- ▶ Théorie neuropsychologique
- ▶ L'activation simultanée de cellules nerveuses augmente la force synaptique liant ces cellules
- ▶ Modélisation naïve.
 - ▶ deux neurones n_i et n_j ,
 - ▶ x_i l'entrée de n_i et x_j l'entrée de n_j
 - ▶ $w_{i,j}$ le poids de la connexion entre n_i et n_j
 - ▶ pour chaque exemple d'apprentissage, $w_{i,j}$ est mis à jour en fonction des valeurs de x_i et x_j : $w_{i,j}$ augmente si $x_i = x_j$ et diminue sinon.
 - ▶ lorsque $w_{i,j}$ est élevé, l'activation de n_i "pré-active" n_j

Thématique générale

- ▶ Comprendre la dynamique de la mémorisation de séquences.
- ▶ Trois questions :
 - ▶ What is the relation between the spacing between two repetitions of the same sequence and the development of a memory trace of that sequence ?
 - ▶ How does sequence encoding vary with sequence length, number of repetitions and context ?
 - ▶ How are small, regular sequences that are embedded in larger sequences encoded ?

Partie expérimentale

- ▶ Acquisition de données linguistiques et non linguistiques chez l'humain
- ▶ Acquisition de données chez le babouin
- ▶ Modélisation à l'aide de modèles psychologiques.

Partie modélisation

- ▶ Traitement des données acquises avec le modèle PARSEUR
- ▶ Implémentation de l'algorithme de PARSEUR à l'aide de modèles hebiens
- ▶ Traitement des données acquises à l'aide de modèles neuronaux bio inspirés (mixed coding neurons)

Moyens

- ▶ 6 mois de stages