



Master Sciences Cognitives – 2017-2018

Etude et simulation informatique de scénarios d'interaction entre un bébé virtuel et une mère qui lui apprend à parler

La question des relations entre mécanismes de perception et d'action s'est installée au cœur des réflexions et des travaux à la fois en sciences cognitives, avec la découverte des neurones miroir et les avancées en neuroimagerie, et dans le domaine de la robotique autour des enjeux d'apprentissage et d'adaptabilité.

Nous étudions cette question dans le cadre de la communication parlée, en développant des modèles computationnels d'agents capables à la fois de percevoir et de produire des stimuli de parole simples (voyelles, consonnes, syllabes) et en analysant sur ces modèles en quoi les capacités de production peuvent interagir avec les capacités perceptives, les contraindre, les enrichir ou les compléter.

Ceci a conduit à la mise au point d'un modèle bayésien d'agent communiquant, que nous avons baptisé COSMO, pour *Communicating about Objects using SensoriMotor Operations*. Ce modèle probabiliste provient de la modélisation de la situation de communication, dans laquelle deux agents veulent communiquer à propos d'un objet de l'environnement. Le modèle COSMO est basé sur l'hypothèse fondamentale que cette situation de communication peut être internalisée et émulée dans le cerveau de chaque agent, qui est alors en mesure d'agir aussi bien en tant que locuteur qu'auditeur.

Dans ce contexte, l'apprentissage de ce modèle combine l'apprentissage des capacités perceptives et motrices. Il repose sur plusieurs étapes. Dans la première, l'agent apprenant construit un système de catégorisation des signaux acoustiques produits par les agents maîtres (apprentissage auditif). Dans la seconde, l'agent apprend, par accommodation, un système de prédiction des conséquences acoustiques de ses gestes moteurs (apprentissage sensorimoteur).

Enfin, dans la troisième étape, l'agent apprend un système de catégorisation des gestes moteurs lui permettant d'atteindre les cibles acoustiques (apprentissage moteur). Cette séquence d'apprentissages est consistante avec, et inspirée par les données de la littérature sur l'apprentissage de la parole chez le bébé.

Objectifs

Nous n'avons jusqu'à présent jamais exploré de **scénarios d'interaction réalistes entre un agent apprenant et son tuteur**. Or les mères, lorsqu'elles parlent à leurs enfants, utilisent des types d'interaction très variés selon les cultures. On peut ainsi observer des interactions très fortes utilisant souvent un format de parole très particulier, le « mamanais » (ou « motherese », ou Infant Directed Speech, IDS), dans lequel elles mettent en jeu une prosodie et des trajectoires motrices très exagérées. On peut au contraire trouver des cultures où les interactions sont extrêmement faibles. Or, les conséquences de ces variabilités de situation d'interaction sur l'apprentissage du langage n'ont jamais été étudiées computationnellement. Il s'agira ici **d'implémenter sur COSMO des scénarios d'apprentissage réalistes, permettant pour la première fois de coupler des agents COSMO (« bébé » et « maman ») dans de réelles situations d'interaction, et de tester les modifications d'apprentissage selon les conditions d'interaction entre un agent apprenant (un « bébé ») et son maître (une mère ou un père).**

Travail proposé

Pour réaliser l'objectif, l'étudiant devra **implémenter un scénario d'interaction à deux agents, un « bébé » et une « maman »**, et réaliser des **simulations d'apprentissage de stimuli simples** – probablement des voyelles – pour **tenter de montrer, pour la première fois, les modifications de l'efficacité de l'apprentissage selon les conditions d'interaction, et notamment dans le cas d'une parole dirigée vers l'enfant.**



Contexte pratique

Le travail sera réalisé au sein d'une **équipe pluridisciplinaire alliant modélisateurs** – spécialistes de la programmation bayésienne – et **spécialistes de la parole et de la cognition**. Il sera encadré conjointement par Jean-Luc Schwartz, chercheur à GIPSA-Lab (parole) et Julien Diard, chercheur au LPNC (modélisation bayésienne). Une collaboration sera établie avec des spécialistes du développement de la parole à l'ENS Paris (Alejandrina Cristia).

Une poursuite en thèse est envisagée, dans l'objectif d'étudier des scénarios d'interaction plus complets, impliquant des **structures hiérarchiques emboîtées** (syllabes, mots, phrases) et permettant à un bébé virtuel **d'apprendre ces hiérarchies** par un ensemble de processus non supervisés (bottom-up) et supervisés (top-down). Cette thèse comportera des **enjeux théoriques** sur les processus d'interaction et d'apprentissage, et des **enjeux technologiques** pour le développement d'agents communicants autonomes, dotés de capacités d'apprentissage et d'interaction vocale.

Compétences demandées

Le candidat devra avoir soit une formation en sciences cognitives et un intérêt pour la modélisation, soit une formation en programmation et simulation informatique, et une affinité pour la modélisation mathématique en sciences cognitives.

Contacts

Jean-Luc Schwartz (GIPSA-Lab)
jean-luc.schwartz@gipsa-lab.grenoble-inp.fr 04 76 57 47 12

Julien Diard (LPNC)
julien.diard@univ-grenoble-alpes.fr 04 76 82 58 93
<http://diard.wordpress.com>

Financement

Indemnités de stage assurées par le projet ERC « Speech Unit(e)s »