



Modélisation probabiliste de l'apprentissage sensorimoteur de la parole.

Cadre théorique

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

L'objectif principal de ce stage est d'explorer la distinction et l'ordonnement entre les trois étapes d'apprentissage. Par exemple, les étapes sensorimotrices et motrices peuvent être reliées, par un passage progressif d'un apprentissage essentiellement sensorimoteur à un apprentissage essentiellement moteur, au fur et à mesure que la distribution de probabilités sur les gestes moteurs est apprise. Cette distribution, qui est un *prior* dans l'apprentissage sensorimoteur, devient une amorce pour l'apprentissage moteur. Nous faisons la prédiction que combiner ces apprentissages les rendrait plus efficaces, et donc plus rapides.

Un objectif complémentaire est d'étudier l'influence du paradigme d'interaction avec le maître sur la vitesse de l'apprentissage. Par exemple, nous comparerons des mécanismes d'imitation avec validation du maître à des mécanismes de babillage autonome, avec cible acoustique ou cible de communication.

grenoble
images
parole
signal
automatique



Travail proposé

Pour réaliser l'objectif, l'étudiant sera en charge de programmer un ensemble de simulations informatiques pour comparer les performances de plusieurs scénarios d'apprentissage de voyelles, dans COSMO. Cette programmation reprendra et étendra une maquette, disponible au laboratoire. Une revue de la littérature sur le cadre théorique des apprentissages sensorimoteurs est également attendue.

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), Julien Diard, chercheur au LPNC (modélisation bayésienne) et Marie-Lou Barnaud, en thèse en sciences cognitives au sein des deux laboratoires et en charge du développement du modèle COSMO.

Une poursuite en thèse est envisageable, pour poursuivre les recherches sur COSMO et sur les interactions perception-action en parole.

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>

Marie-Lou Barnaud (GIPSA-Lab & LPNC)

marie-lou.barnaud@gipsa-lab.grenoble-inp.fr

Financement

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