

TP - Conception d'un moteur de fusion multimodale

Objectifs

Le but de ce travail est de spécifier et concevoir un moteur de fusion multimodale pour interagir avec une palette de dessin ne disposant d'aucun bouton de fonction. Les interactions possibles sont :

- la reconnaissance de geste 2D,
- la reconnaissance de parole,
- le pointage sur la palette de dessin.

A partir de ces trois types d'interaction, il vous est demandé de réaliser un moteur de fusion multimodale qui permette d'effectuer les actions suivantes sur la palette de dessin.

- Créer objet [Position] [Couleur]
 - La désignation de l'**action** (créer un rectangle ou une ellipse) se fera à l'aide d'un geste 2D. Un geste différent peut être prévu pour chaque type d'objet (rectangle ou ellipse);
 - La désignation de la **position** et de la **couleur** est optionnelle. Ils peuvent être présents dans n'importe quel ordre, mais forcément après la désignation de l'action ;
 - La **position** devra être spécifiée par un mot clé donné à la voix (« ici », « là », « à cette position ») et complétée par un pointage sur la palette désignant la position.
 - La **couleur** pourra être définie soit en donnant une couleur à la voix (« rouge », « noir », « vert », etc.) ; soit en désignant sur la palette un autre objet et en annonçant à la voix « de cette couleur ».
- Déplacer [Objet] [Position]
 - La désignation de l'**action** (déplacer) se fera à l'aide d'un geste 2D ;
 - La désignation de l'**objet** et de la **position** est obligatoire. Ils peuvent être présents dans n'importe quel ordre, mais forcément après la désignation de l'action ;
 - La désignation de l'**objet** se fera par le pointage de l'objet à déplacer, et complété à la voix par la désignation de l'objet (« cet objet », « ce rectangle » ou « cette ellipse » pour être plus précis). Cette désignation pourra aussi être complétée par la couleur de l'objet à déplacer (à la voix) ;
 - La **position** devra être spécifiée par un mot clé à la voix (« ici », « là », « à cette position ») et complétée par un pointage sur la palette désignant la position.

Principales étapes

- Prendre en main les différents agents IVY mis à disposition ;
- Préparer la grammaire utilisée par la reconnaissance de parole ;
- Intégrer, sous forme d'agent IVY, l'algorithme de 1\$ Recognizer pour pouvoir reconnaître les gestes effectués sur la palette ;
- Préparer le dictionnaire de gestes nécessaires pour la réalisation des différentes actions ;
- Faire un agent IVY « Fusion » qui recevra les événements des différents agents IVY (palette, geste et voix) et renverra la commande à effectuer sur la palette : Créer, déplacer ou supprimer un objet.

Travail à rendre

Ce travail peut être réalisé en binôme.

Pour le **dimanche 18 Février au plus tard**, vous devez envoyer :

- un document d'une dizaine de pages contenant :
 - le langage d'interaction utilisé pour la reconnaissance de geste et de parole,
 - vos choix de conception (ordre des instructions, gestion du temps, CARE, etc.),
 - une description de la structure de données utilisée,
 - le contrôleur de dialogue (liste des événements, liste des actions, automate)
 - un exemple d'utilisation de votre application.
- Une vidéo contenant une succession de séquences illustrant le fonctionnement du moteur de fusion dans les différents cas envisagés

Le document, les vidéos et les codes sources sont à envoyer par mail à mathieu.raynal@irit.fr