

Master Recherche 2<sup>e</sup> année  
Intelligence, Interaction, Information

*Projet présenté par Céline Coutrix  
Sous la responsabilité de Laurence Nigay et Philippe Renevier*

# Systemes mixtes : Modèle d'interaction mixte

Présenté et soutenu le 22 juin 2005

## **Jury**

Joëlle Coutaz, *Professeur Université Joseph Fourier*

Yves Denneulin, *Maître de conférences INPG*

Jérôme Euzenat, *Directeur de recherches INRIA*

## **Examineurs extérieurs**

Sabine Coquillart, *Directeur de recherches INRIA*

Laurence Pasqualetti, *Ingénieur de recherche FT R&D*

# Table des matières

<b>1</b>	<b>Introduction</b>	<b>3</b>
1.1	Sujet et ses motivations . . . . .	3
1.2	Objectifs . . . . .	4
1.3	Structure du mémoire . . . . .	4
<b>I</b>	<b>Espace Conception</b>	<b>6</b>
<b>2</b>	<b>Éléments de conception</b>	<b>7</b>
2.1	Réalité mixte . . . . .	7
2.1.1	Définitions . . . . .	7
2.1.2	Exemples . . . . .	8
2.1.3	Caractérisation . . . . .	11
2.2	Modalité d'interaction et multimodalité . . . . .	16
2.2.1	Définition d'une modalité $m$ . . . . .	16
2.2.2	Caractérisation d'une modalité . . . . .	16
2.2.3	Composition de modalités . . . . .	18
2.3	Objet de la tâche et Outil de la tâche . . . . .	20
2.3.1	Outil et Réalité Mixte . . . . .	21
2.3.2	Approche instrumentale des interfaces graphiques . . . . .	24
2.4	Synthèse et conclusion du chapitre 2 . . . . .	27
<b>3</b>	<b>Nouveau modèle d'interaction mixte : point de vue modalité</b>	<b>28</b>
3.1	Définition . . . . .	28
3.1.1	Objet mixte . . . . .	28
3.1.2	Modalité d'interaction mixte . . . . .	33
3.1.3	Objet de la tâche . . . . .	45
3.1.4	Lien avec les modélisations existantes . . . . .	45
3.2	Espace de caractérisation . . . . .	49
3.2.1	Compatibilités, Continuités . . . . .	49
3.2.2	Degré d'intégration . . . . .	50
3.2.3	Parties absentes dans une modélisation . . . . .	51
3.2.4	Synthèse des espaces de caractérisation . . . . .	51
3.3	Classification des systèmes existants . . . . .	51
3.3.1	Systèmes mixtes et interface graphique/réalité virtuelle . . . . .	51

3.3.2	Réalité augmentée et virtualité augmentée . . . . .	53
3.3.3	Exécution augmentée et évaluation augmentée . . . . .	54
3.3.4	Objet mixte . . . . .	55
3.3.5	Outil d'interaction . . . . .	55
3.3.6	Synthèse des taxonomies . . . . .	57
3.4	Conclusion . . . . .	57
<b>4</b>	<b>Synthèse et conclusion la partie I</b>	<b>58</b>
<b>II</b>	<b>Espace Applicatif</b>	<b>59</b>
<b>5</b>	<b>RAZZLE : un jeu de RA mobile et collaboratif</b>	<b>60</b>
5.1	Description . . . . .	60
5.1.1	Matériel . . . . .	61
5.1.2	Interface de RAZZLE dans le casque semi transparent . . . . .	62
5.2	Techniques d'interaction pour le ramassage d'une pièce . . . . .	62
5.2.1	Description . . . . .	62
5.2.2	Étude du pouvoir comparatif du modèle d'interaction mixte . . . . .	63
5.3	Évaluation ergonomique en magicien d'Oz . . . . .	66
5.3.1	Description du protocole expérimental . . . . .	66
5.3.2	Résultats . . . . .	67
<b>6</b>	<b>Aspect génératif du Modèle de l'Interaction Mixte</b>	<b>69</b>
6.1	Conception avec le modèle d'interaction . . . . .	69
6.2	Modalités d'interaction conçues pour RAZZLE . . . . .	69
6.3	Architecture logicielle . . . . .	72
6.3.1	Modèle d'architecture PAC-Amodeus . . . . .	72
6.3.2	Architecture de RAZZLE . . . . .	78
6.3.3	Programmation . . . . .	78
<b>7</b>	<b>Synthèse et Conclusion la partie II</b>	<b>82</b>
<b>8</b>	<b>Conclusion générale</b>	<b>83</b>
8.1	Résumé des contributions . . . . .	83
8.2	Perspectives de développement . . . . .	85
8.2.1	Extensions . . . . .	85
8.2.2	Prolongements : plateforme logicielle de conception et développement . . . . .	86

# Chapitre 1

## Introduction

### 1.1 Sujet et ses motivations

Ce projet de recherche s'inscrit dans le domaine de l'Interaction Homme-Machine et a trait à la conception ergonomique et à la réalisation logicielle de techniques d'interaction sur supports indifféremment mobiles ou fixes, et qui reposent sur le paradigme de réalité mixte (RM). La réalité mixte vise la fusion harmonieuse des mondes physique et numérique (ordinateur). Il s'agit de développer des techniques d'interaction utilisables. Puisque l'être humain a l'habitude d'interagir dans le monde physique de plusieurs façons différentes, les systèmes informatiques de réalité mixte doivent pouvoir leur proposer un éventail pertinent de modalités d'interaction en prenant en compte des éléments physiques dans l'interaction.

Ce faisant, nous cernons notre étude aux modalités d'interaction et aux formes de multimodalité pour la réalité mixte. Nous constatons que les travaux à l'origine de nouvelles modalités d'interaction et formes de multimodalité reposent sur un ensemble de dispositifs physiques dédiés à l'interaction et centrés autour d'un même espace d'interaction : l'ordinateur de bureau, comme un microphone combiné avec une souris définissent des formes d'interaction multimodale visant à offrir une communication plus efficace et plus naturelle avec l'ordinateur. Or les formidables progrès accomplis dans la miniaturisation des microprocesseurs et dans les réseaux informatiques sans fil permettent d'envisager que l'ordinateur de bureau ne soit plus le seul lieu d'interaction entre l'utilisateur et le monde numérique. Ceci s'inscrit dans le mouvement de l'ordinateur ubiquitaire et évanescent. La multiplication des ordinateurs de poche et des assistants personnels, la généralisation de l'utilisation des cartes à puces et la multiplication des systèmes embarqués dans des objets d'usage courant (automobile, télévision, etc.) constituent des témoins de cet essor. Nous retenons de cet essor que l'espace d'interaction devient plus vaste, il comprend l'environnement physique et ne se limite plus seulement à un ordinateur sur un bureau. Ainsi les objets physiques deviennent des supports à l'interaction. Par exemple les Media Blocks [17] reposent sur l'utilisation d'icônes physiques (des cubes) pour manipuler des fichiers informatiques. Quant aux systèmes de chirurgie assistée par ordinateur, ils utilisent des localisateurs 3D pour situer des outils chirurgicaux dans des images médicales acquises en phase préopératoire. L'espace d'interaction sort alors définitivement de l'espace composé par le bureau et l'ordinateur.

Dans ce contexte, la réalité mixte est un paradigme d'interaction qui est né de cette volonté de fusionner les capacités de traitements informatiques et l'environnement physique. En réponse à l'avancée des nouvelles technologies d'interaction, les applications reposant sur le paradigme de

RM sont de plus en plus diversifiées. Nous projettons donc de considérer d'une part les modalités d'interaction en entrée (de l'utilisateur vers le système) et en sortie (du système vers l'utilisateur) reposant sur des objets de l'environnement qui peuvent être physiques : l'environnement fait alors partie du support à l'interaction avec un ordinateur. Enfin nous visons la conception de modalités qui permettent à l'utilisateur de manipuler des entités à la fois physiques et numériques afin de favoriser la fluidité d'actions entre les deux mondes, physique et numérique.

Dans ce contexte, le projet consiste à étudier les techniques d'interaction qui concourent à la fusion harmonieuse des mondes physique et numérique. L'approche que nous adoptons est centrée sur l'interaction.

## 1.2 Objectifs

Notre étude consiste à étudier l'interaction au sein d'un système de réalité mixte au regard des modalités d'interaction et formes de multimodalité. Nous visons donc un modèle d'interaction qui

- définit au sein d'un canevas cohérent les concepts utiles à la conception
- permet de classer les systèmes existants
- sert de support à la conception en guidant les choix de conception et en permettant de décrire de nouvelles formes d'interaction, que nous souhaitons développer pour les évaluer.

Face à l'éventail des possibilités et à la complexité de conception, l'objectif de notre étude est l'identification de concepts, l'élaboration de modèles et la mise en œuvre de techniques d'interaction (modalités d'interaction et formes de multimodalité) dans le but de favoriser la fusion harmonieuse des deux mondes physiques et numériques.

Les objectifs des travaux présentés ici sont :

- Définition d'une technique d'interaction : caractérisation par rapport aux mondes physique et numérique et au contexte d'interaction.
- Conception de techniques d'interaction pour manipuler des objets numériques ancrés dans le monde physique
- Réalisation logicielle des techniques d'interaction conçues
- Intégration des techniques dans un système de Réalité Mixte mobile, le jeu RAZZLE, en vue de leur évaluation expérimentale.

## 1.3 Structure du mémoire

La structure en chapitres de ce mémoire reflète notre démarche de recherche : Nous avons d'abord étudié les aspects de conception, puis nous avons mis à l'épreuve nos contributions pour la conception d'une application. La première partie du mémoire compte quatre chapitres.

- Dans un premier temps, nous étudions les trois facettes de notre travail au chapitre 2 : les systèmes mixtes, les modalités d'interaction et l'interaction multimodale, et enfin les travaux qui distinguent objet et outil dans la réalisation de la tâche.
- Le chapitre 3 présente notre modèle d'interaction mixte, qui repose sur une capitalisation des approches étudiées au chapitre précédent.
- Le chapitre 4 nous permet de faire le point sur les contributions conceptuelles de notre modèle de l'interaction mixte présenté au chapitre précédent.
- La deuxième partie consacrée à la conception d'un système mixte compte trois chapitres. Le chapitre 5 décrit le système sur lequel nous avons appliqué concrètement notre modèle

d'interaction mixte.

- Le chapitre 6 illustre sur un cas concret les capacités génératives de nouvelles techniques d'interaction de notre modèle.
- Nous capitalisons les résultats issus de l'étude du cas pratique au chapitre 7.
- Nous concluons notre mémoire au chapitre 8 par une synthèse de nos contributions et nous élargissons ensuite le cadre d'étude par l'esquisse de perspectives conceptuelles et techniques.

Première partie

**Espace Conception**

## Chapitre 2

# Éléments de conception

À l'origine de cette étude, nous trouvons des travaux sur la réalité mixte, l'interaction multimodale, et l'interaction instrumentée. Toutes ces approches de l'interaction homme-machine servent de fondement à notre travail, qui vise l'établissement d'un modèle d'interaction avec les systèmes mixtes. En gardant cet objectif à l'esprit, nous présentons ici ces approches fondatrices en commençant par la réalité mixte, puis nous présentons l'interaction multimodale, et enfin l'interaction "outillée" ou instrumentée.

### 2.1 Réalité mixte

La réalité mixte est le paradigme d'interaction qui est l'objet de cette étude. Nous commençons par les aspects terminologiques, puis nous présentons quelques exemples concrets, pour ensuite étudier les espaces de caractérisation.

#### 2.1.1 Définitions

La réalité mixte vise la fusion harmonieuse du monde réel ou physique de l'utilisateur avec le monde virtuel ou numérique, et ses possibilités de stockage ou de traitement des données. À l'origine de la réalité mixte, on trouve le Digital Desk [33] : l'utilisateur travaille sur un bureau classique augmenté de nouvelles capacités grâce un ordinateur. Il dessine avec ses crayons habituels mais peut aussi effectuer des opérations telles que *copier/coller* ou *redimensionner*. Ces systèmes associent ainsi le monde réel et le monde informatique, pour affranchir les utilisateurs des barrières séparant les deux mondes. Au contraire, la réalité virtuelle tend à immerger l'utilisateur dans le monde numérique et à l'isoler de la réalité, s'opposant par là à la réalité mixte.

Le terme communément utilisé de "réalité augmentée" ne décrit que partiellement ces systèmes. De façon plus adaptée, les systèmes qu'elle comprend sont appelés "systèmes mixtes". D'autres termes désignant les systèmes mixtes ont déjà été identifiés [29] :

**Bit/Atome** désigne le couple objet numérique (composé de bits) et objet physique (composé d'atomes). Cette appellation est liée aux *interfaces tangibles*, où le bit informatique est rendu préhensible (*tangible bits*).

**Environnement augmenté** est un terme qui se rattache à l'informatique ubiquitaire ou pervasive. C'est l'environnement physique, c'est-à-dire le bâtiment, la pièce, etc. qui sont augmentés de capacités informatiques.

**Interaction augmentée** comprend les approches visant à faciliter les manipulations informatiques faites par l'utilisateur, grâce aux données extraites de l'environnement. Du point de vue de l'utilisateur, la tâche s'effectue dans le monde réel, par manipulation directe, non plus par métaphore de manipulation directe.

**Vidéo augmentée** désigne la fusion d'une représentation vidéo de la réalité avec des informations numériques.

**Virtualité augmentée et Réalité augmentée** Une distinction a été faite [13] entre la *Virtualité augmentée* — où la tâche appartient au monde numérique mais où l'interaction s'effectue en contexte physique — et la *Réalité augmentée* — où la tâche réside dans le monde physique mais son accomplissement est facilitée par un système informatique. Les systèmes mixtes forment un continuum entre ces deux extrémités.

Tous ces termes désignent un seul et même type de système informatique dont il est question ici : les systèmes mixtes. Afin d'étoffer notre compréhension de ces systèmes informatiques, nous en présentons quelques exemples au paragraphe suivant.

### 2.1.2 Exemples

Pour concrétiser la notion de système mixte et pour illustrer les démonstrations qui vont suivre, nous présentons ici quelques systèmes mixtes que nous jugeons représentatifs, au regard des taxonomies qui seront présentées par la suite.

#### CASPER

Le système CASPER [13] (Computer ASSisted PERicardial puncture) est un système de chirurgie assistée par ordinateur. La tâche du chirurgien (utilisateur de CASPER) consiste en une ponction péricardiale. Pour cette opération, le chirurgien fait une incision dans la poitrine juste assez grande pour introduire l'aiguille de ponction vers le fluide en excès près du cœur. Une trajectoire de ponction optimale est planifiée avant l'opération à partir d'images échographiques du cœur. En temps réel pendant l'opération, le chirurgien voit la position courante de l'aiguille par rapport à la trajectoire optimale planifiée. Un système de localisation (caméra et diodes attachées à l'aiguille) permet de connaître la position courante de l'aiguille, selon la partie émergée de l'aiguille. Plusieurs versions du système ont été développées : dans la première, les positions courantes et prévues sont affichées sur un écran comme le montre la figure 2.1 ; dans une autre version du système, la trajectoire prévue est affichée dans un casque à vision semi transparent. Enfin une troisième version de CASPER consiste à augmenter l'aiguille d'un bras nommé PADyC : PADyC n'agit pas de manière autonome avec l'aiguille sur le patient, mais permet au chirurgien de percevoir la trajectoire à suivre, grâce à un retour d'effort.

#### AmbientROOM

Les travaux présentés par Ishii et al. [17] [18] s'inscrivent dans une démarche appelée "tangible interface", que nous avons introduites au paragraphe 2.1.1. AmbientROOM est une pièce augmentée de la façon suivante :

- En tâche de fond, certaines données sont rendues observables par le système, comme
  - le taux d'activité d'une personne distante par des ondes affichées sur le plafond (métaphore des ondes à la surface de l'eau),



FIG. 2.1 – Le système CASPER avec affichage sur un écran d'ordinateur

- une donnée numérique (comme le nombre de courriers électroniques non lus) par une ambiance sonore (pluie, oiseaux, etc.),
- Au premier plan, il est possible pour l'utilisateur de contrôler et d'agir sur ces informations :
  - l'ouverture d'une bouteille en verre permet aussi d'observer le débit sur le réseau, sous la forme sonore (bruit de trafic routier),
  - l'horloge de la pièce permet de naviguer dans les différents états du système, en bougeant les aiguilles manuellement, au cas où l'utilisateur serait sorti de la pièce et souhaiterait à son retour observer un état passé.

### Phicons

Les Phicons [17] (de l'anglais physical icons, icônes tangibles) sont implémentés par plusieurs systèmes : les mediablocks [17] avec la technologie infrarouge et les Phicons [32] réalisés avec la technologie RFID [RFID]. Les Phicons sont des cubes en bois dans [17] et des objets fonctionnels dans [32]. Les applications sont diverses :

- Le transport de fichier : il suffit alors d'amener le Phicon sur une imprimante pour imprimer le fichier [17].
- Le montage vidéo : l'utilisateur associe des séquences d'un film à des Phicons différents (qu'elle/il peut annoter physiquement) et elle/il réalise le montage du film en ordonnant les cubes [17].
- L'augmentation d'une carte de visite (par exemple en l'équipant d'un tag RFID) : si l'utilisateur approche la carte de visite de son dispositif numérique personnel, alors la carte communique des informations comme l'adresse du propriétaire de la carte, sa page Web, etc. [32].
- L'extension des fonctionnalités de document : par exemple, instrumenter un dictionnaire de langue étrangère d'un tag RFID peut permettre de lancer une procédure de traduction automatique à l'approche du dictionnaire près d'un document [32].

### TROC

TROC [29] est un système mixte collaboratif et mobile développé à la suite du projet MAGIC [29]. Le but du jeu TROC est de collecter une liste d'objets ou vignettes, répartis sur un terrain, à l'aide de cubes.

Les vignettes sont des objets numériques augmentés, car ancrés dans le monde physique. Ils appartiennent à une famille d’animaux. Elles émettent un son spatialisé qui correspond à leur famille (un chat miaule), pour aider l’utilisateur à les trouver. Le cube est l’outil utilisé pour ramasser les vignettes. C’est un objet physique qui est augmenté. Pour s’en servir l’utilisateur doit le mettre devant la vignette et spécifier une commande, par exemple oralement. La figure 2.2 illustre cette utilisation du cube. Une fois la vignette collectée, celle-ci se trouve dans le cube. Les objets ne sont pas directement visibles : pour apercevoir la famille des chats, c’est-à-dire l’ensemble des vignettes “chats”, l’utilisateur doit posséder le filtre de vision correspondant. De plus l’utilisateur dépose les objets collectés et récupère des points de magie utiles pour le jeu en interagissant avec *sa* “borne”. C’est un emplacement du terrain de jeu où des actions sont déclenchées quand le joueur y entre. Il s’en approche et celle-ci le reconnaît afin d’effectuer la mise à jour automatiquement.

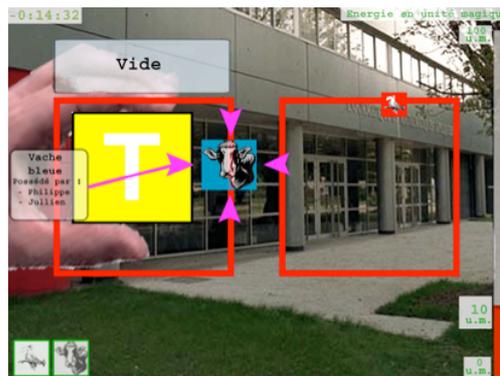


FIG. 2.2 – TROC

### Table Magique

La Table Magique [4] est un tableau blanc horizontal augmenté par des techniques de vision par ordinateur. Les utilisateurs utilisent le tableau blanc normalement, et écrivent donc au feutre sur la table. S’ils veulent utiliser les traitements informatiques pour ces écritures, ils les scannent pour pouvoir ensuite les manipuler. Cette manipulation s’effectue grâce à un ensemble de gestes prédéfinis, pour créer, transformer, et supprimer un motif ou entité. Les gestes sont effectués avec des jetons et sont reconnus par une caméra placée au dessus de la table.

1. Pour créer une entité, l’utilisateur désigne la zone à scanner sur la table. Pour ce faire, il prend deux jetons, les mets en contact l’un contre l’autre sur la table. Le système projette alors un rectangle élastique qui représente la selection entre les deux jetons, comme l’illustre la figure 2.3. Si l’utilisateur cache un des jetons avec sa main, la selection est annulée. En laissant les jetons immobiles deux secondes, le système scanne la zone voulue : dès que la capture de la zone commence, le rectangle élastique n’est plus affiché, et dès qu’elle est terminée, l’écriture numérisée est affichée par dessus son homologue physique.
2. Chaque entité créée par l’utilisateur peut subir trois transformations différentes : translation, rotation, et redimensionnement. Pour appliquer une transformation, il faut attacher un ou deux jeton(s) à une entité, en les posant au bord de l’entité. Quand un jeton est sur le bord, le système en informe l’utilisateur en affichant le contour de l’entité. Le point de contact

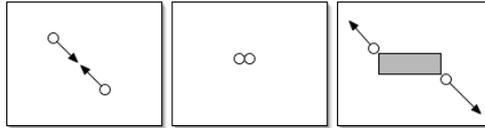


FIG. 2.3 – Définition d’une zone de sélection sur la table augmentée par manipulation de deux jetons

définit le point de contrôle de l’entité. Pour détacher un jeton, l’utilisateur doit le cacher. Un seul jeton de contrôle suffit pour translater une entité, et deux permettent de simultanément translater, faire tourner, et redimensionner une entité.

3. Supprimer une entité se fait en la redimensionnant jusqu’à une taille nulle.

Bien que faisant partie de la classe des systèmes mixtes, tous ces exemples présentés ont des différences majeures dans la façon dont on interagit avec eux. Il est donc nécessaire de les caractériser plus précisément que par l’appellation “système mixte”. Pour cela, de nombreuses caractérisations des systèmes mixtes ont été élaborées : nous les exposons dans le paragraphe suivant.

### 2.1.3 Caractérisation

En tant que représentants d’un paradigme d’interaction récent, ces systèmes mixtes font l’objet de multiples tentatives de caractérisation. Afin d’en avoir une meilleure description et donc une meilleure compréhension. Les approches présentées se complètent et permettent en définitive d’avoir une compréhension plus fine de la réalité mixte.

#### Cible de l’augmentation

Une première approche [22] que nous retenons distingue trois types d’augmentation, selon sa cible. Une augmentation peut viser

- l’utilisateur (qui portent alors des équipements),
- l’objet de la tâche (alors équipé de dispositif(s)),
- l’environnement.

Par exemple, lorsque l’utilisateur de CASPER porte un casque de vision semi transparent, le système augmente l’utilisateur. Lorsque le plafond de l’AmbientROOM est augmenté d’ondes rappelant celles à la surface de l’eau, pour rendre observable l’activité d’un individu distant, c’est l’environnement de l’utilisateur qui est augmenté.

Cette approche ne suffit pas pour caractériser finement les systèmes mixtes, car elle ne distingue pas les systèmes du point de vue de l’interaction. C’est pourquoi nous présentons au paragraphe suivant une étude qui focalise sur l’interaction.

#### Interaction avec l’augmentation

Une dimension pertinente pour la caractérisation des systèmes mixtes est la qualification de l’interaction du point de vue de l’utilisateur. Dans cette optique, Dubois a identifié deux axes orthogonaux de classification des systèmes mixtes [13] :

1. Réalité augmentée et Virtualité augmentée,
2. Exécution augmentée et Evaluation augmentée.

Le premier axe distingue les systèmes dont l'objet de la tâche de l'utilisateur est physique (cas de la réalité augmentée RA), à la différence de ceux dont l'objet de la tâche est numérique (cas de la virtualité augmentée VA). Par exemple, un système mixte concoure à une opération chirurgicale d'un *patient*, comme CASPER [13], est un système de RA, au contraire s'il permet de déplacer un *fichier*, c'est un système de VA.

Néanmoins, la distinction RA/VA n'est pas toujours facile à appliquer. Considérons par exemple la métaphore de bureau. La métaphore est une projection de concepts du monde physique dans le monde numérique (cf. 3.18). Dans le cadre de la métaphore du bureau, la tâche (alors numérique) s'effectue dans un système de compréhension qui relève de la réalité. Un système mixte tel que le Digital Desk peut être qualifié de VA si l'on considère que la métaphore a été augmentée, ou de RA si l'on fait abstraction de la métaphore du bureau et si l'on considère que ce sont les objets physiques (comme un ticket de métro) qui ont été augmentés. Cet axe est donc un continuum et la limite entre RA et VA peut être difficile à fixer.

Orthogonalement au premier critère de distinction (RA et VA), Dubois a aussi identifié deux sortes d'augmentation, conformément à la théorie de l'action de Norman : l'exécution augmentée et l'évaluation augmentée présentées à la figure 2.4. L'évaluation augmentée qualifie un système qui ajoute des éléments de l'autre monde pour aider la perception et l'interprétation de l'état interne du système. L'exécution augmentée, de façon analogue, qualifie les systèmes dont l'augmentation facilite l'exécution de la tâche.

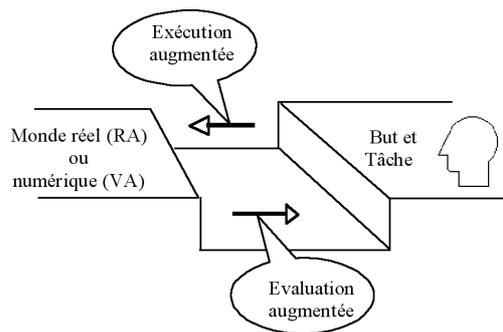


FIG. 2.4 – Espace de Dubois

Par exemple, cet axe de caractérisation des systèmes mixtes permet de qualifier les versions de CASPER. Dans chacune, le chirurgien effectue sa tâche (geste chirurgical), tout en étant aidé dans la correction de la trajectoire par le système. Aussi CASPER est un système de RA, puisque l'objet de la tâche est dans le monde physique. Dans la version 2 de CASPER avec un casque, le champ opératoire est augmenté de la trajectoire planifiée. L'évaluation est donc augmentée. De même dans la version 3 avec retour d'effort, l'aiguille de ponction est augmentée : l'évaluation est donc ici aussi augmentée.

Nous constatons qu'un système de RA est rarement augmenté par l'exécution, car il faudrait alors que l'accomplissement de la tâche, dans le monde physique, passe par le système, ce qui n'est possible que si le système contient un effecteur (imprimante, robot, etc.). L'intérêt d'un tel système ne vaut que si l'exécution par le système de la tâche a un avantage par rapport à l'exécution directe

par l'utilisateur, comme par exemple si l'utilisateur et l'objet de la tâche sont distants (comme les projets de télé-échographie).

Les axes RA/VA et exécution/évaluation augmentées permettent donc de classer les systèmes mixtes. Au delà de ces dimensions de classification des systèmes, de nouvelles propriétés d'utilisabilité propres aux systèmes mixtes ont été identifiées [13] pour qualifier l'interaction avec un objet augmenté. Celles-ci reposent sur une extension des propriétés d'observabilité et d'honnêteté [13] :

- La compatibilité perceptuelle, extension de la propriété d'observabilité au cas de  $N$  concepts, traduit la possibilité pour l'utilisateur de percevoir les données relatives à  $N$  concepts en prenant en compte leur dispersion géographique dans l'environnement et les sens perceptifs mis en jeu. La compatibilité perceptuelle implique donc l'observabilité de chacun des concepts considérés isolément.
- La compatibilité cognitive, extension de la propriété d'honnêteté au cas de  $N$  concepts, traduit la capacité de l'utilisateur à interpréter correctement les données relatives à  $N$  concepts en considérant leurs différents modes de représentation. Là encore, la compatibilité cognitive implique l'honnêteté des représentations de chacun des concepts considérés isolément.

Cas particulier de la perception et interprétation de plusieurs concepts, l'utilisateur est amené à percevoir plusieurs représentations relatives à un même concept. Ce cas particulier de la compatibilité a été appelé par le terme continuité. Comme précédemment, cette propriété s'applique aux niveaux perceptuel et cognitif, ce qui donne lieu à deux nouvelles propriétés :

- La continuité perceptuelle, adaptation de la compatibilité perceptuelle, traduit la possibilité pour l'utilisateur de percevoir les données relatives à *un* concept en prenant en compte leur dispersion géographique dans l'environnement de l'utilisateur et les sens perceptifs mis en jeu.
- La continuité cognitive, adaptation de la compatibilité cognitive, traduit la capacité de l'utilisateur à interpréter correctement les données relatives à *un* concept en considérant leurs différents modes de représentation.

Ainsi, dans la première version de CASPER où l'aiguille de ponction est représentée par rapport à une trajectoire pré-calculée sur l'écran du PC, nous concluons à une discontinuité perceptuelle puisque l'aiguille physique se trouve dans le champ opératoire et l'aiguille numérique affichée sur l'écran. Le concept concerné par cette discontinuité perceptuelle est l'aiguille "présente" dans deux lieux distincts (l'écran et le champ opératoire).

Toutes ces notions rapellées ici caractérisent la façon dont l'utilisateur interagit avec les entités augmentées d'un système mixte. Cette notion d'entité mixte ou d'objet augmenté est récurrente dans la littérature. Elle nécessite elle-aussi la caractérisation plus détaillée qui suit.

## **Augmentation d'un objet**

Au sein d'un système mixte, nous trouvons des objets dont la nature est à la fois physique et numérique, telle l'horloge dans l'AmbientROOM ou l'aiguille dans CASPER. Plusieurs travaux caractérisent l'objet augmenté par la quantité d'information qui transite pour définir l'augmentation. Cette caractérisation peut être appelée *perméabilité* dans [21] et "quantité d'information physique ou numérique ajoutée à l'autre monde" dans [13]. Nous présentons ces deux approches.

Dans [21], des propriétés de la frontière entre mondes physique et numérique sont identifiées. Il est important de noter que ces propriétés caractérisent les collecticiels mixtes, i.e. des systèmes permettant aux utilisateurs, distants et distribués sur différents espaces physiques et numériques,

de communiquer ou d'effectuer un travail en collaboration. Certaines propriétés sont donc propres au caractère collaboratif des systèmes mixtes étudiés. Néanmoins d'autres propriétés sont valides dans le cadre général des systèmes mixtes. Parmi les propriétés identifiées, nous trouvons celles liées à la *perméabilité*. Elle précise quelle information d'un monde peut passer dans l'autre à travers la frontière et caractérise aussi l'action de cette frontière sur les informations sensorielles (visuelles, auditives, tactiles, etc.) de l'utilisateur. La *visibilité* caractérise l'information visuelle qui peut passer à travers la frontière. À cette propriété sont liées des notions de résolution et de champ de vision. L'*audibilité* caractérise l'information auditive qui traverse. Pour ces deux propriétés, quatre effets sont identifiés : l'atténuation (réduction de la résolution visuelle ou du volume sonore), l'amplification, la transformation (distorsion visuelle ou sonore), l'agrégation (résumé de l'information et présentation dans un autre "format"). La *solidité* est la dernière propriété de perméabilité : elle caractérise la capacité d'une entité d'un des deux mondes à traverser la frontière. Une frontière solide, i.e. avec une *solidité* maximale, ne peut pas être traversée.

Dans [13], une notion similaire a été identifiée, à savoir la "quantité d'information physique ou numérique ajoutée à l'autre monde". La figure 2.5 schématise cette notion. Liée aux capacités interactionnelles du système mixte, comme le décrit la figure 2.6, l'information transite entre les deux mondes.

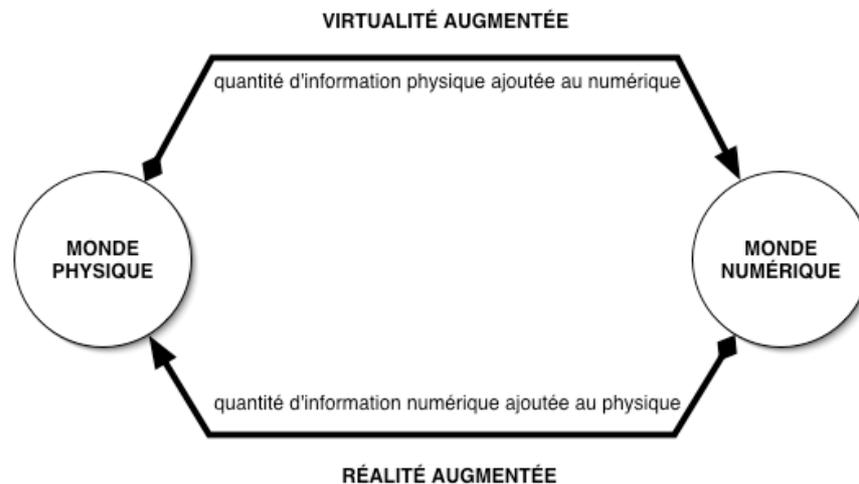


FIG. 2.5 – Quantité d'information transisant dans l'augmentation

Aucune des approches complémentaires pour caractériser les systèmes mixtes que nous avons exposées ne considère la dimension temporelle. Toutes concourent à une caractérisation statique du système mixte. Dans le paragraphe suivant, nous abordons donc les aspects dynamiques de l'augmentation.

### Dynamique de l'augmentation

Nous explorons ici la caractérisation de l'augmentation d'un des deux mondes physique ou numérique vis-à-vis de sa dimension temporelle. Cet aspect de l'augmentation concerne l'évolution de son état au travers du temps, sur intervention explicite de l'utilisateur ou non. Dans [21], une des propriétés identifiées est la *dynamique*. La *dynamique* comprend la durée de vie et l'instant de

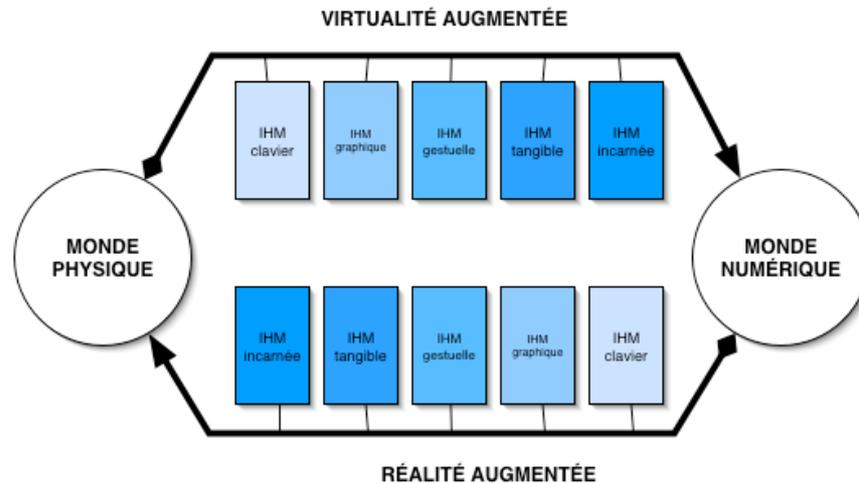


FIG. 2.6 – Deux continua de systèmes mixtes

création de l'augmentation. Une méta-propriété, la *configurabilité*, concerne également la dimension dynamique de l'augmentation, car elle caractérise la possibilité de changement dynamique des propriétés.

Cette notion de dynamique de l'augmentation est également prise en compte dans [29], par les termes de *temporalité* et *mode d'interaction*.

- La *temporalité* d'une augmentation définit la durée de vie de l'augmentation, et prend les valeurs *éphémère* ou *persistant*.
- Le *mode d'interaction* définit la façon dont l'utilisateur définit une entité augmentée. Ce mode d'interaction peut être passif, et l'augmentation est alors déterminée automatiquement sans intervention de l'utilisateur. À l'inverse, le mode d'interaction peut être actif et l'augmentation est alors spécifiée explicitement (intégralement ou partiellement) par l'utilisateur. Ces modes d'interaction peuvent être perçus comme un continuum allant de la passivité de l'utilisateur à son intervention explicite et intégrale. Cependant, il est difficile de mesurer le degré d'intervention de l'utilisateur, aussi ces modes d'interaction ne sont caractérisés que par deux valeurs : passif ou actif. Deux situations peuvent être caractérisées par un mode d'interaction :
  - la création de l'augmentation,
  - la modification de l'augmentation.

Par exemple, les cubes de TROC sont augmentés de façon persistante, tout comme les bouteilles de l'AmbientROOM ou l'aiguille du chirurgien dans CASPER. Néanmoins seul dans TROC, des objets augmentés sont définis et modifiables dynamiquement. Dans AmbientROOM et CASPER les augmentations sont figées à la conception. Dans le cas de TROC, lorsque le joueur explicitement ramasse une vignette numérique pour la ranger dans un cube, le mode d'interaction pour créer cette augmentation du cube physique est actif. À l'opposé lorsque le joueur revient à sa base pour y déposer es vignettes collectées, les cubes sont vidés et le mode d'interaction (modification des cubes augmentés) est passif, la dépose étant automatique dès l'entrés dans la zone.

Nous avons présenté jusqu'ici les systèmes mixtes et leur caractérisation, selon la cible de l'aug-

mentation, la façon dont on interagit avec eux, la manière d’augmenter un objet et enfin l’aspect temporel de l’augmentation. Dans notre étude dont l’objectif est un modèle d’interaction mixte, nous visons la caractérisation fine de l’interaction avec des systèmes mixtes. Dans ce contexte, outre l’existant sur les systèmes mixtes, il convient de nous tourner vers les modalités d’interaction et la multimodalité. Aussi le paragraphe suivant dresse un panorama des travaux sur les modalités d’interaction et la multimodalité.

## 2.2 Modalité d’interaction et multimodalité

Avant d’aborder la multimodalité, nous rapellons d’abord la définition d’une modalité d’interaction, puis les différentes approches pour caractériser les modalités existantes. Nous pouvons ensuite aborder la composition des modalités d’interaction.

### 2.2.1 Définition d’une modalité $m$

Une *modalité d’interaction*  $m$  a été définie [24] formellement en notant :

$d$  le *dispositif* physique d’interaction, en entrée ou en sortie, c’est-à-dire l’objet qui permet l’acquisition ou le rendu de l’information par le système,

$l$  le *langage* d’interaction, un ensemble d’expressions bien formées, composées de symboles, et qui ont un sens.

Une modalité  $m$  se définit de la façon suivante :

$$m \longrightarrow (d, l) | (m, l)$$

Par exemple, (*clavier, langage pseudo – naturel*) est une modalité élémentaire.

### 2.2.2 Caractérisation d’une modalité

Une modalité d’interaction étant définie par le couple  $(d, l)$ , caractériser une modalité revient à caractériser les éléments qui la composent, à savoir son dispositif et son langage.

#### Caractérisation du dispositif

La plupart des taxonomies des dispositifs d’interaction sont liées à la technologie et aux aspects matériels de l’informatique, et sont donc sujettes à ses fluctuations. Nous notons néanmoins une synthèse de ces taxonomies dans [25] dont nous retenons ici les caractéristiques suivantes :

- sens humain (comme le toucher) ou plus précisément la grandeur captée (comme la pression, la température),
- spatial (lieu d’interaction comme l’écran),
- temporel (éphémère ou non)
- sens (capteur ou effecteur, i.e. entrée ou sortie) et persistance,
- débit d’information,
- nombre de dimensions (par exemple deux dimensions captées par la souris)

## Caractérisation du langage

Pour caractériser le langage d'une modalité d'interaction, nous retenons le travail de Bernsen [5]. Il identifie un ensemble de propriétés s'appliquant au langage  $l$  d'une modalité  $(d, l)$ , et qui peuvent être de nature :

- statique ou dynamique,
- linguistique ou non-linguistique.
- analogique ou non-analogique,
- arbitraire ou non-arbitraire.

Un langage d'interaction fait partie d'une des 16 classes définies par cette caractérisation des langages. Ces propriétés du langage  $l$  de la modalité peuvent être plus amplement détaillées :

Une modalité linguistique désigne un langage, c'est-à-dire un système structuré de signes remplissant une fonction de communication. Par exemple, la langue parlée et écrite, le langage gestuel, sont des modalités représentationnelles linguistiques : tous s'appuient sur l'acquis d'un lexique et d'une syntaxe et tous, pour remplir leur fonction de communication, recouvrent une sémantique en situation (ou pragmatique).

Le caractère dynamique ou statique d'une modalité repose sur la présence de la dimension temporelle dans la représentation. Par exemple, la modalité langue naturelle écrite est linguistique, non analogique, non arbitraire, et *statique*. Le langage parlé, quant à lui, présente le même profil que la langue naturelle écrite à l'exception de la dimension temporelle : il est *dynamique*.

Une modalité analogique entretient un rapport de ressemblance avec la réalité : elle fonctionne par ressemblance à l'élément signifiée. Par exemple, le dessin d'une maison est une représentation analogique par rapport à la maison qu'il représente. On parle aussi de représentation iconique ou isomorphique. Il convient de noter que les icônes que l'on conçoit en IHM graphique ne sont pas toujours des représentations analogiques. Une icône est qualifiée d'abstraite lorsque le dessin qui la constitue n'est pas analogique mais contient un indice sémantique (par exemple, une flèche pour désigner une direction) ; elle est arbitraire lorsqu'elle ne présente aucun indice sémantique sur le signifié. Une icône qui opère par analogie est nommée icône représentationnelle. Les icônes semi-abstraites résultent de la composition d'éléments abstraits et représentationnels

Une modalité arbitraire fonctionne en dehors d'un système conventionnel. A l'inverse, les modalités non arbitraires, tels les langages, reposent, pour remplir leur rôle, sur l'existence d'un système sémantique appris. Il convient cependant de noter qu'à un niveau fin, certains termes ne sont pas arbitraires et revêtent en plus un caractère analogique.

Dans la cadre d'une étude de la multimodalité en sortie de grandes quantités d'information [30], des caractéristiques additionnelles du langage d'interaction ont été proposées :

- partiel ou global,
- précis ou vague,
- présence ou absence d'une fonction de déformation.

Par exemple, pour représenter un grand espace d'information à l'écran (comme un grand graphe), le focus d'attention de l'utilisateur doit être observable via une modalité *précise* afin que l'utilisateur puisse effectuer sa tâche, et le reste peut se contenter de rester *vague*. Le système Vitesse [30] permet d'illustrer les modalités en sortie dans le cadre d'une interface graphique. Par exemple à la figure 2.7, deux modalités sont utilisées pour représenter un grand espace d'information : l'une précise et déformante présente le focus de l'utilisateur tandis qu'une autre (vue à plat de l'espace) présente le reste de l'espace de façon non déformée ou à plat, et vague.

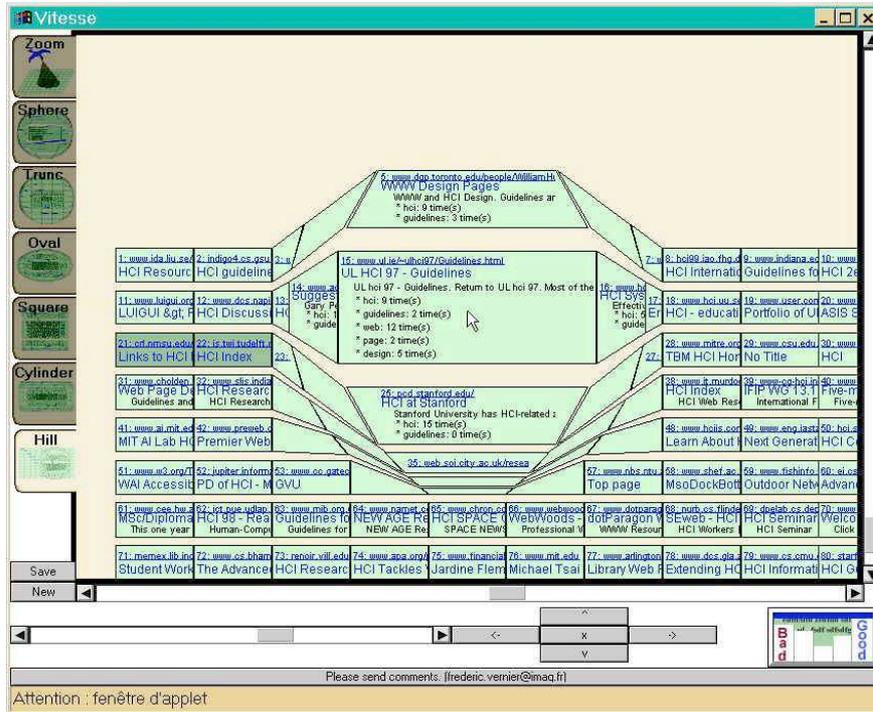


FIG. 2.7 – Modalités de sortie dans Vitesse

Nous venons de caractériser une modalité d’interaction en considérant des propriétés de son dispositif et de son langage. Nous étudions maintenant la composition de modalités pour définir des modalités dites composées, à l’opposé des celles qualifiées de ”pures” [5] ou d’élémentaires.

### 2.2.3 Composition de modalités

Une modalité ( $d, l$ ) est caractérisée, dans le cas d’une modalité d’interaction élémentaire, par son dispositif et son langage d’interaction. C’est le cas que nous venons d’étudier. Nous étudions leurs compositions pour définir des modalités dites composées, qui définissent une interface multimodale. L’interaction multimodale constitue un axe de recherche important de l’Interaction Homme-Machine, et nous ne présentons ici que les résultats que nous avons utilisés dans notre étude.

#### Propriétés CARE et CARE étendues

Les quatre propriétés CARE de compositions des modalités — la *Complémentarité*, l’*Assignation*, la *Redondance*, ou l’*Équivalence* de deux modalités — définissent l’aspect sémantique de la composition. La complémentarité de deux modalités implique que les deux sont essentielles pour que l’expression véhiculée ait un sens, comme le “Met-ça là” de Richard Bolt. L’assignation d’une modalité à une tâche implique que seule cette modalité doit être utilisée pour effectuer cette tâche ; en d’autres termes il n’y a pas de choix possible. La redondance dans l’usage de deux modalités signifie que les deux modalités utilisées expriment le même contenu sémantique, comme dire “Grenoble” et cliquer sur Grenoble sur une carte affichée à l’écran. L’équivalence de deux modalités signifie que

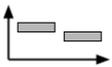
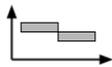
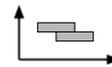
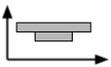
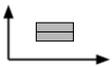
l'une ou l'autre peut être indifféremment utilisée pour effectuer une même tâche. Nous notons que la redondance implique l'équivalence.

CARE ne considère que l'aspect sémantique dans la composition. Adoptant une approche plus globale dans [30] nous trouvons d'autres aspects dans la composition :

- temporel,
- spatial,
- articulatoire,
- syntaxique,
- sémantique.

Les compositions spatiales reflètent la volonté du concepteur de placer les modalités à des endroits spécifiques de l'environnement de l'utilisateur. La composition temporelle est également considérée : elle caractérise les enchaînements possibles de modalités. La composition articulatoire des modalités dénote une composition à un niveau d'abstraction plus bas, entre les flux d'information transitant par les dispositifs associés aux modalités. La composition syntaxique considère les relations entre les langages d'interaction utilisés. Enfin, la composition sémantique concerne la composition des informations véhiculées par les modalités (propriétés CARE exposées ci-dessus).

Pour un aspect, il y a quatre schémas différents de composition. Le tableau 2.1 récapitule les cas identifiés en considérant les quatres aspects ci-dessus et les quatres schémas de composition.

	<i>schéma</i>				
<i>aspect</i>					
<b>temporel</b>	anachronisme	séquence	concomittance	coïncidence	parallélisme
<b>spatial</b>	disjonction	adjacence	intersection	imbriquation	recouvrement
<b>articulatoire</b>	indépendance	fission	fission et duplication	duplication partielle	duplication
<b>syntaxique</b>	différence	complétion	divergence	extension	jemellité
<b>sémantique</b>	concurrence	complémentarié	complémentarité et redondance	redondance partielle	redondance totale

TAB. 2.1 – Compositions de modalités (CARE étendues)

Le système Vitesse [30] permet d'illustrer les compositions de modalités en sortie dans le cadre d'une interface graphique. Par exemple à la figure 2.7, deux modalités sont utilisées pour représenter un grand espace d'information : l'une déformante présente le focus de l'utilisateur tandis qu'une autre (vue à plat de l'espace) présente le reste de l'espace de façon non déformée ou à plat. Si l'on considère la composition de ces deux modalités nous pouvons la caractériser comme suit :

- temporellement parallèle,
- spatialement adjacente,
- articulatoirement dupliquée,
- syntaxiquement étendue,
- sémantiquement complémentaire.

## Niveaux de composition

En relation avec les 5 aspects de composition qui viennent d'être exposés, plusieurs niveaux de compositions de modalités ont été identifiés dans [11] pour le cas des modalités d'entrée. Ces compositions concernent les relations entre langages et tâches, ou, à un autre niveau, les relations entre dispositifs et langages. Plusieurs langages sont composés pour une unique tâche élémentaire<sup>1</sup> t, selon un schéma de composition pour chaque aspect (cf. 2.1). De même, plusieurs dispositifs sont composés pour un unique langage, selon un schéma de composition pour chaque aspect. Les deux niveaux de compositions sont schématisés à la figure 2.8.

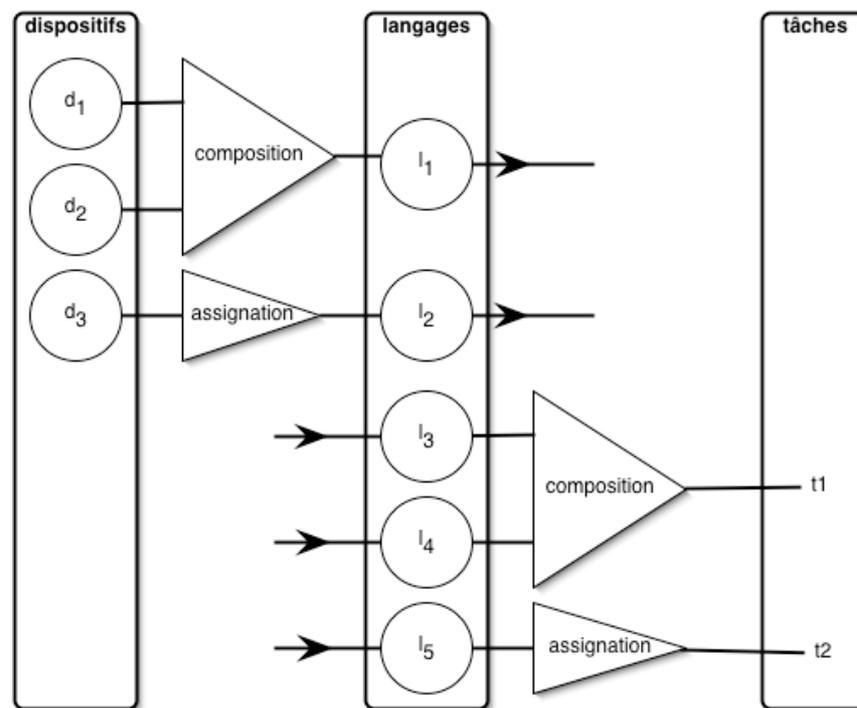


FIG. 2.8 – Niveaux de compositions de modalités

Dans ce chapitre, nous avons d'abord étudié les systèmes mixtes puis affiné l'interaction en considérant les modalités d'interaction et leurs compositions. Nous étudions enfin des approches récentes qui visent à distinguer au sein de l'interaction l'instrument ou l'outil manipulé par l'utilisateur de l'objet lié à la tâche.

### 2.3 Objet de la tâche et Outil de la tâche

Après avoir longtemps considéré principalement l'utilisateur et sa tâche, l'interaction homme-machine a intégré dans certains modèles un nouvel objet qui participe à l'interaction avec le système informatique : l'outil. Comme dans le monde physique, l'interaction avec l'objet qui est le focus de

<sup>1</sup>Une tâche élémentaire ne peut être décomposée que sous forme d'actions physiques.

la tâche peut être médiée par un outil, de la même façon qu'on plante un clou avec un marteau. Les différents modèles basé sur cette distinction sont expliqués maintenant.

### 2.3.1 Outil et Réalité Mixte

#### ASUR

Plusieurs approches ont distingué l'objet de la tâche de l'outil utilisé dans la réalisation de la tâche. La première présentée ici est ASUR (Adaptateur, Système, Utilisateur, entités Réelles), une notation de conception qui est détaillée dans [13] et étendue dans [14]. Cette notation est dédiée aux systèmes de réalité mixte. L'interaction au sein d'un système mixte est modélisé selon ASUR par quatre composants qui sont mis en relation :

#### L'utilisateur $U$

Le système  $S$  comprend :

- les éléments du système utiles à l'interaction en entrée  $S_{tool}$ ,
- les éléments du système utiles à la perception de l'état interne du système,  $S_{representation}$ .  
Ils sont eux-mêmes décomposés en :
  - l'objet de la tâche  $S_{object}$ ,
  - les attributs de l'objet de la tâche  $S_{info}$ .

Les objets réels  $R$  comprennent :

- les outils  $R_{tool}$ ,
- les objets de la tâche  $R_{task}$ .

Les adaptateurs  $A$  transfèrent les données d'un monde à l'autre, et sont de deux types :

- les adaptateurs en entrée du système  $A_{in}$ ,
- les adaptateurs en sortie  $A_{out}$ .

Ces composants sont ensuite mis en relation. Trois types de relation sont possibles :

1. l'échange de données entre 2 composants, symbolisé par une flèche  $\longrightarrow$ ,
2. la proximité physique, symbolisé par un double trait  $=$ ,
3. le déclenchement, symbolisé par  $\implies$ . Le déclenchement est toujours lié à un échange de données, et traduit que l'échange de données n'est possible que si la condition spatial du déclenchement est satisfaite.

Prenons l'exemple de CASPER pour illustrer ASUR. Pendant l'opération, le chirurgien (composant  $U$ ) déplace l'aiguille (composant  $R_{tool}$ ) et peut l'observer, ce qui est décrit par la double flèche  $U \longleftrightarrow R_{tool}$ . La position de l'aiguille est repérée dans l'espace par le localisateur (composant  $A_{in}$ ) :  $R_{tool} \longrightarrow A_{in}$  ; qui transmet la position de l'outil au système informatique :  $A_{in} \longrightarrow S$ . Le système informatique traite cette donnée et en définit la forme graphique en vue de l'affichage à l'écran (composant  $A_{out}$ ) :  $S \longrightarrow A_{out}$ . L'affichage peut alors être perçu par l'utilisateur :  $A_{out} \longrightarrow U$ . Enfin, en permanence, le chirurgien peut percevoir le patient :  $R_{task} \longrightarrow U$  ; et l'aiguille de ponction est au contact du patient, ce qui est représenté par la relation mécanique  $R_{tool} - R_{task}$ . La figure 2.9 illustre cette modélisation de CASPER.

La notation ASUR repose donc sur la distinction entre objet et outil de la tâche.

ASUR possède également l'avantage d'avoir été étendu aux systèmes mobiles, pour devenir un outil de conception pour les systèmes mixtes mobiles. L'approche IRVO décrite dans le paragraphe suivant repose aussi sur la distinction entre objet et outil de la tâche.

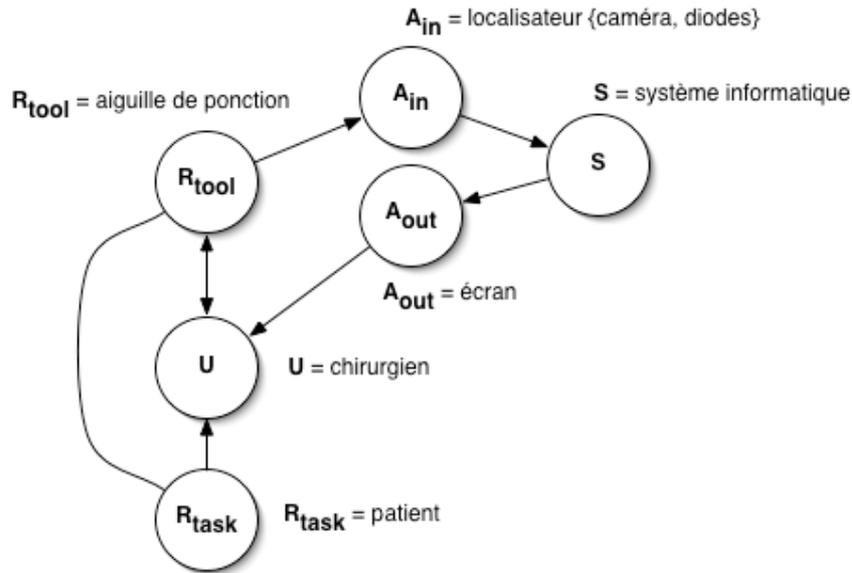


FIG. 2.9 – CASPER version écran, décrite selon la notation ASUR

## IRVO

Le modèle IRVO (Interacting with Real and Virtual Objects) est aussi dédié à la réalité mixte, et explicite aussi la notion d’outil dans l’interaction. IRVO comprend plusieurs catégories d’entités, illustrées à la figure 2.10 :

**L’utilisateur  $U$** , représenté avec ses canaux sensoriels d’entrée/sortie ( $KH$  pour le toucher, le geste ;  $V$  pour la vue ;  $A$  pour l’ouïe et la voix) ;

**Les objets** sont de deux types :

- l’objet de la tâche  $O$ , qui peut être purement physique  $O_r$ , ou physique et augmenté  $O_{r+v}$  (cas de la RA), purement numérique  $O_v$ , ou numérique et augmenté  $O_{v+r}$  (cas de la VA),
- les outils  $T$ , qui peuvent se décliner comme les objets de la tâche en  $T_r, T_{r+v}, T_v, T_{v+r}$ .

**Les transducteurs** sont de type “senseur”  $S$  ou “effecteur”  $E$ .

**Le “modèle interne de l’application”  $M$** , c’est-à-dire le système informatique sans les objets, qui peuvent être mixtes ou numériques. Il n’est pas forcément présent dans un schéma IRVO, puisqu’il n’interagit pas directement avec l’utilisateur.

Dans l’exemple de CASPER de la figure 2.10, l’objet mixte de la tâche  $O$  est formé d’un objet physique  $O_r$ , le patient, incluant le liquide à ponctionner, et un objet numérique qui l’augmente, à savoir la trajectoire idéale. L’opérateur  $\oplus$  permet de schématiser la fusion perceptuelle : l’utilisateur perçoit l’objet mixte comme un tout. Les flèches représentent un échange de données entre deux entités, et celles en pointillés signifient que cet échange est secondaire vis-à-vis de la tâche.

La notation IRVO est proche de celle d’ASUR. Nous retrouvons les mêmes entités de base Utilisateur, objet de la tâche, outil de la tâche, adaptateur (ou transducteur en IRVO).

Tandis qu’IRVO associe les canaux sensoriels humains au niveau du composant Utilisateur, ASUR les décrit au niveau des échanges (flèches) qui sont liées à l’utilisateur.

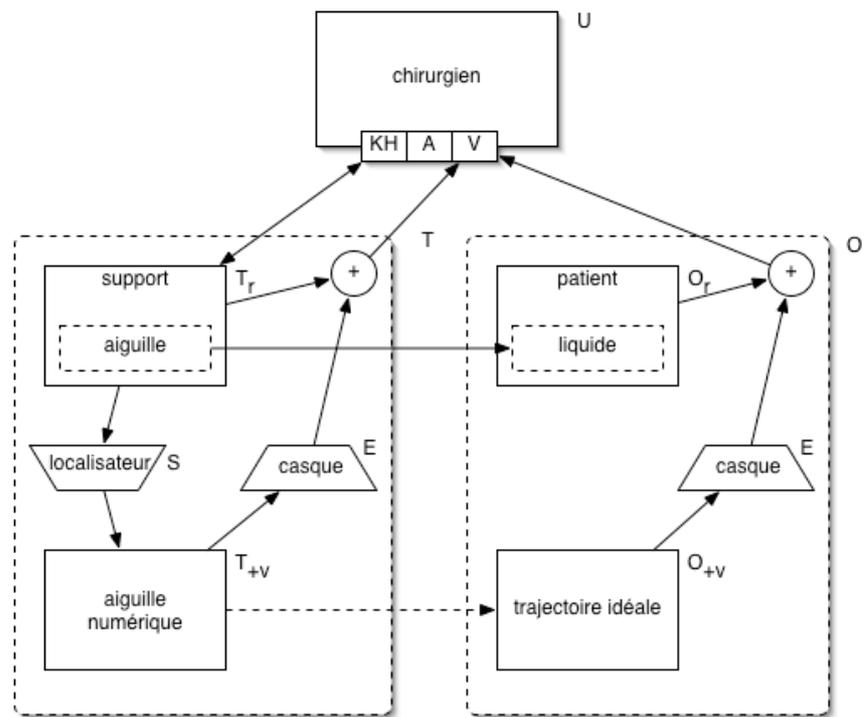


FIG. 2.10 – CASPER, version casque, décrit selon la notation IRVO

Pour notre étude, nous retenons de ces deux notations, la distinction importante entre objet et outil de la tâche, ces derniers pouvant être physiques, numériques ou mixtes. Nous étudions maintenant le modèle instrumental qui repose sur cette distinction outil ou instrument et objet de la tâche.

### 2.3.2 Approche instrumentale des interfaces graphiques

Nous retrouvons cette distinction entre objet et outil (ou instrument) dans le modèle d'interaction instrumentale [1][2], pour décrire les techniques d'interaction graphiques, du point de vue de l'utilisateur, dans le cadre du paradigme de l'ordinateur-outil [2]. Ce modèle est présenté à la figure 2.11.

L'utilisateur agit sur l'instrument ou outil, et ce dernier transforme l'action en *commande* ciblée sur l'objet d'intérêt du domaine ou objet de la tâche. La *réaction* de l'instrument permet à l'utilisateur de contrôler son action sur l'instrument. L'objet du domaine fournit une *réponse* vers l'instrument et un *retour d'information* vers l'utilisateur. Les deux types de réaction participent à l'observabilité du système et permettent à l'utilisateur d'évaluer si ses actions ont bien eu l'effet escompté.

#### Instrument

L'instrument lui-même se compose d'une partie *logique* (par exemple, une barre de défilement) et d'une partie *physique* (par exemple la souris). Ce point de vue sur l'instrument est à rapprocher de la définition d'une modalité comme un couple  $(d, l)$  (paragraphe 2.2.1) et de la définition de la réalité mixte (dispositif physique et traitement informatique) (paragraphe 2.1.1). L'*activation* d'un instrument se produit quand les parties logique et physique ont été associées. Cette association peut être spatiale (en amenant la partie physique dans l'espace d'une partie logique) ou temporelle (une action précédente a activé l'instrument).

#### Caractérisation

Les propriétés de ces instruments permettent de les comparer. Elles sont au nombre de trois : le *degré d'indirection*, le *degré d'intégration*, et le *degré de compatibilité*.

1. Le *degré d'indirection* se définit comme suit :
  - (a) Le *décalage spatial* est la distance entre l'instrument logique et l'objet du domaine ciblé.
  - (b) Le *décalage temporel* est la différence temporelle entre l'action de l'utilisateur et la réponse de l'objet ciblé.
2. Le *degré d'intégration* mesure le rapport entre les degrés de liberté de l'instrument physique et ceux que permet l'instrument logique.
3. Le *degré de compatibilité* mesure la similitude entre l'action de l'utilisateur sur l'instrument physique et la réponse de l'objet ciblé.

Reprenons l'exemple issu de [1]. Dans cet exemple, l'utilisateur fait défiler le contenu d'un document. Comme l'illustre le schéma de la figure 2.12, l'utilisateur agit sur l'instrument physique, c'est-à-dire la souris, de façon à ce que sa position à l'écran coïncide avec la flèche de la barre de défilement. Ceci active l'instrument logique, c'est-à-dire la barre de défilement, qui applique la commande à l'objet du domaine, c'est-à-dire le document. Le retour d'information consiste à mettre

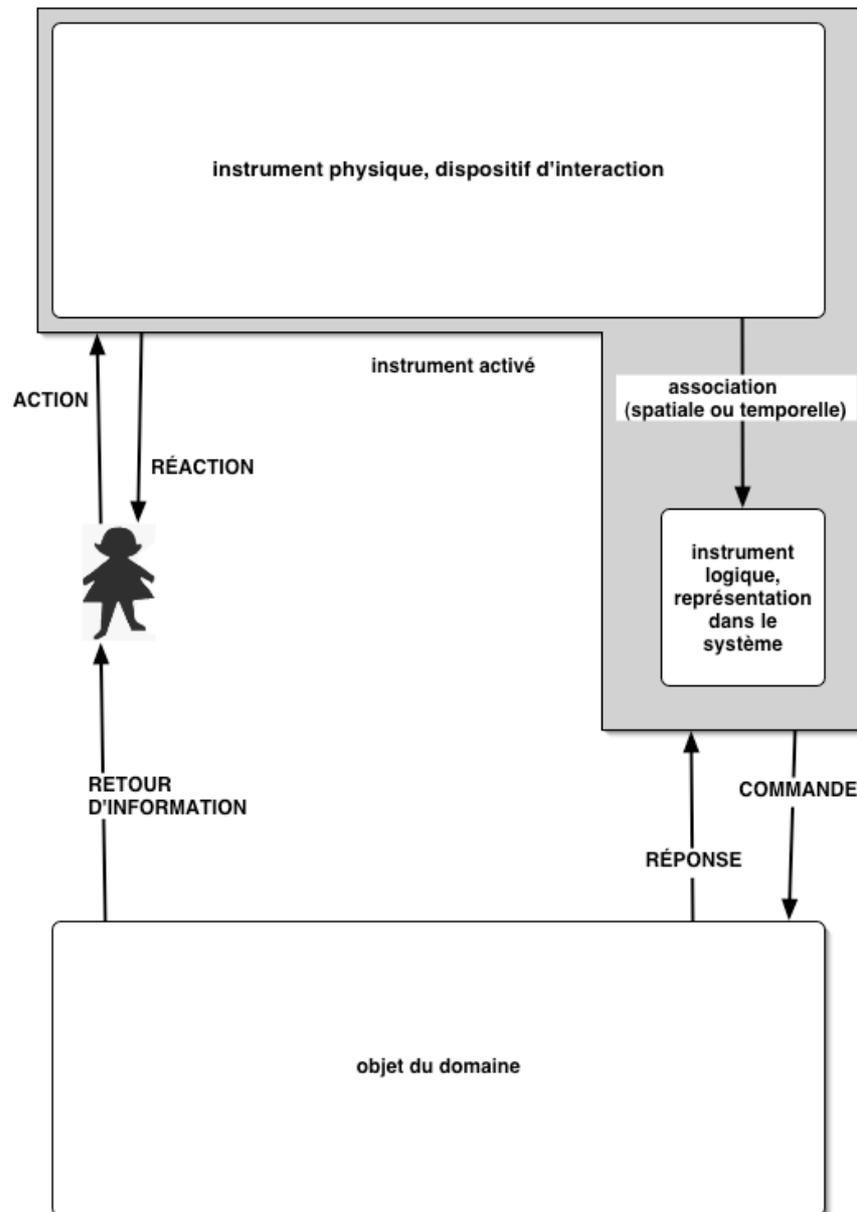


FIG. 2.11 – Modèle instrumental de l'interaction

à jour la vue sur le document, et la réponse du document permet de mettre à jour l'ascenseur de la barre de défilement.

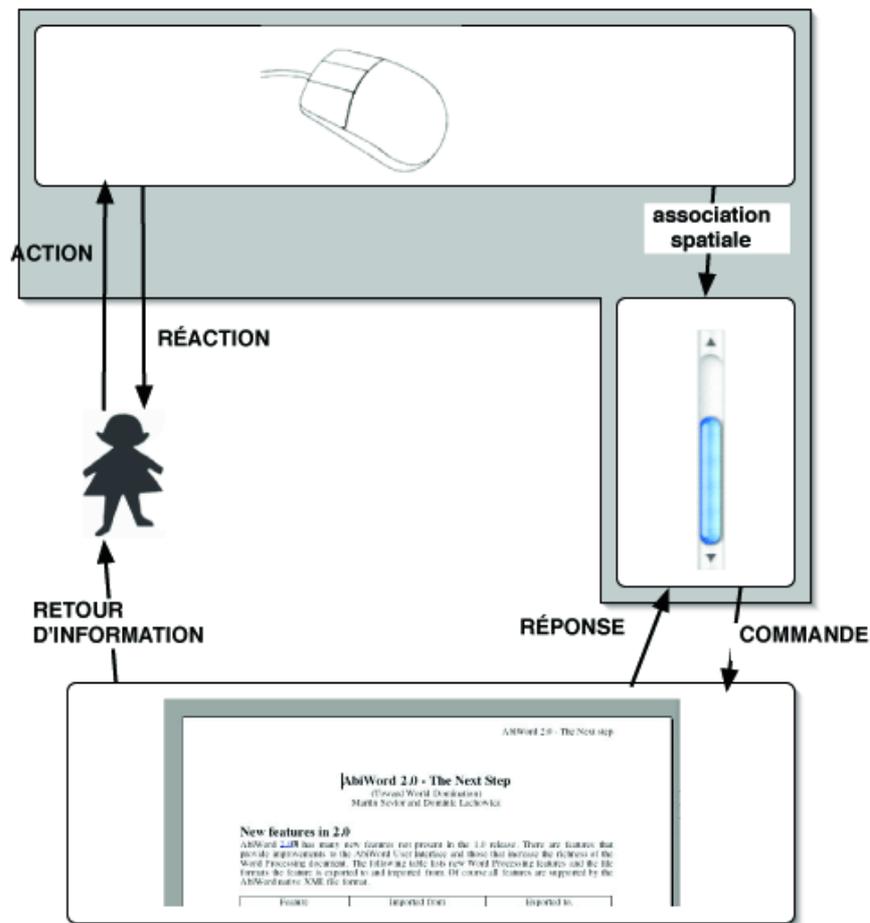


FIG. 2.12 – Modèle instrumental de la barre de défilement

Nous caractérisons alors l'instrument manipulé composé de la souris et de la barre de défilement comme suit :

- Le décalage spatial de la barre de défilement est nul, et son décalage temporel est nul. Néanmoins certaines barres de défilement attendent que le bouton de la flèche soit relâché. Dans ce cas, le décalage temporel est grand et par conséquent, le degré d'indirection est aussi plus grand. En conclusion, le degré d'indirection est faible.
- Comme la souris capture deux dimensions et la barre de défilement n'en a qu'une, le degré d'intégration de l'outil est égal à  $1/2$ .
- Le degré de compatibilité mesure la similarité entre l'action de l'utilisateur sur l'instrument et l'action de l'instrument sur l'objet du domaine. Comme l'utilisateur bouge l'ascenseur vers le bas pour déplacer le document vers le haut, ce degré est faible.

## 2.4 Synthèse et conclusion du chapitre 2

Ce chapitre expose les approches et modèles qui servent de fondements à notre étude. Nous avons d'abord étudié les systèmes mixtes pour mieux cerner leurs objectifs. Nous avons ensuite présenté des travaux dédiés aux modalités d'interaction et à la multimodalité afin d'affiner l'interaction. Nous avons enfin focalisé sur les approches récentes qui reposent sur la distinction entre outil et objet de la tâche.

Tous ces travaux ont permis des avancées notables, mais ces axes d'étude, bien que complémentaires, ne sont pas encore rejoints. Il est donc nécessaire maintenant de proposer une unification des modèles proposés. En exploitant chacun de ces modèles et leurs contributions, nous visons la définition d'un outil de conception des systèmes mixtes. Dans cette optique, nous allons dans le chapitre suivant présenter notre modèle de l'interaction entre l'utilisateur et les systèmes mixtes.

## Chapitre 3

# Modèle d'Interaction Mixte : la réalité mixte du point de vue des modalités d'interaction et de la multimodalité

Un modèle d'interaction a pour but de fournir un cadre de travail utilisable aux concepteurs de systèmes informatiques interactifs. L'utilisation d'un modèle de l'interaction doit avoir trois intérêts [2] pour le concepteur : support à la description détaillée du système, support à l'évaluation prédictive, et enfin support à la génération de nouvelles techniques d'interaction. Le modèle de l'interaction avec les systèmes mixtes que nous définissons dans ce chapitre prend ses racines, à différents niveaux de définition du modèle, dans les travaux qui concernent la réalité mixte, la multimodalité et l'interaction instrumentale. Après avoir présenté ces approches fondatrices dans le chapitre précédent, nous introduirons, de façon incrémentale dans ce chapitre notre modèle, en définissant d'abord un objet mixte puis une modalité d'interaction mixte.

### 3.1 Définition

Notre modèle d'interaction repose sur la notion d'objet mixte que nous définissons au paragraphe suivant. Un objet mixte étant défini, nous modélisons ensuite une technique d'interaction mixte.

#### 3.1.1 Objet mixte

Nous considérons trois types d'objets qui prennent part à l'interaction : les objets purement physiques, les objets purement numériques, et les objets mixtes. Un objet physique se caractérise par un ensemble de propriétés. Ces propriétés sont perceptibles par les cinq sens humains. Un objet numérique, de façon similaire, contient un ensemble de propriétés, qui peuvent être stockées et traitées en machine. Un objet mixte est un objet qui appartient aux deux mondes : il est composé d'une partie physique et d'une partie numérique comme le schématise la figure 3.1. Nous reprenons la définition de [9] : un objet mixte est défini par le lien entre un objet physique et un objet numérique. Notre modèle vise à définir ce lien qui unit ces deux parties en une entité mixte.

La partie physique de l'objet mixte comprend un ensemble de propriétés, appelées *propriétés physiques* de l'objet mixte. De même, la partie numérique de l'objet mixte comprend un ensemble

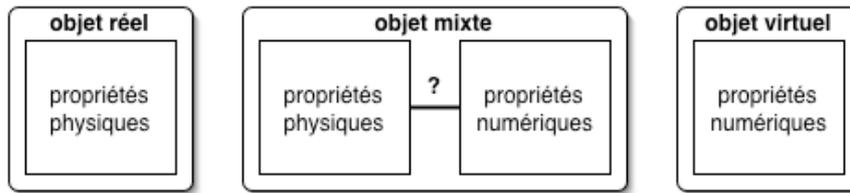


FIG. 3.1 – Objet physique, objet mixte et objet numérique

de propriétés, appelées *propriétés numériques* de l'objet mixte. En considérant la définition d'une modalité d'interaction comme un couple  $(d, l)$ , l'association entre ces deux ensembles de propriétés peut ensuite se définir de la façon suivante :

- En entrée du système, une modalité  $(d_o^e, l_o^e)$  permet de :
  1. capter certaines données parmi les propriétés physiques, à l'aide d'un dispositif  $d_o^e$ , dispositif en entrée de l'objet,
  2. puis traduire ces données physiques captées en propriétés numériques, grâce à un langage  $l_o^e$ .
- En sortie du système, une autre modalité permet de :
  1. générer certaines données à partir des propriétés numériques, à l'aide d'un langage  $l_o^s$ ,
  2. rendre ces données physiques générées perceptibles par des propriétés physiques, grâce à un dispositif  $d_o^s$ .

Le lien entre les propriétés physiques et numériques d'un objet mixte peut donc se définir par deux modalités que nous qualifions de modalités de liaison pour les opposer aux modalités d'interaction : une modalité de liaison en entrée du système et une modalité de liaison en sortie. La figure 3.2 présente notre définition de l'objet mixte. Dans ce schéma, nous avons choisi de représenter l'information par des rectangles, et les cercles entre ces rectangles représentent une transformation de cette information. Traditionnellement, quand la définition de modalité  $(d, l)$  est utilisée, l'information, qui correspond aux rectangles ici, n'est pas explicitée. Nous l'avons fait par soucis de clarté dans les figures 3.2 et suivantes.

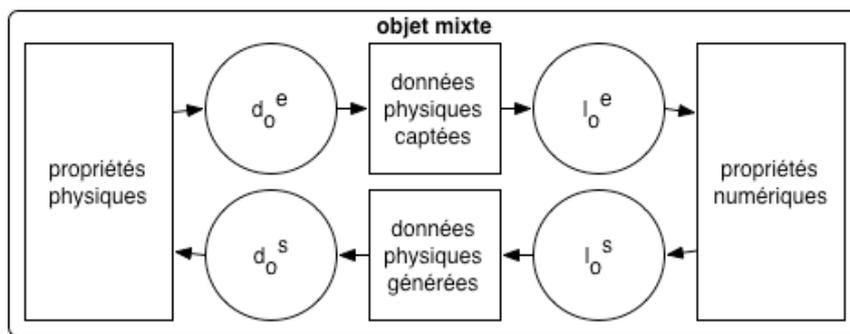


FIG. 3.2 – Objet Mixte

Pour illustrer notre définition d'un objet mixte, considérons les versions du système CASPER.

Dans la version de CASPER où les informations de guidage sont visuelles affichées à l'écran du PC, l'aiguille est un objet mixte. Parmi ses propriétés physiques, nous trouvons sa position dans l'espace en trois dimensions. Ses propriétés numériques sont la position dans le référentiel. Entre ces deux ensembles de propriétés, nous définissons deux modalités. En entrée du système, le dispositif  $d_o^e$  est l'ensemble caméras et diodes, les données physiques captées sont numériques. Le langage  $l_o^e$  transforme les données numériques en leur donnant un sens relatif au référentiel utilisé. En sortie du système, dans la version de CASPER avec le bras PADyC attaché à l'aiguille, la position de l'aiguille mixte par rapport à la trajectoire pré-calculée (propriétés numériques) est interprétée par le langage  $l_o^s$  pour générer un ordre de commande pour le bras PadyC, le bras  $d_o^s$  rendant tangible la trajectoire.

### **Liaison multimodale**

On applique exactement la définition des compositions de modalités d'interaction défini au paragraphe 2.2.3 au cas des modalités de liaison, puisque les deux types de modalités utilisent la même définition  $(d, l)$ . Un objet mixte peut avoir deux modalités de liaison en entrée, et les deux modalités seront composées, selon les 5 aspects et les 5 schémas présentés, avant de former les propriétés numériques de l'objet mixte. De la même façon, un objet mixte peut avoir plusieurs modalités de liaison en sortie, et ces modalités de liaison seront composées avant de participer aux propriétés physiques de l'objet mixte. Dans la modélisation d'un objet mixte, on représente la composition de deux modalités par un triangle qui permet de joindre les flèches issues de deux modalités.

### **Niveaux d'abstraction**

Il convient de remarquer que la modélisation d'un objet mixte autorise plusieurs niveaux d'abstraction. Le choix du niveau d'abstraction est laissé au concepteur. D'après la définition d'une modalité introduite dans le paragraphe 2.2, une modalité de liaison  $m$  est définie de la façon suivante :

$$m \longrightarrow (d, l)|(m, l)$$

En d'autres termes, une modalité peut être la modalité considéré ou le dispositif d'une autre modalité au niveau d'abstraction supérieur. Comme un objet mixte est défini par le lien qui unit ses parties physique et numériques, et que ce lien vient d'être défini comme une modalité en entrée et une modalité de sortie, nous pouvons alors reformuler  $m \longrightarrow (d, l)|(m, l)$  ainsi :

Un objet mixte peut être l'objet mixte considéré, ou le dispositif d'un autre objet mixte.

De cette façon, une souris peut entrer dans la modélisation d'un objet de deux façons différentes : soit elle est elle-même l'objet mixte considéré, soit elle est le dispositif captant en entrée certaines propriétés physiques d'un autre objet (la main par exemple). Les figures 3.3 et 3.4 montrent ces deux modélisations possibles. Un autre exemple d'objet qui peut être considéré indifféremment comme un objet mixte ou comme un dispositif est l'écran. L'écran peut bien être un objet mixte, puisqu'il appartient à la fois au monde physique et au monde numérique ; mais il est également un dispositif qui permet à d'autres objets, tel un document, de rendre leurs propriétés observables physiquement. Les figures 3.5 et 3.6 schématisent ces deux appréciations différentes de l'écran.

Nous constatons donc que même si ces exemples d'objets peuvent être traditionnellement considérés comme les dispositifs d'une modalité d'interaction, ils sont aussi des objets mixtes.

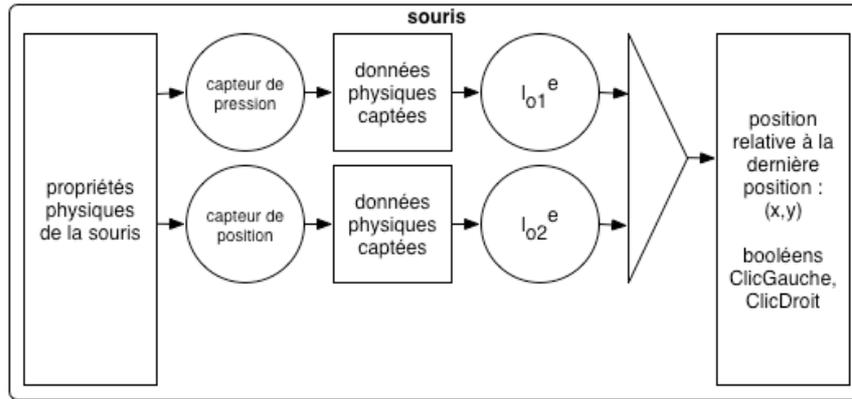


FIG. 3.3 – La souris modélisée comme un objet mixte

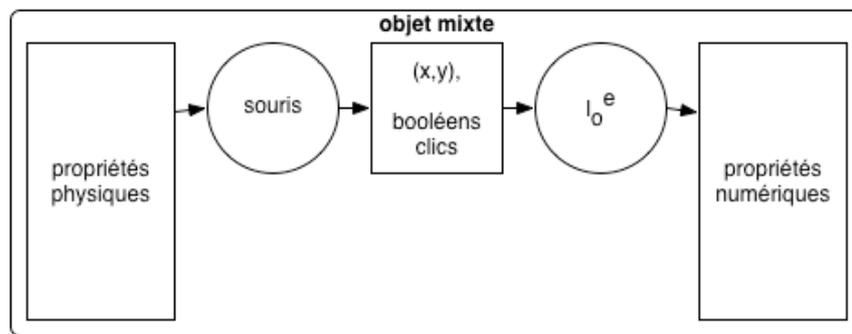


FIG. 3.4 – La souris modélisée comme un dispositif

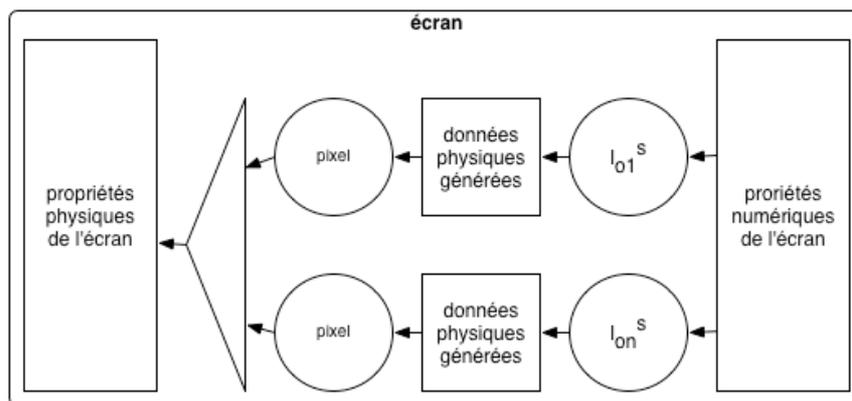


FIG. 3.5 – L'écran modélisé comme un objet mixte

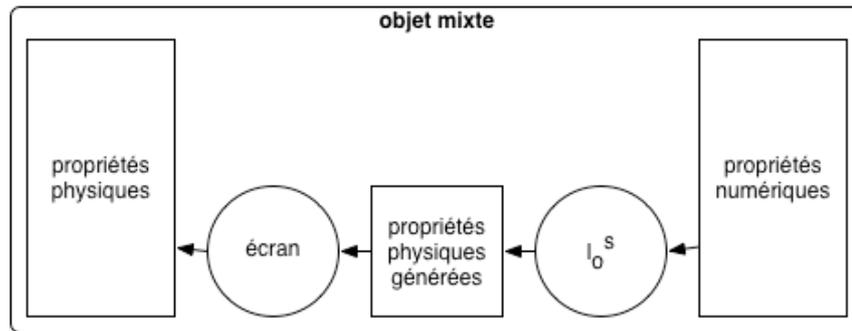


FIG. 3.6 – L'écran modélisé comme un dispositif

### Cas de la non-continuité

La figure 3.2 présente le cas le plus général d'un objet mixte. En effet, il convient de noter que la modélisation d'un objet mixte peut comporter deux parties physiques spatialement distinctes. Par exemple dans le cas de la version de CASPER de la figure 2.1, où les informations de guidage incluant la position de l'aiguille sont affichées sur un écran, l'aiguille de ponction est un objet mixte qui comporte deux parties physiques, l'une l'outil physique (aiguille) associé à une modalité de liaison en entrée ( $d_o^e, l_o^e$ ), l'autre les pixels à l'écran associés à une modalité de liaison en sortie ( $d_o^s, l_o^s$ ), où  $d_o^s$  désigne l'écran. Au contraire, la version de CASPER avec un casque à vision semi transparente évite cette discontinuité spatiale. Au paragraphe 3.2, nous étudions un ensemble de propriétés dont les différentes formes de discontinuité. Aussi le cas d'un objet mixte qui contient deux parties physiques (que nous modélisons comme un ensemble de propriétés physiques) n'est alors qu'une dimension d'un espace de caractérisation présenté au paragraphe 3.2.

### Cas de l'augmentation unilatérale

La figure 3.2 est un cas général, duquel nous pouvons encore isoler un autre cas particulier : la modélisation d'un objet mixte peut ne comporter qu'une seule modalité : en entrée ou bien en sortie. Considérons par exemple la version de CASPER avec un casque à vision semi transparent dans lequel n'est affichée que la trajectoire pré-calculée. L'aiguille de ponction est alors un objet mixte augmenté qu'en "entrée", sa modélisation ne comportant qu'une modalité d'entrée ( $d_o^e, l_o^e$ ). Ceci ne signifie donc pas nécessairement que l'interface est mal conçue : dans le cas de CASPER, nous notons qu'augmenter l'évaluation de l'aiguille est inutile puisque celle-ci est déjà dans le champ de vision de l'utilisateur. Au contraire, dans la première version de CASPER, augmenter l'évaluation de l'aiguille par un affichage à l'écran était essentiel. Nous constatons donc que l'interface dont l'évaluation de l'aiguille n'est pas augmentée est pourtant la plus utilisable.

### Objet mixte et métaphore

Les systèmes modélisables dans le cadre défini ici s'étendent des interfaces homme machine simples comme des terminaux de commandes aux systèmes mixtes. Le modèle dépasse donc la notion de réalité mixte. Il souligne de plus que la plupart des systèmes informatiques sont mixtes, car l'interaction avec l'homme est nécessairement dans le monde physique. Pourtant, il est parfois

implicitement supposé que la réalité mixte implique la manipulation d’objets physiques non dédiés à l’informatique. Or l’approche de [16] montre que la différence entre un objet comme une souris et un autre comme un cadre photo [32] n’existe que par le sens qu’on associe à l’objet, et donc par la notion de métaphore. La notion de métaphore restant très subjective, un modèle qui décrit l’interaction du point de vue du système ne peut pas, de façon réaliste, prendre en compte la métaphore et ses aspects culturels.

Le modèle permet donc de décrire indifféremment tous les objets mixtes, dans le sens rigoureux du terme. Le domaine de définition de notre modèle est donc étendu. La notion de métaphore sera discutée plus longuement au paragraphe 3.18.

Une fois la notion d’objet mixte définie dans le cadre de notre modèle, nous pouvons en apprécier le pouvoir descriptif. Comme nous l’avons démontré avec les trois versions de CASPER, le modèle autorise une compréhension plus fine de l’augmentation, parce qu’il intègre des “modalités de liaison” entre les parties physiques et numériques. Ces modalités de liaison explicitent deux niveaux d’abstraction entre les parties physiques et numériques d’un objet mixte, grâce au dispositif et au langage. Cette nouvelle définition de l’objet mixte posée, elle constitue la notion fondamentale sur laquelle se base la technique d’interaction mixte que nous définissons au paragraphe suivant.

### 3.1.2 Modalité d’interaction mixte

Cette partie a pour but de décrire l’interaction avec un objet de la tâche, qu’il soit purement physique, mixte ou purement numérique. Pour analyser l’interaction, nous généralisons le modèle de l’interaction instrumentale présenté au paragraphe 2.3.2 aux systèmes mixtes en exploitant la définition d’une modalité d’interaction du paragraphe 2.2.

Dans le cas de l’interaction en entrée, selon le modèle d’interaction instrumentale, nous considérons donc l’utilisateur, un instrument ou outil et un objet de la tâche. En affinant ce modèle instrumental de base, nous considérons que l’instrument *physique* originel (figure 2.11) est un outil *mixte*, parce que cet outil est nécessairement lié au monde physique, puisque l’être humain interagit dans le monde physique, mais aussi au monde numérique, pour pouvoir échanger des informations avec la partie logique de l’outil de la figure 2.11. L’outil est donc formé d’un outil mixte et d’un outil logique, et il est noté outil d’interaction.

Nous distinguons successivement deux types d’outil d’interaction.

- Ceux où l’outil mixte peut être associé à plusieurs outils logiques. C’est le cas le plus général. L’outil mixte sert donc à réaliser des actions grâce à des outils logiques variés. Nous pouvons métaphoriquement comparer ce cas à un tournevis (outil mixte) dont les pointes sont changeables (outil logique).
- Ceux où l’outil mixte est assigné (au sens des propriétés CARE du paragraphe 2.2.3) à l’outil logique.

#### Outil mixte non dédié

Nous rencontrons le cas de l’outil mixte non dédié quand un même outil mixte permet d’effectuer des tâches élémentaires différentes, selon la situation rencontrée. L’outil mixte est alors associé à différents outils logiques, un pour chaque tâche élémentaire. C’est le cas décrit dans le modèle initial [1][2]. L’outil d’interaction est alors composé de deux parties, comme dans [1][2] :

- un outil mixte, objet mixte, comme une souris ou la position de l'utilisateur dans le jeu TROC pour joueur mobiles, présenté au paragraphe 2.1.2,
- un outil logique, objet numérique, composé uniquement de propriétés numériques.

L'utilisateur fait une *action* modifiant les *propriétés physiques* de l'outil mixte, et celui-ci lui permet d'ajuster son *action* grâce sa *réaction*. Les *propriétés numériques* de l'outil mixte sont interprétées par un *langage d'interaction de l'outil en entrée*  $l_{io}^e$  pour mettre à jour les *propriétés numériques* de l'outil logique. L'état interne de l'outil logique peut être observable via l'outil mixte, grâce à la modalité de sortie ( $l_{io}^s, (l_{outil}^s, d_{outil}^s)$ ). Les *propriétés numériques* de l'outil logique sont traduites en une *tâche élémentaire* par un *langage d'interaction en entrée*  $l_i^e$ , et celle-ci s'applique à l'objet de la tâche. L'état interne de l'objet de la tâche est alors modifié, et celui-ci en rend compte à l'utilisateur par un *retour d'information* (si l'objet de la tâche est mixte et possède une modalité de liaison en sortie) et à l'outil mixte par une *réponse*. Cette *réponse* peut éventuellement participer à l'élaboration par le système de la *réaction* de l'outil d'interaction. La figure 3.7 schématise cette interaction avec un objet de la tâche indifféremment mixte ou numérique selon une modalité d'interaction modélisée comme un couple associant un outil mixte à un outil logique.

Pour illustrer cette définition, considérons la Table Magique du paragraphe 2.1.2, un système mixte où l'objet de la tâche, une écriture sur la table, est mixte. La modélisation de la figure 3.7 s'applique donc. Une modalité d'interaction se compose d'un outil mixte *et* d'un outil logique, puisque les jetons, qui constituent les deux outils mixtes, permettent d'effectuer des tâches différentes selon le contexte. Ils ne sont donc pas dédiés. Il y a trois types de tâches donc trois outils logiques potentiellement activables par les outils mixtes : un pour la création, un pour la transformation, et un pour la suppression. Les figures 3.8 et 3.9 montre le schéma associé à cette modélisation pour la création. Nous notons la présence des deux autres outils logiques pour la transformation et la suppression.

L'outil mixte est aussi invariant pour chaque tâche élémentaire : ce sont le ou les jeton(s). Les jetons se modélisent de la façon suivante. Un jeton constitue un objet mixte, soit une modalité d'interaction ; ces deux modalités sont ensuite composées pour agir sur l'outil logique. Chaque jeton est augmenté par une caméra et un langage qui traduit l'image du jeton en une position dans le référentiel de l'espace de travail. Ceci forme un outil mixte. Leurs *langages d'interaction de l'outil en entrée* sont composés, mettent à jour les informations de l'outil logique adéquat.

Pour la création, les propriétés de l'outil logique se composent de la position et de l'aire du rectangle, ainsi que de du temps passé à position et aire inchangées. Le langage  $l_{io}^s$  permet de re-traduire ces propriétés en forme élastique. Le langage  $l_i^s$  permet quant à lui d'observer quand la tâche élémentaire est en train d'être effectuée, en faisant disparaître la forme élastique. Le langage  $l_{objet}^s$  permet de rendre observable l'état interne de l'objet de la tâche quand celui-ci a été modifié : l'image qui a été numérisée est projetée.

## Outil mixte dédié

L'outil d'interaction qui vient d'être décrit constitue la situation la plus générale : l'association d'un outil mixte et d'un outil logique. Cette situation décrit le cas où l'outil mixte peut être associé à plusieurs outils logiques, comme dans [1][2]. Nous considérons maintenant un cas particulier en modélisant un outil mixte dédié à une tâche élémentaire. Dans ce cas l'outil mixte est toujours assigné au même outil logique. Par exemple, les bouteilles de l'AmbientROOM que l'on ouvre et ferme est un exemple d'outil mixte dédié au même outil logique qui gère l'ouverture/fermeture. L'assignation de l'outil mixte avec l'outil logique étant permanente, nous simplifions la modélisation en

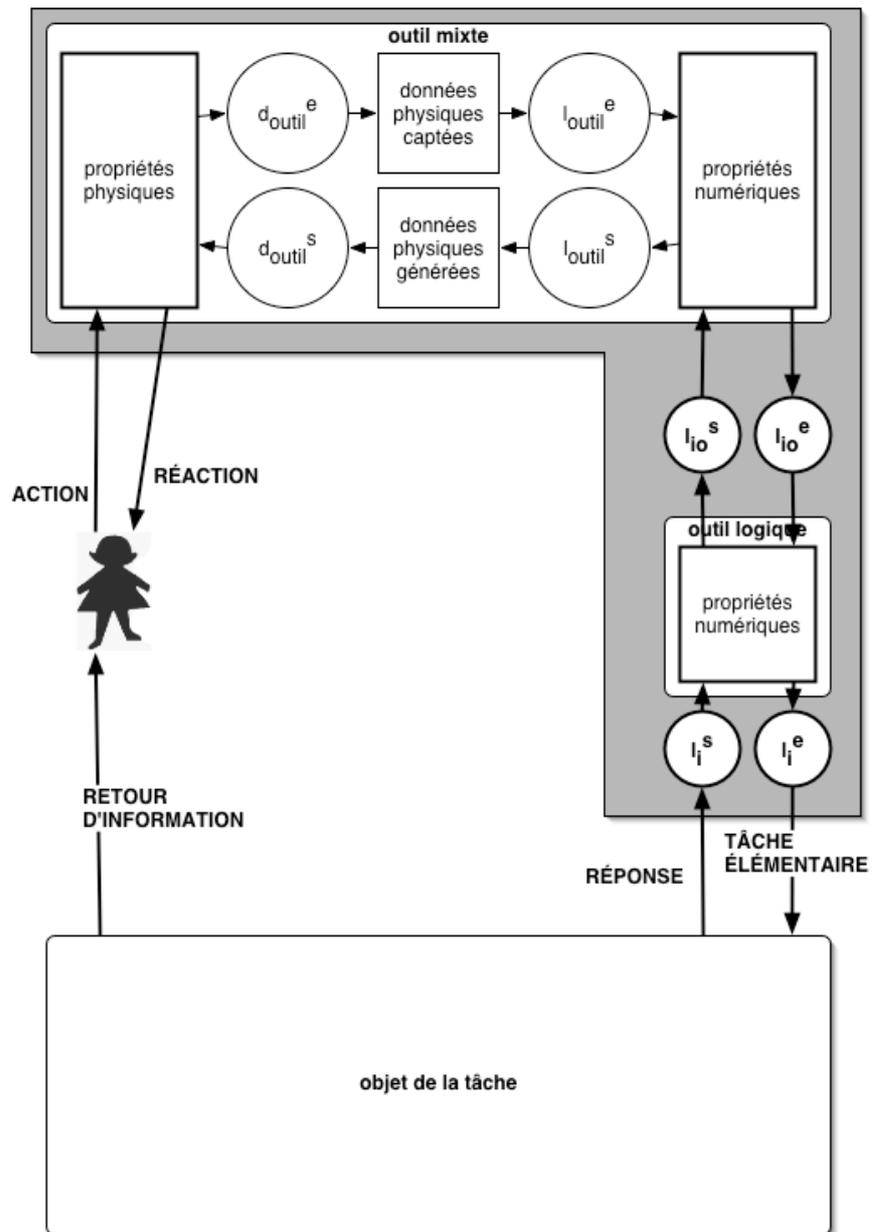


FIG. 3.7 – Outil d'interaction : outil mixte et outil logique

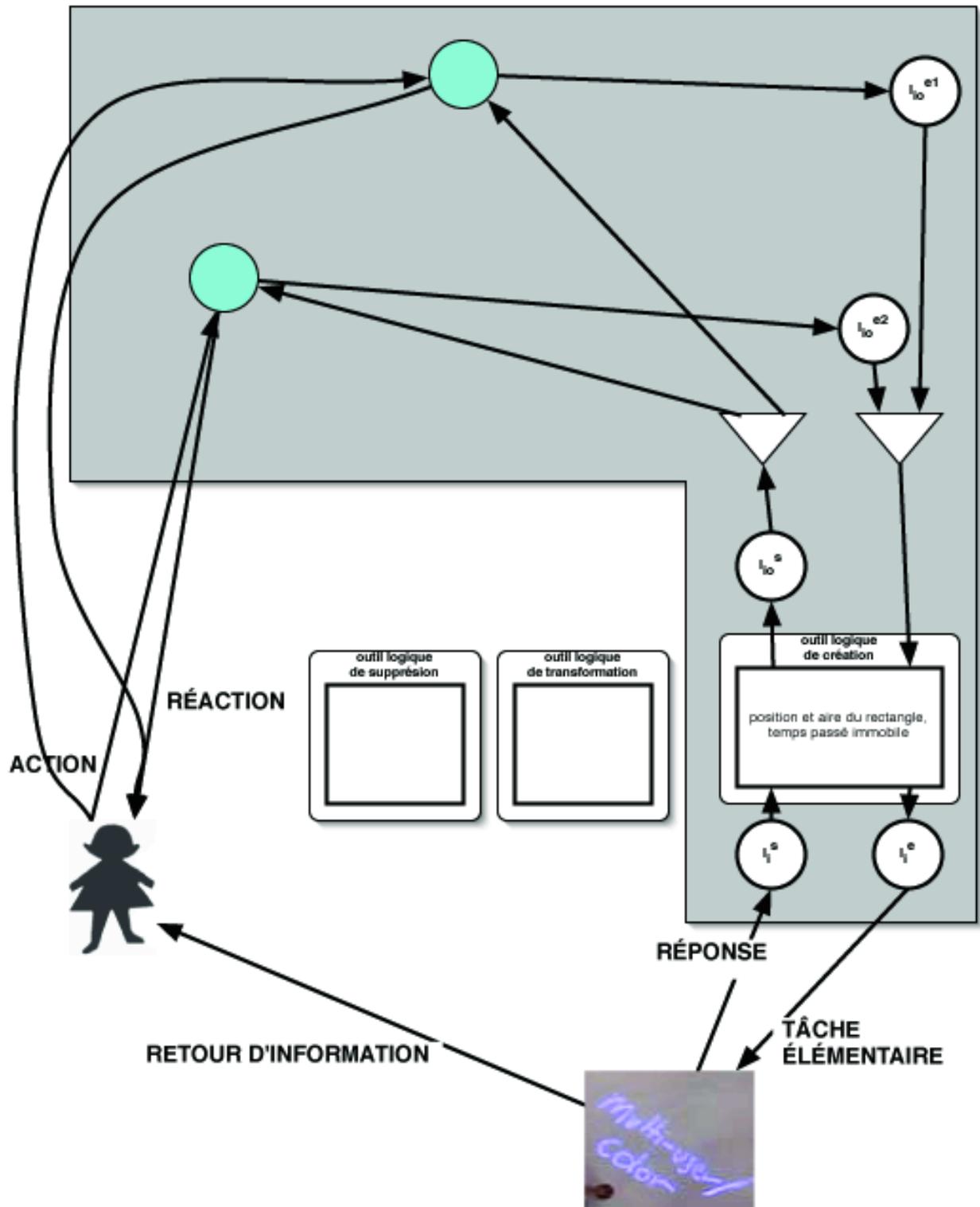


FIG. 3.8 – Outil d'interaction de la Table Magique lors de la création d'une entité (1)

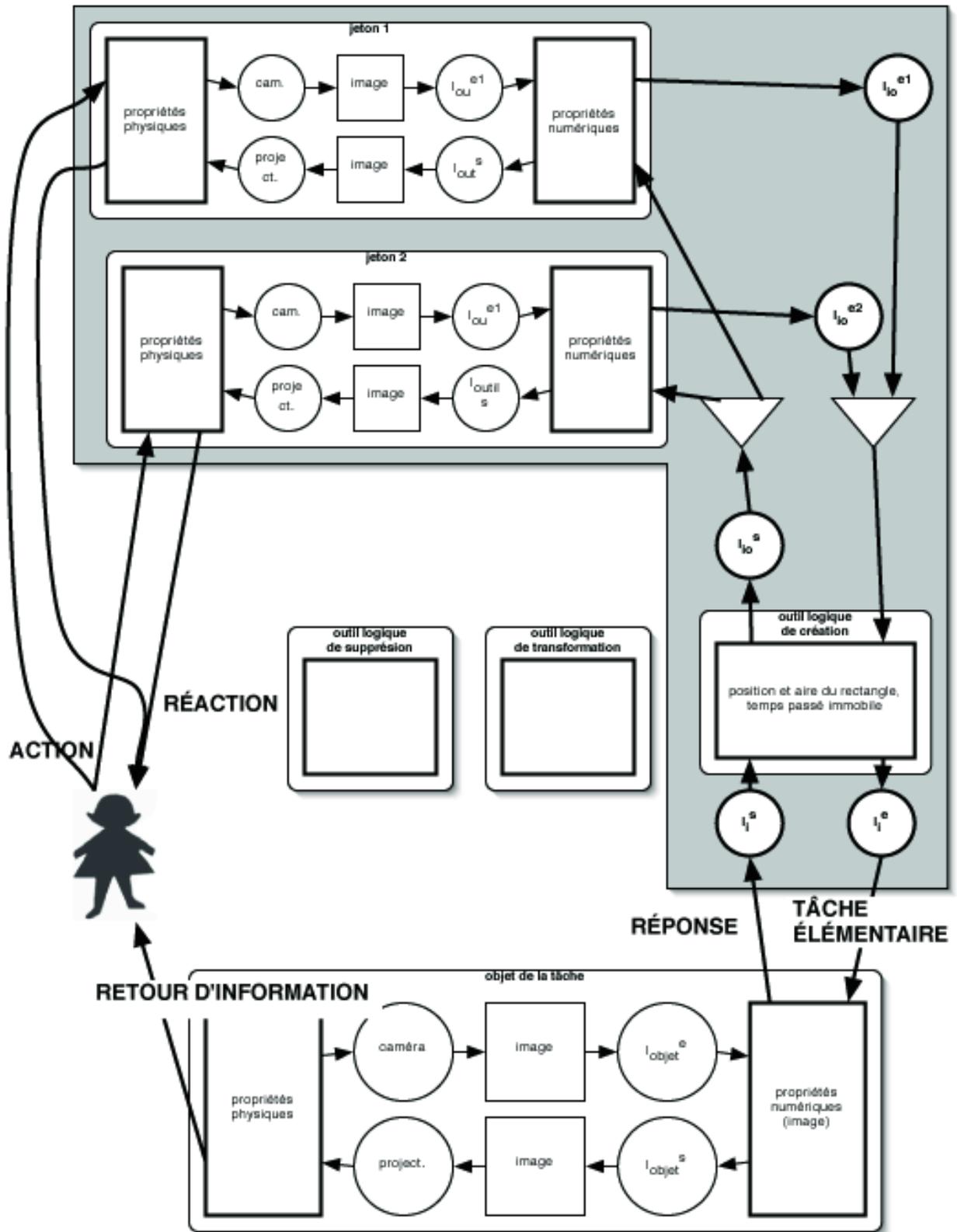


FIG. 3.9 – Outil d'interaction de la Table Magique lors de la création d'une entité (2)

confondant l’outil logique avec les propriétés numériques de l’outil mixte. Ce processus de simplification est schématisée en observant successivement les figures 3.10, 3.11 et 3.12.

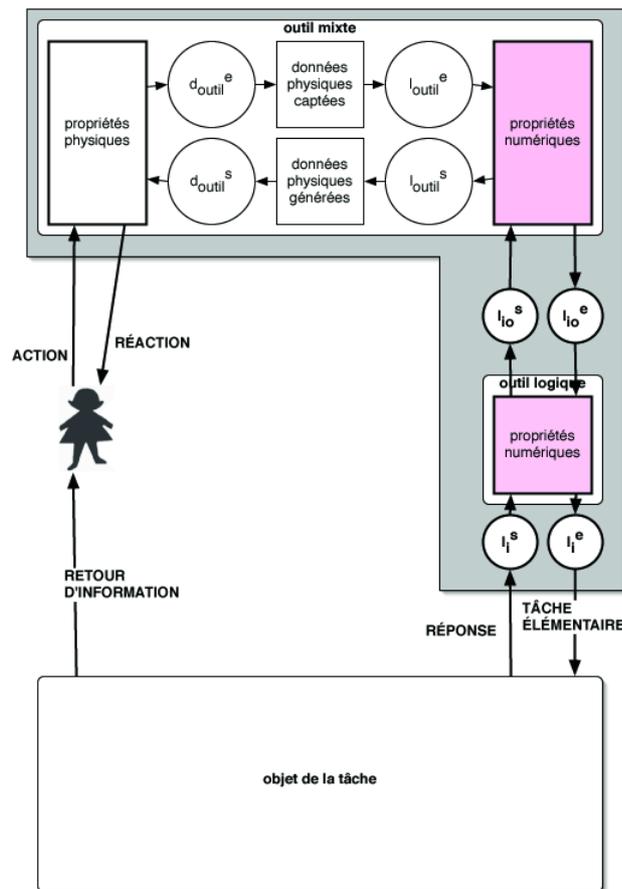


FIG. 3.10 – Simplification de l’outil d’interaction dédié (1)

Ainsi,

- les propriétés numériques de l’outil mixte et de l’outil logique sont alors confondues dans les propriétés numériques de l’outil mixte,
- les langages  $l_{outil}^e$  et  $l_{io}^e$  sont alors confondues dans  $l_{outil}^e$ ,
- les langages  $l_{outil}^s$  et  $l_{io}^s$  sont alors confondues dans  $l_{outil}^s$ .

La figure 3.13 (respectivement 3.14) présente la modélisation de l’interaction entre l’utilisateur et le système respectivement dans le cas général (respectivement pour la tâche élémentaire ”augmenter le volume du son”).

### Exemple d’outil mixte dédié

Considérons maintenant le système TROC, un système de VA où l’objet de la tâche, la borne où l’utilisateur peut recevoir de nouveaux filtres ou des unités magiques, est mixte. Notre modèle s’applique donc aussi. Cette application est présentée à la figure 3.15. La technique d’interaction

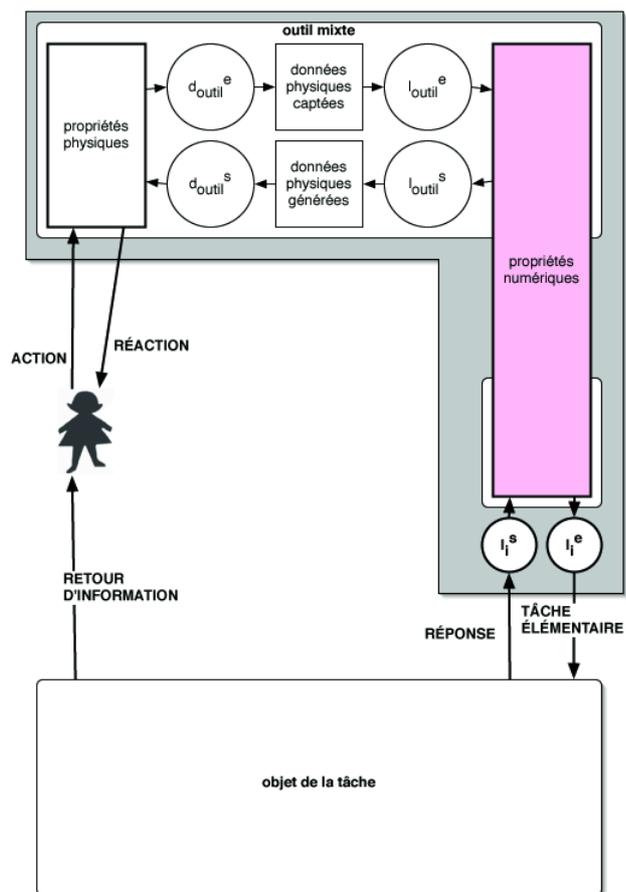


FIG. 3.11 – Simplification de l’outil d’interaction dédié (2)

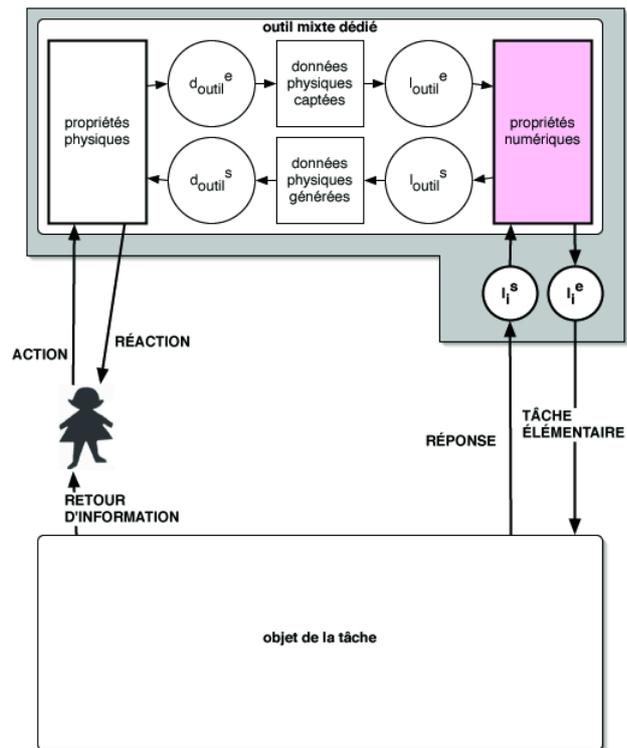


FIG. 3.12 – Simplification de l'outil d'interaction dédié (3)

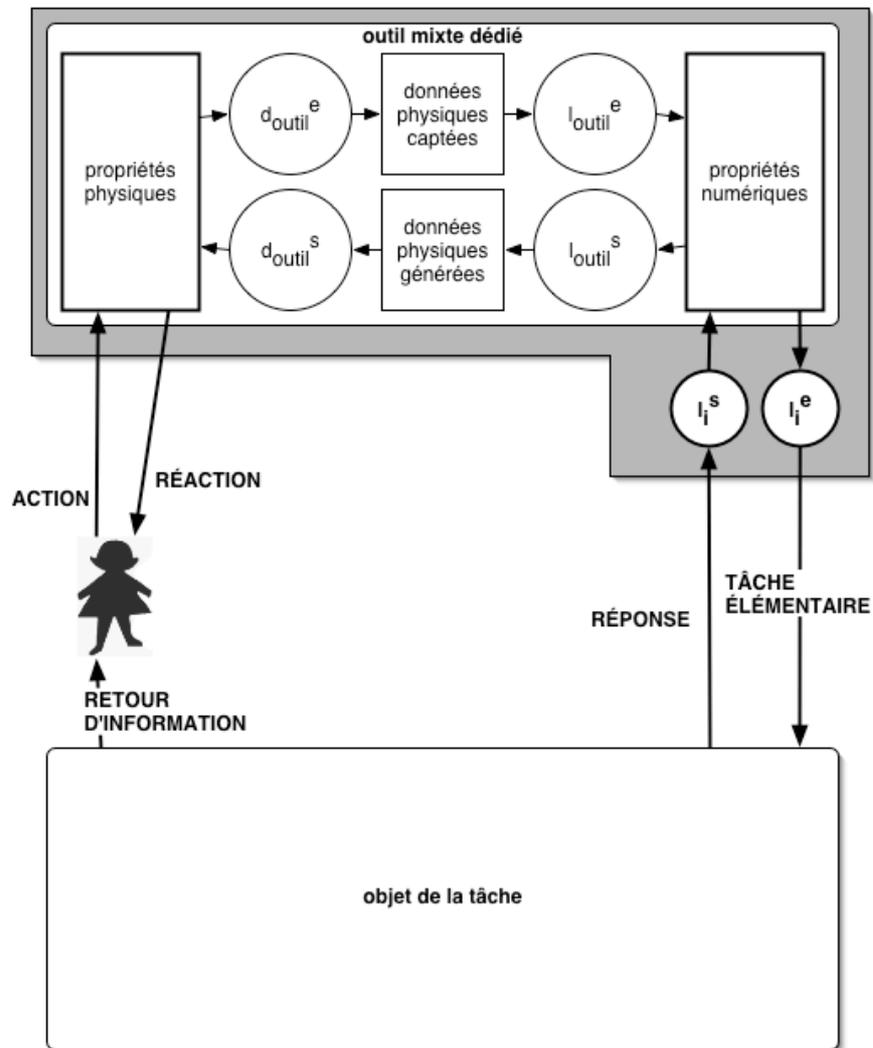


FIG. 3.13 – Outil d'interaction : outil mixte dédié

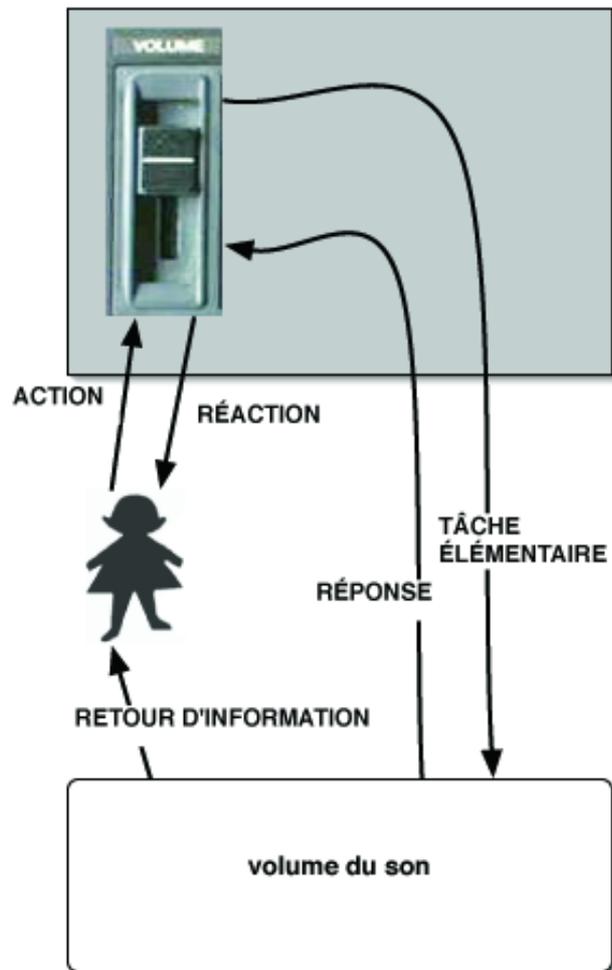


FIG. 3.14 – Outil d'interaction : exemple d'outil mixte dédié

avec la borne est d'abord composée d'un outil mixte. Les propriétés physiques de l'outil mixte comprennent la position du joueur. Celle-ci est captée par le capteur de position, et la donnée ainsi obtenue est interprétée par le langage qui lui donne un sens dans le référentiel du jeu. Ces coordonnées de position font ainsi partie des propriétés numériques de l'outil. On remarque alors que cet outil mixte est dédié à l'interaction avec la borne : Utilisée seule, la position de l'utilisateur ne permet que d'interagir avec la borne. Dans ce cas, les coordonnées de position font donc partie des propriétés numériques de l'outil dédié. Alors, le langage d'interaction  $l_i^e$  traduit cette propriété, si la position coïncide avec celle de la borne, en une tâche élémentaire "mettre le contenu de la borne dans les réserves de l'utilisateur". L'objet de la tâche ne véhicule pas de retour d'information, puisque la borne ne rend pas compte à l'utilisateur qu'elle a été vidée. Il n'y a pas non plus de réaction de l'outil mixte : une autre version de TROC aurait pu être envisagée, où l'utilisateur serait informé sur le déplacement à faire pour effectivement entrer à l'intérieur de la borne.

Par contre, nous notons que la position, en tant que modalité d'interaction passive<sup>1</sup> et composée avec la parole et le geste, peut permettre de ramasser une vignette, par exemple. Ce n'est alors pas le même outil mixte que pour interagir avec la borne.

### Outil mixte et artéfact

Nous soulignons ici qu'une technique d'interaction n'implique pas forcément la manipulation d'un artéfact. Nous avons décrit l'interaction avec les systèmes mixtes en termes de modalités d'interaction. Contrairement au modèle instrumental initial, on peut donc décrire l'interaction sans qu'il y ait nécessairement d'artéfact manipulé. Un microphone et le langage pseudo-naturel, ou la main, une caméra et la manipulation directe peuvent faire partie d'une modalité d'interaction. Ainsi la notion d'outil d'interaction dépasse celle traditionnellement admise. Un exemple déjà modélisé peut être trouvé à la figure 3.15, où l'utilisateur de TROC s'approche de sa "borne" pour mettre à jour ses filtres de vision, ses cubes, etc. La position captée constitue alors l'outil mixte de l'outil d'interaction sans artéfact. Ce cas, certes aux limites du modèle d'interaction instrumentale, est donc intégré à notre modèle, car ce dernier est basé sur la notion de modalité d'interaction et non d'artéfact.

### Absence de partie numérique

Dans le cas de la RA, où la tâche est réelle, nous trouvons des cas aux limites du modèle, où certaines parties du modèle général étant alors absentes. Si l'interaction se passe dans le monde physique, alors le propos n'est pas de la décrire ici. Dans ce cas, l'interaction n'est pas décrite par le modèle, mais la partie numérique de l'objet mixte est exploitée par le système comme une donnée et non une technique d'interaction. C'est le cas de KARMA [15] ou NaviCam [27]. Les outils de ces systèmes sont purement physiques, et ne sont donc pas modélisés. Seuls l'objet de la tâche, c'est-à-dire une partie de l'imprimante dans le cas de KARMA, est mixte. Seule l'évaluation de l'objet de la tâche est augmentée. Nous modélisons le cas de KARMA à la figure 3.16, avec la cartouche d'encre comme objet de la tâche.

---

<sup>1</sup>Une modalité d'interaction est active lorsqu'elle est utilisée consciemment par l'utilisateur, et passive dans l'autre cas.

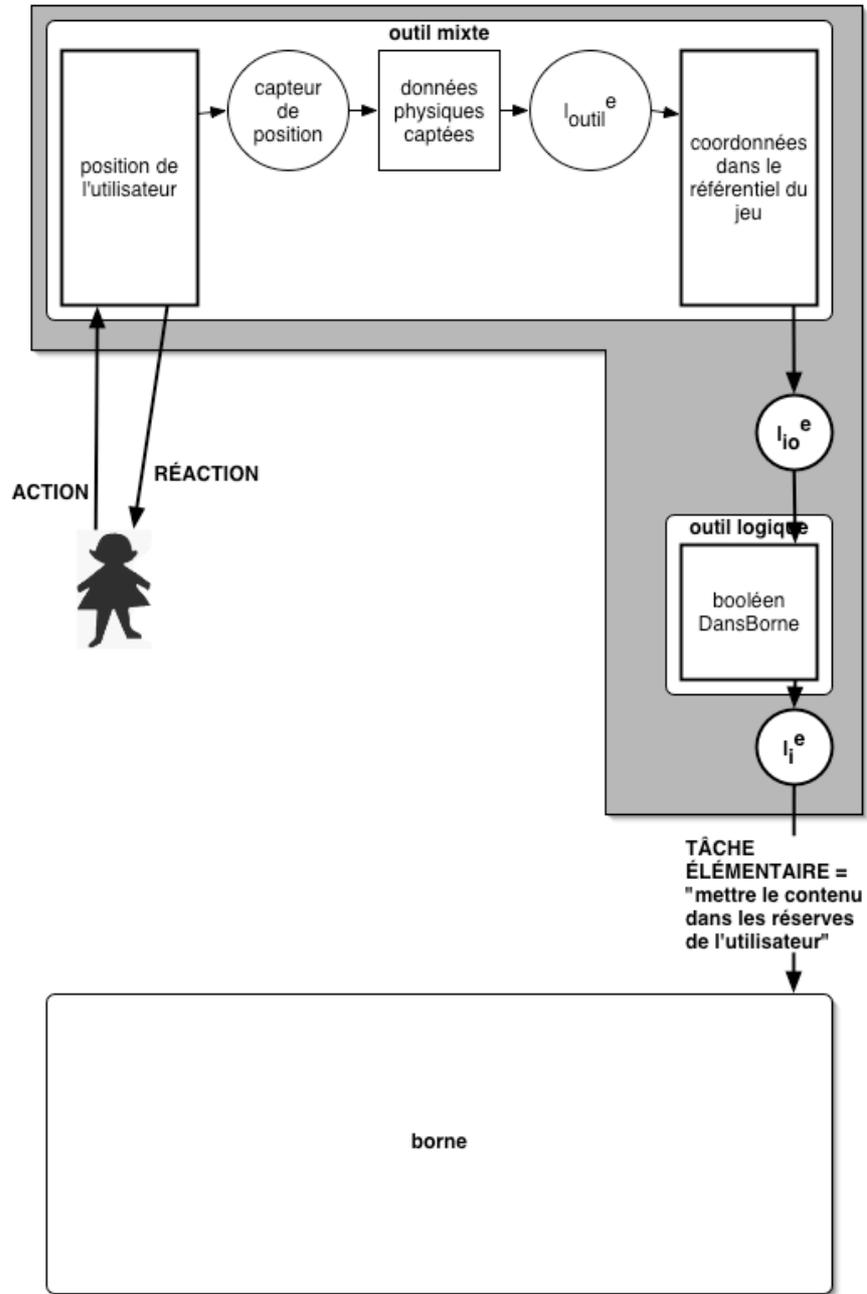


FIG. 3.15 – Outil d’interaction de TROC lors de l’interaction avec la borne

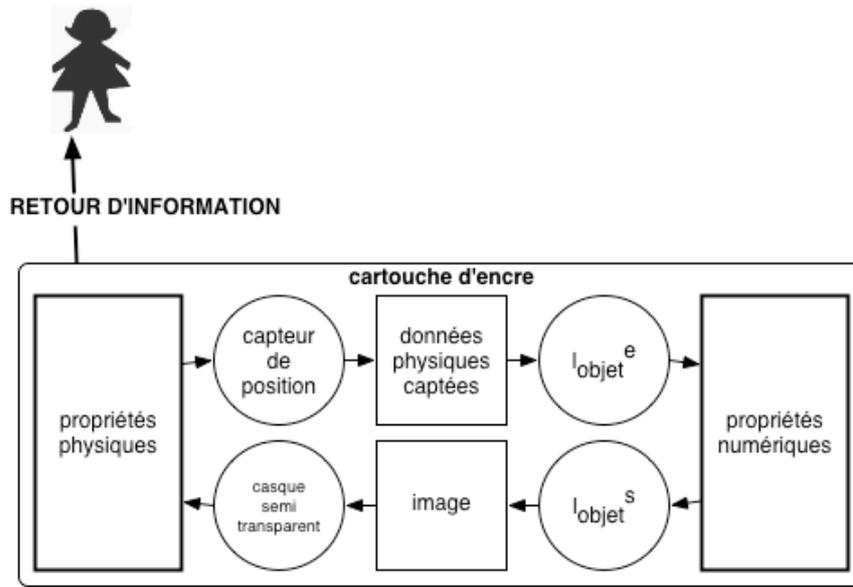


FIG. 3.16 – Modélisation de KARMA

### 3.1.3 Objet de la tâche

Après avoir focalisé sur l'outil mixte (figure 3.7), nous considérons maintenant l'objet de la tâche. En toute rigueur, celui-ci peut être physique, numérique ou mixte. Si l'objet de la tâche est physique, alors les propriétés numériques éventuelles de l'outil d'interaction ne participe pas à l'interaction avec l'objet physique de la tâche. Dans les deux autres cas, pour qu'il y ait un retour d'information de l'objet vers l'utilisateur, comme le montre la figure 3.7, il est nécessaire que l'objet de la tâche soit mixte *et* qu'il ait une modalité de liaison en sortie, c'est-à-dire qu'il ait un langage interprétant ses propriétés numériques en propriétés physiques, via le dispositif de l'objet en sortie. Dans le cas où l'objet de la tâche n'est pas purement physique, si l'objet n'est pas augmenté en sortie, son état interne n'est pas observable par l'utilisateur. Seules les propriétés physiques sont perceptibles par l'utilisateur. La figure 3.17 étend la modélisation de la figure 3.7 avec un objet de la tâche mixte.

### 3.1.4 Lien avec les modélisations existantes

Si le lien avec le modèle instrumental de l'interaction est immédiat, puisqu'il est le fondement même de notre modèle, le lien avec ASUR ou IRVO décrits aux paragraphes 2.3.1 et 2.3.1 peut quant à lui être précisé avec plus de rigueur, afin de vérifier que notre modèle permet l'unification des modèles précédents.

### Modèle de l'interaction instrumentale

Le modèle instrumental de base [1] [2] se retrouve ici entièrement. Nous en avons détaillé certaines parties, en utilisant la notion de modalité. En particulier, les entités qui entrent en compte dans l'interaction, tels que les deux parties de l'outil et l'objet de la tâche, sont explicités en terme

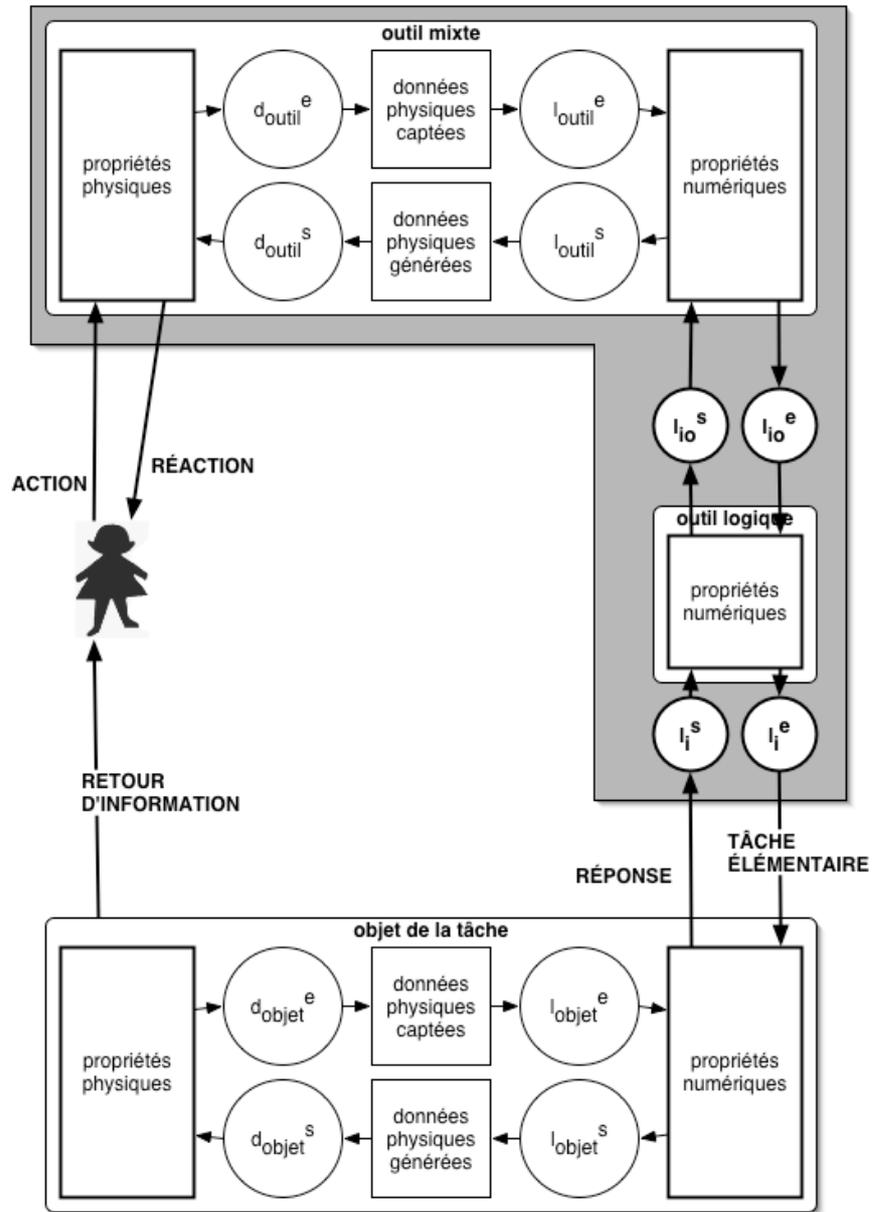


FIG. 3.17 – Technique d'interaction mixte sur un objet mixte

de propriétés physiques et numériques liées par des modalités de liaison. De plus, la partie de l'outil qui était avant proposée comme physique dans le modèle de l'interaction instrumentale a été identifiée comme mixte, puisqu'elle fait le lien entre le monde physique et le monde numérique. Enfin, l'association entre outil mixte et outil logique est aussi détaillée : elle se fait selon un langage  $l_{io}^e$ , ce qui nous permet d'exprimer l'association (spatiale ou temporelle) [2] d'un outil mixte à un outil logique en termes plus formels guidés par le langage. Les termes employés dans [1] [2], comme l'*action*, la *réaction*, le *retour d'information*, sont explicités ici en terme de modalités de liaison ou d'interaction. Nous synthétisons la correspondance entre les termes des deux modèles dans la table 3.1.

<b>interaction instrumentale</b>	<b>interaction mixte</b>	<b>signification</b>
objet du domaine	objet de la tâche	objet physique, mixte ou numérique
instrument physique	outil mixte	objet mixte
instrument logique	outil logique	
association spatiale/temporelle		
action	action	utilisation de la modalité d'interaction en entrée $((d_{outil}^e, l_{outil}^e), l_{io}^e, l_i^e)$
réaction	réaction	utilisation de la modalité d'interaction en sortie $((d_{outil}^s, l_{outil}^s), l_{io}^s, l_i^s)$
retour d'information	retour d'information	utilisation de la modalité d'interaction en sortie $(d_{objet}^s, l_{objet}^s)$
commande	tâche élémentaire	unité d'information véhiculée par $l_i^e$
réponse	réponse	unité d'information véhiculée par $l_i^s$

TAB. 3.1 – Correspondance des termes de notre modèle d'interaction mixte avec celui d'interaction instrumentale [2]

Dans [2], le niveau de précision d'un modèle d'interaction est étudié au regard de trois propriétés d'un modèle :

- pouvoir descriptif
- pouvoir évaluatif
- pouvoir génératif

Il est constaté qu'un modèle d'interaction à haut niveau d'abstraction a un pouvoir descriptif fort au détriment des pouvoirs évaluatif et génératif. Aussi notre modèle d'interaction mixte affinant le modèle d'interaction instrumentale a donc un pouvoir évaluatif et génératif plus élevé que celui de l'interaction instrumentale.

Enfin, il convient de noter que tandis que le modèle d'interaction instrumentale est dédié aux interfaces graphiques, nous focalisons dans notre étude sur les systèmes mixtes.

## ASUR

Le lien avec la notation ASUR est moins direct qu'avec le modèle précédent. Les composants  $R$  d'ASUR sont formés de *propriétés physiques* des objets ; les composants  $A$  correspondent aux dispositifs des modalités de liaison des objets ; et enfin, les composants  $S$  sont décomposés en entités utiles à l'interaction en entrée, ce qui correspond aux modalités d'interaction en entrée de notre modèle, et en  $S_{objet}$  et  $S_{info}$  qui correspondent aux propriétés numériques de l'objet de la tâche. La correspondance entre ASUR et notre modèle est résumée dans la table 3.2.

ASUR	modèle d'interaction mixte
$A_{in}$	$d^e$
$A_{out}$	$d^s$
n'existe pas	$l_{objet}^e$
n'existe pas	$l_{objet}^s$
n'existe pas	$l_{outil}^e$
n'existe pas	$l_{outil}^s$
n'existe pas	$l_{io}^e$
n'existe pas	$l_{io}^s$
n'existe pas	$l_i^e$
n'existe pas	$l_i^s$
$S_{outil}$	propriétés numériques de l'outil mixte et de l'outil logique
$S_{objet}$	propriétés numériques de l'objet de la tâche
$S_{info}$	propriétés numériques de l'objet de la tâche
$U$	$U$ est source ou destination des flèches <i>action, réaction, retour d'information</i>
$R_{outil}$	propriétés physiques de l'outil mixte
$R_{tache}$	propriétés physiques de l'objet de la tâche

TAB. 3.2 – Correspondance des termes de notre modèle d'interaction mixte avec ceux de la notation ASUR

Notre modèle d'interaction mixte affine les notions déjà identifiées dans ASUR en explicitant les modalités de liaison et d'interaction. Ainsi notre modèle permet une modélisation plus fine de l'interaction qu'ASUR. En particulier, les langages des modalités de liaison et d'interaction ne sont pas présents dans ASUR qui ne distingue pas l'outil mixte de l'outil logique.

Enfin notre modèle autorise une description à plusieurs niveaux d'abstraction (comme une souris qui peut être modélisé comme un outil mixte ou un dispositif au sein d'un outil mixte). Ces différents niveaux d'abstraction ne sont pas possibles avec la notation ASUR.

## IRVO

La différence fondamentale entre IRVO et le modèle de l'interaction mixte est, là encore, la notion de modalité (de liaison ou d'interaction). IRVO, contrairement aux modèles précédents de l'interaction, propose une décomposition de l'objet mixte en partie physique et partie numérique. En cela IRVO et notre modèle se retrouve. Néanmoins notre description d'un objet mixte selon des

modalités de liaison est beaucoup plus fine que la simple juxtaposition de propriétés physiques et numériques. Les correspondances entre les termes d'IRVO et le modèle de l'interaction mixte sont explicitées dans la table 3.3.  $S$  est un senseur, soit un dispositif d'entrée;  $E$  un effecteur, soit un dispositif de sortie;  $O$  est l'objet de la tâche;  $T$  est l'outil;  $U$  est l'utilisateur; et  $M$  est la partie du système qui ne participe pas à l'interaction.

IRVO	modèle
$S$	$d^e$
$E$	$d^s$
$O$	objet de la tâche
$T$	outil d'interaction
$U$	U est source ou destination des flèches <i>action, réaction, retour d'information</i>
$M$	M ne rentre pas en compte dans notre modèle de l' <i>interaction</i> mixte

TAB. 3.3 – Correspondance des termes de notre modèle avec ceux du modèle IRVO

Nous notons que la distinction dans IRVO entre  $O_{v+r}$  et  $O_{r+v}$  n'est pas faite dans notre modèle. Néanmoins il nous semble difficile de qualifier un objet mixte en disant qu'il est plutôt un objet physique augmenté qu'un objet numérique augmenté. Nous adhérons plutôt à l'approche de dire qu'il existe des objets purement physiques et purement numériques et des objets mixtes.

D'autre part, l'opérateur de fusion perceptuelle est inutile dans notre modèle, car il s'agit en réalité du cas d'une propriété déjà identifiée, la continuité perceptuelle [13].

Nous avons pu constater dans ce paragraphe que le modèle de l'interaction mixte permet d'unifier les cadres de travail précédents, puisqu'il ne les conteste pas, mais plutôt les enrichit. Il permet de plus d'unifier des axes de recherche plus larges encore, tels que l'interaction multimodale et la réalité mixte. Ses qualités en terme de description des systèmes interactifs sont donc intéressantes. Pourtant, même s'il est essentiel qu'un modèle d'interaction ait un fort pouvoir descriptif pour faciliter la compréhension de l'interaction, outre ce pouvoir descriptif, il convient d'analyser le pouvoir évaluatif du modèle par rapport aux espaces de caractérisation existants.

## 3.2 Espace de caractérisation

Nous venons de vérifier que ce modèle n'entre pas en contradiction avec les modèles préexistants. Les espaces de caractérisation existants sont donc capitalisés et utilisables ici. Les propriétés des modèles de base nous servent donc à construire une première version d'un espace de caractérisation des systèmes mixtes propre à notre modèle, espace de caractérisation que nous enrichissons grâce aux notions de modalité et multimodalité.

### 3.2.1 Compatibilités, Continuités

Les définitions de compatibilités perceptuelles et cognitives ont été rapellées dans le paragraphe 2.4 : la compatibilité perceptuelle est l'observabilité conjointe de  $N$  concepts, et la compatibilité cognitive est l'honnêteté de la représentation de ces  $N$  concepts. Forts de notre rapprochement entre

réalité mixte et modalités d'interaction, nous pouvons reformuler ces deux propriétés et trouver leur équivalent en terme de modalités et multimodalité : la compatibilité perceptuelle met  $N$  concepts en jeu, or ceux-ci sont rendus observables par des dispositifs d'interaction  $d$ . La compatibilité perceptuelle est donc la caractérisation de la composition spatiale, temporelle, et articulatoire des  $N$  modalités de sortie propres à chaque concept. Par cet énoncé, nous étendons donc la notion de compatibilité perceptuelle aux dimensions temporelles et articulatoires. De même, la compatibilité cognitive est la caractérisation de la composition syntaxique et sémantique des  $N$  modalités de sorties propres à chacun des  $N$  concepts.

La continuité correspond à la représentation multiple d'un *unique* concept. Ce sont donc là aussi la composition de  $N$  modalités. Par exemple, l'état interne de l'objet de la tâche peut constituer la source unitaire d'information représentée plusieurs fois. Dans le cadre de notre modèle, il est alors intéressant d'étudier en particulier les deux modalités en sortie qui véhiculent l'état interne de l'objet de la tâche :  $(d_{objet}^s, l_{objet}^s)$  pour le *retour d'information* et  $((d_{outil}^s, l_{outil}^s), l_{io}^s, l_i^s)$  pour la *réponse* et la *réaction*. C'est donc la composition de ces deux modalités particulières que l'on devra étudier avec précaution lors de la conception de systèmes mixtes.

Nous notons que la notion de continuité perceptuelle a aussi été définie avec ASUR [13] en étudiant la relation spatiale entre l'outil et l'objet de la tâche. Nous pouvons rapprocher cette notion de continuité perceptuelle de celle du degré d'indirection et plus précisément du décalage spatial, dans le modèle d'interaction instrumental [2] : la relation spatiale entre l'outil logique et l'objet de la tâche est présentée comme une propriété de l'outil. Cette propriété permet de guider la conception d'un système.

En conclusion, compatibilité et continuité sont des propriétés utiles lors de la conception d'un système. Ces propriétés ont été étudiées dans le contexte de d'autres approches [1] [13]. Notre modèle permet une étude systématique de ces propriétés en identifiant le choix de conception (choix du  $d_{outil}^s$  par exemple) qui impactent sur ces propriétés. De plus nous étendons ces propriétés en exploitant des résultats issus de travaux sur la multimodalité et en particulier les cinq aspects de la composition de modalités.

### 3.2.2 Degré d'intégration

Le degré d'intégration [1] [2] mesure le rapport entre les degrés de liberté de l'instrument physique et ceux de l'instrument logique, soit ici entre les "degrés de liberté" de l'outil mixte et de l'outil logique.

Dans le modèle nous trouvons plusieurs niveaux pour caractériser cette propriété. En effet l'outil mixte contient un dispositif, et celui-ci est caractérisé par le nombre de dimensions captées pour chaque grandeur. De plus, cette propriété peut aussi être caractérisée, toujours pour l'outil mixte, par le pouvoir d'expression du langage de l'outil mixte  $l_{outil}$ . Ces deux propriétés mises ensemble forment le numérateur du rapport identifié sous le nom de *degré d'intégration* dans [1][2].

Le dénominateur, c'est-à-dire le nombre de "degrés de liberté" de l'outil logique, peut être mesuré par le degré d'expression du langage  $l_{io}$ . Afin de préserver un rapport cohérent, nous posons une nouvelle définition du *degré d'intégration* dans le cadre de notre modèle de l'interaction mixte : c'est le rapport entre le pouvoir expressif du langage  $l_{outil}$  et le langage  $l_{io}$ .

En termes de travaux futurs, il serait donc nécessaire d'introduire une mesure des pouvoirs d'expression des langages d'interaction, ou au moins un ordre total sur les langages, pour pouvoir les classer et permettre une évaluation prédictive moins empirique que celle déjà faite [2].

### 3.2.3 Parties absentes dans une modélisation

Le modèle de la figure 3.7 est un cas général. Dans la pratique, au sein de la modélisation d'un système mixte particulier, certaines parties des modalités d'interaction en entrée ou en sortie seront absentes. Par exemple, dans le modèle de TROC de la figure 3.15, toutes les modalités de sorties sont absentes. Quantifier ces parties absentes du modèle nous permet d'apprécier le degré d'observabilité du système. Parfois une modalité se révèle inutile, par exemple afficher la position de l'aiguille dans le casque semi transparent de la seconde version de CASPER, alors que celle-ci est déjà observable dans le monde réel. Pourtant il est souvent nécessaire de rendre *honnêtement observable* l'état interne de l'objet de la tâche et de l'outil d'interaction. On peut donc caractériser un système par sa complétude de modélisation.

### 3.2.4 Synthèse des espaces de caractérisation

La définition de ces propriétés au sein de notre modèle est donc cohérente. Ces propriétés ne sont pas nouvelles. Néanmoins, grâce aux résultats issus de travaux sur la multimodalité, nous les avons étendues.

Ces propriétés sont utiles en phase de conception et font l'objet d'évaluation prédictive. Il convient néanmoins de noter qu'une propriété n'est pas forcément le garant d'une bonne interaction et donc il n'est pas forcément souhaitable qu'un système vérifie une propriété. Ce point a été illustré avec le système CASPER par exemple.

Ces propriétés comme guide à la conception participent au pouvoir comparatif de notre modèle. D'autres capacités sont aussi essentielles à un modèle d'interaction, comme son pouvoir taxinomique.

## 3.3 Classification des systèmes existants

Si un modèle d'interaction a un fort pouvoir de description, il doit aussi permettre de les différencier finement. Afin d'apprécier l'intérêt du modèle du point de vue de son pouvoir taxinomique, nous allons étudier son pouvoir de classification au regard des taxonomies déjà établies.

### 3.3.1 Systèmes mixtes et interface graphique/réalité virtuelle

La virtualité augmentée se différencie de la réalité virtuelle, car cette dernière immerge l'utilisateur dans le monde numérique pour qu'il effectue sa tâche. De l'ordinateur de bureau jusqu'aux systèmes totalement immersifs, l'attention de l'utilisateur est requise par le monde numérique. Le monde réel, quant à lui, n'est pas pris en compte dans la conception du système interactif.

Il convient de remarquer que le modèle présenté ici, même si ce n'est pas son propos, comprend également les modalités d'interaction avec les systèmes informatiques à interface graphique classiques ou les systèmes immersifs. Pour illustrer ce modèle et en particulier les propriétés de l'instrument, nous reprenons l'exemple présenté dans [1]. En entrée, l'outil mixte qu'est la souris a certaines propriétés physiques, captées par les capteurs de pression sous les boutons et le capteur de position,  $d_o^e$ . Ces données sont interprétées et deviennent une position et des valeurs booléennes, grâce au langage associé à la souris  $l_o^e$ . Enfin, à leur tour, ces propriétés numériques sont interprétées en propriétés numériques de la barre de défilement par le langage  $l_{i_o}^e$ . L'état interne de la barre est traduit en tâche élémentaire à appliquer au document, objet de la tâche. En sortie, les propriétés

numériques de la barre de défilement sont interprétées et la réaction de l’outil d’interaction mixte, outre la réaction physique de la souris, mettra en surbrillance la flèche de la barre de défilement sur l’écran. La réponse de l’objet de la tâche sera le décalage de la vue du document. Cette réponse participera, par l’intermédiaire de l’outil logique, à la réaction de l’instrument, par la mise à jour de l’ascenseur de la barre à l’écran. La souris est donc modélisée comme un outil mixte, celui-ci comprenant deux parties physiques spatialement disjointes : la souris objet physique et le curseur affiché l’écran. Cet outil mixte correspond donc au cas de l’aiguille dans CASPER lorsque sa position par rapport à la trajectoire pré-calculée est affichée sur l’écran (figure 2.1). Ces deux systèmes, le premier étant dit “mixte” et l’autre pas, sont indifféremment modélisés ici. Or ils sont de natures différentes et ne font pas partie du même paradigme d’interaction. Il est donc nécessaire de les distinguer.

Dans [6], il est énoncé que “les concepts établis pour l’ordinateur de bureau ne sont pas adaptables, comme le concept de la manipulation directe qui, historiquement, a servi de principal moyen de compréhension de l’interface graphique. [...] La réalité augmentée est une approche de conception des systèmes informatiques qui augmente les objets au lieu de les remplacer et de les représenter [...]” En d’autres termes, ce qui permet de faire la différence entre un ordinateur de bureau avec une interface graphique et une souris, et un système de réalité mixte, c’est la signification que l’utilisateur associe à l’objet qu’il/elle manipule. Dans les termes de notre modèle de l’interaction mixte, le langage de l’interaction dans le cas de l’interface graphique fonctionne par analogie avec le bureau physique, mais pas l’outil (dont l’outil mixte est la souris). Dans le cadre de la réalité mixte, l’analogie existe dès les propriétés physiques de l’outil mixte, comme avec le Digital Desk [33] ou CASPER.

La différence entre interface graphique classique et réalité mixte s’exprime donc en terme de métaphore. À ce stade de l’argumentation, précisons la notion de métaphore. La métaphore remplace un concept A par un concept B. La métaphore et l’analogie (et donc le langage analogique de Bernsen [5] présenté au paragraphe 2.2.2) sont des notions très proches. La métaphore en informatique est la projection d’un domaine de concepts vers un autre connu de l’utilisateur, comme le système de fichier vers le bureau. Un domaine de concept est constitué des concepts et des relations qu’ils entretiennent entre eux. Dans le cadre du modèle de l’interaction mixte, cela signifie que la métaphore s’applique aux propriétés, physiques ou numériques, mais aussi aux langages qui régissent la façon dont elles se comportent en entrée ou en sortie.

Dans [16], Fishkin distingue deux types de métaphore :

- la *métaphore de nom* se définit par l’analogie suivante : “un *A* dans le système correspond à un *B* dans le monde physique”. Il s’agit ici de remplacer un concept par un autre. En appliquant notre modèle nous identifions plusieurs niveaux : en effet les métaphores de nom peuvent s’appliquer respectivement au niveau des :

- propriétés physiques de l’outil mixte,
- propriétés numériques de l’outil mixte,
- propriétés numériques de l’outil logique,
- propriétés numériques de l’objet de la tâche,
- propriétés physiques de l’objet de la tâche.

L’application d’une métaphore aux propriétés physiques implique que le système est un système mixte. En revanche, l’application d’une métaphore aux propriétés numériques ne permet pas de faire la différence entre systèmes mixtes et de RV.

- la *métaphore de verbe* se définit par “faire l’action *A* dans le système correspond à faire l’action

$B$  dans le monde physique”. Tandis que la métaphore de nom s’applique aux propriétés au sein de notre modèle, nous étudions les métaphores de verbe aux niveaux des langages :

- $l_{outil}$
- $l_{io}$
- $l_i$
- $l_{objet}$

Ce type de métaphore ne permet pas de distinguer les systèmes mixtes et de RV. Il est néanmoins souligné dans [16] que ce type de métaphore est utilisé parmi les systèmes mixtes.

La figure 3.18 met en évidence les éléments de notre modèle où l’application de métaphore est possible selon que l’on considère un système mixte ou de RV.

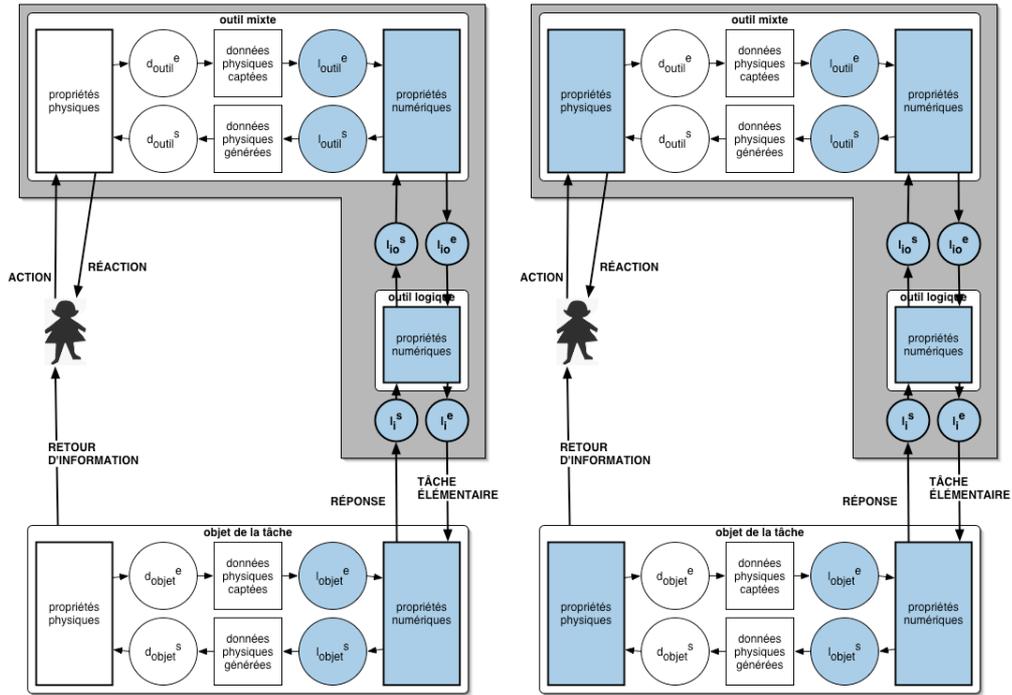


FIG. 3.18 – Domaine potentiel d’applications de métaphores dans le cas (1) des systèmes informatiques graphiques ou de réalité virtuelle et (2) des systèmes mixtes

En conclusion, nous avons vu que ce qui permet de distinguer les systèmes mixtes de ceux qui ne le sont pas, c’est le niveau d’application de la métaphore. Nous ne considérons maintenant que les systèmes mixtes. La réalité augmentée et la virtualité augmentée avaient été distinguées dans des taxonomies précédentes. Nous allons étudier dans la partie suivante dans quelle mesure ces classes sont distinguables dans notre modèle.

### 3.3.2 Réalité augmentée et virtualité augmentée

La distinction entre la réalité augmentée (RA) et la virtualité augmentée (VA) repose sur le type de l’objet de la tâche. Cette distinction est immédiate dans notre modèle si l’objet de la tâche

est soit réel (RA), soit numérique (VA).

Dans le cas d'un objet mixte, un système de RA reposera sur une tâche purement physique (interaction physique entre l'outil et l'objet de la tâche). Au contraire un système de VA contiendra dans sa modélisation une tâche numérique et donc un langage  $l_i$  pour l'exprimer.

Par exemple dans CASPER (RA), la réalisation de la tâche "faire la ponction" repose sur une interaction physique entre l'outil mixte qu'est l'aiguille de ponction et l'objet de la tâche, le patient. Ce dernier peut être augmenté (comme avec un casque de vision transparent) et sera donc modélisé comme un objet de la tâche mixte. Dans la modélisation de CASPER, aucun langage  $l_i^e$  ne véhiculera la tâche "faire la ponction". Au contraire dans TROC (VA), comme le montre la figure 3.15, un langage  $l_i^e$  véhicule l'expression de la tâche "mettre le contenu de la borne dans les réserves de l'utilisateur".

### 3.3.3 Exécution augmentée et évaluation augmentée

Outre la distinction RA/VA l'appréciation du deuxième axe de classification [13], l'exécution/évaluation augmentée, est affinée par notre modèle d'interaction mixte. En effet, le modèle autorise une caractérisation de l'exécution/évaluation fine par l'identification de plusieurs langages qui définissent l'augmentation.

#### Exécution

Au sein de notre modèle, l'augmentation de l'exécution est défini à plusieurs niveaux : le niveau de l'outil mixte ( $l_{outil}^e$ ), celui de l'interaction entre l'outil mixte et l'outil logique ( $l_{io}^e$ ) et celui de l'interaction entre l'outil et l'objet de la tâche ( $l_i^e$ ). L'exécution est augmentée si les langages  $l_{io}^e$  et  $l_i^e$  véhiculent des informations qui participent à l'interaction en entrée et ne sont pas uniquement une donnée exploitée par le système en vue d'augmenter l'évaluation.

#### Évaluation

Nous pouvons aussi apprécier plus finement l'évaluation augmentée grâce à sa caractérisation en terme de langage d'une modalité. Nous identifions les niveaux suivants :

- l'évaluation de l'état interne de l'objet de la tâche ( $l_{objet}^s$ ),
- l'évaluation de l'état interne de l'outil mixte ( $l_{outil}^s$ ),
- l'évaluation du contrôle de l'outil mixte sur l'outil logique ( $l_{io}^s$ ),
- l'évaluation du contrôle de l'outil logique sur l'objet de la tâche ( $l_i^s$ ).

Il semble à première vue manquer, dans ces unités *élémentaires* de l'évaluation de l'état interne du système, l'évaluation de l'état interne de l'outil logique. Celle-ci se fait au travers des deux langages  $l_{io}^s$  et  $l_{outil}^s$ , et donc s'exprime selon l'évaluation du contrôle de l'outil mixte sur l'outil logique et de l'état interne de l'outil mixte.

Ces quatre facettes de l'évaluation affinent la classification selon l'évaluation augmentée et fournit un cadre utile à la conception. Par exemple pour TROC, lorsqu'on ramasse une vignette,

- L'évaluation de l'état interne de l'objet de la tâche est assurée par un langage  $l_{objet}^s$  non nul, qui affiche l'image dans le casque à vision transparente à l'intérieur du cube, et l'affiche sur la tablette PC lorsqu'il est ramassé.
- L'évaluation des états internes des outils mixtes ( $l_{outil}^s$  de la parole et du cube) n'est pas possible. D'une part il manque un retour lexical de ce que le système a reconnu de la parole, et d'autre part le cube devrait être marqué comme reconnu.

- L'évaluation du contrôle de l'outil mixte sur l'outil logique ( $l_{io}^s$ ) n'existe pas,
- L'évaluation du contrôle de l'outil logique sur l'objet de la tâche ( $l_i^s$ ) n'existe pas non plus : l'utilisateur n'a aucun moyen de savoir quelle commande est appliquée avant d'observer le changement d'état de la vignette.

Nous notons enfin que l'évaluation est liée aux propriétés d'observabilité et d'honnêteté de l'interface. Aussi notre affinement de l'évaluation permet en phase de conception d'étudier précisément l'observabilité et l'honnêteté (espace de caractérisation au paragraphe 3.2).

### 3.3.4 Objet mixte

Là où les approches précédentes décrivaient le lien entre la partie physique et la partie numérique d'un objet mixte de façon minimale, notre modèle permet de caractériser en terme de modalités de liaison en entrée et en sortie la manière dont les deux parties sont assemblées pour former l'objet mixte. Une modalité de liaison comme une modalité d'interaction se caractérise en terme de dispositif et de langage. Nous réutilisons alors les caractéristiques des dispositifs et langages issues de travaux sur la multimodalité pour les appliquer au cas d'un objet mixte.

Les dispositifs des modalités de liaison sont caractérisés par les grandeurs captées ou restituées : débit d'information et nombre de dimensions.

Au niveau des langages, les propriétés de Bernsen peuvent caractériser l'objet mixte : si, par exemple, le langage est analogique, la compréhension de l'augmentation et de l'objet mixte dans son entier par l'utilisateur pourra potentiellement être meilleure : l'analogie, comme la métaphore, réduit le fossé sémantique de la théorie de l'action de Norman. Nous pouvons aussi caractériser les langages en précisant s'ils sont précis/vagues, partiels/globaux, et s'ils contiennent ou non une fonction de déformation.

Ainsi l'objet mixte lui-même, objet de base de notre modèle, est décrit et caractérisé précisément en utilisant des résultats liés à la multimodalité. De plus nous pouvons également étudier au sein d'un objet mixte la notion de fusion de modalités élémentaires de liaison. La fusion de modalités fait l'objet de nombreux travaux en interaction multimodale. Par exemple la souris comprenant plusieurs capteurs ou dispositifs  $d_{objet}^e$  est décrite par la fusion des deux modalités de liaison comme le montre la figure 3.3.

Ainsi nous classons les objets mixtes en caractérisant les modalités élémentaires de liaison en entrée et en sortie, et en considérant les formes de composition des modalités de liaison. Le tableau 3.4 synthétise cette classification.

Notons que cette classification s'applique à tout objet mixte qu'il soit outil mixte ou objet de la tâche.

### 3.3.5 Outil d'interaction

Pour caractériser finement les outils d'interaction nous appliquons la même méthode que pour un objet mixte puisqu'un outil mixte est décrit par des modalités qui cette fois sont des modalités d'interaction  $((outilMixte, l_{io}), l_i)$ . Comme le montre le tableau 3.5, nous pouvons classer les outils d'interaction en considérant ces trois niveaux. De plus comme pour les objets mixtes du paragraphe précédent, il convient aussi de caractériser les fusions qui peuvent intervenir à deux niveaux :

- fusion d'outils mixtes pour un langage  $l_{io}$ ,
- fusion de langages  $l_{io}$  pour un langage  $l_i$ .

<b>modalité de liaison en entrée</b>			
absente			
élémentaire			
	caractéristique du dispositif		
	caractéristique du langage		
composée			
	spatial		
	temporel		
	articulatoire		
	syntaxique		
	sémantique		
	modalité simple (1)		
		caractéristique du dispositif	
		caractéristique du langage	
	modalité simple (n)		
		caractéristique du dispositif	
		caractéristique du langage	
<b>modalité de liaison en sortie</b>			
absente			
élémentaire			
	caractéristique du dispositif		
	caractéristique du langage		
composée			
	spatial		
	temporel		
	articulatoire		
	syntaxique		
	sémantique		
	modalité simple (1)		
		caractéristique du dispositif	
		caractéristique du langage	
	modalité simple (n)		
		caractéristique du dispositif	
		caractéristique du langage	

<b><i>Outil mixte</i></b>		
cf. classification d'un objet mixte		
<b><i>Langage d'interaction de l'outil</i></b>		
entrée		
	caractéristique du langage	
sortie		
	caractéristique du langage	
<b><i>Langage d'interaction</i></b>		
entrée		
	caractéristique du langage	
sortie		
	caractéristique du langage	

TAB. 3.5 – Classement d'un outil d'interaction

### 3.3.6 Synthèse des taxonomies

Nous avons démontré le fort pouvoir taxinomique de notre modèle (lié à son pouvoir descriptif) et en comparaison des taxonomies existantes. En effet, notre modèle permet de distinguer des systèmes qui dans les taxonomies appartiennent à la même classe. Par exemple, considérons les systèmes KARMA et CASPER. Tous deux sont classés parmi les systèmes de RA à dont l'évaluation est augmentée. Selon notre modèle, le premier n'a pas d'outil d'interaction, au contraire du second. De même les systèmes NaviCam et KARMA n'ont pas été distingués jusqu'à aujourd'hui. Selon notre modèle, ils le sont par simple comparaison des objets de la tâche : les modalités de liaison  $y$  sont différentes. En effet en étudiant les langages de l'objet en sortie  $l_{objet}^s$ , nous notons que celui de l'objet de la tâche dans KARMA est graphique, alors que celui de l'objet de la tâche dans NaviCam est le langage pseudo naturel écrit. On constate ainsi que des systèmes très proches sont distingués grâce à notre modèle.

## 3.4 Conclusion

Ce chapitre a été consacré à l'exposé de notre modèle d'interaction mixte. Nous avons d'abord modélisé l'élément de base, un objet mixte, puis l'interaction homme-machine où des objets mixtes interviennent.

Nous avons ensuite souligné les apports du modèle en étudiant son pouvoir descriptif, comparatif et taxinomique.

Tout au long de ce chapitre nous avons souligné l'intérêt de notre approche qui consiste à exploiter des résultats issus de travaux sur l'interaction instrumentale et multimodale pour modéliser les systèmes mixtes.

## Chapitre 4

# Synthèse et conclusion la partie I

Comme expliqué en introduction, notre démarche originale consiste à étudier les systèmes mixtes comme des systèmes mettant en œuvre des interactions multimodales et instrumentales. Aussi dans le chapitre 2 consacré à l'existant, outre un état de l'art des systèmes mixtes, nous avons présenté les résultats les plus importants sur l'interaction multimodale et instrumentale. Toutes les approches exposées au chapitre 2 nous servent de fondement pour notre modèle d'interaction mixte qui fait l'objet du chapitre 3. Notre modèle unifie donc des approches issues d'axes de recherche distincts du domaine de l'Interaction Homme-Machine.

L'élément de base de notre modèle d'interaction mixte est l'objet mixte. Une fois l'objet mixte défini et ses cas remarquables exposés, nous avons modélisé l'interaction homme-machine faisant intervenir des objets mixtes. Ce modèle de l'interaction, comme tout modèle de l'interaction, a été discuté à la lumière des requis fondamentaux d'un modèle de l'interaction. Le premier pouvoir requis est comparatif. Nous avons examiné, par le biais de propriétés ergonomiques, dans quelle mesure le modèle de l'interaction mixte pouvait contribuer en phase de conception d'un système mixte, en permettant au concepteur de choisir entre plusieurs techniques d'interaction alternatives, et ceci avant même le développement du système. Ensuite, son pouvoir taxinomique, pour classer les systèmes existants et mieux comprendre l'espace solution des systèmes mixtes, a été étudié. Nous avons montré que des systèmes mixtes non distingués par les taxonomies existantes le sont grâce à ce modèle. Ceci est un résultat important en vue d'une meilleure compréhension de cet axe de recherche en pleine effervescence.

Nous concluons donc que le mariage des points de vue "interaction instrumentale", "interaction multimodale" et "interaction mixte" permet de faire contribuer chaque domaine à l'avancée de l'autre mais surtout d'envisager une méthode de conception des interfaces homme-machine regroupant l'ensemble de ces paradigmes d'interaction.

Dans la partie suivante, nous nous focaliserons sur l'étude de RAZZLE, un système mixte sur supports mobiles, selon notre modèle de l'interaction mixte. L'objectif visé est double : mettre à l'épreuve le pouvoir comparatif et étudier le pouvoir génératif de notre modèle.

Deuxième partie

**Espace Applicatif**

## Chapitre 5

# RAZZLE : un jeu de RA mobile et collaboratif

À la suite de MAGIC [28] [29] et de TROC [26] [29], deux systèmes mixtes, collaboratifs et mobiles, le jeu RAZZLE a été développé et exposé comme démonstrateur de ce type de système lors de la fête de la science les 15 et 16 octobre 2004 où la photo 5.1 a été prise. Son développement a donc précédé le modèle de l'interaction mixte. Par conséquent nous avons modélisé et étudié RAZZLE a posteriori. Ceci nous a permis d'étudier le pouvoir comparatif du modèle d'interaction mixte, et de le valider expérimentalement grâce aux évaluations ergonomiques que nous avons menées. Le pouvoir génératif du modèle a aussi été mis à l'épreuve pour produire les spécifications de la seconde version de RAZZLE. C'est donc un système dont nous allons retracer l'évolution grâce au modèle d'interaction mixte. Nous décrivons d'abord la première version de RAZZLE, puis nous modélisons les techniques d'interaction de cette première version de RAZZLE. Ensuite, nous décrivons l'évaluation ergonomique qui a été menée et qui a confirmé nos hypothèses faites grâce au pouvoir comparatif du modèle. Puis, au chapitre 6 nous démontrons le pouvoir génératif du modèle en détaillant les techniques d'interaction conçues pour la seconde version de RAZZLE. Nous donnons enfin des précisions quant à l'implémentation de ces techniques d'interaction conçues grâce à notre modèle d'interaction mixte.

### 5.1 Description

RAZZLE est un jeu mixte, mobile et collaboratif pour deux joueurs. Le but du jeu est de ramasser les pièces d'un puzzle numérique qui sont ancrés dans le monde physique. Les joueurs doivent donc se déplacer pour explorer le terrain de jeu. RAZZLE appartient à la classe des systèmes de virtualité augmentée. Les pièces mixtes sont dispersées partout sur le terrain de jeu. L'utilisateur peut interagir avec les pièces mixtes grâce à la technique du "terrain augmenté" [29]. Cette technique permet aux utilisateurs de voir des objets mixtes avec une localisation dans l'espace physique, s'ils sont bien orientés et suffisamment proches de l'objet. Un joueur ne peut interagir qu'avec les quatre pièces qu'il cherche, identifiées par la couleur, et il est le seul joueur à chercher ces pièces. Une fois ramassées et donc retirées du terrain de jeu, les pièces sont "portées" par le joueur, jusqu'à ce qu'il les dépose dans une zone particulière du monde physique, commune aux deux joueurs, et appelée "zone de dépose". Les joueurs se déchargent alors automatiquement de leur pièces de puzzle en entrant tout simplement dans la zone de dépose. Puis, ces pièces sont ajoutées au puzzle, pour

pouvoir voir le puzzle résultant. Le jeu s'arrête quand le puzzle est complet.

Nous décrivons dans ce qui suit les plate-formes matérielle et logicielle de RAZZLE, et ensuite nous détaillons les modalités disponibles pour interagir avec les pièces de puzzle mixtes.

### 5.1.1 Matériel

Chaque joueur dans RAZZLE porte un ordinateur portable où sont installés Java et la bibliothèque jogl (Open GL pour Java [20]). Le terrain de jeu est couvert par un réseau sans fil, ce qui permet à l'ordinateur portable du joueur d'être connecté au serveur central de RAZZLE. Un casque à vision semi transparente est également connecté à l'ordinateur portable. De plus, l'ordinateur portable possède un port série RS-232 pour connecter un capteur d'orientation XSENS MT9. La figure 5.1 est le plan d'un joueur équipé du matériel décrit. À l'ordinateur portable est connecté une trackball, qui permet à l'utilisateur de bouger un curseur affiché dans le casque par dessus le monde physique, et la trackball permet aussi de cliquer. Enfin, nous utilisons la technique du magicien d'Oz pour simuler l'information sur la position du joueur. Dans la figure 5.2, un compère suit le joueur sur le terrain de jeu pour mettre à jour sa position.



FIG. 5.1 – Un enfant jouant avec RAZZLE lors de la fête de la science 2004



FIG. 5.2 – Une scène du jeu RAZZLE avec le magicien d'Oz au premier plan et un joueur au fond

### 5.1.2 Interface de RAZZLE dans le casque semi transparent

La technique du "terrain augmenté" [29] est basé sur la semi transparence du casque. À travers le casque, les pixels noirs sont transparents, et plus les pixels sont sombres, plus ils sont transparents. Le programme affiche donc dans le casque une scène 3D, utilisant jogl, qui représente le terrain de jeu. Pourtant, comme les objets du monde physique (portes, murs, etc.) sont déjà visibles à travers le casque, ils ne sont pas représentés. Seules les pièces de puzzle le sont.

La première partie de l'interface utilisateur en sortie est la superposition des informations numérique au monde physique. C'est donc ici l'affichage des pièces de puzzle selon le point de vue de l'utilisateur. La seconde partie de l'interface utilisateur en sortie est le tableau de bord affiché en bas du casque à vision semi transparent. Il indique au joueur les quatre pièces qu'il doit trouver. Si la pièce n'a pas encore été trouvée, elle est visible sur la tableau de bord, et entourée de flèches qui indiquent la direction dans laquelle elle se trouve. Ces flèches ne clignotent seulement quand un des deux joueurs a vu la pièce. Si une pièce a été ramassée, elle est grisée sur le tableau de bord, et n'est plus entourée de flèches. Si la pièce a été déposée dans la zone de dépose, la place qu'elle occupait dans le tableau de bord est laissée vide, afin d'assurer la stabilité d'affichage du tableau de bord. Enfin, à droite du tableau de bord, on indique le temps de jeu. L'interface dans le casque de RAZZLE est montrée à la figure 5.3. Les pixels noirs sont transparents dans le casque et donc le monde physique est visible à travers les pixels noirs.

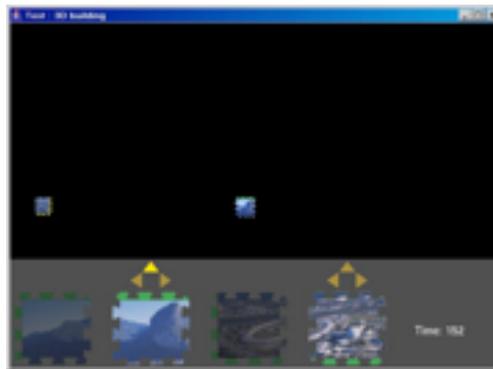


FIG. 5.3 – Interface de RAZZLE dans le casque semi transparent

Après avoir présenté le matériel et la plate-forme logicielle de RAZZLE, nous allons maintenant détailler les modalités pour interagir avec les pièces de puzzle mixtes.

## 5.2 Techniques d'interaction pour le ramassage d'une pièce

### 5.2.1 Description

Plusieurs tâches vont être effectuées lorsque l'utilisateur interagit dans RAZZLE. Le joueur cherche les pièces de puzzle, dépose les pièces de puzzle, ou ramasse une pièce de puzzle. Dans le premier et le dernier cas, la pièce de puzzle mixte est l'objet de la tâche. Nous la modélisons grâce au modèle de l'interaction mixte. Le schéma 5.4 présente cette modélisation. Parmi les propriétés physiques de la pièce, sa position et son orientation sont captées, et sont ensuite interprétées pour

participer aux propriétés numériques de la pièce mixte, dont le point de vue de la pièce. Parmi ces propriétés numériques, l'identité de la pièce et son statut sont interprétés par le langage  $l_o^s$  pour générer une image adéquate. Cette image sera ensuite affichée parmi les propriétés physiques de la pièce grâce au casque semi transparent.

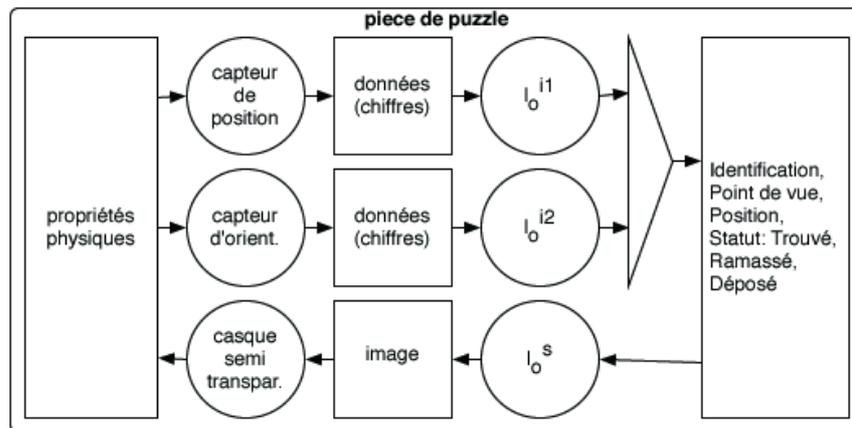


FIG. 5.4 – Modélisation de la pièce de puzzle mixte dans RAZZLE

Nous venons donc de modéliser la pièce de puzzle mixte de RAZZLE. Cette pièce est l'objet de la tâche lorsque l'utilisateur veut ramasser une pièce ou cherche une pièce. Dans cette partie, on choisit la tâche "ramasser une pièce" sur laquelle on va appliquer le modèle et apprécier son pouvoir comparatif et génératif. Nous examinons tout d'abord les trois modalités d'interaction proposées à l'utilisateur pour ramasser une pièce dans la première version de RAZZLE.

1. La première modalité ( $m_1$ ) proposée est inspirée du pouvoir magique du monde des jeux vidéos. L'utilisateur s'approche de la pièce, et s'il reste suffisamment longtemps, c'est-à-dire quelques secondes, à proximité dans un rayon de un mètre autour de la pièce, alors la pièce change de statut et est "ramassée". Le schéma 5.5 illustre cette modalité de ramassage d'une pièce.
2. La seconde modalité ( $m_2$ ) proposée est quant à elle inspirée du monde physique, où l'utilisateur, pour ramasser un objet, s'en approche et s'en saisit d'un geste de la main. Cette technique n'a pas été implémentée et est simulée par le même magicien d'Oz. Cette modalité de ramassage d'une pièce mixte est schématisée à la figure 5.6.
3. La troisième et dernière modalité ( $m_3$ ) pour ramasser une pièce est inspirée des interfaces graphiques, où l'utilisateur, pour sélectionner un objet, amène le curseur dessus et clique avec un bouton de la souris. Dans RAZZLE, l'utilisateur, s'il voit une pièce dans son casque, peut amener le curseur affiché dans le casque sur la pièce en manipulant la trackball, et cliquer pour la ramasser. Cette modalité est modélisée à la figure 5.7.

### 5.2.2 Étude du pouvoir comparatif du modèle d'interaction mixte

Le pouvoir comparatif d'un modèle d'interaction permet au concepteur de comparer plusieurs alternatives de conception alors que le système est encore en phase de conception. Nous avons discuté

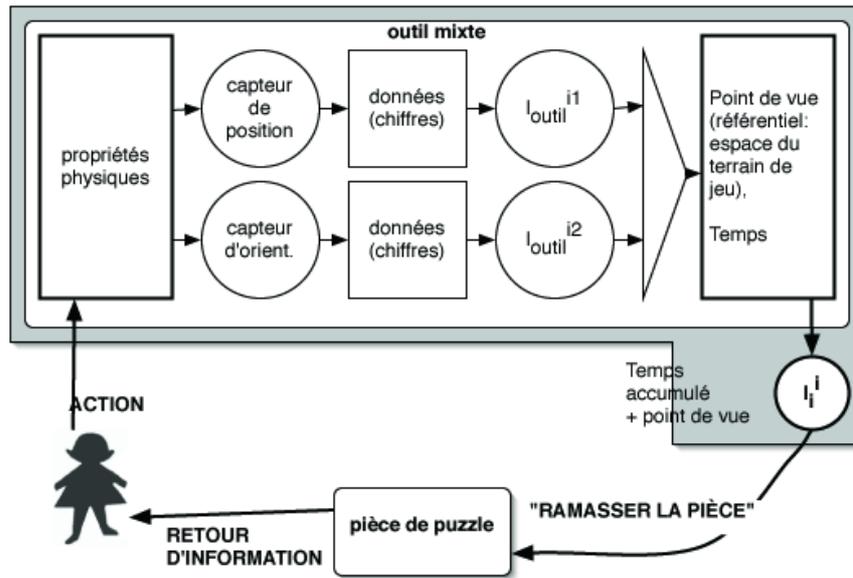


FIG. 5.5 – ( $m_1$ ) : Ramasser une pièce en s'approchant à moins d'un mètre pendant quelques secondes

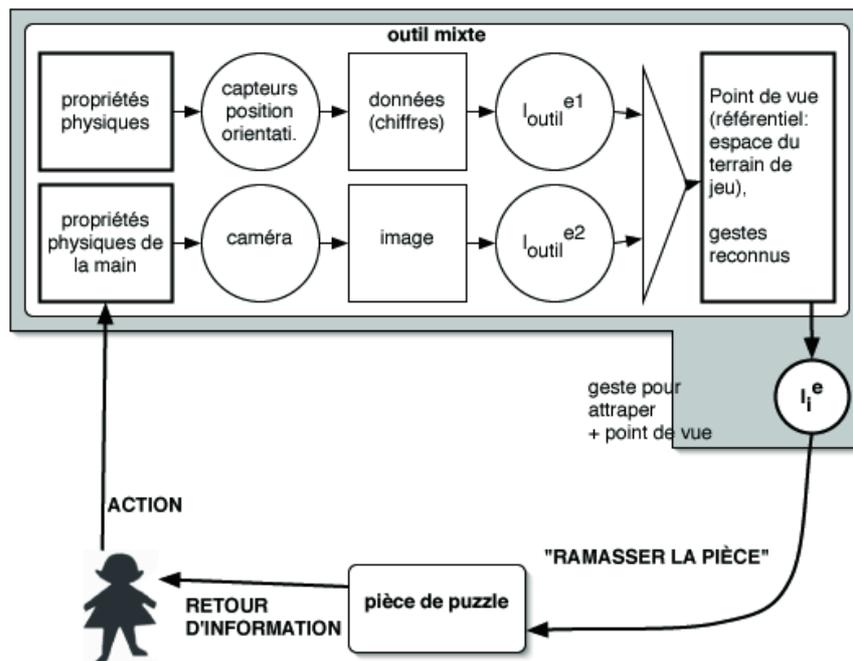


FIG. 5.6 – ( $m_2$ ) : Ramasser une pièce en s'approchant et l'attrapant d'un geste de la main

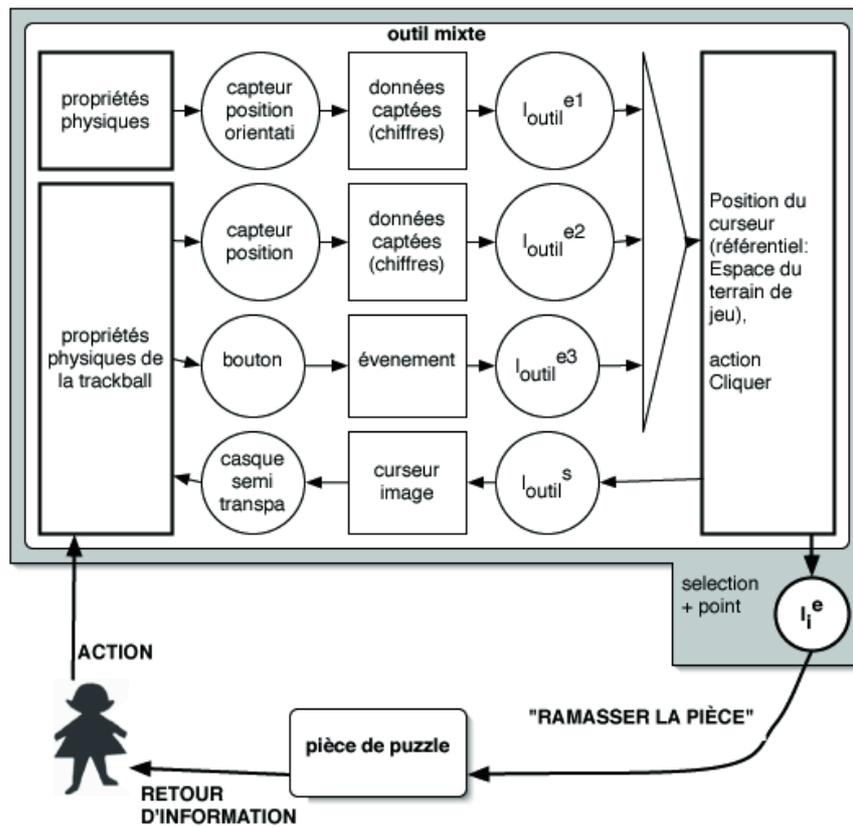


FIG. 5.7 – ( $m_3$ ) : Ramasser une pièce en positionnant le curseur de la souris dessus et en cliquant

celui-ci dans le paragraphe 3.2 du chapitre précédent au regard de propriétés ergonomiques. Grâce au modèle d'interaction mixte, nous pouvons faire certaines hypothèses concernant l'utilisabilité du système. Les trois modalités décrites ci-dessus sont celles qui ont été proposées dans la première version de RAZZLE. Nous évaluons à présent ces modalités à la lumière du modèle d'interaction mixte.

L'observabilité est une des propriétés que nous avons identifiée dans la partie précédente. Nous examinons les trois modalités précédentes de ce point de vue. Nous remarquons que la modélisation de la première technique ( $m_1$ ) de la figure 5.5 pour ramasser une pièce n'a aucune modalité de sortie, hormis le retour d'information qui permet d'évaluer si la pièce a été ramassée. Le contrôle de l'outil d'interaction sur l'objet de la tâche peut donc être difficile à évaluer pour l'utilisateur.

La continuité est une autre propriété que nous avons explicitée dans le modèle de l'interaction mixte. On peut donc remarquer l'absence de continuité spatiale parmi les propriétés physiques de l'outil mixte "trackball". Ceci est susceptible d'avoir une influence sur l'interaction entre le système RAZZLE et l'utilisateur.

Une fois ces hypothèses faites, elles nécessitent une évaluation ergonomique pour être validée. Nous décrivons les expérimentations que nous avons menées.

### 5.3 Évaluation ergonomique en magicien d'Oz

Les modalités décrites dans la partie précédente ont été évaluées sur le plan ergonomique pendant plusieurs séries d'expérimentations. Le but de l'évaluation ergonomique était d'observer comment les utilisateurs utilisaient chacune des trois modalités pour ramasser une pièce de puzzle dans RAZZLE. Les trois modalités disponibles étaient : ramasser la pièce en s'en approchant et en attendant quelques secondes ( $m_1$ ), ramasser la pièce en s'en approchant et en faisant un geste de la main pour l'attraper ( $m_2$ ), et enfin ramasser la pièce en plaçant le curseur de la souris dessus et en cliquant ( $m_3$ ). Cette évaluation expérimentale a été menée avec l'aide d'un ergonome extérieur au laboratoire et embauché pour mettre en place le protocole selon nos objectifs et analyser les résultats observés.

#### 5.3.1 Description du protocole expérimental

Deux phases de tests ont été définies : individuel et par pair. Des critères ont été définis pour mener l'analyse, dont les nombres d'utilisateur minimum pour avoir des données fiables, mais aussi la façon dont les données ont été enregistrées et analysées. Deux techniques ont été utilisées pour l'analyse des tests utilisateurs. La première, le protocole "Penser tout haut" ("Thinking Aloud"), consiste à demander aux utilisateurs de parler à haute voix et de formuler leurs pensées, leurs sentiments ou leurs opinions en même temps qu'ils sont en train d'accomplir leur tâche. Durant les expérimentations de RAZZLE, les utilisateurs étaient invités à parler de leurs prises de décision et leurs difficultés lors des tests. Ceci nous a permis de détecter les problèmes avec précision. La seconde technique utilisée consistait à interroger les utilisateurs après les tests. Grâce à ces entretiens, nous avons pu déterminer

- les caractéristiques des participants (âge, sexe, catégorie socio-professionnelle, expérience des jeux vidéos),
- l'opinion des utilisateurs sur le système.

Toutes les expérimentations ont été enregistrées à l'aide d'une caméra, et celles-ci incluaient les remarques des utilisateurs. Ainsi tout le cheminement des utilisateurs dans le terrain de jeu, et le temps passé pour accomplir la tâche ont été enregistrés. Enfin ces informations ont été analysées a posteriori grâce à une reconstitution de l'interaction par un logiciel dédié à l'analyse de RAZZLE.

Les sujets pour l'expérimentation ont été recrutés suivant deux profils : pour moitié, ils sont "joueurs", et l'autre moitié "non joueurs". Les "joueurs" ont l'habitude de jouer avec des jeux en 3D, alors que les "non joueurs" n'avaient que peu d'expérience dans l'utilisation de ce type de système. Pour les deux groupes de participants, c'était leur première expérience avec système mixte. En tout, 10 utilisateurs ont pris part aux expérimentations en jouant seuls, et 14 ont joué par pair. Les utilisateurs étaient jeunes pour la plupart, tous entre 18 et 40 ans, et la moyenne d'âge était de 26 ans. La proportion d'hommes est plus importante : 15 hommes parmi 24 utilisateurs. De plus, 18 utilisateurs étaient informaticiens, cette forte proportion s'expliquant par le fait qu'ils aient été recrutés parmi le personnel du laboratoire.

Pour contrôler le processus expérimental, deux personnes étaient présentes lors des tests avec les utilisateurs : une ergonome et un informaticien pour assurer le travail du compère. Tous les utilisateurs ont effectué la même tâche : chercher 4 pièces d'un même puzzle parmi les 16 pièces accrochées sur les murs du terrain de jeu, puis les déposer dans la zone de dépôt. Dans les expérimentations par pair, les utilisateurs avaient une contrainte supplémentaire : ils devaient respecter une limite de temps.

### 5.3.2 Résultats

Les données issues des expérimentations nous ont permis de comparer les techniques d'interaction disponibles et modélisées précédemment grâce au modèle d'interaction mixte. Le tableau 5.1 montre comment les 40 pièces ont été ramassées pendant les expérimentations, c'est-à-dire 20 pièces par profil ("joueur" et "non joueur").

<i>profil</i>	$(m_1)$ attendre	$(m_2)$ attraper à la main	$(m_3)$ cliquer
<i>joueur</i>	12	8	0
<i>non joueur</i>	11	6	3

TAB. 5.1 – Utilisation des techniques d'interaction

Ces chiffres laissent supposer que la plupart des participants ont préféré attendre près d'une pièce ( $m_1$ ), et que beaucoup ont choisi d'attraper la pièce avec la main ( $m_2$ ), au lieu de cliquer sur la pièce ( $m_3$ ), avec une légère différence entre les "joueurs" et les "non joueurs". Pourtant, les commentaires des utilisateurs nous forcent à nuancer ces résultats. Ils ont expliqué qu'ils n'avaient parfois aucune intention de ramasser une pièce en attendant à proximité ( $m_1$ ), et qu'ils ont été surpris en s'apercevant que la pièce était déjà ramassée. Parfois ils avaient même commencé un geste pour attraper la pièce ( $m_2$ ), et dans ce cas, ils disent explicitement qu'ils avaient l'intention d'utiliser  $m_2$ , mais qu'ils n'ont pas eu le temps. Ceci confirme donc l'hypothèse faite plus haut quant à l'observabilité de  $m_1$ .

De plus, nous remarquons que les utilisateurs ont significativement évité de se servir de la trackball pour cliquer sur les pièces ( $m_3$ ). En effet, ils ont expliqué que changer leur point de vue pour prendre l'outil en main et le contrôler les perturbait. Pourtant, un joueur a préféré  $m_3$ , parce

que cette technique d'interaction lui permettait de minimiser ses déplacements dans le terrain de jeu. Pour la plupart d'entre eux,  $m_3$  est une modalité qui nécessite de changer l'objet de son attention, du terrain de jeu vers l'outil, ce qui s'apparente à l'opérateur "Homing" de GOMS [8]. En effet, pour utiliser la trackball, la main de l'utilisateur doit être positionnée explicitement, ce qui nécessite l'attention de l'utilisateur. Ce problème avait été identifié sur la modélisation de la modalité  $m_3$  en considérant la continuité spatiale de l'outil mixte trackball. L'hypothèse faite grâce au modèle est donc confirmée par l'étude expérimentale.

Nous avons mis à l'épreuve jusqu'ici le pouvoir comparatif du modèle, qui permet, grâce à l'étude de propriétés, de mettre à jour certaines caractéristiques du système. Le pouvoir requis en phase de conception pour un modèle d'interaction n'est pas seulement comparatif, mais aussi génératif de nouvelles modalités d'interaction. Nous allons donc continuer notre étude de RAZZLE en concevant de nouvelles modalités d'interaction mixte grâce à notre modèle. L'objet du chapitre suivant est donc la conception et la réalisation de nouvelles modalités d'interaction dans RAZZLE.

## Chapitre 6

# Aspect génératif du Modèle de l'Interaction Mixte : conception et réalisation de techniques d'interaction

Dans ce chapitre nous expliquons d'abord succinctement comment utiliser notre modèle en phase de conception puis nous modélisons deux nouvelles techniques d'interaction dans RAZZLE. Nous abordons ensuite leur développement en considérant leurs conception et réalisation logicielles.

### 6.1 Conception avec le modèle d'interaction

Dans le cas général de la modélisation, présenté à la figure 3.17, les paramètres ou choix de conception sont les suivants :

- les dispositifs,
- les langages,
- choisir un outil dédié ou assigné à une tâche ou un outil générique invariant qui s'associe à différents outils logiques selon le contexte.
- De plus, on peut remplacer une unique modalité du cas général par une modalité composée, issue de la fusion de deux modalités. Les paramètres de ce mécanisme de fusion sont identifiés par des travaux sur la multimodalité rappelés au paragraphe 2.2.3.

Outre ces choix de conception, l'absence de certaines dimensions du modèle dans la description d'un système sert de support à la génération d'une nouvelle technique d'interaction. Comblé les lacunes permet d'explorer l'étendue des possibilités concernant les modalités d'interaction.

Nous avons identifié dans le cas le plus général de modélisation les choix de conception à faire. Nous montrons maintenant comment grâce au modèle et au choix de conception induits nous concevons de nouvelles techniques d'interaction dans le cas de RAZZLE.

### 6.2 Modalités d'interaction conçues pour RAZZLE

Ayant identifié conceptuellement grâce au modèle deux problèmes potentiels d'interaction et ayant vérifié expérimentalement ces deux problèmes, la conception de nouvelles techniques visent à les fixer.

Ainsi nous concevons deux nouvelles techniques d'interaction pour minimiser les problèmes d'observabilité et de continuité spatiale.

Tout d'abord, nous considérons le manque d'observabilité du contrôle de l'outil mixte sur l'objet de la tâche et le manque d'observabilité de l'état interne de l'outil mixte. Pour concevoir une nouvelle technique, nous nous basons sur la modélisation initiale de  $m_1$  de la figure 5.5. Nous étudions alors l'ajout d'une modalité de sortie. Plusieurs solutions sont possibles en considérant des dispositifs et des langages différents.

On permet donc à l'utilisateur d'apprécier le contrôle de l'outil mixte sur l'objet de la tâche en ajoutant la *réponse* de l'objet de la tâche via le langage  $l_i^s$ , qui transforme le status "sélectionné" de la pièce en quatre flèches jaunes qui désignent la pièce localisée dans l'espace du terrain de jeu. On aurait pu tout aussi bien utiliser un autre langage, et entourer la pièce sélectionnée, ou écrire sur la pièce "sélectionnée". On aurait pu aussi changer le dispositif, et plutôt que d'écrire via le casque semi transparent, utiliser la modalité (*ecouteurs, langage – pseudo – naturel*) pour dire à l'utilisateur "pièce sélectionnée".

De plus, on ajoute une réaction de l'outil mixte, qui permet à l'utilisateur d'évaluer le temps passé à proximité de la pièce, grâce à un compte à rebours localisé sur la pièce. Là aussi, les alternatives peuvent être énumérées, comme par exemple dire à l'utilisateur le compte à rebours via les écouteurs, ou encore afficher une barre de progression sur la pièce via le casque semi transparent. Le schéma de cette nouvelle modalité d'interaction est présenté à la figure 6.1. La figure 6.2 présente la modélisation de la technique d'interaction développée. Cette nouvelle spécification de  $m_1$  définissent la seconde version de RAZZLE que l'on a développée.



FIG. 6.1 – schéma de la nouvelle modalité  $m_1$  corrigeant les défauts d'observabilité

Le défaut de continuité spatiale parmi les propriétés physiques de l'outil mixte qui a été décelé pour la modalité  $m_3$  nous pousse à trouver une nouvelle technique d'interaction, où le curseur pourra être déplacé avec la tête et le regard, pour ne pas gêner la continuité. On choisit donc de mettre un curseur fixe au centre de l'affichage dans le casque comme le montre la figure 6.3, que l'utilisateur déplacera avec les mouvements de sa tête. Le clic pour ramasser la pièce occasionnait

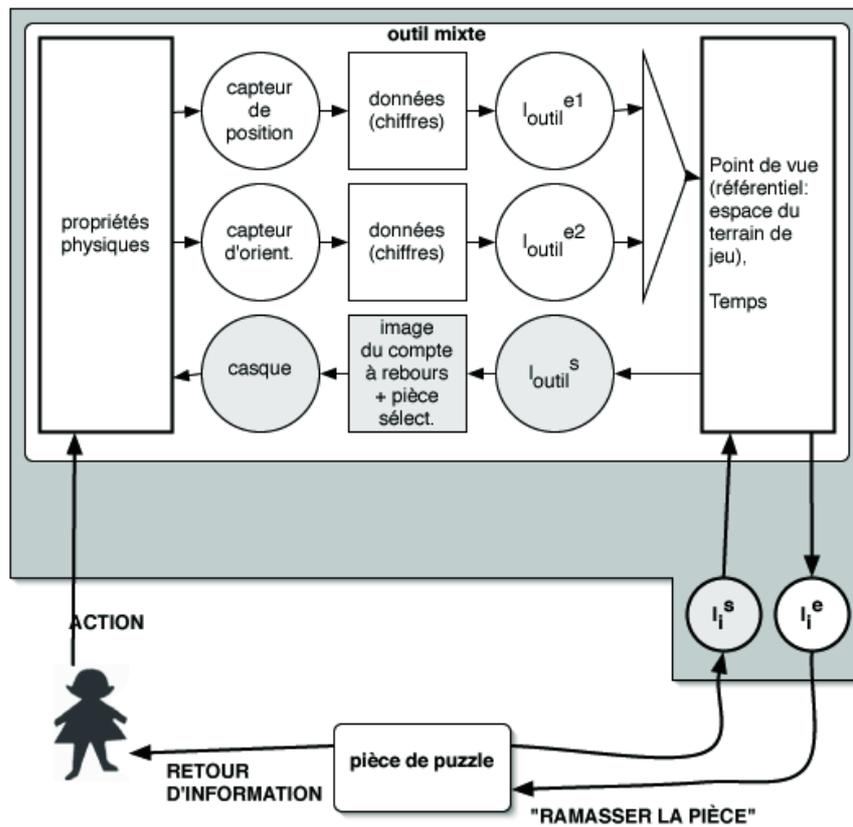


FIG. 6.2 – Nouvelle version de  $m_1$  : état sélectionné de la pièce observable et état de l'outil (temps passé) observable.

lui aussi une discontinuité spatiale. La modalité d'entrée conçue à la place n'est pas localisée : l'utilisateur spécifie en parlant qu'il veut ramasser la pièce sélectionnée par le curseur. Les modalités d'interaction (*capteur – de – localisation, référentiel*), (*capteur – d'orientation, référentiel*) et (*micro, langage – pseudo – naturel*) sont fusionnées en entrée. La nouvelle spécification de  $m_3$ , présenté à la figure 6.4, illustre donc le cas d'un outil mixte donc la modalité d'entrée n'est pas élémentaire.

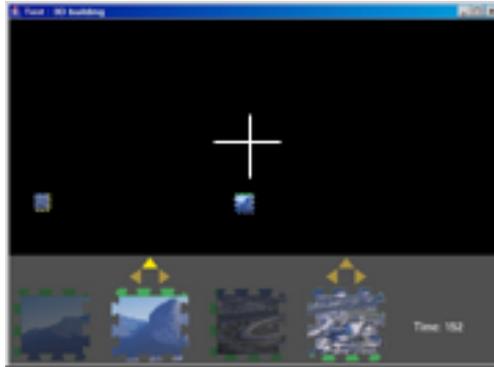


FIG. 6.3 – schéma de la nouvelle modalité  $m_3$  corrigeant les défauts de continuité spatiale

Dans le cycle de conception logicielle, après avoir conçu l'interaction grâce à notre modèle nous abordons les phases de conception et réalisation logicielles. Nous suivrons deux étapes pour exposer la conception et réalisation logicielles des techniques d'interaction conçues et leur intégration au sein du jeu RAZZLE. D'abord, nous décrivons l'architecture logicielle de l'application développée, et enfin nous exposons des éléments de programmation.

## 6.3 Architecture logicielle

La définition du terme "architecture" est l'art de la construction d'édifices selon un ensemble de règles établies. L'architecture d'un édifice, c'est aussi sa forme, sa structure. Dans le domaine de l'Interaction Homme Machine, il s'agit de la façon dont les composants logiciels ou matériels sont organisés et assemblés pour réaliser un système interactif. Un modèle d'architecture logicielle est utilisé lors de la phase de conception logicielle dans le cycle de vie du système interactif pour en concevoir l'architecture. La phase de conception logicielle globale consiste à traduire les spécifications externes (le résultat de la conception ergonomique selon notre modèle) sous forme logicielle en appliquant un modèle d'architecture logicielle. Pour RAZZLE, nous avons appliqué le modèle PAC-Amodeus [23] dont nous rapellons les points principaux au paragraphe suivant avant de présenter notre architecture de RAZZLE.

### 6.3.1 Modèle d'architecture PAC-Amodeus

Le modèle PAC-Amodeus est un modèle hybride combinant les modèles d'architecture Arch et PAC. Nous présentons dans un premier temps ces deux modèles d'architecture, pour ensuite décrire le modèle PAC-Amodeus.

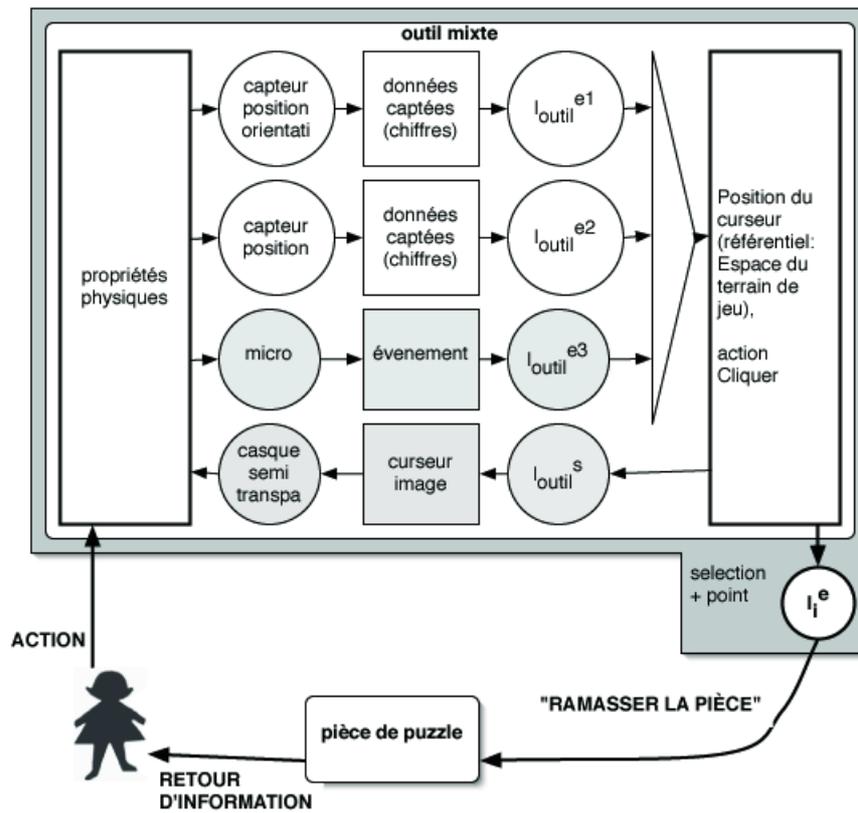


FIG. 6.4 – Nouvelle version de  $m_3$  : curseur fixe au centre du casque pour sélectionner une pièce et ramassage par la parole

## Le modèle Arch

Le modèle Arch [3] est un modèle d'architecture qui offre une décomposition fonctionnelle d'un système interactif. Le principe véhiculé par le modèle, comme le montre la figure 6.5, est de séparer l'interface utilisateur du Noyau Fonctionnel (logique de l'application). Ceci favorise la réutilisation et la portabilité du logiciel. Les composants de ce modèle d'architecture sont :

- Le Noyau Fonctionnel (NF) implémente les fonctionnalités et les concepts du domaine indépendamment de leur présentation.
- L'Adaptateur du Noyau Fonctionnel (ANF) joue un rôle de médiateur entre le Noyau Fonctionnel et le Contrôleur de Dialogue.
- Le Contrôleur de Dialogue (CD) est la pierre angulaire de cette architecture puisque ce composant a la charge de gérer l'enchaînement des tâches.
- La Présentation Logique (PL) joue aussi un rôle de médiateur entre le Contrôleur de Dialogue et la Présentation Physique. Ce composant, généralement assimilé à une boîte à outils graphique abstraite, permet l'indépendance du CD vis-à-vis des boîtes à outils graphique du niveau de la Présentation Physique.
- La Présentation Physique (PP) représente les interacteurs logiciels de bas niveaux (boîte à outils graphiques du système d'exploitation, contrôleur de périphérique) et les périphériques matériels (souris, écran, etc.).

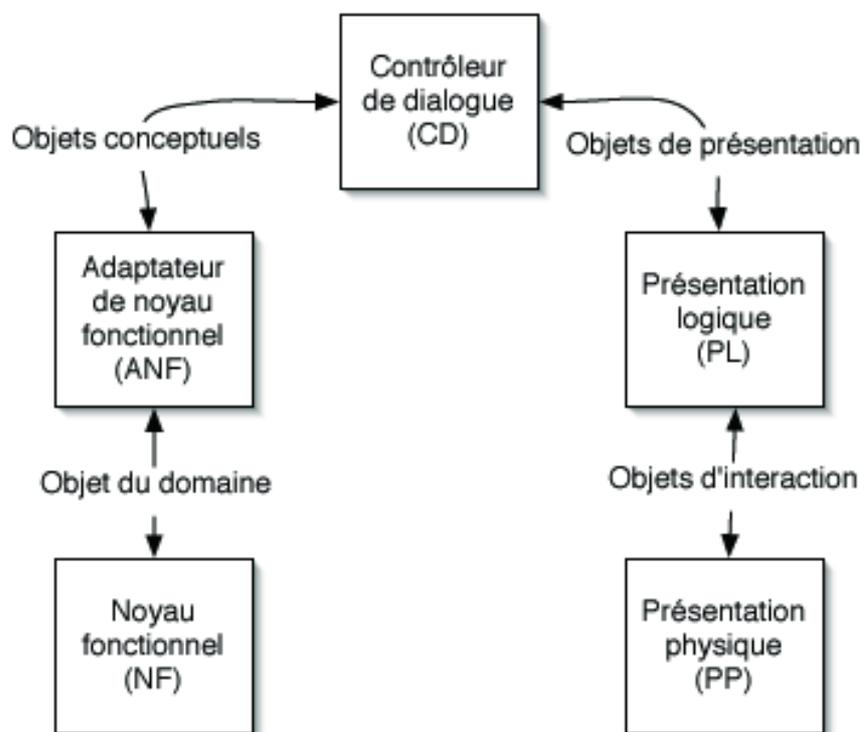


FIG. 6.5 – Le modèle d'architecture Arch

## Le modèle multi-agent PAC

Le modèle PAC (Présentation Abstraction Contrôleur) [10] est un modèle multi-agent pour la conception logicielle des systèmes interactifs. Il repose sur deux principes directeurs : le concept d'agents réactifs à facettes (système de traitement de l'information) et l'organisation hiérarchique de ces agents. Un agent PAC est constitué de trois facettes représentées à la figure 6.6 :

- La facette A, Abstraction, définit la compétence propre de l'agent, indépendamment de toute représentation.
- La facette P, Présentation, définit le comportement de l'agent pour l'utilisateur. Cette facette gère à la fois les communications de l'utilisateur vers l'agent et celles de l'agent vers l'utilisateur.
- La facette C, Contrôle, a deux rôles : exprimer les dépendances et assurer la traduction de formalismes entre les facettes A et P. Cette facette gère aussi la communication avec l'environnement de l'agent, c'est-à-dire d'autres agents PAC. Dans la hiérarchie d'agents PAC, un agent ne communique (par messages) qu'avec son agent père (celui auquel il est attaché et dont il dépend) et ses agents fils (ceux dont il est l'agent père).

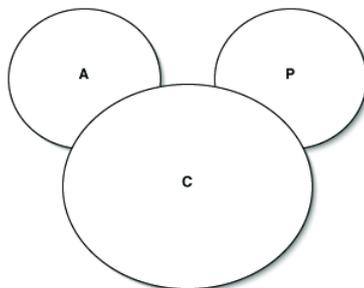


FIG. 6.6 – Le modèle d'architecture multi-agent PAC

## Le modèle hybride PAC-Amodeus

Le modèle d'architecture PAC-Amodeus [23] est l'intégration du modèle multi-agent PAC au modèle fonctionnel Arch. Dans le modèle PAC-Amodeus, le Contrôleur de Dialogue d'Arch est organisé en une hiérarchie d'agents PAC, comme l'illustre la figure 6.7. La facette A de chaque agent est en relation avec un objet du domaine situé dans l'adaptateur fonctionnel. La facette P de chaque agent pointe sur un objet de présentation dans la Présentation Logique. Les messages échangés entre les facettes Contrôle permettent d'articuler l'activité de l'utilisateur, de réguler les interactions.

## Le modèle PAC-Amodeus pour les collecticiels

L'architecture du jeu RAZZLE est une extension de PAC-Amodeus pour les collecticiels. Comme le montre la figure 6.8, les utilisateurs partagent un Noyau Fonctionnel commun, et chacun peut disposer d'un noyau fonctionnel privé, pour limiter les échanges sur le réseau.

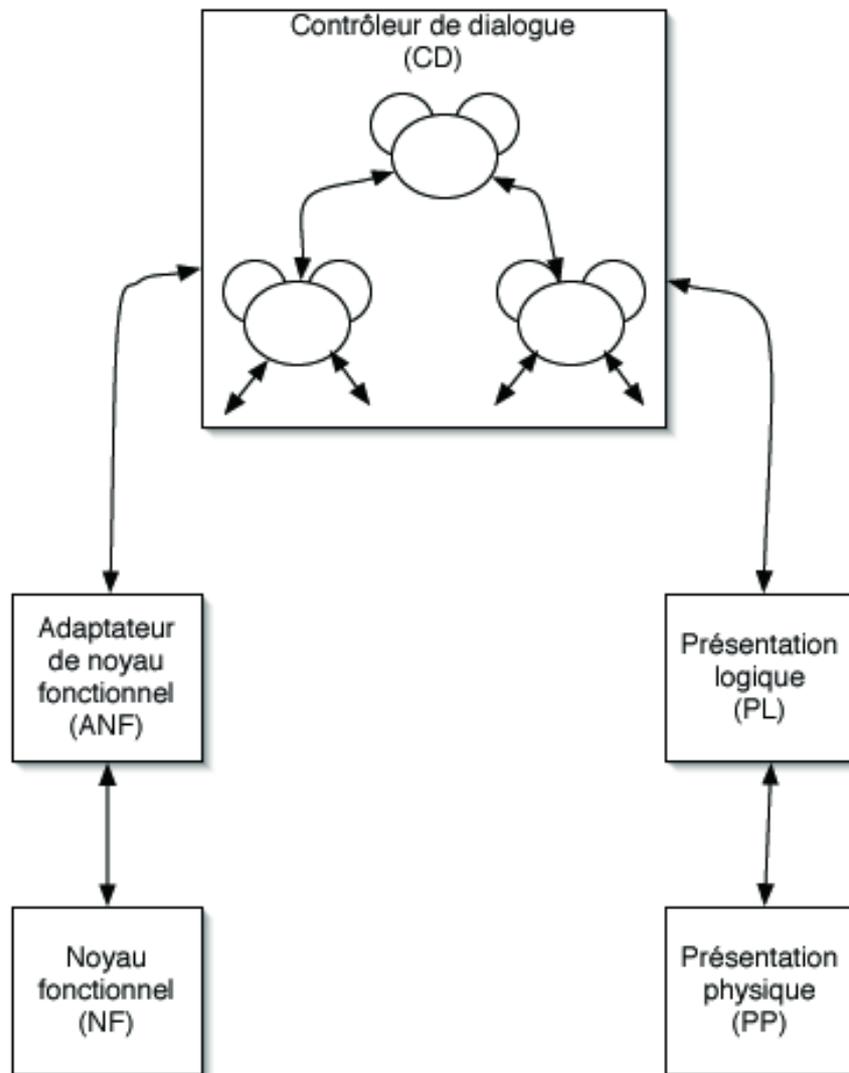


FIG. 6.7 – Le modèle d'architecture PAC-Amodeus

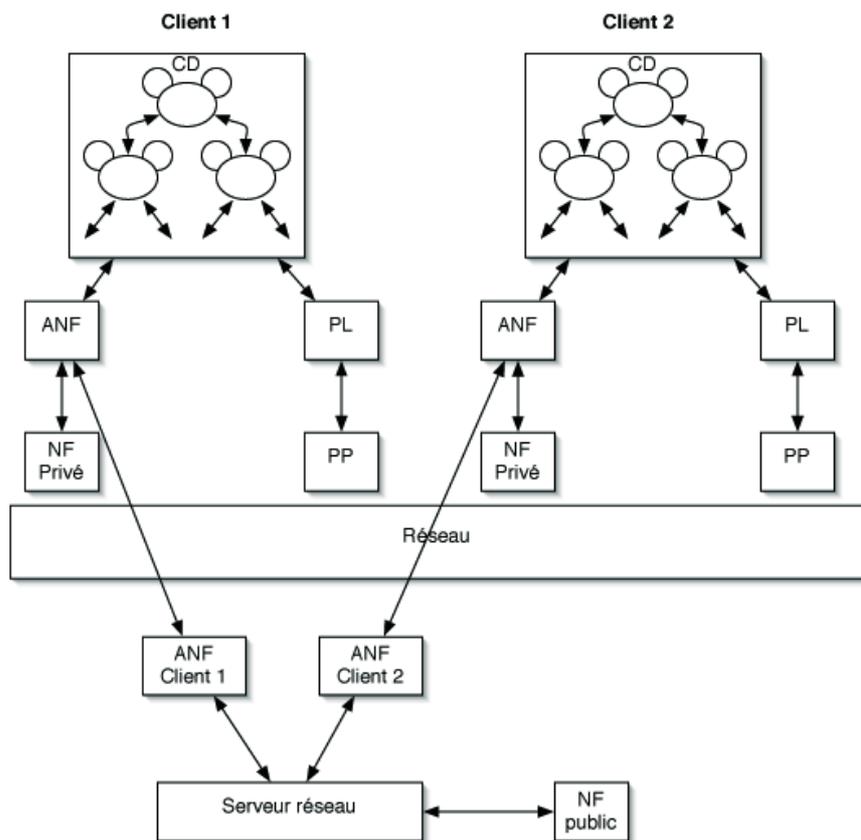


FIG. 6.8 – Le modèle d'architecture PAC-Amodeus pour les collecticiels

### 6.3.2 Architecture de RAZZLE

L'architecture de RAZZLE est présentée à la figure 6.9. Il y a un client par joueur et le serveur.

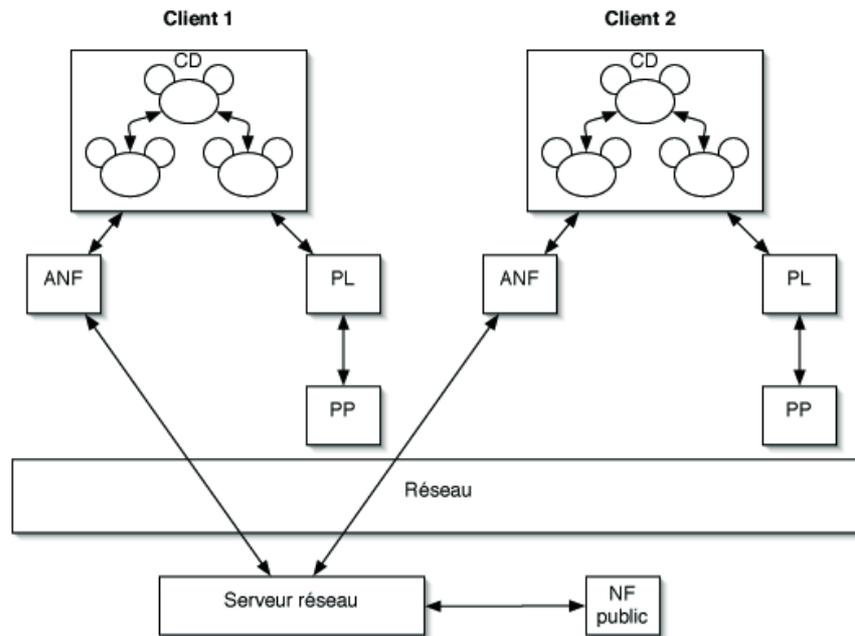


FIG. 6.9 – Le modèle d'architecture PAC-Amodeus de RAZZLE

Les clients n'ont pas de noyau fonctionnel (NF) privé, et leurs adaptateurs de noyau fonctionnel (ANF) font des appels aux méthodes distantes du NF public localisées sur le serveur. Le NF public gère les objets du domaine, c'est-à-dire ceux qui correspondent aux pièces de puzzle et au jeu.

Nous détaillons à la figure 6.10 l'architecture selon PAC du contrôleur de dialogue (CD). Cette hiérarchie d'agents PAC suit la description de l'interface décrite au paragraphe 5.1.2. L'agent "Tableau de bord" gère la partie du logiciel qui concerne le tableau de bord, c'est-à-dire qu'il reçoit des messages venant de l'adaptateur du noyau fonctionnel précisant si les quatre pièces du tableau de bord ont été vues, ramassées ou déposées. L'agent "Terrain augmenté" gère la partie qui concerne le terrain augmenté. Pour l'interaction en entrée, c'est donc lui qui reçoit et transmet le clique de l'utilisateur sur une pièce mixte de puzzle. Pour l'interaction en sortie, c'est lui met à jour l'affichage des pièces selon leur position dans le terrain de jeu et leur statut (ramassées ou non), et selon la position de l'utilisateur. L'agent père de la hiérarchie PAC gère la partie commune aux deux agents fils.

La présentation physique (PP) utilise la bibliothèque jogl (Open GL pour Java [20]) qui gère l'affichage des objets dessinés.

### 6.3.3 Programmation

#### Partie de l'architecture modifiée

Les modifications des modalités de RAZZLE ont été faites dans les parties présentations du modèle. Plus précisément, notre intervention dans le code de RAZZLE a consisté essentiellement

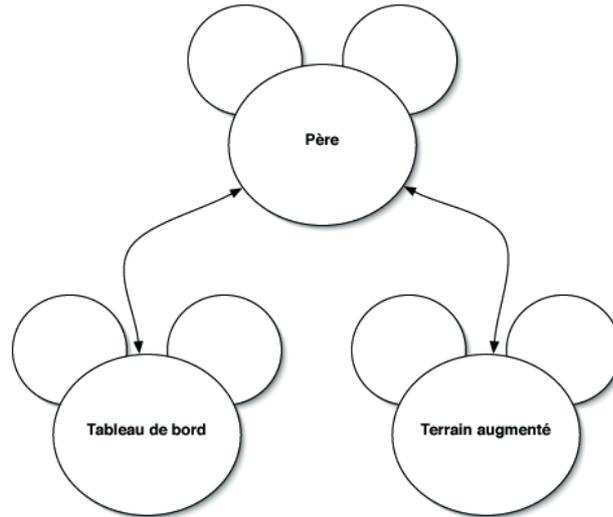


FIG. 6.10 – Le modèle d’architecture PAC pour le contrôleur de dialogue (CD) de RAZZLE

à compléter et/ou modifier le composant présentation logique (PL). De plus, l’ajout d’une modalité nouvelle, la reconnaissance de la parole repose sur des composants ICARE. Nous décrivons brièvement le modèle ICARE avant d’expliquer comment nous l’avons utilisé.

### Architecture ICARE

ICARE [7], pour Interaction-CARE (Complémentarité, Assignation, Redondance, Équivalence [24]), est une approche à composants pour développer des systèmes multimodaux. La place des composants ICARE au sein du modèle d’architecture Arch est rappelée à la figure 6.11. Il y a deux types de composants :

- les composants élémentaires qui permettent de définir des modalités élémentaires.
- les composants de composition qui permettent de définir l’usage combiné des modalités. Les composants de composition sont indépendants des modalités à composer.

Seul le premier type de composants, élémentaire, a été utilisé jusqu’à présent dans RAZZLE Basés sur la définition d’une modalité d’interaction comme étant l’association d’un dispositif  $d$  avec un langage d’interaction  $l$ ,  $(d, l)$ , deux types de composants ICARE élémentaires sont identifiés : les composants Dispositifs et les composants Langues d’interaction.

Un composant Dispositif représente une couche supplémentaire du pilote d’un dispositif physique. Il s’agit du niveau physique d’une modalité. Par exemple dans RAZZLE, le composant Dispositif MT9 (capteur d’orientation) abstrait les données fournies par le pilote du capteur. De même pour le clavier, le composant Dispositif correspondant abstrait les données capturées. Un composant Dispositif peut être lié à un composant Langage d’interaction afin de former une modalité élémentaire.

Un composant Langage d’interaction correspond au niveau logique d’une modalité d’interaction. Par exemple, un composant Langage d’interaction peut abstraire les données d’un composant Dispositif MT9 en une commande correspondant à l’orientation de l’utilisateur selon un système de coordonnées donné. Les composants ICARE Dispositifs et Langues d’interaction constituent donc

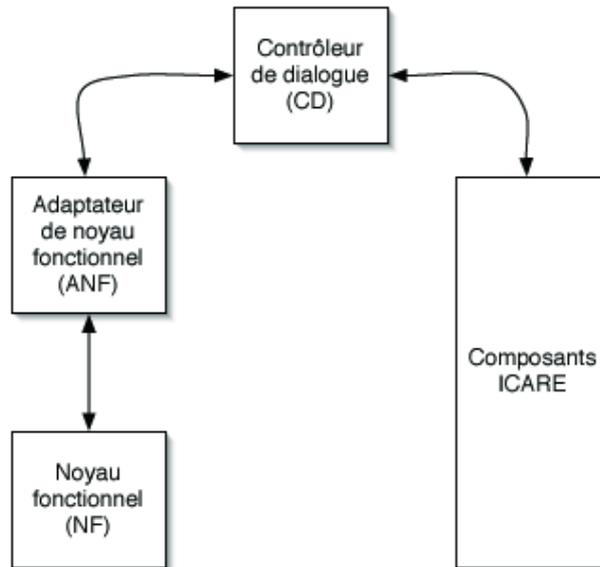


FIG. 6.11 – ICARE et Arch

des blocs élémentaires de construction pour définir des modalités.

Pour programmer les composants ICARE, la technologie à composants JavaBeans de Sun Microsystems [19] a été utilisée. La communication entre les composants est faite par génération d'événements et appels de méthodes. L'événement généré véhicule les données à traiter. Les propriétés des composants ICARE sont des attributs de classe qui peuvent être accessibles ou modifiés (getter/setter). Pour assembler deux composants ICARE, il est nécessaire que l'un des composants soit abonné aux événements générés par l'autre composant ICARE.

### Utilisation de ICARE dans RAZZLE

Le développement de la nouvelle modalité (*micro, langage pseudo naturel*) s'est fait selon ICARE, grâce aux composants élémentaires ICARE Dispositif et Langage d'Interaction. Pour cela, nous avons réutilisé et adapté du code utilisant deux composants Dispositif et Langage d'Interaction.

Le système RAZZLE existant n'étant que partiellement développé avec ICARE, nous projettons à l'avenir de compléter le développement de RAZZLE en suivant le modèle ICARE, et notamment pour composer les modalités [7]. Par contrainte de temps, la composition de la modalité parole avec la position du regard a été programmée directement sans utiliser le composant de composition "complémentarité" d'ICARE.

### Progression de la programmation

Notre programmation dans le code de RAZZLE se situe essentiellement dans la partie présentation logique, nous y avons modifié l'implémentation existante des modalités d'interaction, en entrée et en sortie, en suivant les nouvelles spécifications. Nous avons fait des modifications mineures dans

les autres parties, soulignant l'apport de l'architecture PAC-Amodeus choisie pour RAZZLE. Nous décomposons la phase de codage en quatre étapes.

1. La première partie de l'implémentation a consisté à afficher les quatre triangles jaunes présentés à la figure 6.1. Ceci a été implémenté dans une méthode, qui est appelée dans le code original lorsque la position de l'utilisateur est à moins d'un mètre de la pièce qui est en train d'être dessinée par le programme. De plus, nous avons implémenté une autre solution alternative aux flèches jaunes : en appuyant sur une touche dédiée du clavier, nous pouvons entourer d'un cercle la pièce sélectionnée à la place des triangles.
2. Dans cette même méthode nous affichons un chiffre pour rendre le compte à rebours observable dès que la position de l'utilisateur est à moins d'un mètre de la pièce, comme le présentait le schéma de la figure 6.1. Comme l'affichage de chaînes de caractères est limité par la bibliothèque `jogl`, nous avons affiché via une texture les caractères stockés sous forme d'une image, ce qui a nécessité un travail non prévu à l'origine. En effet, les tailles de police de base qui sont proposées étaient trop petites pour que le compte à rebours puisse être facilement observable par l'utilisateur. Nous avons donc contourné cette difficulté grâce aux textures OpenGL, qu'on peut étirer et donc afficher à la taille souhaitée.
3. Ensuite nous avons implémenté l'affichage d'un curseur fixe en forme de croix au centre du casque, comme nous le présentions à la figure 6.3. Quand le programme reçoit l'évènement "*clique souris*", la pièce qui se trouve sous le curseur en forme de croix est ramassée. Ceci est géré facilement grâce à la méthode de rappel `MousePressed(MouseEvent e)`.
4. Nous nous sommes ensuite familiarisé avec ICARE et des composants Dispositif et Langage d'Interaction qui ont été implémentés à l'origine pour une application utilisant la modalité parole. Le composant Dispositif utilise la technologie ViaVoice [31] pour reconnaître la parole, et le composant Langage d'Interaction, couplé avec un vocabulaire et une grammaire, nous permet de reconnaître le mot "*collect*" pour ramasser la pièce de puzzle. Ceci a nécessité l'importation de nouvelles classes dans le code de RAZZLE. Comme nous n'avons pas encore utilisé les composants de composition ICARE, nous implémentons manuellement la complémentarité (fusion) de la parole avec la position du regard ainsi que l'équivalence des modalités *clique souris* et *dire "collect"*.

Pour les deux premières parties de l'implémentation, nous avons découvert à la fois OpenGL et `jogl`, que nous avons dû maîtriser afin d'implémenter ces modalités. De même, nous avons découvert les composants ICARE pour implémenter la nouvelle modalité composée  $m_3$  utilisant la parole.

Comparées à la taille importante du code de RAZZLE, les modifications apportées sont légères et soulignent l'apport de l'architecture PAC-Amodeus. Néanmoins il convient de noter la complexité pour prendre en main un tel projet, l'installer, l'exécuter et le recompiler, mais aussi tracer les flux d'informations au sein de l'architecture pour localiser les modifications. En phase de développement, nous avons utilisé une version de jeu vidéo classique de RAZZLE, qui présentait l'avantage d'être moins contraignante. Puis, une fois les quatre étapes citées accomplies, nous avons intégré le code développé à la version mixte de RAZZLE pour obtenir la version 2 de RAZZLE.

## Chapitre 7

# Synthèse et Conclusion la partie II

Cette partie était consacrée à l'étude d'un système mixte particulier, RAZZLE. Notre objectif était de mettre le modèle de l'interaction mixte à l'épreuve sur un cas concret de système mixte. Nous avons pu modéliser les techniques d'interaction proposées, les étudier grâce au pouvoir comparatif du modèle de l'interaction mixte et ainsi émettre des hypothèses grâce aux propriétés identifiées que nous avons validées expérimentalement. Enfin, nous avons apprécié le pouvoir génératif du modèle de l'interaction mixte en concevant de nouvelles techniques d'interaction pour les inclure dans RAZZLE. L'étape suivante consiste maintenant à mener un deuxième ensemble de tests expérimentaux pour valider notre conception selon le modèle.

Outre la mise à l'épreuve de notre modèle sur un cas concret de conception, l'étude de RAZZLE et en particulier la réalisation logicielle nous a permis de mesurer l'apport de la plateforme ICARE pour le développement d'un système mixte. Le fait que la plateforme ICARE repose sur les concepts de dispositifs et langages explicites dans notre modèle nous permet d'entrevoir une perspective à long terme de plateforme de développement des systèmes mixtes qui expliciterait notre modèle d'interaction mixte.

# Chapitre 8

## Conclusion générale

### 8.1 Résumé des contributions

Malgré la multiplicité des études et démonstrateurs de réalité mixte, le cadre de compréhension des systèmes mariant les mondes numérique et physique reste flou. Dans notre étude, nous proposons d'adopter un point de vue unificateur sur la réalité mixte en considérant l'interaction instrumentale et multimodale. Nous étudions les modalités mises en jeu par le système mixte, focalisant ainsi sur l'interaction.

Notre travail contribue au domaine de l'Interaction Homme-Machine par la définition d'un nouveau modèle d'interaction, noté modèle de l'interaction mixte.

L'approche adoptée s'appuie sur une analyse conceptuelle des systèmes mixtes mais aussi de l'interaction multimodale et instrumentale (chapitre 2). En effet, l'originalité de notre approche réside dans l'étude de l'interaction mise en œuvre dans les systèmes mixtes du point de vue de l'instrument et de la multimodalité.

S'appuyant sur les résultats exposés dans le chapitre 2, nous décrivons ensuite notre modèle d'interaction mixte au chapitre 3. Les éléments constitutifs saillants de notre modèle sont :

- un objet mixte modélisé par des modalités de liaison, un type de modalités nouvelles que nous avons définis,
- un outil d'interaction défini par le couplage d'un outil mixte avec un outil logique,
- une modalité d'interaction définie par le couple (*outil d'interaction, langage d'interaction*).

Comme tout modèle d'interaction, nous avons aussi étudié le pouvoir descriptif, taxinomique et comparatif de notre modèle, pour mettre en exergue ses apports.

Nous concluons au chapitre 4 cette première partie de nos travaux dédiée à l'espace de conception des systèmes mixtes en soulignant l'apport du mariage des points de vue "interaction multimodale", "interaction instrumentale" et "interaction mixte".

La deuxième partie de notre mémoire est consacrée à l'étude d'un système mixte mobile, RAZZLE un jeu qui consiste à collecter des pièces de puzzle numériques ancrées dans le monde physique. L'objet de cette partie composée de 3 chapitres est de prouver la validité de nos résultats par la conception, la réalisation et l'évaluation ergonomique d'un système mixte. En particulier nous voulions éprouver le pouvoir comparatif du modèle en phase de conception et montrer son pouvoir génératif. La figure 8.1 résume les étapes de travail liées à RAZZLE vis-à-vis de notre modèle d'interaction mixte.

Notre modèle a fait l'objet d'une publication longue à la conférence UBIMOB 2005 (ACM

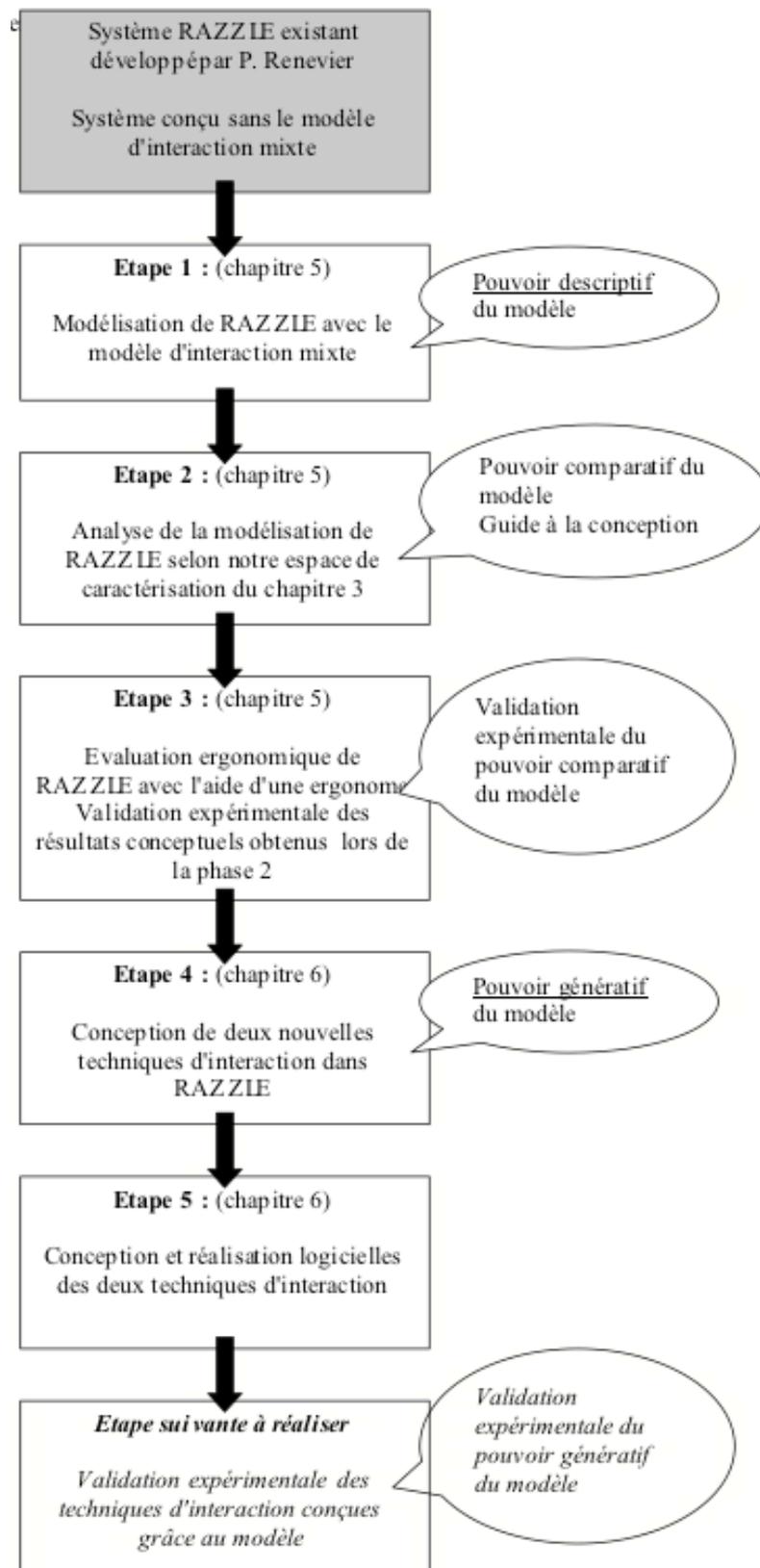


FIG. 8.1 – RAZZLE : Etapes de travail et modèle d'interaction mixte

International Conference Proceedings Series), modèle que j'ai présenté lors de la conférence. Nous avons soumis un article long en cours d'évaluation sur l'application et l'apport du modèle au cas du système mixte RAZZLE à la conférence ACM-ICMI 2005 dédiée à l'interaction multimodale.

Notre modèle de l'interaction mixte, qui tire parti des approches concernant la réalité mixte d'un côté, et l'interaction instrumentale et multimodale de l'autre côté. Ces perspectives n'avaient pas été étudiées conjointement jusqu'à aujourd'hui. Nous avons montré que la précision de notre modélisation implique une meilleure compréhension et classification des systèmes existants, mais aussi qu'elle permet de concevoir de nouvelles techniques d'interaction mixte. Les perspectives qu'offre cette capitalisation de plusieurs approches n'ont pas encore été toutes explorées, notamment l'étude de la fusion de modalités et l'utilité du modèle pour la comparaison et l'évaluation de l'utilisabilité (souplesse et robustesse) de l'interaction. Ceci constitue nos perspectives de recherche.

## 8.2 Perspectives de développement

Un premier travail envisagé à court terme est l'évaluation ergonomique de la nouvelle version de RAZZLE. Comme souligné à la figure 8.1, cette évaluation nous permettra de valider expérimentalement le pouvoir génératif de notre modèle d'interaction mixte dans le contexte de la conception de RAZZLE.

Outre cette évaluation ergonomique, nous entrevoyons pour nos travaux de multiples perspectives. La dualité, conception et réalisation d'un système, apparaît dans nos propositions. Ces dernières s'organisent en deux parties : les extensions et les prolongements à plus long terme.

### 8.2.1 Extensions

Les extensions à notre modèle d'interaction mixte concernent deux axes qui concourent à augmenter son pouvoir descriptif et comparatif.

#### **Composition des modalités de liaison et d'interaction**

Tout d'abord il convient d'étudier plus en détail la composition des modalités de liaison et d'interaction au sein de notre modèle. Dans notre étude, nous avons simplement distingué la notion de modalités élémentaires et composées et modélisé la complémentarité de modalités de liaison. Or la composition de modalités a été très étudiée dans le domaine de l'interaction multimodale et des espaces de composition comme celui exposé au chapitre 2 qui distingue cinq aspects et cinq schémas de composition nous semblent importants d'être pris en compte par notre modèle.

#### **Espace de caractérisation : propriétés ergonomiques**

Nous envisageons aussi d'étendre notre espace par de nouvelles propriétés ergonomiques exprimées au sein de notre modèle et en particulier celles liées à la composition de modalités. Pour viser à plus long terme à un guide de conception, il convient d'analyser dans quel contexte les propriétés doivent être vérifiées. Pour cela le cycle conception/évaluation expérimentale de la figure 8.1 est incontournable.

### 8.2.2 Prolongements : plateforme logicielle de conception et développement

Dans le cadre d'une approche de conception et réalisation de l'interaction qui par nature est dirigée par les modèles (MDE, Model Driven Engineering), nos prolongements à long terme vise la mise en place d'une plateforme de conception et de développement basée sur notre modèle d'interaction. Dans le chapitre 7, pour le développement de RAZZLE, nous avons souligné l'intérêt des composants logiciels ICARE. Les composants ICARE s'accompagnent d'un éditeur qui permet de concevoir graphiquement par assemblage de composants une interaction multimodale. Le modèle à composants ICARE sous-jacents repose sur la distinction entre dispositif et langage. Aussi notre modèle d'interaction mixte basé aussi sur les notions de dispositif et langage pourrait donc faire l'objet d'un éditeur graphique : le concepteur manipulerait alors des composants logiciels pour spécifier graphiquement l'interaction mixte selon notre modèle d'interaction. A partir de cette spécification, des propriétés ergonomiques pourraient être automatiquement vérifiées (pouvoir comparatif du modèle) avant de générer automatiquement le code de l'interaction mixte.

# Bibliographie

- [1] Beaudoin-Lafon *Instrumental Interaction : An Interaction Model for Designing Post-WIMP User Interfaces* CHI'00, CHI Letters, p. 446-453.
- [2] Beaudoin-Lafon *Designing Interaction, not Interfaces* AVI'04, p. 15-22.
- [3] Bass, Faneuf, Little, Mayer, Pellegrino, Reed, Seacord, Sheppard, Szczur *A Metamodel for the Runtime Architecture of an Interactive System* ACM Journal Special Interest Group Computer-Human Interface bulletin (SIGCHI), 1992, vol 24, n 1, ACM, p. 32-37.
- [4] Bérard *The Magic Table : Computer-Vision Based Augmentation of a Whiteboard for Creative Meetings*, IEEE Workshop on Projector-Camera Systems (PROCAM'03), CD-ROM proceedings of the International Conference on Computer Vision (ICCV).
- [5] Bernsen *Taxonomy of HCI Systems : State of the Art*, ESPRIT BR GRACE, 1993, deliverable 2.1.
- [6] Bertelsen, Nielsen *Augmented Reality as a Design Tool for Mobile Interfaces*, DIS'00, ACM.
- [7] Bouchet , Nigay *ICARE : A Component-Based Approach for the Design and Development of Multimodal Interfaces* Extended Abstracts of CHI'04, ACM, p. 1325-1328.
- [8] Card, Moran, Newell *The psychology of Human-Computer Interaction*. Lawrence Erlbaum Associates, 1983.
- [9] Chalon *Réalité mixte et Travail Collaboratif : IRVO, un modèle de l'interaction Homme-Machine*, Thèse de doctorat, décembre 2004.
- [10] Coutaz *PAC : an Implementation Model for Dialog Design*. Interact'87, p. 431-436.
- [11] Coutaz, Nigay *Les propriétés "CARE" dans les interfaces multimodales*, IHM'94.
- [12] Dubois, Nigay, Troccaz, Chavanon, Carrat *Classification space for augmented surgery, an augmented reality case study*, Acte de la conférence INTERACT'99, Chapman and Hall, p. 353-359.
- [13] Dubois *Chirurgie Augmentée, un Cas de Réalité Augmentée; Conception et Réalisation Centrées sur l'Utilisateur*, Thèse de Doctorat, juillet 2001.
- [14] Dubois, Mansoux, Bach *Un modèle préliminaire du domaine des systèmes mixtes*, IHM'04, ACM, p. 61-68.
- [15] Feiner, MacIntyre, Seligman *Knowledge-based augmented reality*. Communications of the ACM, vol. 36, no. 7, July 1993, p. 53-62.
- [16] Fishkin *A Taxonomy for and Analysis of Tangibles Interfaces* Personal and Ubiquitous Computing, Volume 8 Issue 5, September 2004.
- [17] Ishii, Ullmer *Tangible Bits : Towards seamless Interfaces between People, Bits and Atoms*, CHI'97, ACM, p. 234-241.

- [18] Ishii, Wisneski, Brave, Dahley, Gorbett, Ullmer, Yarin *ambientROOM : Integrating Ambient Media with Architectural Space*, CHI'98, ACM, p. 173-174.
- [19] *JavaBeans1.01 specification*, Sun Microsystems, 1997, <http://java.sun.com/products/javabeans/docs/>
- [20] <https://jogl.dev.java.net>
- [21] Koleva, Benford, Greenhalgh *The Properties of Mixed Reality Boundaries*, ECSCW'99, Kluwer.
- [22] Mackay *Augmenting Reality : Linking real and virtual worlds A new Paradigm for interacting with computers*, AVI'98, ACM, p. 1-9.
- [23] Nigay *Conception et modélisation logicielles des systèmes interactifs : Application aux interfaces multimodales*. Thèse de Doctorat, janvier 1994.
- [24] Nigay, Coutaz *A Generic Platform for Addressing the Multimodal Challenge* CHI'95, ACM.
- [25] Nigay, Coutaz *Espaces conceptuels pour l'interaction multimédia et multimodale*, TSI spécial multimédia et collectif, AFCET et Hermes Publ., vol. 15(9), p. 1195-1225.
- [26] Nigay, Dubois, Renevier, Pasqualetti, Troccaz *Mixed Systems : Combining physical and digital worlds*, HCI'03, Theory and Practice (Part I), Vol 1, Edited by Julie Jacko, Constantine Stephanidis, LEA, p. 1203-1207.
- [27] Rekimoto, Nagao *The World through the Computer : Computer Augmented Interaction with Real World Environments*. UIST'95, p. 29-36.
- [28] Renevier, Nigay *Mobile Collaborative Augmented Reality : the Augmented Stroll*, EHCI'01, IFIP WG2.7 (13.2) Working Conference, Toronto, Lecture Notes in Computer Science (LNCS), Vol. 2254, M. Reed Little, L. Nigay Eds, Springer-Verlag Publ., p. 315-334.
- [29] Renevier *Systèmes mixtes collaboratifs sur supports mobiles : Conception et Réalisation*, Thèse de Doctorat, juin 2004.
- [30] Vernier, Nigay *A framework for the Combination and Characterization of Output Modalities* DSVIS'95.
- [31] <http://www-306.ibm.com/software/voice/viavoice/>
- [32] Want, Fishkin, Gujar, Harrison. *Bridging Physical and Virtual Worlds with Electronic Tags*. CHI'99, ACM, p. 370-377.
- [33] Wellner *The Digital Desk Calculator : tangible manipulation on a desk top display*, UIST'91, p. 27-33.

# Table des figures

2.1	Le système CASPER avec affichage sur un écran d'ordinateur . . . . .	9
2.2	TROC . . . . .	10
2.3	Définition d'une zone de sélection sur la table augmentée par manipulation de deux jetons . . . . .	11
2.4	Espace de Dubois . . . . .	12
2.5	Quantité d'information transitant dans l'augmentation . . . . .	14
2.6	Deux continua de systèmes mixtes . . . . .	15
2.7	Modalités de sortie dans Vitesse . . . . .	18
2.8	Niveaux de compositions de modalités . . . . .	20
2.9	CASPER version écran, décrite selon la notation ASUR . . . . .	22
2.10	CASPER, version casque, décrit selon la notation IRVO . . . . .	23
2.11	Modèle instrumental de l'interaction . . . . .	25
2.12	Modèle instrumental de la barre de défilement . . . . .	26
3.1	Objet physique, objet mixte et objet numérique . . . . .	29
3.2	Objet Mixte . . . . .	29
3.3	La souris modélisée comme un objet mixte . . . . .	31
3.4	La souris modélisée comme un dispositif . . . . .	31
3.5	L'écran modélisé comme un objet mixte . . . . .	31
3.6	L'écran modélisé comme un dispositif . . . . .	32
3.7	Outil d'interaction : outil mixte et outil logique . . . . .	35
3.8	Outil d'interaction de la Table Magique lors de la création d'une entité (1) . . . . .	36
3.9	Outil d'interaction de la Table Magique lors de la création d'une entité (2) . . . . .	37
3.10	Simplification de l'outil d'interaction dédié (1) . . . . .	38
3.11	Simplification de l'outil d'interaction dédié (2) . . . . .	39
3.12	Simplification de l'outil d'interaction dédié (3) . . . . .	40
3.13	Outil d'interaction : outil mixte dédié . . . . .	41
3.14	Outil d'interaction : exemple d'outil mixte dédié . . . . .	42
3.15	Outil d'interaction de TROC lors de l'interaction avec la borne . . . . .	44
3.16	Modélisation de KARMA . . . . .	45
3.17	Technique d'interaction mixte sur un objet mixte . . . . .	46
3.18	Domaine potentiel d'application de métaphores dans le cas (1) des systèmes informatiques graphiques ou de réalité virtuelle et (2) des systèmes mixtes . . . . .	53
5.1	Un enfant jouant avec RAZZLE lors de la fête de la science 2004 . . . . .	61

5.2	Une scène du jeu RAZZLE avec le magicien d'Oz au premier plan et un joueur au fond . . . . .	61
5.3	Interface de RAZZLE dans le casque semi transparent . . . . .	62
5.4	Modélisation de la pièce de puzzle mixte dans RAZZLE . . . . .	63
5.5	( $m_1$ ) : Ramasser une pièce en s'approchant à moins d'un mètre pendant quelques secondes . . . . .	64
5.6	( $m_2$ ) : Ramasser une pièce en s'approchant et l'attrapant d'un geste de la main . . .	64
5.7	( $m_3$ ) : Ramasser une pièce en positionnant le curseur de la souris dessus et en cliquant	65
6.1	schéma de la nouvelle modalité $m_1$ corrigeant les défauts d'observabilité . . . . .	70
6.2	Nouvelle version de $m_1$ : état sélectionné de la pièce observable et état de l'outil (temps passé) observable. . . . .	71
6.3	schéma de la nouvelle modalité $m_3$ corrigeant les défauts de continuité spatiale . . .	72
6.4	Nouvelle version de $m_3$ : curseur fixe au centre du casque pour sélectionner une pièce et ramassage par la parole . . . . .	73
6.5	Le modèle d'architecture Arch . . . . .	74
6.6	Le modèle d'architecture multi-agent PAC . . . . .	75
6.7	Le modèle d'architecture PAC-Amodeus . . . . .	76
6.8	Le modèle d'architecture PAC-Amodeus pour les collecticiels . . . . .	77
6.9	Le modèle d'architecture PAC-Amodeus de RAZZLE . . . . .	78
6.10	Le modèle d'architecture PAC pour le contrôleur de dialogue (CD) de RAZZLE . . .	79
6.11	ICARE et Arch . . . . .	80
8.1	RAZZLE : Etapes de travail et modèle d'interaction mixte . . . . .	84

# Liste des tableaux

2.1	Compositions de modalités (CARE étendues) . . . . .	19
3.1	Correspondance des termes de notre modèle d'interaction mixte avec celui d'interaction instrumentale [2] . . . . .	47
3.2	Correspondance des termes de notre modèle d'interaction mixte avec ceux de la notation ASUR . . . . .	48
3.3	Correspondance des termes de notre modèle avec ceux du modèle IRVO . . . . .	49
3.4	Classification d'un objet mixte . . . . .	56
3.5	Classement d'un outil d'interaction . . . . .	57
5.1	Utilisation des techniques d'interaction . . . . .	67

**Mots-clés :** Modalités d'interaction, Multimodalité, Réalité Augmentée, Informatique Mobile, Interaction Homme Machine.

## Résumé

Le paradigme de réalité mixte en Interaction Homme-Machine (IHM) repose sur la fusion des mondes physique et numérique, affranchissant ainsi les utilisateurs des barrières séparant les deux mondes. Le monde ainsi obtenu est qualifié de mixte. Malgré la multiplicité des études et démonstrateurs de réalité mixte, le cadre de compréhension des systèmes mariant les mondes numérique et physique reste flou. Dans mon étude, j'adopte un point de vue unificateur sur la réalité mixte en considérant les modalités d'interaction mises en jeu par le système, focalisant ainsi sur l'interaction. Je présente un nouveau modèle d'interaction, noté modèle de l'interaction mixte. Comme tout modèle d'interaction, j'étudie son pouvoir descriptif et comparatif mais aussi génératif de nouvelles formes d'interaction mixte. De nombreux systèmes de réalité mixte existants sont décrits selon mon modèle, puis comparés. De plus j'illustre l'intérêt du modèle en phase de conception avec le système RAZZLE, un jeu de réalité mixte collaboratif et mobile, selon deux axes : (1) en expliquant comment étudier de façon prédictive des aspects ergonomiques du système à développer (résultats validés par des évaluations expérimentales) et, (2) en définissant de nouvelles formes d'interaction mixte. Les nouvelles formes d'interaction mixte ainsi conçues ont enfin été développées dans la deuxième version de RAZZLE et devront maintenant être expérimentalement testées.