

Un simulateur d'usage pour l'évaluation des systèmes interactifs multimodaux

Using simulation for Automating Usability Evaluation of Multimodal Systems

Pierre DE LOOR (1), Laurent LE BODIC (1), Guillaume CALVET (2) et Jacques TISSEAU (1)

(1) LISYC EA3883 UBO/ENIB
Centre Européen de Réalité Virtuelle
25 rue Claude Chappe
F-29280 Plouzané, France
{deloor; lebodic; tisseau}@enib.fr

(2) GRIC-IRIT
UPS-CNRS UMR(5505)
118, route de Narbonne
F-31062, Toulouse France
calvet@irit.fr

Résumé. Cet article présente un simulateur pour l'évaluation des systèmes interactifs multimodaux mobiles. Il commence par un état de l'art des méthodes d'évaluation par simulation. Celui-ci mène à la proposition d'une architecture en trois modèles pour reproduire au sein d'un environnement virtuel, les interactions et les décisions d'un utilisateur de système interactif. L'accent est mis sur la nécessité de rendre compte des gênes sensorielles et d'obtenir des modèles ouverts pour le prototypage interactif. Chacun de ces modèles est ensuite détaillé en montrant ses spécificités, le travail que doit fournir l'ergonome pour l'utiliser, ainsi que les souplesses qu'il offre pour le prototypage interactif. Un exemple illustratif de l'approche est ensuite décrit. L'article se termine par une discussion montrant les pistes à explorer.

Mots-clés. Simulation, évaluation, multimodalité, contexte, intelligence artificielle.

Abstract. This article introduces a simulator for the evaluation of multimodal systems. It draws a state of the art of simulation methods which leads to an architecture based on three models in order to simulate the interactions between the user, the artefact and the environment. The aim is to integrate the effect of perceptual troubles due to the environment in the decision making of the user. An illustrative example follows this presentation and a discussion of the work in progress ends the article.

Keywords. Simulation, evaluation, multimodalities, context, artificial intelligence.

1 Introduction

Les systèmes informatiques mobiles sont dotés de fonctionnalités croissantes et munis d'interfaces de plus en plus complexes. Pour pallier ce problème de complexité, les techniques d'interaction multimodales ont été proposées (Oviatt, 1996). Elles ne sont pas une solution à part entière car des études empiriques montrent que la multitude de modalités offertes en entrée et en sortie à un utilisateur est source de confusion et de difficulté (Oviatt, 1999). Il est donc nécessaire de mener des études précises d'usage des systèmes multimodaux afin d'analyser les interactions et de détecter les problèmes avant ou pendant leur conception. Cependant, la mobilité de l'utilisateur, généralement associée à la multimodalité, ainsi que la multimodalité elle-même donnent à l'environnement physique entourant l'utilisateur et le système, un statut nouveau. En effet, prenons l'exemple d'une interaction mixte vocale/visuelle en mobilité. Au gré des déplacements de l'utilisateur, selon que l'environnement sera bruyant ou lumineux, les modes d'interaction vont évoluer. Ceci amène à enrichir ou adapter les techniques d'évaluation de l'utilisabilité des interfaces multimodales.

Rappelons que l'utilisabilité rend compte de la facilité d'utilisation et non de la quantité de fonctionnalités offertes (Nielsen, 1993). Cette dimension est difficile à évaluer car fortement liée au comportement de l'homme qui est complexe, sujet à débats et à de nombreuses variations interindividuelles¹ (Aykin et Aykin, 1991 ; Dillon et Watson, 1996 ; Gwizdka et Chignell, 2004). Pour répondre à cette difficulté, les ergonomes recourent des évaluations empiriques, basées sur des questionnaires automatisés tels que WAMMI (Kirakowski *et al.*, 1998), QUIS (Harper et Norman, 1993) ou NetRaker², ou sur des traces d'utilisation avec DRUM (Macleod et Rengger, 1993), MIKE UIMS (Olsen et Halversen, 1988), AMME (Rauterberg, 1995), MRP (Siochi et Hix, 1991), QUIP (Helfrich et Landay, 1999), KALDI (Al-Qaimari et McRostie, 1999), UsAGE (Uehling et Wolf, 1995), VISVIP (Cugini et Scholtz, 1999) et des approches analytiques basées sur des modèles prédictifs telles que TAG (Payne et Green, 1986), GOMS-like (John et Kieras, 1996), CTA (May et Barnard, 1995), PUM (Young *et al.*, 1989).

Une dernière famille utilise des simulations (Kieras *et al.*, 1995 ; Kieras *et al.*, 1997b ; Chi *et al.*, 2000 ; Tollinger *et al.*, 2005) s'appuyant sur des *architectures cognitives* qui sont un modèle exécutable d'utilisateur. L'intérêt des simulations est de pouvoir considérer rapidement différents scénarios en introduisant des événements à des instants donnés et en faisant varier des hypothèses psychologiques d'usage. Il est par exemple intéressant de simuler l'appropriation de l'interface par l'utilisateur, de tester de nombreuses hypothèses fonctionnelles sur cette interface et de procéder à des analyses statistiques. C'est parmi ce type d'approche que nous situons notre travail. Par ailleurs, nous conseillons au lecteur intéressé par un panel et une comparaison détaillée des différentes méthodes d'évaluation, la synthèse faite par (Ivory et Hearst, 2001) ainsi que la lecture de (Byrne, 2003) qui fait une synthèse des architectures cognitives liées au développement d'interfaces homme-machine.

Il apparaît que la spécificité de la multimodalité est relativement peu traitée : outre des difficultés d'ordre logistique pour mener des expérimentations, c'est bien l'importance du rôle de l'environnement, sa variabilité et sa dynamique qui rendent ces études délicates. Pour cela, les approches par simulation mentionnées

¹ Des différences intra individuelles peuvent également apparaître au sein d'activités pourtant *a priori* bien analysées et normalisées. Par exemple, (Noizet et Amalberti, 2000) ont mis en évidence l'influence des activités routinières sur la tâche à effectuer.

² The NetRaker suite. Disponible sur <http://www.netraker.com/info/applications/index.asp>

précédemment, semblent être un moyen bien adapté, car elles permettent d'intégrer le comportement d'un utilisateur avec un environnement simulant les phénomènes qui influent sur son comportement. C'est dans ce cadre que nous proposons une approche utilisant la réalité virtuelle et la simulation. Nous nous sommes concentrés sur la spécification de modèles simulables permettant de rendre compte des modifications d'usage induites par l'environnement lors de l'utilisation d'une interface multimodale. Nous proposons les liens entre ces modèles menant à une simulation complète de l'environnement jusqu'à l'utilisateur en passant par le système interactif. Comme souvent dans ce type d'approche, la validation de la proposition reste très difficile. Néanmoins certaines comparaisons ont pu être effectuées grâce à une expérimentation menée parallèlement par des ergonomes.

Cet article commence par un état de l'art sur les méthodes d'évaluation analytiques et par simulation des systèmes interactifs afin de positionner notre proposition. Les trois parties qui suivent détaillent respectivement le modèle de l'environnement, celui du système interactif et celui de l'utilisateur virtuel. La partie 6 présente l'outil résultant de cette approche au travers d'un exemple avant d'aborder une discussion et les perspectives en partie 7.

2 De l'évaluation prédictive à la simulation

L'évaluation prédictive d'usage s'appuie sur des modèles spécifiant le comportement de l'utilisateur. A partir de ces modèles, il est possible de déduire des caractéristiques telles que ses performances ou les temps d'apprentissage d'une fonctionnalité. Dans cette section, nous proposons une synthèse des méthodes d'évaluation prédictive analytique et de celles basées sur des simulations. Nous présenterons également les principaux obstacles à franchir pour aborder l'évaluation de systèmes interactifs multimodaux

2.1 Evaluations prédictives analytiques

La plupart des évaluations analytiques se basent sur l'utilisation de la méthode GOMS qui s'inspire des modèles cognitivistes du « processeur humain » (Card *et al.*, 1983). Une analyse faite avec GOMS part d'une description de la connaissance procédurale que doit posséder un utilisateur et d'une description des séquences d'actions qui doivent être exécutées pour accomplir une tâche particulière³. Ces éléments sont interprétés, et un mécanisme de chaînage élabore les plans d'actions permettant d'atteindre un but. GOMS est une base solide qui fait référence, mais qui est souvent enrichie pour améliorer la compréhension de la description et la crédibilité comportementale de l'utilisateur avec respectivement NGOMSL (Kieras 1994), GOMSL (Kieras, 1997a) et CPM-GOMS (John *et al.*, 2002). Les enrichissements comportementaux reflètent les compétences de l'utilisateur, la possibilité d'erreur, l'apprentissage de l'interface ou le parallélisme. Les mécanismes associés s'appuient sur des théories psychologiques. Par exemple, NGOMSL est une modélisation GOMS à un niveau d'abstraction élevé et basée sur la Théorie de la Complexité Cognitive⁴ (Polson et Kieras, 1985) qui établit les temps d'exécution et d'apprentissage de tâches en fonction de critères de complexité et de surcharge cognitive et de temps d'erreur typiques. NGOMSL introduit des opérateurs associés à des concepts tels que la manipulation de la mémoire de travail et permet alors des

³ GOMS signifie : Goal-Operator-Method-Selection. Il permet de découvrir les méthodes (séquences d'opérateurs) à partir des **buts** que peut atteindre l'utilisateur, d'**opérateurs** représentant ce que peut exécuter l'interface et de **règles d'utilisation**.

⁴ Généralement abrégé sous le terme CCT.

prédictions d'apprentissage ou de transfert asymétrique (passage d'une interface à une autre). CPM-GOMS (Critical-Path-Method) considère des opérateurs 'perceptifs', 'cognitifs' et 'contrôles moteurs', distincts associés à une sémantique parallèle. Il est ainsi possible d'évaluer des performances d'un utilisateur capable « d'agir tout en décidant ».

L'élaboration de ces modèles reste une difficulté identifiée. C'est pourquoi d'autres approches sont associées à une méthodologie de construction qui fait partie du modèle lui-même. C'est le cas de CTA (Cognitive Task Analysis, cf. May et Barnard, 1995) et PUM (Programmable User Model, cf. Young *et al.*, 1989). Enfin, nous pouvons mentionner les travaux de (Rauterberg, 1995) et (John *et al.*, 2004) qui tentent d'élaborer un modèle respectivement par l'observation d'interactions ou par l'étude d'une interface réelle, au travers d'une théorie cognitive.

Plus on affine le modèle de l'utilisateur et plus les hypothèses sont nombreuses et discutables. Il est alors souhaitable de faire de nombreux tests concernant des cas d'usages particuliers. Il devient alors nécessaire d'intégrer une simulation du modèle de l'utilisateur et de son interaction avec l'environnement. C'est la raison pour laquelle les approches par simulation ont été envisagées.

2.2 Approches par simulation

Les approches par simulation se différencient par l'utilisation de modèles comportementaux exécutables qui peuvent être interfacés avec un environnement. En effet, si l'on peut décrire en GOMS le temps que met un opérateur pour reconnaître un objet, le modèle analytique n'est pas capable d'exécuter cette reconnaissance lors de l'apparition de l'objet et de tester des scénarios où il y aurait plusieurs objets apparents par exemple. Il existe de très nombreuses architectures qui tentent de simuler un comportement humain, certaines sont dédiées à l'étude d'IHM. C'est le cas par exemple de CAPS (Just et Carpenter, 1992) de LICAI/CoLides (Kitajima et Polson, 1996) et de l'Info Scent Simulator (Chi *et al.*, 2003). D'autres sont plus générales comme SOAR⁵ (Newell, 1990) et ACT-R (Anderson, 1993), elles sont adaptées et intégrées dans des outils supportant des modules sensori-moteurs comme GLEAN (Kieras *et al.*, 1995), ACTR-PM (Byrne et Anderson, 1998) et EPIC (Kieras *et al.*, 1997b) qui s'intéresse plus précisément à l'évaluation d'interfaces multimodales. Nous tentons dans la suite de ce paragraphe d'évaluer cet existant en décomposant cette analyse selon les axes de modélisation de la cognition, de l'apprentissage et de mécanismes sensori-moteurs.

Cognition :

La plupart des architectures sont basées sur un système de production de règles. C'est le cas de CAPS, SOAR, ACT-R et ACTR-PM, EPIC et GLEAN. Ces règles reflètent les enchaînements nécessaires pour atteindre des buts. Les nuances portent sur la granularité de ces règles et l'apport de critères psychologiques dans leur interprétation. Par exemple, SOAR permet de pondérer les règles à l'aide de préférences entre des opérateurs. Spécialisé dans la résolution de problèmes, il est basé sur un cycle de décision de 50 ms mais les critères psychologiques restent assez pauvres. A l'inverse, le moteur cognitif d'EPIC est parallélisé. Il repose sur un modèle de mémoire de travail plus élaboré partitionnant celle-ci en mémoire auditive, visuelle ou encore tactile. Les inférences parallèles rendent alors compte d'activités multitâches tout en s'appuyant sur le modèle CCT pour évaluer les temps d'exécution. ACT-R est également un système de production de règles. Il intègre

⁵ Initié par Newell pour mettre en œuvre sa Théorie Unifiée de la Cognition, ce modèle est basé sur le principe P9 (Problem Space Principle) du Modèle du Processeur Humain.

des notions de psychologie cognitive axées sur la mémoire déclarative telles que le degré d'activation d'un symbole fixant le temps d'accès à l'information par le 'processeur humain'. LICAI n'est pas un système de règles mais poursuit les mêmes objectifs. Cette approche est basée sur la construction d'un réseau d'informations disponibles pour l'utilisateur, issu de l'étude de l'interface. Ce réseau contient des éléments nécessaires à l'accomplissement de buts ainsi que des liens entre eux. Un mécanisme de propagation à l'intérieur du réseau permet de sélectionner des éléments et des chemins dans le réseau pour atteindre des buts. Avec LICAI, on rend par exemple compte du fait que si les objets d'une interface sont bien labellisés (en fonction des buts) alors le nombre d'objets présents n'est pas un frein à la vitesse de décision. CoLiDes (Comprehension-based Linked model for Deliberate Search), issu de LICAI est utilisé pour traiter les pages web, avec l'intégration d'un mécanisme sophistiqué de mesure de similarité sémantique. Enfin, l'Info Scent Bloodhound Simulator simule un utilisateur virtuel surfant sur une page web. Il utilise la notion de *piste informationnelle*⁶ simulant la perception par l'utilisateur, du coût d'accès à une information. Ce coût est utilisé pour évaluer si l'utilisateur va utiliser le lien correspondant ou y renoncer. Une des originalités de cette approche est donc l'intégration de l'abandon consécutif à la pénibilité de la tâche à accomplir.

Apprentissage et oubli :

De nombreuses architectures intègrent une modélisation de l'apprentissage. Il s'agit de refléter une évolution dans l'utilisation d'un système, au fur et à mesure que celui-ci est utilisé. Ainsi, en cas de difficulté pour résoudre un but, SOAR explore en utilisant des méta-règles et apprend le résultat de ces explorations en créant de nouvelles règles. De ce fait, il fait un apprentissage "envahissant" et n'oublie rien implicitement. L'apprentissage de ACT-R n'est pas invasif. Il est basé sur le concept d'analyse rationnelle de Anderson : à chaque production est associée une probabilité d'atteinte du but et un coût en temps d'exécution de la règle associée. Avec ACT-R, on peut alors rendre compte par exemple que plus un élément a été utilisé récemment, plus il sera facile à retrouver en mémoire. Ceci nous renvoie à la notion d'oubli intégrée dans CAPS (Collaborative Activation-based Production System) qui s'articule autour de l'analyse de la charge cognitive en mémoire de travail. Cette approche permet également de mettre en évidence des phénomènes tels que l'oubli de cartes bleues laissées dans un distributeur si les billets sont donnés avant la carte. Enfin, GLEAN qui est initialement un outil destiné à simplifier la modélisation et l'analyse de modèles GOMS permet de calculer des performances d'exécution ou des temps d'apprentissage estimés issus de ce modèle.

Mécanisme sensori-moteurs :

La perception d'une interface par un utilisateur peut être simplifiée à l'extrême en considérant qu'un élément perçu est un symbole présent ou non. Ces hypothèses sont souvent faites lors de l'utilisation d'architectures cognitives telles que SOAR ou ACT-R. Cependant, GLEAN3 intègre des modules sensori-moteurs qui simulent par exemple les temps d'acquisition de l'information en mémoire de travail. Ce modèle permet également grâce à une interface logicielle (API⁷ C++) de « brancher » un simulateur de système interactif externe quelconque, voire un système réel. Les difficultés soulignées dans (Kieras et Santoro, 2004) portent sur la

⁶ Traduction libre de « Information Scent »

⁷ Application Programming Interface

modélisation du système interactif. Ces difficultés sont accrues si l'on songe à intégrer les aspects multimodaux qui nécessitent également un modèle de l'environnement. Est mentionnée également la nécessité de rendre GOMS « interruptible », pour rendre compte de changements de tâche ou de stratégies d'exécution de tâches liés par exemple à une modification de l'état du système. Là encore cette difficulté est accrue par l'introduction d'aspects multimodaux. ACTR-PM permet d'associer à ACTR les fonctionnalités d'un contrôle moteur parallèle. Un effort particulier porte sur le contrôle visuel. Le modèle peut être connecté directement à une IHM (Kieras et Meyer, 1996). Une expérience très parlante est la simulation d'utilisateurs devant mémoriser des associations visuelles. On leur demande de mémoriser les boutons qui permettent d'afficher une couleur. Dans certaines expériences, ces boutons ont la couleur en question, d'autres fois, ils ont le nom de la couleur ou une icône quelconque. Le modèle permet de reproduire assez fidèlement des résultats réels. Enfin, EPIC met l'accent sur la modélisation du système sensori-moteur multimodal et son lien avec un système à base de règles (basé sur CCT). Le modèle est doté de nombreux paramètres (par exemple, la reconnaissance d'une forme bleue prend 250 ms ...), plusieurs « processeurs » fonctionnent en parallèle. Chacun d'eux s'occupent d'une unité de contrôle (processeur auditif, visuel, contrôle oculaire, vocal, gestuel et tactile ...etc.) ; ces unités de contrôle prédisent des temps d'exécution ou de perception. Le niveau de raffinement des processeurs peut être illustré au travers du processeur visuel : celui-ci ne s'intéresse pas à imiter le processus sensoriel tel que la reconnaissance de caractères mais renseigne sur des informations telles que : « est-ce qu'un objet est visible, si oui à quel niveau de détails, combien de temps cela met-il pour transférer l'information en question dans la mémoire déclarative ».

	SOAR	EPIC/SOAR	EPIC	GLEAN	ACT-R	ACTR/PM	Info Scent
Théorie cognitive	MPH	MPH/CCT	MPH/CCT	CCT	ACT	ACT	Info scent
Modélisation de l'apprentissage	oui	oui	non	non	oui	oui	non
Modélisation du contrôle moteur	non	oui	oui	non	non	oui	oui
Modèle du Système interactif	non	oui	oui	oui	non	oui	oui
Traitement de la multimodalité	non	oui	oui	oui	non	oui	non
Intégration de l'environnement	non	non	non	non	non	non	non

Tableau 1. *Comparaison d'approches de simulation de l'utilisateur pour l'évaluation*

Le contrôle moteur prend en compte l'exécution d'un mouvement : il le décompose en temps de préparation/temps d'exécution, sachant que le temps de préparation va diminuer si l'on répète un mouvement souvent ou un mouvement similaire. Le Tableau 1 résume les caractéristiques des approches mentionnées précédemment.

Résultats, intérêts et limites :

Au cours de cette description, nous avons mentionné différents exemples applicatifs de ces architectures. Même s'il est ancien et critiqué sur ses hypothèses cognitives, SOAR est le projet le plus abouti. Il en existe de nombreuses applications dans divers domaines et dans celui de l'interaction homme-machine (voir Ritter *et al.*, 2000). Les travaux de (Altmann et John, 2001) sont un exemple d'utilisation de SOAR pour simuler un programmeur découvrant un gros programme informatique. Ce modèle permet entre autres d'analyser la conséquence du désordre sur l'activité cognitive. Mais c'est EPIC qui est la première tentative de modélisation d'activités multitâches grandement nécessaire pour évaluer les interfaces mobiles. EPIC a été utilisé pour rendre compte de tests psychologiques perceptifs (appuyer sur une lettre dès qu'une couleur apparaît) ou de l'énumération articulative (retenir un texte oralement). EPIC est assez souple pour intégrer des modélisations *à la GOMS* et des stratégies psychologiques différentes comme cela est illustré dans (Kieras et Meyer, 2000). Enfin, notons l'intégration de notions d'apprentissage avec EPIC-SOAR (Chong et Laird, 1997).

Comme nous l'avons présenté en introduction, l'intérêt de ces approches réside dans le gain quant à la maîtrise de la prévisibilité des interactions possibles. Celles-ci sont de plus en plus complexes et multiples, rendant l'usage d'expérimentations également difficile. Les simulations permettent de détecter un problème d'interaction ou une difficulté d'apprentissage ou de compréhension. Elles sont alors intégrables à la phase de conception pourvu que celle-ci autorise des allers et retours comme c'est le cas dans les cycles de développement en spirale ou en étoile (Kolski *et al.*, 2001).

Pendant, ceci n'évite pas de nombreuses difficultés. Parmi celles-ci, l'impossibilité ou l'illusion de simuler un comportement exact, la difficulté de procéder à une validation par expérimentation et la difficulté de trouver les éléments pertinents à modéliser. Ceci a d'ailleurs donné naissance à des outils tels que *X-PTR* (Tollinger *et al.*, 2005) modélisant des activités cognitives en partant du 'design' d'une interface et permettant de spécifier de nombreuses hypothèses psychologiques très simplement à l'aide de contraintes pour passer de l'une à l'autre rapidement. La majorité des approches basées sur des règles sont issues de théories cognitivistes qui ont des difficultés à rendre compte de l'ambiguïté du choix et des aspects parallèles de la perception durant l'action, tandis que les autres rendent mieux compte de ces aspects perceptifs dans la simulation, en intégrant la notion d'abandon de la tâche, même si elles ne traitent pas de multimodalité et s'intéressent davantage à des activités de navigation hors contexte. C'est sur ces deux points : l'intégration de la multimodalité en contexte et la modélisation de la notion de motivation que notre travail apporte une contribution.

Positionnement de notre approche :

Notre approche repose sur l'outil SIHMM⁸ permettant de simuler, grâce à la réalité virtuelle, un modèle comportemental d'utilisateur en interaction avec son environnement. Ce travail ne prétend évidemment pas répondre à l'ensemble des problèmes mentionnés précédemment. Par exemple, il ne recherche pas une exactitude comportementale fine de « bas niveau » : nous nous concentrons sur l'impact de l'environnement en terme de prise de décision et non sur la simulation du système sensori-moteur. En effet, il est possible de rendre compte du contexte d'usage perçu par l'utilisateur au travers d'abstractions symboliques et qualitatives

⁸ Simulation d'Interaction Homme Machine Multimodale

beaucoup plus faciles à manipuler par un simulateur environnemental. SHIMM se distingue des outils précédents par les caractéristiques suivantes :

- La modélisation du contexte environnemental et de sa dynamique qui peut être spécifiée à l'aide d'un éditeur de scénarios.
- La possibilité d'un prototypage interactif, favorisant les tests d'hypothèses à la fois sur le moteur décisionnel de l'utilisateur virtuel, sur les fonctionnalités du système interactif et sur l'environnement lui-même.
- Une modélisation des Systèmes Interactifs Multimodaux orientée tâche facilitant la simulation et l'analyse. Cette modélisation met en évidence les modalités d'interaction ainsi que la multimodalité.
- Un modèle d'utilisateur couplant un modèle qualitatif parallèle de la perception basé sur la notion de gêne et de motivation à la tâche, proche des modèles connexionnistes, avec un modèle cognitif à base de règles pour la prise de décision intégrant les concepts associés au modèle de la faillible machine de J. Reason (Reason, 1990). Selon certains points de vue, il est moins riche que les architectures cognitives précédentes mais il est hautement configurable et peut être enrichi durant les expérimentations.

La décomposition de SHIMM en trois modèles en interaction est présentée en figure 1. Ceux-ci font l'objet des trois sections suivantes. Les scénarios ne constituent pas un modèle à part entière mais plutôt un module de contrôle pouvant interagir avec chacun des modèles en fixant des buts à l'utilisateur, en imposant des déplacements et des rendez-vous aux objets mobiles de l'environnement et en interrompant le système interactif.

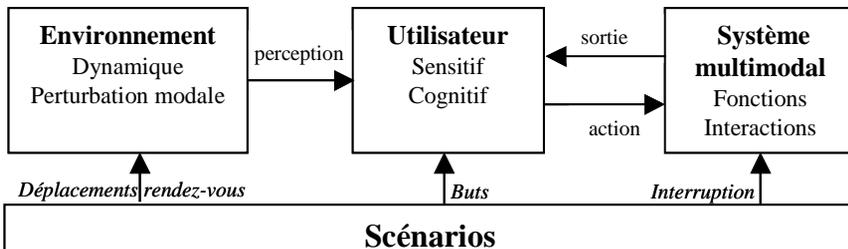


Figure 1. Les trois modèles de SHIMM

3 Modélisation de l'environnement d'interaction

L'environnement virtuel d'interaction doit permettre de simuler des scènes riches et complexes correspondant aux lieux d'usage des systèmes interactifs tels qu'un hall de gare ou un centre commercial. La crédibilité de ces environnements nécessite une grande dynamicité : il faut les peupler d'entités en interaction, ayant une influence sur l'usage du système par le biais de perturbations physiques ou sociales. Il faut également définir une granularité comportementale permettant de spécifier des scénarios à un niveau d'abstraction en rapport avec la mise en évidence de ces perturbations. Notre objectif n'est pas dans une recherche de consistance scénaristique formelle, mais plutôt dans l'obtention d'évolutions crédibles aux yeux de l'ergonome.

3.1 Modélisation

La modélisation s'appuie sur une approche multi-agents et sur la plateforme de simulation d'agents situés oRis (Harrouet, 2000). Le paradigme agent permet de s'abstraire de la réduction globale de la complexité au profit d'une approche phénoménologique du comportement de chaque acteur/entité de l'environnement. Les agents possèdent des capacités de perception, de décision et d'action qui les dotent de possibilités d'interaction. L'évolution et l'interaction simultanée des agents fait émerger un comportement global complexe qui n'est pas explicité. Nous avons doté les agents de la capacité de définir leurs vitesses de déplacement ou leurs trajectoires en fonction d'un plan de haut niveau et d'un univers virtuel informé. Ceci permet, entre autres, de s'affranchir des problèmes de collisions ou de définition explicite de vitesse pour obtenir un rendez-vous (la rencontre topologique de agents). Ce modèle environnemental est nommé DPE pour Dynamic Perturbing Environment.

Gêne sensorielle :

Les agents possèdent des attributs de perturbations modales associés aux sens de l'homme.

Interférence – gêne	Description
Auditive	Interférence auditive, bruit, source sonore. Ex : bruit de véhicule, etc
Vocale	Perturbation de la boucle phonologique de l'utilisateur, particulièrement sensible en cas de communication orale avec le système interactif Ex : une foule de badauds
Lumineuse	Perturbation lumineuse. Ex : un réverbère dans une rue sombre.
Extéroceptive	Perturbation gestuelle. Ex : un espace restreint
Olfactive	Interférence olfactive, peut diminuer par exemple la motivation de l'utilisateur. Ex : échappement d'un véhicule
Kinesthésique	Interférence avec les capacités sensorielles de l'utilisateur, au niveau kinesthésique (vibrations, tactile...) Ex : un escalator
Affective	Interférence d'ordre sociologique. Ex : un restaurant, une foule stressante.

Tableau 2. *Interférence et sens perceptif*

Le simulateur s'occupe de transmettre ces perturbations à l'utilisateur virtuel. La figure 2 illustre les mécanismes que nous avons implémentés en oRis. Ceux-ci permettent, une fois que les agents perturbateurs sont configurés, de rendre compte quantitativement des valeurs des perturbations perçues par un agent placé à une certaine distance de sources. Pour simplifier ce calcul, nous approximations les fonctions de dispersion à l'aide d'équations linéaires. Actuellement, les interférences possibles sont celles données dans le tableau 2.

Usage par l'ergonome :

L'utilisateur de SHIMM peut configurer l'environnement 'à la souris'. Pour cela, une librairie de composants standard est à sa disposition. Ceux-ci possèdent des paramètres graphiques (taille, couleur) et les paramètres spécifiques à l'évaluation (gêne associée, puissance de cette gêne). Ils sont positionnés sur une scène tridimensionnelle. Il est ensuite possible d'associer à ces composants des caractéristiques dynamiques : les actions. Certaines actions sont également standard, telles que les déplacements dans l'environnement. D'autres devront être décrites avec le langage oRis.

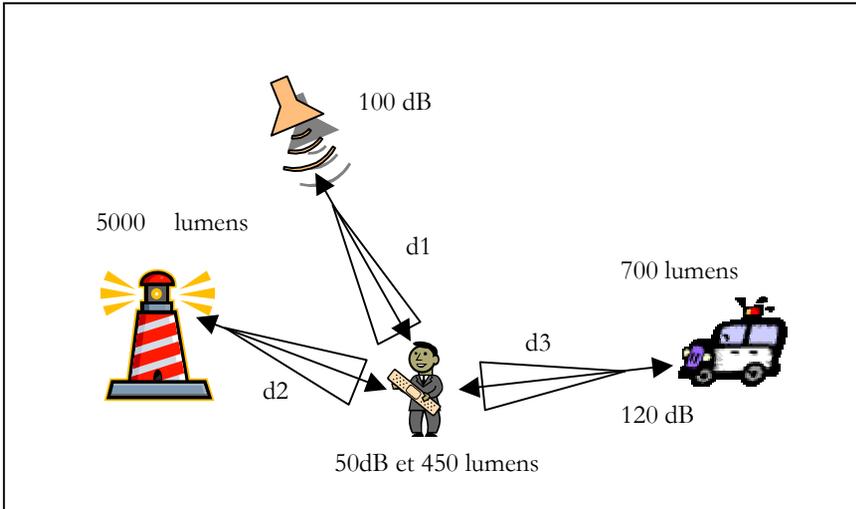


Figure 2. Le DPE quantifie l'ensemble des perturbations sur l'utilisateur

C'est le cas par exemple de l'escalator qui possède un comportement cyclique émettant parfois du bruit. Ceci constitue une des limites de la souplesse de l'outil car oRis est un langage de programmation proche du C++ plus accessible aux informaticiens qu'aux ergonomes. Cependant, il est tout à fait possible d'étendre la librairie de composants à un nombre très élevé d'éléments de la vie courante. Une fois qu'un tel composant est défini en oRis, il est récupéré et inspecté par SHIMM qui propose à l'utilisateur, sous forme de liste, toutes les actions potentielles que peut exécuter le composant.

L'éditeur de scénarios permet de spécifier des déclenchements ou arrêts d'actions associées à chaque élément de la scène à des instants précis. La figure 3 représente l'équivalent de ce qu'il permet de saisir : la première *ligne de vie* décrit un déplacement d'un véhicule d'un point d'ordonnée x_1 et d'abscisse y_1 à un point d'ordonnée x_2 et d'abscisse y_2 en un temps précis, le tout en émettant un bruit de 60 décibels. Enfin, il peut également provoquer explicitement des rencontres entre l'utilisateur virtuel et les éléments de l'environnement. Pour cela il propose le lieu et le moment de la rencontre, SHIMM calcul alors la vitesse des déplacements nécessaire pour que celle-ci ait lieu. La figure 4 est une représentation de l'environnement tel qu'il est obtenu avec notre outil.

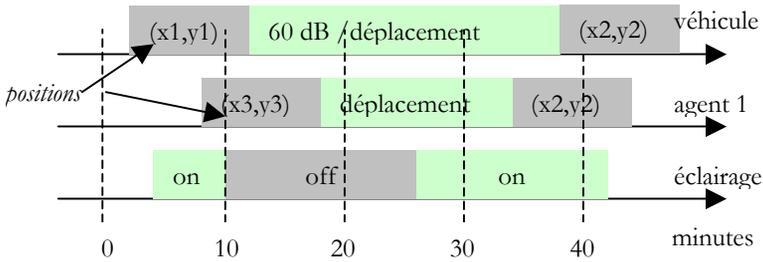


Figure 3. Ce que permet de spécifier l'éditeur de scénarios pour le DPE

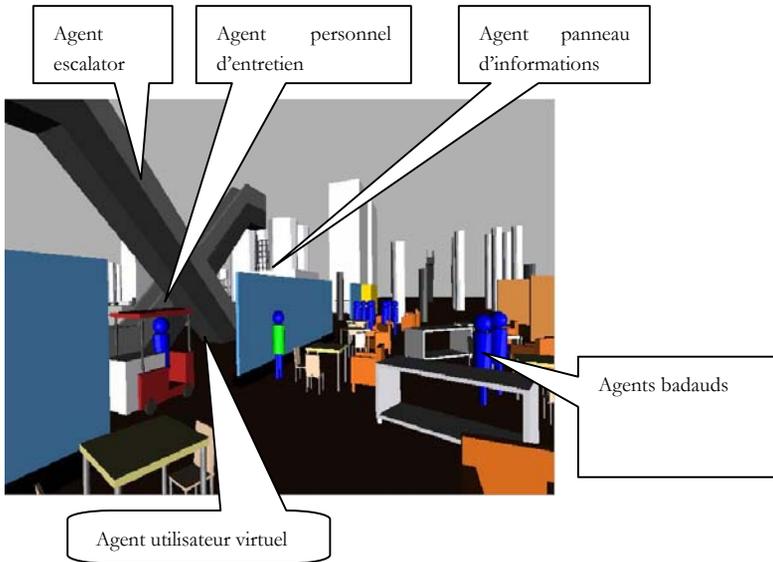


Figure 4. Modélisation de l'environnement à base d'agents perturbateurs

Durant la simulation, il est possible de provoquer des changements pour tester des hypothèses précises rapidement : on peut, par exemple, déplacer à la souris une source lumineuse pour voir l'impact de cette modification sur les choix de l'utilisateur à un instant particulier.

4 Modélisation du système interactif multimodal

La modélisation du système interactif simulable pose le problème du rapport entre l'expressivité, la simplicité d'expression et les temps d'exécution. Il existe des modèles proches des ergonomes, centrés sur les besoins et l'activité des opérateurs tels que MAD (Scapin et Pierret-Golbriech, 1989) HTA (Carey *et al.*, 1989) ou CTT (Paterno, 2001). Mais ceux-ci ne sont pas adaptés aux systèmes multimodaux et rarement simulables. Une autre famille de modèles s'intéresse à la description fonctionnelle destinée à la conception. Ils sont simulables. Un exemple est l'utilisation de réseaux de Petri (Tabary et Abed, 2002 ; Pontico *et al.*, 2004 ou Navarre *et al.*, 2002). Cependant les modèles existants ne traitent pas des modalités.

Par conséquent nous avons été amenés à élaborer un modèle, inspiré des précédents et des statecharts, qui intègre les interactions multimodales ainsi qu'une sémantique opérationnelle. Ce modèle, UMAR (pour User Modalities Artefact Representation ; Lebodic, 2005) tente d'apporter les trois points suivants : 1) une modélisation des modalités d'interaction 2) une expressivité et une lisibilité apportées par le support du parallélisme, de l'abstraction et par la granularité de la représentation orientée « modalité d'interaction » 3) la prise en compte de possibilités d'interaction évoluées, comme les propriétés CARE (Complémentarité, Assignment, Redondance, Equivalence) de la multimodalité explicitées par (Coutaz et Niguy, 1995) et rappelées dans la suite du document.

4.1 Les modalités

Au cours d'une interaction, nous identifions trois types de modalités distinctes : la *modalité d'action* de l'utilisateur sur le système, la *modalité de contrôle*, qu'il utilise pour réguler cette action, ainsi que la *modalité de sortie* renvoyée par le système, suite à l'exécution de l'action. Prenons l'exemple de la saisie de nombres par pression sur des touches. Une pression correspond à la *modalité d'action*. La *modalité de contrôle* correspond au retour proprioceptif du clavier, quant à la *modalité de sortie*, elle peut être mise en œuvre à travers l'émission d'un signal sonore ou visuel. La qualité d'une interaction dépendra de la possibilité d'utiliser ces trois modalités. Par conséquent elles sont une composante essentielle d'UMAR (voir figure 5).

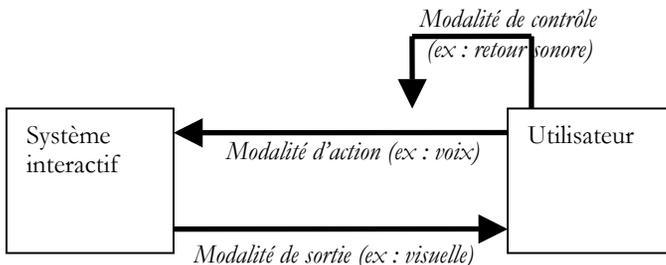


Figure 5. Illustration des modalités d'action/contrôle et sortie à la base d'UMAR

Pour leur notation, nous nous sommes appuyés sur le couple $\langle P, R \rangle$ défini par Laurence Nigay (Nigay, 2001) où P désigne un dispositif physique et R dénote un système représentationnel. Ce couple permet de faire les liens symboliques entre l'utilisateur et le système interactif. Pour représenter les modalités d'action, nous avons ajouté un symbole caractérisant la modalité de contrôle de l'interaction. Ceci nous amène à décrire une modalité d'action par le triplet $\langle P, R, C \rangle$. Le tableau 3 donne quelques exemples de ce triplet : P et R reprennent les notions respectives de dispositif physique et représentationnel, tandis que C caractérise la modalité de contrôle de la modalité d'interaction décrite. Les modalités de sortie sont émises par les systèmes qui ne les contrôlent pas encore. Leur notation ne nécessite donc pas la représentation du contrôle.

Par contre, les temps de réaction du système peuvent influencer sur la qualité de l'interaction et doivent être simulés. C'est pourquoi ces modalités de sortie sont formalisées par un triplet $\langle P, R, Tr \rangle$ où Tr caractérise le temps avant que l'émission de la sortie ait lieu.

Exemple - Type	Physique	Représentation	Contrôle
Tactile	Doigt Stylet	Touche Up Symbole Up	Visuel Visuel, Proprioceptif (PK)
Vocale	Microphone	Annuler	Auditif
Tangible	Main	Onglet composition	Visuel, PK
Geste 2D	Doigt, main	Tabulation	Visuel, PK

Tableau 3. Exemples de modalités d'interaction utilisateur/système/contrôle

4.2 Granularité

UMAR s'articule autour de la notion d'états, assimilables à des fonctions offertes par le système multimodal à l'utilisateur. L'accès aux différentes fonctions est lié à l'usage d'une ou plusieurs modalités d'action. Le système interactif multimodal est donc discrétisé en fonctions offertes à l'utilisateur. La fonction peut être affinée selon les choix de l'ergonome mais nous avons fait l'hypothèse qu'une granularité pertinente consistait à considérer chaque changement de modalité de sortie comme un changement d'état. Ceci est lié à l'objectif des simulations, axées sur la prise de décision de l'utilisateur et dépendant du contexte. Or le contexte est lié en partie à la perception du système interactif lui-même défini principalement par les modalités de sortie.

Ainsi, Les modalités d'action font partie de la notation des arcs d'un diagramme UMAR tandis que les modalités de sortie caractérisent les états. UMAR reprend le raffinement d'état et le parallélisme des statecharts. Ceci est résumé sur la figure 6 où la notation <modalité d'action <modalité de contrôle>> permet de préciser textuellement la modalité de contrôle d'une modalité d'action (voir l'exemple en figure 7).

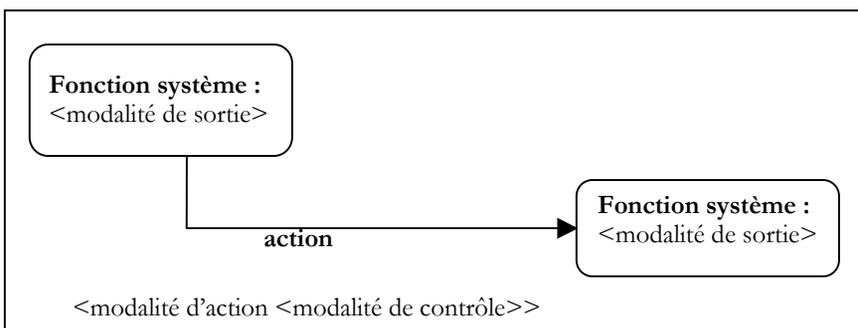


Figure 6. La granularité d'UMAR est relative aux modalités

Les diagrammes UMAR qui fonctionnent en parallèle se synchronisent par le biais de variables d'état et de gardes. UMAR ne supporte pas pour l'instant certains mécanismes tels que la mémorisation de l'état d'un graphe lors de sa préemption. Pour rendre compte de la multimodalité du système et des propriétés CARE sans expliciter toutes les combinaisons d'enchaînement qu'elles impliquent, nous utilisons une notation particulière. Celle-ci est basée sur les opérateurs °, & et ||, représentant :

- La complémentarité, notée « A ° B ». Elle implique que les modalités A puis B, ou B puis A doivent être utilisées pour réaliser la commande.
- La redondance, notée « A & B ». Elle implique que les modalités A et B doivent être utilisées.
- L'équivalence, notée « A || B ». Elle implique que la modalité A ou B peut être utilisées indifféremment.
- L'assignation n'a pas de notation spécifique, il s'agit en effet d'un choix d'implémentation. Elle pourra par contre être observée dans la description.

4.3 Exemple

Nous allons décrire un exemple pour rendre compte des possibilités du formalisme UMAR. La figure 7 décrit un système interactif de consultation de messages, à partir d'une liste de courriels. Le diagramme E2 montre les différentes manipulations relatives au courriel : l'utilisateur peut ouvrir un mail pour le lire, y répondre, ou déclencher l'action réponse directement de l'état précédent. Dans cet exemple, les modalités d'interaction peuvent être combinées (M5). L'exemple montre également l'usage du parallélisme par le biais du diagramme E1 qui montre les différents modes d'utilisation du système (tactile seul, vocal seul et dual) ainsi que la dépendance de E2 sur E1. Par exemple, l'interaction M5 de E2 n'est possible que si le système est en mode dual, mode défini par l'état actif de E1. Le symbole ∞ indique que le temps d'interaction n'est pas limité (il est en fait dépendant des décisions de l'utilisateur).

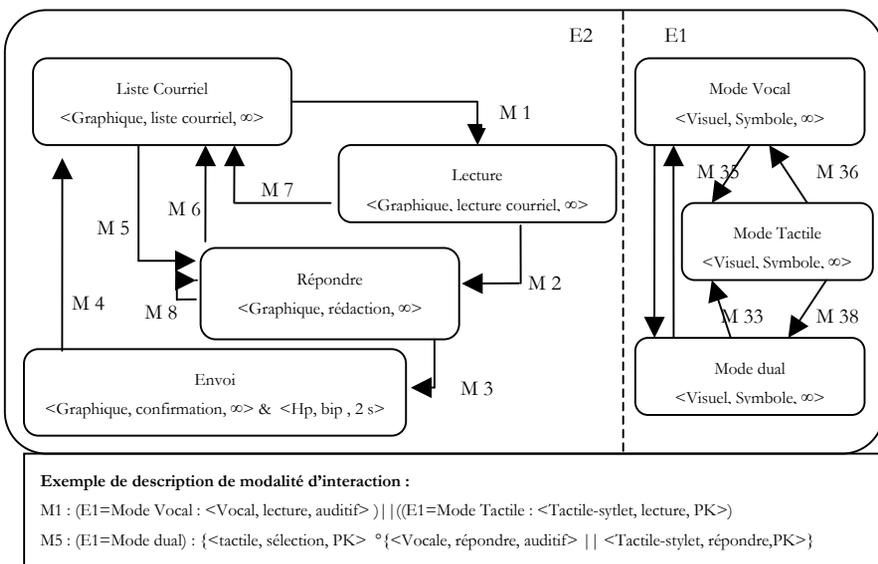


Figure 7. Exemple de description UMAR

La modalité M8 représente une abstraction fonctionnelle : la tâche de réponse à un courriel est complexe, composée de nombreuses sous-tâches d'interaction avec le clavier, si on ne souhaite pas raffiner cette description, il suffit d'indiquer les modalités qu'elle occupe par des arcs rebouclant. Par conséquent, le simulateur ne pourra qu'exhiber ce comportement abstrait.

Nous avons utilisé UMAR pour décrire différents systèmes interactifs tels que les systèmes immersifs de réalité virtuelle des PDAS ou des distributeurs de billets.

Usage par l'ergonome :

L'utilisation d'UMAR au sein de SHIMM est effectuée par le biais d'éditeurs permettant de renseigner les états, les arcs et les modalités. L'éditeur de scénarios permet quant à lui de spécifier soit les états du système à atteindre par l'utilisateur, soit les états que l'on force pour provoquer par exemple une sonnerie téléphonique. Ce sont les modules composant le mécanisme de décision de l'utilisateur virtuel qui se chargeront de traiter ces informations en calculant par exemple le chemin à parcourir pour atteindre un état du système interactif. Cet utilisateur émet les modalités d'actions qui sont interprétées pour exhiber les modalités de sorties correspondantes.

5 Modélisation de l'utilisateur virtuel

Comme cela a été expliqué au paragraphe 2, nous nous concentrons sur l'impact de l'environnement en terme de prise de décision et non sur la simulation du système sensori-moteur de l'utilisateur. Le modèle de l'utilisateur virtuel doit permettre d'exhiber des tendances décisionnelles crédibles et diverses sans s'attacher à une modélisation précise et figée telles que celles rencontrées généralement au sein des travaux résumés au paragraphe 2.2. Rappelons que l'obtention d'une telle précision est très délicate et sujette au problème du rendu des variations interindividuelles. Nous nous sommes donc concentrés sur l'identification de critères de crédibilité attendus grâce à des remarques et observations d'ergonomes et de spécialistes du comportement. Il faut que les réactions de l'utilisateur virtuel reflètent des notions telles que la motivation, l'agacement, l'erreur ou l'abandon de tâche suite à l'impossibilité d'usage d'une modalité d'action ou de contrôle. Le modèle de l'utilisateur doit rendre compte également d'erreurs liées à l'impossibilité de percevoir une modalité de sortie. Le modèle qui en suit est pragmatique, il évalue les interférences perceptives et leurs impacts sur la prise de décision ; c'est pourquoi nous l'avons nommé SCU (Sensitive and Cognitive User). Ce modèle est synthétisé en figure 8.

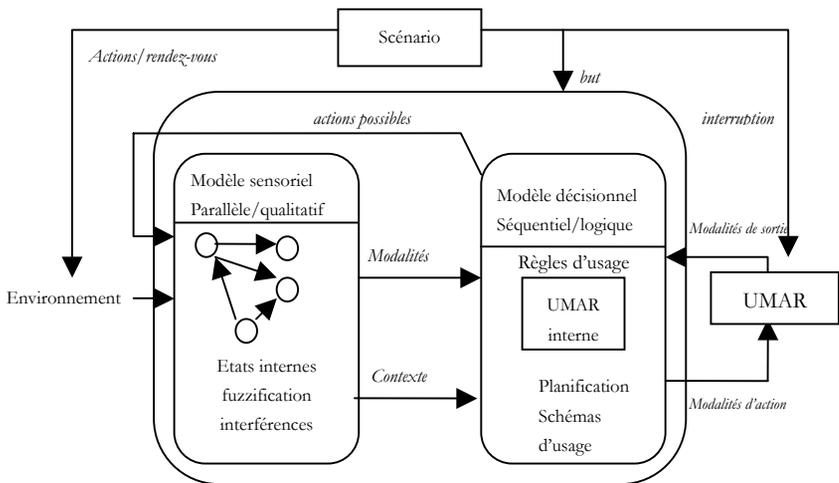


Figure 8. Modèle de décision SCU

Il est composé d'un modèle sensitif qui détermine l'inhibition de modalités en fonction de la perception des perturbations, à l'aide de critères psychologiques qualitatifs. Il renvoie également une abstraction de l'environnement et de l'état mental de l'utilisateur que nous synthétisons sous le terme de *contexte*.

Le modèle décisionnel utilise ces informations pour rechercher et planifier des séquences d'actions pertinentes. Pour cela, il analyse les buts imposés à l'utilisateur par le scénario, recoupe ceux-ci avec sa connaissance du système qui est une description plus ou moins partielle du modèle UMAR du système interactif, appelée « UMAR interne » et le *contexte* qui fera l'objet du paragraphe 5.2. Il en déduit le choix d'une action à un instant donné destinée au système interactif multimodal. Cette action est synthétisée par l'émission d'une modalité d'action vers UMAR.

5.1 Modèle sensoriel

Ce modèle est basé sur l'utilisation de différentes cartes cognitives floues (Dickerson et Kosko, 1994), ou FCM (Fuzzy Cognitive Map) en anglais. Les FCMs sont des graphes d'influence entre des concepts nommés possédant un degré d'activation continu. Les arcs orientés entre concepts représentent des liens d'excitation ou d'inhibition selon qu'ils sont labellisés d'un + ou d'un - . Comme a pu le montrer (Aguilar, 2005), les cartes cognitives floues, qui devraient plutôt se nommer cartes connexionnistes floues, constituent un véritable langage de prototypage de comportement. Elles ont d'ailleurs été utilisées dans différents domaines : en économie, où elles sont le support d'un modèle issu de la théorie de la décision dans (Axelrod, 1976), pour la modélisation d'un univers virtuel dans (Dickerson et Kosko, 1994) ou pour celui du comportement d'un barreur virtuel dans (Paranthen *et al.*, 2002). Pour la simulation comportementale, les FCMs possèdent des concepts d'entrée issus de l'environnement ou du scénario, des concepts de sorties qui dans notre cas seront utilisés par le modèle décisionnel et des concepts internes propres à rendre compte des mécanismes perceptifs et sensoriels. Nous avons représenté, en figures 9 et 10 et à titre d'exemple, des extraits de la carte cognitive correspondant à un modèle possible d'utilisateur. La description de la carte finale correspondante, possédant 27 concepts et 35 arcs (Lebodic, 2005) serait fastidieuse et n'apporterait pas d'éléments pertinents pour une compréhension globale de l'approche. Il faut garder à l'esprit que l'ergonome peut modifier celle-ci à souhait et que les hypothèses que nous relatons ici, même si elles sont le résultat de collaborations avec des psychologues, peuvent tout à fait être discutées.

Le premier extrait de la carte (figure 9) montre des concepts liés à la notion d'abandon d'une tâche. Il y a quatre concepts d'entrée : deux sont issus du scénario au sein duquel l'ergonome peut spécifier l'importance d'un but (sur une échelle de 0 : peu important à 1 : très important) ainsi que la pression temporelle associée à la scène qui reflète la nature stressante du scénario. Les deux autres concepts d'entrée sont également liés au scénario mais de façon indirecte : il s'agit de la mobilité et de la notion de double tâche. La mobilité rend compte du fait que l'utilisateur virtuel est en déplacement, tandis que la double tâche indique qu'il est déjà en train d'effectuer une autre tâche. Ces deux concepts dépendent de la situation de l'utilisateur à chaque instant et des décisions prises précédemment. Les deux concepts internes sont la motivation et les ressources cognitives. Leurs rôles sont complémentaires : la motivation empêche l'abandon tandis que la surcharge cognitive l'entraîne. Les différents arcs reflètent les postulats suivants : l'importance du but augmente la motivation. Celle-ci à un effet stimulant, c'est pourquoi un arc rebouclant a été ajouté. Informellement, il signifie que « plus on est motivé, plus on se motive ». Notons que pour éviter des valeurs infiniment croissantes, l'algorithme

de simulation des cartes cognitives est défini de sorte que les valeurs d'activation tendent asymptotiquement vers des minima ou des maxima déterminés. La mobilité, l'activité double et la pression temporelle augmentent la quantité de ressources cognitives utilisées, tandis que la motivation diminue celle-ci. En fait nous pourrions plutôt dire qu'elle rend l'utilisateur plus tolérant à la charge cognitive donc qu'elle diminue celle-ci indirectement. Enfin, le concept « abandon » est augmenté par la charge cognitive et diminué par la motivation. Concrètement un seuil est fixé pour définir si l'abandon est effectif ou non.

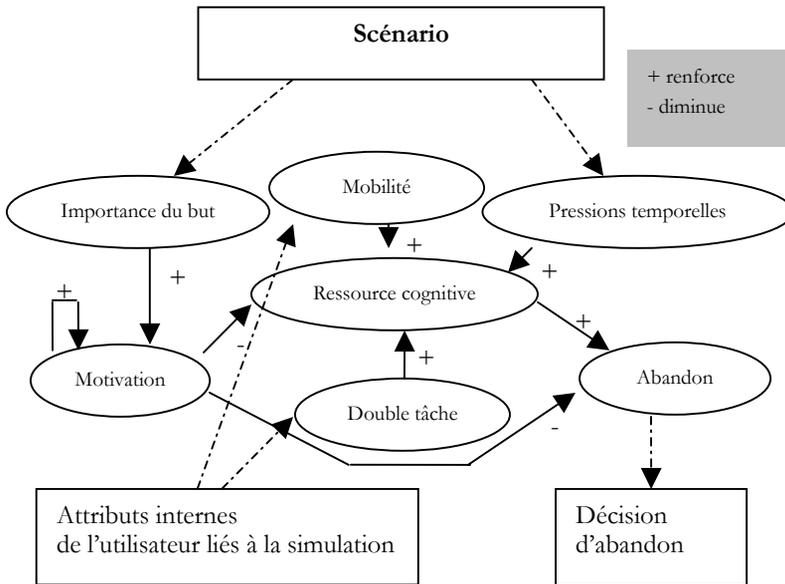


Figure 9. Les influences sur la décision d'abandon

Le deuxième extrait de carte (figure 10) « calcule » les interférences de l'environnement et leurs impacts sur l'élection des modalités de sortie vocale et tactile. Les concepts d'entrée liés à l'environnement sont la perception de bruit, de voix de foule ou de vibration. Il est à noter que, grâce à oRis, l'environnement virtuel rend compte d'une intensité sonore ou d'une présence d'un objet informé (tel que la foule) associé à une distance. Ces valeurs peuvent être réinterprétées pour définir le degré d'activation des concepts par le biais d'une fonction de transformation floue. Ces fonctions ont été définies pour les modalités du tableau 1 mais peuvent être re-définies à souhait par l'ergonome, au travers d'un fichier de configuration. Notons sur cette carte des concepts d'entrée liés aux connaissances ou habitudes de l'utilisateur, utilisés par des schémas d'usage typiques⁹ (tactiles ou vocaux) qui sont présents dans la base de connaissances du modèle décisionnel présenté au paragraphe suivant. La figure 11 illustre l'aspect dynamique du modèle : elle montre une évolution possible de la valeur d'activation du concept « vocal

⁹ un schéma typique peut représenter une habitude : si je connais bien la modalité gestuelle, j'aurai tendance à l'utiliser souvent. À l'inverse, si je ne connais pas la modalité vocale, en cas d'équivalence, j'aurai tendance à ne pas la sélectionner.

interdit » en fonction de la perception d'un bruit et de l'activation d'un schéma typique utilisant la modalité vocale. Pour plus de détails sur les spécifications de la dynamique des cartes cognitives, nous renvoyons le lecteur à (Kosko, 1986).

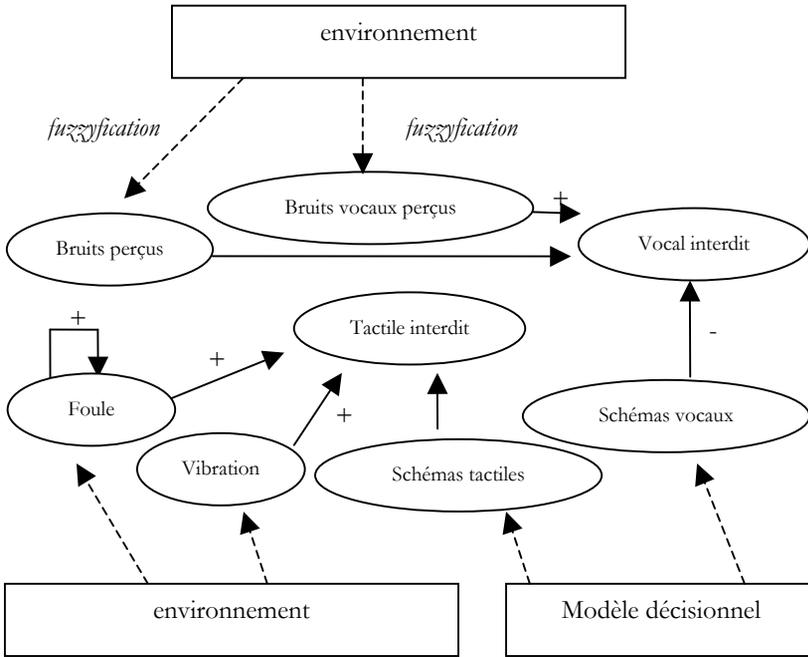


Figure 10. Modélisation de quelques interférences environnementales

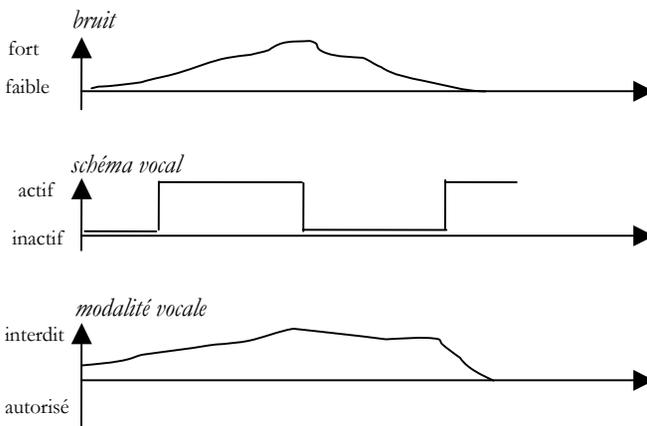


Figure 11. Exemple de l'influence de deux concepts d'entrée sur l'usage d'un modalité

Usage par l'ergonome :

Dans SHIMM, un outil graphique de description des FCMs permet à l'ergonome de tester rapidement de nombreuses hypothèses sensibles en modifiant les arcs ou les concepts. Il peut également redéfinir les niveaux des fonctions floues et faire apparaître les comportements correspondants. Quantitativement, le modèle utilisateur est actuellement traduit en un fichier de quelques centaines de lignes et la spécification des différentes fonctions de *fuzzification* en une soixantaine de lignes. La simulation des FCMs est ensuite effectuée grâce à oRis.

5.2 Le contexte

Le contexte est le moyen d'articuler le lien entre le modèle sensitif et le modèle décisionnel. La notion de contexte fait l'objet de débats et de recherches en intelligence artificielle (Brézillon, 1996). Synthétiquement, nous pouvons le définir comme l'ensemble des éléments pris en compte à un instant donné pour prendre une décision. Une des difficultés réside en une description pertinente de ces éléments qui permette un traitement aussi générique que possible afin de faire un modèle de décision également générique. Pour cela nous proposons d'utiliser neuf des concepts du modèle sensitif qui synthétisent l'état mental de l'utilisateur ainsi que sa perception de l'environnement.

Concept	Signification	Valeurs possibles
L	<i>La luminosité ambiante</i>	{faible, moyenne, forte}
B	<i>Bruit ambiant</i>	{nul, faible, moyen, fort}
V	<i>Voix interférentes</i>	{absent, présent}
A	<i>Contexte affectif</i>	{seul, accompagne, surveillé}
T	<i>Pression temporelle</i>	{faible, forte}
Bu	<i>Importance du but</i>	{faible, moyen, fort}
M	<i>Mobilité</i>	{oui, non}
D	<i>Activité double</i>	{oui, non}
C	<i>Ressource cognitive</i>	{faible, moyenne, forte}

Tableau 4. Signification des variables de modélisation du contexte

A chacun de ces concepts sont associées des valeurs floues. Ainsi, le contexte, base d'entrée du modèle décisionnel est le 9-uplet : {L,B,V,A,T,Bu,M,D,C} dont les valeurs et la signification sont référencées dans le tableau 4 : même si une telle définition peut sembler arbitraire, il s'avère qu'à l'usage, elle permet de distinguer des cas d'utilisation caractéristiques avec le moteur décisionnel et s'avère pertinente pour la description de schémas typiques et l'apprentissage contextuel (voir paragraphe suivant). Enfin, l'utilisation d'un langage déclaratif pour la description du modèle décisionnel permet facilement d'ajouter des concepts ou de modifier les valeurs qui les caractérisent pour modifier la définition du contexte.

5.3 Modèle décisionnel

Le modèle décisionnel est un système à base de règles. Etant donnés nos objectifs de prototypage interactif, donc d'ouverture et de malléabilité, et n'ayant pas l'ambition de redéfinir une architecture cognitive exhaustive, nous avons opté pour l'utilisation de GnuProlog¹⁰ qui s'interface avec oRis facilement par le biais d'une A.P.I¹¹. La version actuelle de la base de règles du modèle contient environ

¹⁰ <http://gnu-prolog.inria.fr/>

¹¹ Application Programming Interface, développée par Pierre-Alexandre Favier (Favier, 2004).

500 règles. La base de faits est quant à elle très dépendante de l'application cible. Elle contient trois ensembles de faits distincts :

- 1) Les actions possibles associées au système interactif, décrivant les modalités d'entrées et de sortie, ainsi que les arcs définissant le fonctionnement *connu par l'utilisateur* du système interactif. Il peut donc s'agir du diagramme UMAR lui-même ou d'une version simplifiée de celui-ci, que nous avons nommée *UMAR interne* et que l'utilisateur virtuel peut découvrir. La figure 12 illustre la description en Prolog des modalités d'entrée (associées à une modalité de contrôle) et de sortie telles qu'elles ont été présentées au paragraphe 4.1. Elle illustre également la description d'un arc de UMAR. Notons qu'un niveau de connaissance est associé à cet arc. Il s'agit d'une variable numérique reflétant le fait que l'utilisateur ne connaît pas (0) ou connaît très bien (5) la fonctionnalité associée à l'arc. Cette base de faits est générée automatiquement par l'éditeur de UMAR.
- 2) Les schémas typiques (ou prototypiques) (Schank et Abelson, 1977) reflètent les habitudes et tendances d'utilisation généralement reliées au contexte d'utilisation. Un schéma typique dépend du contexte représenté par les valeurs des neuf variables proposées en 5.2. Y sont associés les états de départ et d'arrivée du système interactif connus par l'utilisateur, ainsi que la ou les modalités d'usage associées et favorites. Enfin, le succès ou l'échec récent de l'utilisation de ce schéma peut être référencé et utilisé par la base de règles. Le modèle sensitif est influencé par les schémas issus de cette base de faits (voir paragraphe 5.1).

```
% modalité d'entrée : <Physique, Représentation, Contrôle, Durée>
action(styilet, scrollhb, [graphique, pk],2).
action(clic, selection_active, [graphique, pk],1).

%modalité de sortie : <Physique, Représentation, Durée>
sortie(graphique, selectionListe, inf).

% arc UMAR : <départ, arrivée, connaissance, multimodalité CARE>
% dans cet exemple, la propriété CARE est une équivalence entre
% deux modalités
arc(etatA, etatB, 3, modalité(action(parole, inserer,[sonore],2),
action(clic, icone, [graphique, pk], 1), equivalence).

% Schema typique : <contexte, etatDépart, étatArrivée, [modalités
d'usages], reussite>
schema(l=faible/b=fort/v=absent/a=seul/t=forte/bu=faible/m=oui/d=
non/
c=faible, etatA, etatC, [tactile, vocale], succes).
```

Figure 12. Exemple de faits Prolog du modèle de décision

- 3) Les routines (Amalberti, 2001), très proches des schémas typiques, reflètent également une habitude d'usage. Cependant, alors qu'un schéma typique va influencer le choix d'une modalité d'action, une routine va être exécutée de façon instinctive sans aucune réflexion, surtout en cas de pression temporelle forte. Celles-ci sont interprétés et peuvent être générées par le système de règles décrit en figure 12. Les prédicats *action/4* et *sortie/3* décrivent les

modalités telles que décrites page 13 (l'atome *inf* représentant le symbole ∞). Le prédicat *schéma/3* caractérise le triplet {contexte, modalité, résultat} constituant les schémas typiques. Le contexte est un terme composé *Prolog* permettant de décrire de façon qualitative symbolique la valeur des 9 variables du tableau 4.

La base de règles implémente les mécanismes de sélection et de traitement de l'information. Ils sont inspirés des études de (Reason, 1990) sur l'interaction homme-machine. Pour lui, les connaissances sont sélectionnées suivants trois protocoles :

- 1) Le mécanisme d'erreur qui provoque la sélection d'actions inopportunes en fonction d'activateurs fréquentiels, de récence et de primauté. En fait, nous avons adapté cette notion en intégrant la notion de distance contextuelle. Nous avons introduit des opérateurs permettant de comparer les neuf concepts constituant le contexte. Cette comparaison est empirique : ainsi si 8 concepts sont identiques pour deux contextes C1 et C2, ceux-ci seront moins distants que deux contextes C3 et C4 pour lesquels seuls 7 concepts ont la même valeur. Si un schéma typique ou une routine ne sont pas trop éloigné du contexte courant, alors il sera sélectionné pour influencer la prise de décision.
- 2) La sélection par similarité : elle est appliquée quand une connaissance coïncide parfaitement avec le but à réaliser, les informations perçues de l'environnement et l'état interne de l'utilisateur virtuel. Il s'agit typiquement de l'utilisation de connaissances de type routine. Notons que le système prend en compte les niveaux de connaissance introduits précédemment pour trancher entre deux actions possibles. Si toutefois celles-ci possèdent le même niveau de connaissance, alors un tirage aléatoire est effectué.
- 3) Le mécanisme de recherche systématique propose un calcul du meilleur chemin en fonction de la disponibilité de ressources cognitives. Ici, c'est le modèle du système interactif connu par l'utilisateur qui est la base de la recherche. Celle-ci demande néanmoins des ressources cognitives importantes. Ce mécanisme peut être utilisé pour simuler un débutant. Le système va générer, au fur et à mesure des simulations, des routines qui constitueront la connaissance apprise au fur et à mesure des expériences.

Usage par l'ergonome

Par défaut, l'ergonome n'a rien à définir pour le modèle décisionnel. Celui-ci peut utiliser le diagramme UMAR pour définir la base de faits minimale sans routines ni schémas typiques. L'introduction de ces derniers demande une spécification explicite au sein d'un fichier proche de la forme visible en figure 11. Ensuite, l'utilisation des différents protocoles de Reason est activable séparément par le biais d'options intégrées au simulateur. L'ergonome a accès à un certain nombre de paramètres tels que le nombre de répétitions d'une interaction avant que celle-ci ne se transforme en routine, ou le nombre de concepts identiques caractérisant des contextes semblables.

6 Usage et Evaluation de SIHMM

L'évaluation d'une telle approche est un problème délicat car elle nécessite d'étudier des scénarios d'usages réels complexes car se réalisant dans des environnements riches et dynamiques. Nous devons reconnaître qu'en l'état, nous ne pouvons montrer formellement des comparaisons statistiques corrélant les choix d'un utilisateur virtuel avec ceux d'un utilisateur réel. De plus, les différences inter-

individuelles sont telles, qu'une étude statistique est périlleuse et demanderait une classification de comportements typiques et un échantillonnage des testeurs selon cette classification, travail de longue haleine que nous n'avons pas pu effectuer. David E. Kieras relate les problèmes similaires qu'il a rencontré et en déduit d'ailleurs que : « *plutôt que de valider un modèle théorique à partir de données expérimentales (difficilement accessibles), il vaut mieux l'exécuter et le donner à observer aux spécialistes qui vont trouver les 'lacunes' de celui-ci de façon évidente* » (Kieras et Santoro, 2004). Pour notre part, nous avons entrepris une démarche complémentaire en développant SIHMM en partenariat avec des ergonomes de France Telecom et des psychologues du CERV et en effectuant tout de même des expérimentations. Ainsi psychologues et ergonomes nous ont fait profiter de leur expérience et de leurs réflexions pour étayer nos modèles. Ils nous ont fourni de nombreux retours concernant les modèles et les simulations associées. Par exemple, ils sont à l'origine de la prise en compte des temps de réaction du système interactif et de l'introduction du concept de motivation au sein du modèle perceptif. Des corrections ont donc été régulièrement apportées et nous permettent d'affirmer que notre approche répond à certains besoins pour rendre compte de l'usage d'interfaces multimodales. Cette démarche permet en tout cas de mieux cerner les problématiques de la validation écologique de ces interfaces.

Les expériences ont été menées en partenariat avec France Telecom. Ces expérimentations sont très lourdes et différents problèmes logistiques ont rendu leur exploitation délicate. En effet, elles ont nécessité l'usage de systèmes expérimentaux qui ont eu de nombreuses défaillances et qui ont largement perturbé la naturalité de l'interaction observée, ne nous laissant que quelques fragments temporels exploitables. Ceci rend très délicat la comparaison virtuel/réel sur des notions telles que l'apprentissage de routines ou l'usage de schémas typiques, corroborant ainsi les affirmations de D.E. Kieras. Néanmoins, ces expériences ont été simulées avec SIHMM et les observations non quantifiées des ergonomes ont pu être en partie reproduites. L'expérimentation consistait à inciter des usagers, par le biais d'un jeu, à dialoguer au travers de PDAs multimodaux en récoltant des informations à la Cité des Sciences de La Villette. La figure 13 montre une copie de quelques vues de SIHMM.

Nous allons illustrer son usage sur le scénario suivant : « *Christine est attirée par l'innovation. Elle utilise facilement la modalité vocale sur son PDA. Elle se promène dans la Cité des Sciences où il y a quelques badauds et surtout une foule bruyante. Elle cherche le numéro d'un autre joueur tout en se déplaçant vers une salle d'exposition pour gagner du temps* ».

Partant de ce scénario, nous pouvons donner une idée de la tâche de modélisation que doit effectuer l'ergonome.

Modélisation de l'environnement

Les éléments de la scène sont tous dans la librairie du DPE (voir §3). Sa construction à la souris ne pose pas de problèmes particuliers. L'ergonome indique simplement les gênes que peuvent occasionner chacun d'eux sur l'utilisateur. En particulier, une gêne sonore pour la foule et une gêne sociale pour les badauds et la foule. La figure 13 montre la vue proposée à l'ergonome.

Modélisation du système interactif

Le fonctionnement du PDA est décrit avec un diagramme UMAP comportant 34 états et 47 arcs. Les états sont regroupés en 3 diagrammes fonctionnant en parallèle et les arcs reflètent des interactions multimodales de type tactile, vocale ou gestuelle. Le fichier XML résultant comporte 3700 lignes.

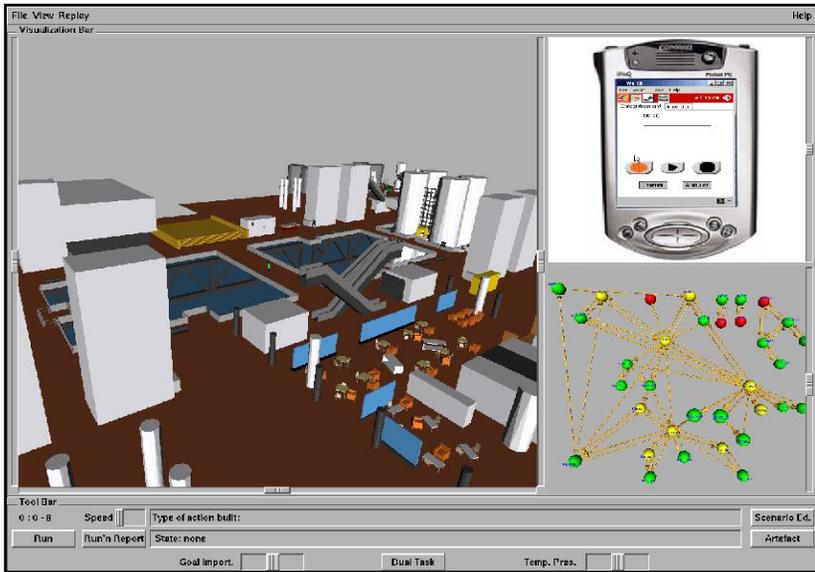


Figure 13. Quelques vues offertes par le logiciel SIHMM

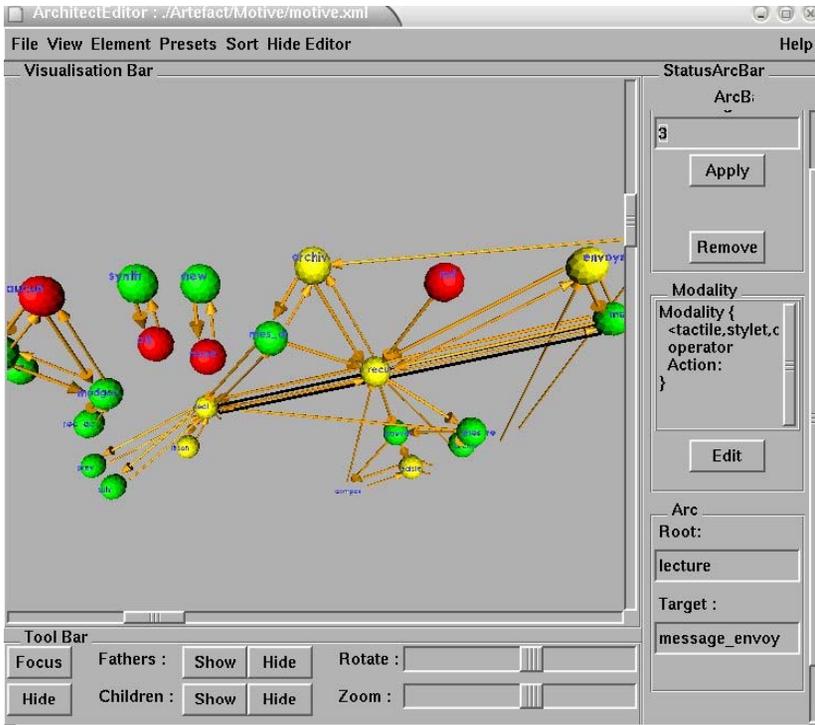


Figure 14. Une vue du diagramme UMAR du scénario

Prenons l'exemple du passage du système vers un état proposant l'envoi d'un nouveau message. Dans l'interface de saisie de cet état (figures 14 et 15), l'ergonome doit renseigner les champs tels que l'état initial, la multimodalité de sortie associée ou le fait que cet état puisse ou non être préempté par le scénario. La multimodalité de sortie est décrite conformément à la formalisation proposée en §4.1 ; pour cela un éditeur dédié est proposé à l'ergonome. Ici, il permet d'indiquer qu'il s'agit d'une redondance entre la modalité graphique de type icône à l'écran et un signal sonore de 2s. Ce qui est noté « modality{<graphique, icône, infini>&<parole, notify, 2>} », où infini caractérise le fait que le temps de l'interaction n'est pas défini. Les arcs sont saisis de façon similaire, la figure 15 montrant l'éditeur associé.

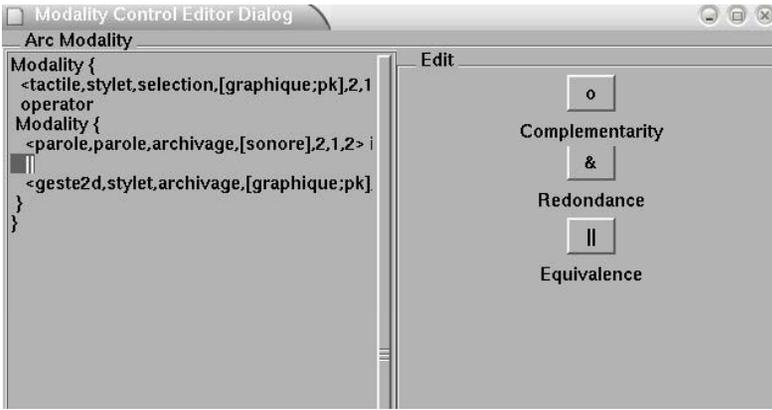


Figure 15. Assistant de saisie de modalité CARE

Modélisation de l'utilisateur

L'ergonome peut se baser sur le modèle sensitif par défaut, de l'utilisateur comportant 27 concepts et 35 arcs (voir §5.1). Il peut travailler à partir d'une représentation graphique ou de l'éditeur illustré en figure 16.



Figure 16. Configuration du modèle sensitif (gauche) et décisionnel (droite)

Cet éditeur permet de spécifier les valeurs floues caractéristiques associées aux concepts d'entrée, ainsi que les symboles associés aux couples (concepts/valeurs) qui seront utilisés pour caractériser le *contexte*. Les valeurs symboliques telles que *faible* ou *fort* sont ici représentées par des nombres. Le modèle décisionnel est paramétrable par le biais de l'éditeur. Il permet d'indiquer l'usage de routines ou celui d'apprentissage de routines, d'indiquer la modalité favorite et d'autoriser la mesure de distance entre deux contextes qui pourra générer des erreurs de confusion. Il permet également de caractériser le nombre d'actions que considèrera le modèle comme actions récentes. Enfin, il permet de spécifier la profondeur maximum autorisée de l'arbre de recherche pour caractériser l'effort cognitif acceptable par l'utilisateur.

Le scénario

L'éditeur de scénarios permet de définir des zones temporelles associées aux actions des agents du DPE, du système interactif ou de l'utilisateur virtuel. Les actions associées au système interactif correspondent à des changements d'état du graphe UMAR. Pour cela, il faut lors de l'édition de ce graphe, spécifier si un état est interrompant (capable de prendre la priorité sur les autres états du graphes) ou interruptible (capable d'être quitté à tout moment). Les actions spécifiques associées à l'utilisateur virtuel sont des buts à atteindre, qui sont en fait des états du graphe UMAR ou des déplacements de zone en zone. Ces déplacements sont également les actions des autres éléments du DPE.

Exploitation des résultats :

Nous distinguons deux classes de résultats, ceux qui correspondent aux comportements observés par les utilisateurs de SIHMM et qui a été la principale base d'enrichissement des modèles et les fichiers de *log* en XML dont l'exploitation statistique confirme les observations mais n'a de sens qu'en comparaison avec les résultats de l'observation de situations réelles.

<pre> <sihmm-trace> <ModalityAction>[tactile, clic, icone_editer, [graphique,pk], 1, 0, 0] </ModalityAction> - <State> <Date>126</Date> <Name>compose_1_mail</Name> - <FeedBackModality> <Phys>graphique</Phys> <Rep>fenetrenewmail</Rep> <Duree>1077693813</Duree> </FeedBackModality> </State> </pre>	<pre> <ModalityAction>[tactile, stylet, boutonback, [graphique,pk], 2, 1, 0]</ModalityAction> - <Gene_Sens> <Sono>77.4351183342</Sono> <Toucher>77.4351183342</Toucher> <Odorat>77.4351183342</Odorat> <Gout>77.4351183342</Gout> <Light>77.4351183342</Light> <Gest>77.4351183342</Gest> <Socio>77.4351183342</Socio> </Gene_Sens> <Mobility>0</Mobility> <Parole>0.393743932247</Parole> <Tangible>0.39211666584</Tangible> <Tactile>0.37521097064</Tactile> <Geste2D>0.0451345108449</Geste2D> <Charge_Cognitive>0.832071840763</Charge_Co gnitive> </pre>
---	--

Tableau 5. Extrait d'un fichier de log de SHIMM

Pour l'analyse durant la simulation, SIHMM fournit une vision en 3D des déplacements, des interactions, associés à un affichage de l'état de l'utilisateur et de ses décisions en terme d'interaction avec le système interactif. Il visualise également une animation du graphe UMAR, à laquelle est associée une représentation graphique de l'interface correspondante (voir figure 13). Il est alors possible de constater que le système interactif est inexploitable en présence de bruit car il n'offre pas de modalité redondante à la modalité vocale. La carte cognitive de l'utilisateur virtuel voit son concept d'abandon augmenter jusqu'à ce que celui-ci soit considéré par le modèle décisionnel.

Pour l'analyse statistique des résultats, les fichiers XML correspondants contiennent une description datée des actions entreprises par l'utilisateur ainsi que la représentation du contexte au moment de chaque prise de décision. Un court extrait d'un tel fichier est représenté tableau 5. Ces résultats vont mettre en évidence des tendances comportementales telles que l'usage de la modalité préférée (vocale) de Christine dans 80% des cas de notre exemple. Les autres cas vont révéler une gêne sociale et vocale forte l'incitant à passer en mode tactile-clavier. Cette gêne est liée à l'approche de la foule qui se déplace aléatoirement dans l'environnement. Puisque les temps d'interaction liés à cette modalité sont plus élevés, il en résulte un délai supplémentaire pour atteindre le but requis. On peut également observer que la multiplication des interactions à effectuer en présence de la foule va progressivement introduire des routines d'usage de la modalité tactile qui sera alors utilisée au détriment du vocal. Ces résultats peuvent sembler triviaux, il est clair que le simulateur rend compte des paramètres qui lui ont été inculqués et que sur des scénarios simplistes, la simulation semble d'un faible apport même s'il était dans un premier temps, nécessaire de retrouver de tels résultats. Ceci nous amène à une discussion sur les perspectives de ce travail.

7 Discussion et perspectives

Nous venons de présenter une architecture de modèles couvrant l'ensemble utilisateur-environnement-système interactif, ainsi qu'un outil support de ces modèles, permettant de les simuler et de les modifier avec facilité pour un prototypage interactif d'interfaces multimodales. Il est clair qu'une telle approche ne peut couvrir l'ensemble des besoins identifiés pour la mesure de l'utilisabilité de ces interfaces. Néanmoins, elle constitue un premier pas vers une intégration globale nécessaire à une approche écologique (Gibson, 1979).

De plus, elle fait une proposition concernant le modèle de l'utilisateur. Ce dernier offre une structure souple (cognitivisme + connexionnisme) mais réfléchit combinant les aspects qualitatifs parallèles pour les perceptions, couplés à un système à base de règles dont plusieurs fonctionnalités peuvent être inhibées ou paramétrées. Cette souplesse permet de ne pas être aussi précis que ce que proposent les architectures cognitives classiques. Ceci est à la fois un avantage et un inconvénient certainement lié à une approche plus connexionniste, voire constructiviste que cognitiviste du comportement humain.

En effet, un des axes fédérateurs du Centre Européen de Réalité Virtuelle s'articule autour de la cré-action (Tisseau, 2001) : la nécessité de tester les modèles pour les comprendre et les améliorer. Ici l'ergonome construit sa simulation, peut modifier la carte cognitive de l'utilisateur, supprimer une modalité pour constater l'influence de l'usage du système ainsi amputé, etc. Néanmoins, cette première expérience nous permet de tirer de nombreuses leçons : pour l'instant, les tests effectués peuvent sembler moyennement pertinents. D'où vient le problème ?

Certainement de la faible complexité rencontrée dans les exemples traités. Le scénario décrit dans cet article a vertu d'explication, il est donc simple et les résultats qu'il permet d'exhiber semblent frappés de bon sens. De façon caricaturale, la question peut se poser de savoir s'il est bien utile de modéliser tout ce système si cela permet de constater que lorsqu'il y a du bruit, il n'est pas possible d'entendre son téléphone ! Pourtant, l'évaluation d'usage est une problématique d'actualité car elle est confrontée à la complexité, complexité des interactions parallèles multimodales, complexité des situations. Les ergonomes travaillent à partir de scénarios d'usage pour imaginer des situations typiques, singulières. Par conséquent, il nous faut rendre compte de cette complexité et de cette singularité pour montrer l'intérêt de l'usage des simulations, exhiber des comportements non triviaux et trouver des failles dans les mécanismes d'interaction.

Il nous semble alors moyennement intéressant de faire des statistiques d'usage sur de nombreuses simulations, sauf s'il s'agit de mesurer une difficulté moyenne pour un utilisateur moyen. A l'inverse, nous cherchons à multiplier les simulations pour trouver un cas particulier d'usage impossible ou difficile. Ces cas sont particuliers car non envisagés trivialement par l'ergonome ou le développeur de système interactif. Ils ne seront certainement pas mis en évidence par une validation formelle qui implique souvent une simplification.

La figure 17 illustre ce propos de la façon suivante : les comportements possibles sont représentés par le cercle A, tandis que les différents cercles C correspondent à ce que peut exhiber un modèle. Un modèle complexe et souple (C_n) permet de couvrir une plus grande surface du cercle A (zone B) mais peut en sortir, un modèle plus précis, ne couvrira qu'une partie de A très fine (C_1) tandis qu'un modèle réductionniste qui simplifie selon un point de vue particulier, en couvrira une partie plus large mais moins complexe (C_3).

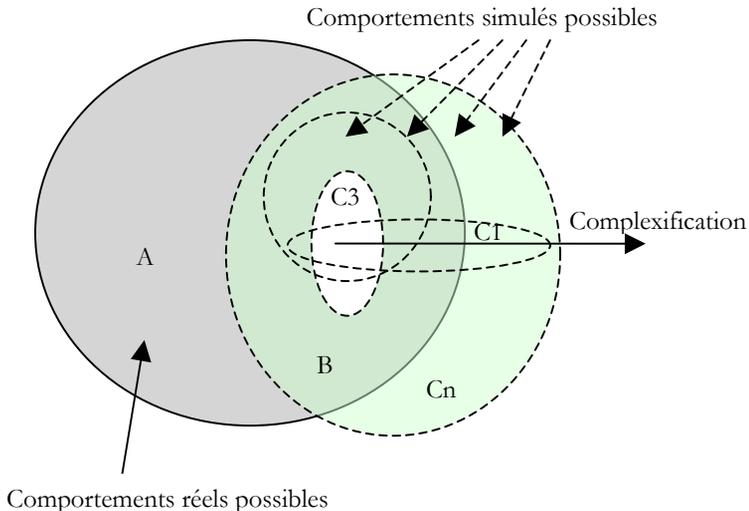


Figure 17. *Complexité simulée versus complexité réelle*

Notre approche consiste à agrandir C au risque de sortir de A. Pour atteindre cet objectif, nous devons complexifier les scénarios, les interactions et les capacités

de l'utilisateur virtuel. Pour les scénarios, il faut pouvoir introduire une arborescence dans la spécification des buts de l'utilisateur.

Concernant l'utilisateur, il faut pouvoir lui fixer plusieurs buts à atteindre en parallèle, pas forcément en rapport avec le système interactif. Il faut augmenter son autonomie, lui permettre d'agir sur d'autres éléments que ce système pour résoudre un problème d'interaction. L'introduction de notions telles que la pression temporelle est une première tentative pour imiter de tels cas. Pour les interactions, il faut implémenter des systèmes plus complexes, nous envisageons par exemple la modélisation de systèmes immersifs de réalité virtuelle. Cependant, même si nous ne cherchons pas une exactitude comportementale illusoire, la modélisation du contrôle moteur proposée par des architectures telles que EPIC ou ACTR-PM nous semble très intéressante pour rendre compte de certains phénomènes, occultés par notre approche (et éviter de sortir du cercle A). Il faut, pour cela, évaluer s'il est nécessaire de construire des modèles dédiés ou si les cartes cognitives peuvent être le support de modélisation de ces phénomènes. De même, le raffinement de la description des interfaces tel qu'il est fait dans (Chi *et al.*, 2000) permet de mieux filtrer la complexité de la situation en évitant de simuler des comportements impossibles qui seraient considérés dans notre approche. Pour cela il faut introduire de nouveaux concepts dans UMAR, qui remettent peut-être en cause la granularité décrite dans ce document. Enfin, concernant le problème de la validation, nous pensons comme D.E. Kieras que l'expertise peut compenser sa difficile mise en œuvre. Cependant, il nous semble très important de travailler à la mise en œuvre de méthodes et de protocoles de validation orientés complexité, quasi inexistantes aujourd'hui, pour étayer nos hypothèses.

8 Remerciement

Nous tenons à remercier les personnes qui ont collaboré avec nous pour garantir la réalisation de ce travail ; nous remercions ainsi particulièrement Julien Kahn, Frédéric Falleta et Moustapha Zouinar de France Telecom R& D ainsi que Pascal Salembier de l'IRIT à Toulouse. Nous remercions également Jean Yves Guinard et Karen Martinaud du CERV pour leur contribution sur les notions psychologiques de ce travail. Enfin, nous tenons à remercier les lecteurs anonymes de la revue RIHM pour les nombreuses remarques constructives qu'ils ont pu nous faire afin d'améliorer la qualité de ce document.

Ces travaux sont l'aboutissement de travaux de doctorat de L. Le Bodic et G. Calvet et ont été principalement financés par FT R&D.

9 Bibliographie

Aguilar, J. (2005). A Survey about Fuzzy Cognitive Maps Papers. *International journal of computational cognition*, 3(2), 27-33.

Al-Qaimari, G., McRostie, D. (1999). KALDI: a computer-aided usability engineering tool for supporting testing and analysis of human-computer interaction. In *Proceedings of the Third international Conference on Computer-Aided Design of User interfaces*, J. Vanderdonckt and A. Puerta (Eds.), Kluwer Academic, 337-355.

Altmann, E.M., John, B.E. (1999). Episodic indexing: A model of memory for attention events. *Cognitive Science*, 23(2), 117-156.

Amalberti, R. (2001). La maîtrise des situations dynamiques. *Psychologie Française*, 46, 2, 107-118.

- Anderson, J.R. (1993). *Rules of the Mind*. Lawrence Erlbaum Associates, Hilldale, NJ.
- Axelrod, R. (1976). *Structure of decision*. Princeton University Press, USA.
- Aykin, N.M., Aykin, T. (1991). Individual differences in Human-Computer Interaction. *Computers & Industrial Engineering*, 20(3), 373-379.
- Brezillon, P. (1996). *Context in human-machine problem solving: A survey*. Technical Report 96/29, Laforia, Université Paris VI, Paris.
- Byrne, M.D., Anderson, J.R. (1998). Perception and Action. In J. R. Anderson & C. Lebiere (Eds.), *Atomic Components of Thought*, 167-200.
- Byrne, M.D. (2003). Cognitive architecture. In *the Human-Computer interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications*, J. A. Jacko and A. Sears, (Eds.), 97-117, Mahwah, NJ: Erlbaum.
- Card, S.K., Moran, T.P., Newell, A. (1983). *The psychology of human-computer interaction*. Lawrence Erlbaum Associates. Hillsdale, NJ.
- Carey, M.S., Stammers, R.B., Astley, J.A. (1989). Human-Computer Interaction design: the potential and pitfalls of Hierarchical Task Analysis. In D. Diaper (Ed.), *Task analysis for human-computer interaction*, 56-74. Chichester, UK: Ellis Horwood.
- Chi Hi, E.H., Pirolli, P., Pitkow, J. (2000). The scent of a site : A system for analyzing and predicting information scent, usage, and usability of a web site. In *Proceedings of the Conference on Human Factors in Computing Systems*, 161-168, NY:ACM Press, New York.
- Chi H., Rosien A., Supattanasiri G., Williams A., Royer C., Chow C., Robles E., Dalal B., Chen J., Cousins S. (2003). *The Bloodhound Project: Automating Discovery of Web, Usability Issues using the InfoScent Simulator*. Disponible à citeseer.ist.psu.edu/655522.html.
- Chong, R.S., Laird, J.E. (1997). Identifying dual-task executive process knowledge using EPIC-Soar. In *Proceedings of the Nineteenth Annual Conference of the Cognitive Science Society*. M. Shafto and P. Langley (Eds.), 107-112, Hillsdale, NJ: Erlbaum.
- Coutaz, J., Nigay, L. (1995). Four Easy Pieces for Assessing the Usability of Multimodal Interaction: The CARE properties. In *Proceedings of IFIP Int. Conf. on Human-Computer Interaction Interact '95*, Chapman & Hall, London, 1995, pp. 115-120.
- Cugini, J., Scholtz, J. (1999). VISVIP: 3D Visualization of Paths through Web Sites. In *Proceedings of the 10th international Workshop on Database & Expert Systems Applications*, IEEE Computer Society, Washington, DC, 259-263.
- Dickerson, J., Kosko, B. (1994). Virtual worlds as fuzzy cognitive maps. *Presence*, volume 3, 173-189.
- Dillon, A. et Watson, C. (1996). User analysis in HCI - the historical lessons from individual differences research. *International Journal Human-Computer Studies*, 45(6), 619-637.
- Favier, P. (2004). *La notion d'intentionnalité dans la spécification de comportements d'agents autonomes situés en environnement virtuel*. Thèse de l'université de Bretagne Occidentale, décembre.

Gibson (1979). *The ecological approach to visual perception*. Lawrence Erlbaum Associates, London.

Gwizdka, J., Chignell, M. (2004). Individual differences and task-based user interface evaluation: a case study of pending task in email. *Interacting with computers*, 16 (4), 769-797.

Harper, P.D., Norman, K.L. (1993). Improving user satisfaction: The questionnaire for user interaction satisfaction version 5.5. In *the 1st Annual Mid-Atlantic Human Factors Conference*, Virginia Beach, 224-228.

Harrouet, F. (2000). *oRis : s'immerger par le langage pour le prototypage d'univers virtuels à base d'entités autonomes*. Mémoire de doctorat, Université de Bretagne Occidentale.

Helfrich, B., Landay, J.A. (1999). *QUIP: Quantitative User Interface Profiling*. Disponible à : <http://home.earthlink.net/~bhelfrich/quip/>

Ivory, M.Y., Hearst, M.A. (2001). The state of the Art in Automating Usability Evaluation of User Interfaces. *ACM Computing Surveys*, 33(4), 470-516.

John, B.E., Kieras, D.E. (1996). The GOMS family of user interface analysis techniques: comparison and contrast. *ACM Trans. Computer-Human Interaction*, 3 (4), 320-351.

John, B., Vera, A., Matessa, M., Freed, M., Remington, R. (2002). Automating CPM-GOMS. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: Changing Our World, Changing Ourselves*, 20-25, ACM Press.

John, B. E., Prevas, K., Salvucci, D. D., Koedinger, K. (2004). Predictive human performance modeling made easy. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 455-462, ACM Press.

Just, M.A., Carpenter, P.A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, 99(1), 122-149.

Kieras, D. (1994). GOMS modeling of user interfaces using NGOMSL. In *Conference Companion on Human Factors in Computing Systems*, C. Plaisant (Ed.), 371-372, ACM Press.

Kieras, D.E., Wood, S.D., Abotel, K., Hornof, A. (1995). GLEAN: A Computer-Based Tool for rapid GOMS Model Usability Evaluation of User Interface Designs. In *Proceedings of UIST*, Pittsburgh, New-York, ACM Press, 91-100.

Kieras, D.E., Meyer, D.E. (1996). An overview of the EPIC architecture for cognition and performance with application to human-computer interaction. *Human-Computer Interaction*, 12(4), 391-438.

Kieras, D.E. (1997a). A guide to GOMS model usability evaluation using NGOMSL. In *Handbook of human-computer interaction (2nd ed.)*, M. Helander, T. Landauer, & P. Prabhu (Eds.), 733-766, Elsevier Science Inc., New York.

Kieras, D.E., Wood, S.D., Meyer, D.E. (1997b). Predictive Engineering Models Based on the EPIC Architecture for a Multimodal High-Performance Human-Computer Interaction Task. *ACM Transactions on Computer-Human Interaction*, 4(3), 230-275.

- Kieras, D.E. (1997a). A guide to GOMS model usability evaluation using NGOMSL. In M. Helander, T. Landauer, & P. Prabhu (Eds.), *Handbook of human-computer interaction* (2nd ed.), 733-766, Elsevier Science Inc., New York.
- Kieras, D.E., Meyer, D.E. (2000). The role of cognitive task analysis in the application of predictive models of human performance. In *Cognitive task analysis*. J.M.C. Schraagen, S.E. Chipman, and V.L. Shalin (Eds), New Jersey: Lawrence Erlbaum, 237-260.
- Kieras, D.E., Santoro, T.P. (2004). Computational GOMS modeling of a complex team task: lessons learned. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 97-104, ACM Press.
- Kirakowski, J., Claridge, N., Whitehand, R. (1998). Human centered measures of success in web site design. In *Proceedings of Fourth Conference on Human Factors & the Web*, Basking Ridge, NJ, USA.
- Kitajima, M., Polson, P.G. (1996). A comprehension-based model of exploration. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: Common Ground*, M. J. Tauber (Ed.), 324-331, ACM Press.
- Kolski, C., Ezzedine H., Abed, M. (2001). Développement du logiciel : des cycles classiques aux cycles enrichis sous l'angle des IHM. In Kolski C. (Ed.), *Analyse et conception d'IHM, Interaction homme-machine pour les SI 1*, 23-46, Hermes, Paris.
- Kosko, B. (1986). Fuzzy cognitive maps. *International Journal of Man-Machine Studies*, 24, 65-75.
- Lebodic, L. (2005). *Approche de l'évaluation des systèmes interactifs multimodaux par simulation comportementale située*. Thèse de doctorat de l'université de Bretagne Occidentale, juin.
- Macleod, M., Rengger, R. (1993). The Development of DRUM: A Software Tool for Video-Assisted Usability Evaluation. In *Proceedings of the Eighth Conference of the British Computer Society Human Computer Interaction Specialist Group - People and Computers VIII*, Alty, J.L., Diaper, D., Guest, D. (Eds.). 293-309, Loughborough University, UK.
- May, J., Barnard, P. (1995). The case for supportive evaluation during design. *Interacting with Computer*, 7, 115-144.
- Navarre, D., Palanque, P., Bastide, R. (2002). Model based interactive prototyping of highly interactive applications. In Kolski C., Vanderdonck J. (Eds.), *Computer-Aided Design of User Interfaces III*, Kluwer Academic Publishers, Dordrecht, 205-216.
- Newell, A. (1990). *Unified theories of cognition*. Harvard University Press.
- Nielsen, J. (1993). *Usability Engineering*. Academic Press Inc.
- Nigay, L. (2001). *Modalité d'interaction et multimodalité*. Habilitation à diriger des recherches, Université Joseph-Fourier.
- Noizet, A., Amalberti, R. (2000). Le contrôle cognitif des activités routinières des agents de terrain en centrale nucléaire: un double système de contrôle. *Revue d'Intelligence Artificielle*, 14, 107-129.
- Olsen, D.R., Halversen, B.W. (1988). Interface usage measurements in a user interface management system. In *Proceedings of the 1st Annual ACM SIGGRAPH*

Symposium on User interface Software (Alberta, Canada, October 17-19, 1988), UIST '88, ACM Press, New York, NY, 102-108.

Oviatt, S. (1996). Multimodal interfaces for dynamic interactive maps. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI), volume 1 of Papers: Multi-Modal Applications*, 95-102, 1996.

Oviatt, S. (1999). Ten myths of multimodal interactions. *Commun. ACM*, 42(11), 74-81.

Parenthoënn M., Tisseaun J., Morineau T. (2002). Believable Decision for Virtual Actors. In *International Conference on Systems, Man and Cybernetics (IEEE-SMC'02)*, Hammamet, Tunisia.

Paternò, F. (2001). *Model-Based Design and Evaluation of Interactive Applications*. Springer.

Payne, S.J., Green T.R.G. (1986). Task Action Grammar: A Model of the Mental Representation of Task Languages. *Human-Computer Interaction*, 2, 93-133.

Polson, P.G., Kieras, D.E. (1985). A quantitative model of the learning and performance of text editing knowledge. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (San Francisco, California, United States), 207-212, ACM Press, New York.

Pontico, F., Farenc, C., Winckler, M. (2004). Une architecture de dialogue basée sur un modèle pour les applications Web. In *XVI conférence sur l'interaction Homme-Machine, IHM2004*, 251-254, ACM press.

Rauterberg, M. (1995). From novice to expert decision behaviour: a qualitative modelling approach with Petri nets. In *Symbiosis of Human and Artifact: Human and Social Aspects of Human-Computer Interaction. (Advances in Human Factors/Ergonomics Y. Anzai, K. Ogawa & H. Mori (Eds.), Vol. 20B, 449-454, Amsterdam: Elsevier.*

Reason, J. (1990). *Human error*. Cambridge University Press.

Ritter, F.E., Baxter, G.D., Jones, G., Young, R.M. (2000). Supporting cognitive models as users. *ACM Trans. Computer-Human Interaction*, 7, 141-173.

Scapin, D.L., Pierret-Golbriech, C. (1989). Towards a method for task description: MAD. In *Work with Display Units 89*, L. Berlinguet and D. Berthelette (Eds.), 371-380, North Holland, Elsevier science.

Schank, R.C., Abelson, R.P. (1977). *Scripts, plans, goals and understanding*. Erlbaum, Hillsdale.

Siochi, A.C. et Hix, D. (1991). A study of computer-supported user interface evaluation using maximal repeating pattern analysis. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: Reaching Through Technology*, S.P. Robertson, G. M. Olson, and J. S. Olson (Eds.), CHI '91, ACM Press, New York.

Tabary, D., Abed, M. (2002). A software environment task object-oriented design (ETOOD). *Journal of Systems and Software*, 60 (2), 129-140.

Tisseau, J. (2001). *Réalité Virtuelle, autonomie in virtuo*. Habilitation à Diriger les Recherches, Université de Rennes 1.

Tollinger, I., Lewis, R.L., McCurdy, M., Tollinger, P., Vera, A., Howes, A., Pelton, L. (2005). Supporting efficient development of cognitive models at multiple skill levels: exploring recent advances in constraint-based modeling. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '05, 411-420, ACM Press.

Uehling, D.L., Wolf, K. (1995). User action graphing effort (UsAGE). In *Conference Companion on Human Factors in Computing Systems*. Katz, R. Mack, and L. Marks (Eds.), 290-291, ACM Press.

Young, R.M., Green, T.R., Simon, T. (1989). Programmable user models for predictive evaluation of interface designs. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: Wings For the Mind*, K. Bice and C. Lewis (Eds.), 15-19, ACM Press.