

# THÈSE

présentée devant

**l'Université Toulouse III : Paul Sabatier**

pour obtenir le grade de :

DOCTEUR DE L'UNIVERSITÉ PAUL SABATIER  
Spécialité : INFORMATIQUE

par

Mehdi EL JED

Titre de la thèse :

*Interactions sociales en univers virtuel*  
*Modèles pour une interaction située*

À soutenir le 26 Septembre 2006 devant la commission d'examen

- |     |      |          |   |
|-----|------|----------|---|
| M.  | L.   | Chaudron | Maître de Recherche à l'Onéra, Toulouse.<br>Examineur                                     |
| M.  | J.P. | Jessel   | Professeur à l'Université Paul Sabatier, Toulouse.<br>Examineur                           |
| M.  | C.   | Licoppe  | Professeur à l'École Nationale Supérieure des<br>Télécommunications, Paris.<br>Rapporteur |
| Mme | S.   | Pesty    | Professeur à l'Université Pierre Mendès France,<br>Grenoble.<br>Rapporteur                |
| M.  | B.   | Pavard   | Directeur de recherche CNRS-IRIT, Toulouse.<br>Directeur de thèse                         |



*A mes parents Kamel et Sihem,  
A mon frère Tarak,  
A Olfa et Selym,  
pour tout leur amour.*



# Remerciements

Je tiens à exprimer ma gratitude à Bernard Pavard, directeur de recherche au CNRS, pour avoir encadré et dirigé mes recherches. Je le remercie de m'avoir offert l'opportunité d'effectuer cette thèse dans les meilleures conditions, et pour la gentillesse et la disponibilité dont il a fait preuve à mon égard. Ses conseils et remarques constructives m'ont permis d'améliorer grandement la qualité de mes travaux et de ce mémoire.

Je remercie également mes deux rapporteurs : Madame Sylvie Pesty, professeur à l'Université Pierre Mendès France de Grenoble et Monsieur Christian Licoppe, professeur à l'École Nationale Supérieure des Télécommunications à Paris (ENST).

Je tiens également à remercier Messieurs Laurent Chaudron et Jean Pierre Jessel pour avoir accepté de participer à mon jury de thèse.

Merci aussi à ceux qui m'ont aidé tout au long de cette aventure, et en particulier à Nico Pallamin pour son aide précieuse lors de la modélisation et l'animation des personnages virtuels, à Béatrice Cahour pour ses conseils avisés lors des séances d'expérimentations et à Pascal Salembier pour son soutien permanent.

Un grand merci également à tous mes amis et collègues de l'équipe GRIC à l'IRIT pour cette magnifique ambiance qu'ils font régner au sein du laboratoire : Lucila, Sandrine, Colin, Guillaume et Nico sans oublier Chantal, Nadhem, Cécile, Julie, Asma et Ariel.

Enfin, je tiens à remercier toute ma famille pour son soutien constant et sa présence malgré les distances. Particulièrement, j'exprime ici toute ma reconnaissance à Olfa pour son courage et sa persévérance.



# Résumé

L'interaction sociale en univers virtuel pose des problèmes qui vont au delà des questions du réalisme de l'environnement ou même du réalisme des personnages virtuels. Comme les études ethnométhodologiques l'ont bien montré, la réflexivité des interactions humaines (ce couplage dynamique entre l'acteur et son environnement) aussi bien que les comportements indexicaux (références implicites aux objets externes de l'environnement) sont des éléments majeurs pour qu'une simulation en univers virtuel soit crédible et donc susceptible de produire la notion de présence sociale.

La recherche proposée s'intéresse au développement d'un modèle d'interaction sociale capable de prendre en compte, en univers virtuel collaboratif, une partie de la dimension émotionnelle et sociale des interactions humaines.

Plusieurs travaux antérieurs ont souligné un certain nombre de difficultés liées aux modalités d'accès à l'information dans les interactions en univers virtuels (comme la réduction du champ de vision, l'incohérence de l'univers sensible, etc.) ainsi que des limites relatives à la cohérence des expressions émotionnelles et corporelles.

Ces difficultés contribuent à maintenir un couplage faible entre les actions effectuées en univers virtuel et leurs significations par rapport à l'interaction (orientant souvent les utilisateurs vers des interactions orientées vers le rétablissement de l'intelligibilité de la situation).

Ces limites peuvent être compensés par une approche qui repose sur une meilleure mise en contexte de l'utilisateur. Dans notre approche, chaque utilisateur contrôle son propre avatar (représentation de l'utilisateur dans l'environnement virtuel) et peut prendre des décisions selon ses propres perceptions, son expertise et historique.

La problématique de recherche devient donc d'offrir une solution pour maintenir un contexte d'interaction 'riche' lors de la collaboration. Les utilisateurs peuvent ainsi accéder aux éléments contextuels de l'univers virtuel qui leur semblent pertinents pour la prise de décision et s'engager dans des modalités d'interactions orientées spécifiquement vers leurs objectifs.

Nous proposons des solutions qui permettent d'enrichir l'interaction sociale en univers virtuel. D'une part, l'interface proposée permet aux interac-

tants d'exploiter leurs références indexicales (par exemple pointer de la main des objets de l'univers, orienter le regard vers une direction, etc.).

D'autre part, notre modèle d'interaction sociale permet de produire automatiquement des comportements chez les avatars qui soient pertinents par rapport au contexte de l'interaction (par exemple distribuer le regard vers ses interlocuteurs, regarder les autres avatars en marchant, effectuer des expressions gestuelles en parlant, etc.).

Nous proposons également un modèle émotionnel pour simuler les états internes des personnages virtuels en interaction.

Ces modèles s'intègrent dans une architecture multi-agents capable de fusionner de façon 'réaliste' les actions intentionnelles décidées par l'acteur humain et les comportements non-intentionnels (produits par le modèle d'interaction sociale) comme les gestes, postures, expressions émotionnelles qui dépendent du contexte dans lequel évoluent les avatars.

Enfin, pour illustrer le fait qu'une interaction 'socialement située' est possible en univers virtuel, et souligner le fait que ce choix de conception se prête de manière naturelle à la simulation collaborative, nous proposons des expérimentations exploratoires dont les résultats préliminaires indiquent que les utilisateurs peuvent rapidement s'engager dans des modalités d'interaction complexes prenant en compte aussi bien l'état émotionnel des personnages virtuels que les indicateurs indexicaux se référant à des objets distants dans leur espace virtuel.

# Table des matières

<b>1</b>	<b>Introduction</b>	<b>13</b>
1.1	Motivations . . . . .	13
1.2	Collaboration en univers virtuel . . . . .	14
1.3	Approche pour des interactions situées . . . . .	15
1.4	Enjeux sociaux importants . . . . .	16
1.5	Contributions . . . . .	17
1.6	Organisation du mémoire . . . . .	19
<b>I</b>	<b>Définitions et état de l’art</b>	<b>21</b>
<b>2</b>	<b>Réalité virtuelle &amp; Systèmes multi-agents</b>	<b>23</b>
2.1	Réalité virtuelle . . . . .	24
2.1.1	Définition de la réalité virtuelle . . . . .	24
2.1.2	Intérêt de la réalité virtuelle pour la formation . . . . .	26
2.1.3	État de l’art des systèmes existants . . . . .	27
2.2	Les systèmes multi-agents . . . . .	31
2.2.1	Notion d’agent . . . . .	32
2.2.2	Action et communication . . . . .	36
2.2.3	Rôle de l’utilisateur . . . . .	37
2.3	Synthèse . . . . .	39
<b>3</b>	<b>Interaction virtuelle située</b>	<b>41</b>
3.1	Interaction située en univers virtuel . . . . .	42
3.1.1	Indexicalité . . . . .	42
3.1.2	La réflexivité des interactions . . . . .	43
3.1.3	La négociation du sens . . . . .	44
3.1.4	Articulation indexicalité-réflexivité en univers virtuel . . . . .	45
3.2	La notion de contexte . . . . .	46
3.2.1	Éléments de définition du contexte . . . . .	47
3.2.2	Structure du contexte en univers virtuel . . . . .	48
3.3	Immersion . . . . .	49
3.3.1	la notion d’Immersion . . . . .	50
3.3.2	Les différentes formes d’immersion . . . . .	51

3.3.3	Immersion émotionnelle et sociale . . . . .	52
3.4	Synthèse . . . . .	52
<b>II</b>	<b>Emotions &amp; Interactions sociales</b>	<b>55</b>
<b>4</b>	<b>Modèle émotionnel</b>	<b>57</b>
4.1	Emotions en univers virtuel . . . . .	59
4.1.1	Définition des émotions . . . . .	59
4.1.2	Qu'est ce qu'un modèle émotionnel . . . . .	59
4.1.3	Quelques modèles émotionnels . . . . .	60
4.2	Description du modèle émotionnel . . . . .	61
4.2.1	Quelques repères psychologiques . . . . .	62
4.2.2	Émotion . . . . .	65
4.2.3	Personnalité . . . . .	67
4.2.4	Relation entre personnalité et émotions . . . . .	68
4.2.5	Humeur . . . . .	69
4.2.6	Relation entre humeur et émotions . . . . .	70
4.3	Processus d'évaluation des émotions . . . . .	71
4.3.1	La perception des évènements . . . . .	72
4.3.2	Sensibilité aux émotions . . . . .	74
4.3.3	Évaluation des émotions . . . . .	76
4.3.4	Mise à jour de l'humeur . . . . .	78
4.3.5	La mémoire des émotions . . . . .	79
4.4	Discussion . . . . .	80
<b>5</b>	<b>Vers des communications situées</b>	<b>83</b>
5.1	Les communications situées . . . . .	84
5.1.1	Introduction . . . . .	84
5.1.2	Approche pour des communications situées . . . . .	85
5.2	Les communications intentionnelles . . . . .	87
5.2.1	Usage du langage naturel . . . . .	87
5.2.2	Usage des gestes indexicaux . . . . .	88
5.2.3	La direction du regard . . . . .	89
5.2.4	La proxémique . . . . .	90
5.3	Les communications non intentionnelles . . . . .	92
5.3.1	Gestualité lors du dialogue . . . . .	92
5.3.2	La notion de stress dans le comportement . . . . .	94
5.3.3	Postures et animations . . . . .	96
5.3.4	Expression faciale . . . . .	97
5.3.5	Hochement de tête . . . . .	98
5.4	Conclusion . . . . .	99

<b>III</b>	<b>Plate-forme de formation</b>	<b>101</b>
<b>6</b>	<b>Architecture de la plate-forme</b>	<b>103</b>
6.1	Architecture d'un agent . . . . .	104
6.1.1	Le module de perception . . . . .	105
6.1.2	Mécanisme de coordination . . . . .	107
6.1.3	Les modules comportementaux . . . . .	109
6.1.4	Adressage et communication . . . . .	111
6.1.5	Architecture d'intégration . . . . .	113
6.2	Aspect fonctionnel : le contexte de l'interaction . . . . .	114
6.2.1	Le contexte de l'avatar . . . . .	114
6.2.2	Approche de modélisation . . . . .	115
6.2.3	Le conflit d'intentionnalité entre l'utilisateur et son avatar	119
6.3	Exemples d'interactions sociales . . . . .	119
6.3.1	Cas des situations de dialogue . . . . .	120
6.3.2	Cas des situations de rencontre . . . . .	121
6.4	Modélisation de l'univers virtuel . . . . .	122
6.4.1	Objets de l'environnement virtuel . . . . .	123
6.4.2	Personnages virtuels et animations . . . . .	124
6.4.3	Géolocalisation du son . . . . .	125
6.5	Conclusion . . . . .	126
<b>7</b>	<b>Expérimentations</b>	<b>127</b>
7.1	Contexte d'évaluation . . . . .	128
7.2	Évaluation de l'ergonomie d'interaction . . . . .	130
7.2.1	Interface de contrôle . . . . .	130
7.2.2	Champ visuel de l'utilisateur . . . . .	132
7.3	Évaluation de quelques aspects de l'engagement émotionnel et social . . . . .	135
7.3.1	Gestion du dialogue et direction du regard . . . . .	135
7.3.2	Prise en compte des expressions émotionnelles . . . . .	138
7.3.3	Proxémique et gestes déictiques . . . . .	142
7.4	Discussions . . . . .	144
<b>IV</b>	<b>Conclusions &amp; Perspectives</b>	<b>147</b>
<b>8</b>	<b>Conclusions</b>	<b>149</b>
8.1	Synthèse . . . . .	149
8.2	Conclusions . . . . .	151
8.3	Perspectives . . . . .	153

<b>V</b>	<b>Annexes</b>	<b>155</b>
<b>A</b>	<b>Application</b>	<b>157</b>
A.1	Implémentation . . . . .	157
A.2	Déploiement de l'application en réseau . . . . .	158
<b>VI</b>	<b>Bibliographie</b>	<b>161</b>

# Table des figures

1.1	Aperçu de notre plate-forme de formation au commandement : Trois avatars en situation d'interaction sociale. . . . .	17
2.1	Définition de la réalité virtuelle d'après Fuchs. . . . .	25
2.2	Définition de la réalité virtuelle d'après Burdea et Coiffet. . . . .	25
2.3	Aperçu d'un village Bosniaque dans le système MRE. . . . .	29
2.4	Aperçu d'une simulation d'un accident de route dans ADMS. . . . .	30
2.5	Plusieurs vues d'une maison en feu sous VectorCommand. . . . .	30
2.6	Aperçu d'une simulation d'un incident avec DiaboloVR. . . . .	31
2.7	Principales caractéristiques d'un agent. . . . .	33
2.8	Architecture typique d'un agent hybride. . . . .	35
2.9	Interaction dans un système multi-agents. . . . .	36
4.1	Schéma d'influences entre les éléments du modèle émotionnel. . . . .	62
4.2	Classes d'émotions du modèle OCC. . . . .	66
4.3	Processus de calcul de l'intensité d'une nouvelle émotion. . . . .	77
4.4	Processus d'oubli des émotions au cours du temps. . . . .	80
5.1	L'émergence du sens selon que l'on s'adresse aux approches représentationnelles ou constructivistes. . . . .	85
5.2	Cadre situé pour les interactions engagées en univers virtuel. . . . .	86
5.3	Exemples d'expressions gestuelles indexicales. . . . .	88
5.4	Contrôle intentionnel de la direction de regard. . . . .	91
5.5	Exemples d'expressions émotionnelles chez un avatar : une émotion de déception et une émotion de satisfaction. . . . .	97
5.6	Exemples d'expressions labiales lors d'une élocution. . . . .	98
6.1	Architecture générale d'un agent représentant un avatar. . . . .	104
6.2	Approche classique pour la perception. . . . .	106
6.3	Approche 'orientée-comportement' pour la perception. . . . .	106
6.4	Exemple d'une architecture d'un module comportemental. . . . .	110
6.5	Processus de communication entre agents. . . . .	112
6.6	Exemple d'une architecture d'intégration pour un agent. . . . .	113
6.7	Contexte de l'avatar versus contexte de l'utilisateur. . . . .	115

6.8	Processus de perception du contexte chez un agent. . . . .	117
6.9	Aperçu d'une situation de dialogue entre plusieurs avatars. . .	120
6.10	Aperçu d'une situation de rencontre entre deux avatars. . . . .	121
6.11	Modélisation de quelques objets 3D dans notre univers virtuel.	124
6.12	Quelques avatars de notre environnement virtuel. . . . .	125
7.1	Extrait de la simulation d'un scénario de feu d'hôtel. . . . .	128
7.2	Interface de contrôle. . . . .	131
7.3	Angle de vue de l'utilisateur selon un référentiel égo-centrique et exocentrique. . . . .	133
7.4	Aperçu de l'angle de vue d'un utilisateur. . . . .	134
7.5	Disposition des avatars durant le dialogue selon l'angle de vue du participant. . . . .	136
7.6	Extrait de la simulation du personnage virtuel dans l'apparte- ment. . . . .	139
7.7	Extrait de la simulation du personnage virtuel devant un hôtel en feu. . . . .	140
7.8	Extrait d'une simulation de l'usage du déictique gestuel. . . . .	142
A.1	Aperçu de l'atelier graphique Virtools. . . . .	158
A.2	Schéma expérimental pour une session collaborative en univers virtuel. . . . .	159

# Liste des tableaux

2.1	Tableau comparatif des principaux travaux existant. . . . .	32
4.1	Format des données de la table des évènements. . . . .	73
4.2	Exemple de données la table des évènements associée à un scénario de Feu d'hôtel. . . . .	73
4.3	Valeurs possibles du facteur de modération. . . . .	78
5.1	Exemple d'une adaptation permanente de la démarche d'un avatar selon son niveau de stress. (Avec : $\psi_2 > \psi_1$ ) . . . . .	96
6.1	Exemple d'ordre de priorités attribuées à certains comportements : (1) faible priorité, (5) forte priorité. . . . .	109
6.2	Exemple d'éléments contextuels dans une situation de dialogue pour un avatar particulier. . . . .	116
6.3	Exemple de comportements selon le contexte d'interaction. . . . .	122



# Introduction



# Chapitre 1

## Introduction

### 1.1 Motivations

L'interaction sociale en univers virtuel collaboratif pose des problèmes qui vont au delà des questions du réalisme de l'environnement ou même du réalisme des personnages virtuels.

Comme les études ethnométhodologiques l'ont bien montré (Garfinkel, 1967; Amiel, 2004), la réflexivité des interactions humaines ainsi que l'indexicalité des comportements (références implicites ou explicites aux objets de l'environnement) sont des éléments majeurs pour qu'une simulation interactive en univers virtuel soit crédible et donc susceptible de produire la notion de présence sociale.

La réflexivité des interactions est basée sur l'idée que les actions humaines modifient le contexte de l'interaction et que, réflexivement, ce nouveau contexte va modifier le processus décisionnel de tous les acteurs en co-présence dans cette même interaction.

C'est la force de ce couplage dynamique entre les actions produites et le contexte partagé qui donne consistance aux échanges sociaux dans une interaction humaine.

Dans nos travaux, nous nous intéressons à l'étude des interactions sociales en univers virtuel. Particulièrement, nous nous focalisons sur les communications situées dans des simulations virtuelles de gestion de crises où la dimension émotionnelle et sociale de l'interaction sont déterminantes dans la gestion de la situation.

Cette dimension émotionnelle et sociale a été, jusqu'à nos jours, surtout prise en compte dans des milieux artistiques comme le cinéma, la télévision ou encore le théâtre. Les spectateurs sont souvent immergés par des senti-

ments d'empathie envers les acteurs qui arrivent à les faire émouvoir, stresser, angoisser ou encore rire. Ils se reconnaissent également dans les différents comportements sociaux et attitudes culturelles exprimées par ces acteurs humains. Les communications humaines sont riches d'actes sociaux très porteurs d'informations dans le dialogue comme l'utilisation des expressions gestuelles pour transmettre des messages implicites ou encore l'utilisation du regard et des expressions émotionnelles pour mieux véhiculer du sens dans le discours.

Les récentes avancées des technologies de réalité virtuelles nous permettent aujourd'hui d'espérer faire le pont entre les technologies déjà à contenu émotionnel mais non interactives comme le cinéma ou la télévision et les technologies interactives classiques mais à faible dimension sociale ou culturelle comme les jeux vidéo, les CSCW, etc. Nous pensons que c'est seulement à ces conditions que la technologie de réalité virtuelle peut prendre une dimension professionnelle car la qualité et l'efficacité des interactions humaines ne saurait exister sans que soient pris en compte les gestes, les activités implicites et leurs contenus émotionnels qui sont porteurs de nombreux messages significatifs pour l'action.

## 1.2 Collaboration en univers virtuel

Comme il a été montré dans de nombreux travaux, la collaboration en univers virtuel pose de nombreux problèmes qui peuvent être attribués soit au réalisme de l'environnement physique, soit à la non pertinence des expressions gestuelles et corporelles des acteurs par rapport au contexte.

Dans un article sur l'usage des artefacts en situation d'interaction en univers virtuel, Hindmarsh et al. ont bien montré que les acteurs ont du mal à identifier et localiser les objets de l'environnement qui sont explicitement ou implicitement référencés au cours d'un dialogue (Hindmarsh et al., 1998).

Ces auteurs ont également mis en évidence que, du fait d'un champ visuel périphérique restreint et des incertitudes de localisation des objets sonores, il peut devenir difficile pour les acteurs de se construire une image cohérente de la scène.

Cette difficulté de maintien de la cohérence de l'univers sensible peut avoir pour conséquence une mobilisation de ressources cognitives spécifiques à ce problème réduisant ainsi la dimension 'naturelle' (ou située) de l'interaction.

Par exemple, les acteurs peuvent avoir recours à un dialogue spécifiquement orienté vers le rétablissement de l'intelligibilité de la situation.

A ces difficultés liées au manque de réalisme de l'univers virtuel, s'ajoutent celles liées à la cohérence des expressions gestuelles et corporelles des avatars.

L'interprétation des actes de communication est étroitement liée à l'interprétation des expressions gestuelles et corporelles. Du fait que les moteurs d'animation ne peuvent que partiellement exprimer ces expressions corporelles, il devient difficile de maintenir un couplage pertinent entre les activités des différents acteurs en situation d'interaction.

Cette recherche a pour objectif d'explorer les possibilités que peut offrir un modèle d'animation des avatars qui prend en compte certaines pratiques d'interaction sociale comme les possibilités de références indexicales au cours des activités collaboratives (comme par exemple, les gestes déictiques ou la direction du regard en fonction du contexte social.) et les expressions émotionnelles (faciales et corporelles).

Nous pensons que ces comportements sociaux peuvent contribuer sensiblement à l'établissement, dans un univers virtuel, d'un couplage pertinent entre les actions produites et le contexte partagé de la situation.

Par exemple, les références indexicales aussi bien verbales (ici, vers là, à côté, etc.) que corporelles (la direction du regard, les gestes déictiques, etc.) dont la valeur sémantique est étroitement liée au contexte de la situation, représentent à la fois un moyen économique de communication (en codant un message de façon partielle) et riche en sens lors d'une activité collaborative.

Ces références indexicales, associées aux expressions émotionnelles, contribuent à la richesse du contexte de la situation en permettant de créer davantage de sens durant les interactions virtuelles. Ils représentent des indices contextuels supplémentaires qui contribuent au maintien d'un couplage réflexif pertinent entre les différents acteurs en interaction.

C'est ce couplage entre les actions produites et l'intelligibilité de la situation virtuelle qui nous semble essentiel pour engager véritablement des utilisateurs dans une interaction socialement et émotionnellement située.

### **1.3 Approche pour des interactions situées**

Notre objectif est de proposer un modèle permettant à des acteurs humains d'interagir de façon émotionnelle, sociale et culturelle dans un espace virtuel.

Contrairement aux approches traditionnelles de l'Intelligence Artificielle qui tentent de modéliser les comportements humains selon un paradigme essentiellement représentationnel, notre approche repose sur la mise en contexte de l'utilisateur afin qu'il puisse prendre des décisions selon ses propres perceptions de l'univers, son expertise et son historique (Pavard and Dugdale, 2002).

Dans cette approche, chaque avatar est dirigé par un utilisateur humain

(comme dans un jeu vidéo) mais ses gestes, expressions faciales ainsi que ses comportements émotionnels sont générés par notre modèle émotionnel et social. Ce modèle, prend en considération les actions intentionnelles de l'utilisateur pour ce qui est du contrôle de l'avatar (par exemple marcher, courir, parler, etc.). Il tente également d'identifier le contexte de l'interaction afin de produire chez l'avatar un comportement non intentionnel à la base d'une interaction sociale la mieux adaptée au contexte de la situation (par exemple regarder les autres avatars en marchant, effectuer des expressions gestuelles en parlant, etc.). L'utilisateur est ainsi "cognitivement et socialement situé" dans ses interactions avec les autres acteurs.

Cette approche de l'action située nous évite de modéliser l'ensemble des processus cognitifs associés à la prise de décision puisque c'est l'utilisateur qui contrôle son avatar selon sa propre intention.

Notre problématique se réduit alors à modéliser uniquement une partie de la cognition relative à la perception du contexte de l'interaction afin de produire chez l'avatar un comportement non intentionnel adapté à la situation et qui soit à la fois émotionnel et social.

Dans cette perspective, nous nous attendons à ce que, avec un modèle émotionnel et social qui 'enrichi' les comportements des avatars, les acteurs arrivent à compenser les écarts d'expression (de leurs avatars) et ainsi s'engager mutuellement dans une véritable interaction orientée vers les objectifs qu'ils ont à atteindre.

## 1.4 Enjeux sociaux importants

Cette approche peut avoir des retombées considérables sur l'apprentissage et la formation par la mise en place d'une collaboration en contexte.

C'est dans un apprentissage en contexte que l'apprenant donnera un sens à sa démarche pour construire ses connaissances et développer ses compétences. L'apprenant régule également sa conduite en fonction des différentes situations auxquelles il est soumis. Il peut ainsi changer de but et d'intention au cours de la résolution d'une tâche, modifiant par la même la signification du contexte dans lequel il est engagé.

L'autre intérêt de cette approche située dans le processus d'apprentissage est la prise en compte des interactions sociales. En effet, l'Intelligence Artificielle traditionnelle a trop mis l'accent sur les modèles expert du domaine se substituant totalement à la cognition humaine et a, de ce fait, sous-estimé l'importance des interactions sociales dans le processus d'interaction.

Offrir la possibilité à des apprenants d'accomplir en univers virtuel des interactions langagières naturelles intégrant, comme dans les communications en

univers réel, des références indexicales (ici, vers là, etc.), des incompréhensions dans les échanges, des émotions véhiculées par la voix, etc. nous apparaît essentiel pour une co-construction du sens en activités collaboratives. Les apprenants peuvent ainsi coordonner leurs actions pour mieux collaborer en univers virtuel afin de réaliser une œuvre commune.

En se basant sur cette approche, nous avons développé une plate-forme de réalité virtuelle pour la formation au commandement de sapeurs pompiers (figure 1.1). Ce projet a été réalisé en étroite collaboration avec le Service Départemental d'Incendie et de Secours de l'Essonne (SDIS 91) afin de bien prendre en compte l'aspect humain dans l'interaction.



FIG. 1.1 – Aperçu de notre plate-forme de formation au commandement : Trois avatars en situation d'interaction sociale.

## 1.5 Contributions

Nous décrivons dans cette section notre démarche de travail et nos contributions dans le cadre de cette recherche.

### Analyse des activités en situation

Cette analyse est effectuée en situation réelle lors de plusieurs sessions de formation des sapeurs pompiers de l'Essonne (Paris) au commandement en situation d'urgence.

Nous avons filmé puis analysé ces sessions de formation. Les étapes de cette analyse sont :

- Analyse des différentes situations d'interventions et de prise de décision

afin d'identifier les scénarios virtuels représentatifs des difficultés rencontrées par les apprenants ;

- Analyse des échanges verbaux entre les différents interlocuteurs afin d'identifier les principaux dysfonctionnements (à l'aide de vidéos et de débriefings suite aux simulations) ;
- Analyse de la communication gestuelle avec constitution de bases de données gestuelles.

Cette analyse ergonomique des activités en situation nous a permis de définir les qualités que doit posséder un environnement virtuel de formation pour permettre à des apprenants de reproduire une partie des mécanismes naturels d'interaction.

Nous identifions particulièrement l'importance de la prise en compte de la dimension émotionnelle et sociale lors des interactions.

### **Modèle émotionnel**

Nous proposons un modèle émotionnel pour simuler les états internes des personnages virtuels (ou agents) en interactions dans un environnement virtuel (émotions, humeur, personnalité et degré de stress).

Ce modèle permet, de façon analytique, d'identifier et de maintenir les états émotionnels d'un agent en interaction, son humeur et son niveau de stress en fonction de sa personnalité déclarée et de son interprétation des événements perçus. Il se base sur un modèle de la tâche de l'agent pour évaluer les événements susceptibles de changer son état émotionnel.

Le modèle d'évaluation des émotions s'inspire du modèle OCC décrit dans les travaux de (Ortony et al., 1988) auquel nous rajoutons quelques éléments nécessaires à l'étude de l'émergence des émotions comme la personnalité et l'humeur.

### **Modèle d'interaction sociale**

Nous proposons également un modèle permettant d'intégrer une partie des pratiques sociale lors des interactions virtuelles.

Comme nous l'avons précédemment évoqué, la difficulté à laquelle se heurte la conception d'un modèle d'interaction sociale est liée à notre capacité de production de couplages réflexifs entre les actions des différents acteurs.

Par exemple, une expression déictique comme la désignation d'un objet par un geste de la main sera obligatoirement suivie par un changement de la direction du regard des différents interlocuteurs.

En partant du modèle émotionnel et des décisions d'actions prises par l'utilisateur, nous proposons d'une part une solution pour effectuer de façon intuitive des expressions indexicales (des déictiques, des mouvements d'exploration

de l'univers par le regard, etc.) et d'autre part des solutions pour accompagner l'utilisateur dans son interaction comme la production automatique de comportements sociaux et émotionnels (par exemple distribuer le regard entre les interlocuteurs lors d'un dialogue).

Ce modèle s'intègre également dans une architecture d'agent qui permet une fusion des comportements intentionnels décidés par les utilisateurs et les comportements non intentionnels ou automatiques (comme les gestes, postures, expressions émotionnelles, etc.) qui dépendent du contexte dans lequel évolue le personnage virtuel.

Les expressions comportementales ainsi obtenues contribuent au caractère situé de l'interaction et permettent d'engager les utilisateurs dans une nouvelle forme d'immersion sociale et émotionnelle.

## Plate-forme de réalité virtuelle

Nous avons réalisé dans le cadre du développement de notre plate-forme de réalité virtuelle une partie de la modélisation graphique de l'environnement virtuel, des objets de la scène 3D, des personnages virtuels et des animations comportementales.

Nous avons ensuite intégré ces éléments dans un même atelier graphique capable de créer de l'interactivité (la suite logicielle Virtools en l'occurrence). Nous avons également mis en œuvre les mécanismes de coordination et de collaboration nécessaires pour garantir à des utilisateurs connectés depuis des sites distants des interactions virtuelles 'situées'.

## Expérimentations

Nous conduisons dans cette recherche une série d'expérimentations exploratoires dans le but d'établir qu'une interaction en univers virtuel peut être socialement et émotionnellement située.

### 1.6 Organisation du mémoire

Ce mémoire est organisé autour de 3 grandes parties :

- La première partie (partie I) est associée à un état de l'art nous permettant de définir le contexte de notre étude et la méthodologie de travail. Cette partie est composée de deux chapitres correspondant à un ensemble de définitions et de concepts concernant deux champs disciplinaires autour desquels s'articule notre travail :
  1. Les travaux sur la réalité virtuelle et les systèmes multi-agents (Chapitre 2).
  2. Les travaux en sciences cognitives sur les théories de l'action située et les travaux en ethnométhodologie sur la réflexivité et l'indexicalité dans les interactions (Chapitre 3).

- La deuxième partie (partie II) présente le modèle d’interaction sociale que nous mettons en œuvre pour reproduire des interactions ‘situées’ en univers virtuel. Cette partie comporte deux chapitres qui traitent du :
  1. Modèle émotionnel pour évaluer et maintenir les états émotionnels des agents en interaction (Chapitre 4).
  2. Modèle d’interaction sociale permettant à nos agents d’exprimer des comportements émotionnels et sociaux adaptés au contexte de l’interaction (Chapitre 5).
- La troisième partie (partie III) est une description de l’implémentation de la plateforme d’exécution. Elle se compose de deux chapitres :
  1. Description de l’architecture logicielle de la plate-forme ainsi que des différents éléments de modélisation mis en œuvre (Chapitre 6).
  2. Présentation des évaluations de notre modèle d’interaction sociale à travers les expérimentations exploratoires qui serviront de base pour des expérimentations à une plus grande échelle avec des professionnels du métier (Chapitre 7).

Première partie

Définitions et état de l'art



## Chapitre 2

# Réalité virtuelle & Systèmes multi-agents

*Non, la science n'est pas une illusion. Mais ce serait une illusion de croire que nous puissions trouver ailleurs ce qu'elle ne peut pas nous donner.  
(Sigmund Freud)*

Après avoir souffert d'une image trop ludique, la réalité virtuelle se professionnalise en toute discrétion<sup>1</sup>. Aujourd'hui, cette technologie pénètre de nombreux domaines comme la formation professionnelle, le médical ou encore la sécurité militaire et civile.

Les raisons profondes de cette expansion professionnelle de la réalité virtuelle sont certes son réalisme de plus en plus convainquant mais aussi ses capacités à induire des situations immersives, des collaborations professionnelles et à intégrer la dimension sociale et émotionnelle dans les interactions virtuelles.

Notre objectif est d'utiliser la réalité virtuelle pour la formation d'acteurs professionnels au travail collaboratif dans des environnements complexes. Il s'agit de permettre à des utilisateurs d'interagir en univers virtuels pour accomplir des tâches coopératives et acquérir un savoir-faire de techniques complexes.

---

<sup>1</sup>D'après une étude publiée en mai 2005 par Cyberedge Information Services, le chiffre d'affaires mondial du domaine aurait atteint 43 milliards de dollars en 2003, en croissance de 8,9 % sur un an.

Cet objectif repose sur une approche multi-agents pour représenter les comportements des entités en interactions dans le système de réalité virtuelle.

Dans ce chapitre, nous commençons par présenter la notion de réalité virtuelle tel que nous l'envisageons dans notre cadre d'étude. Nous présentons ensuite ses apports et son intérêt pour la formation ainsi qu'un tour d'horizon des principaux systèmes de formation existants.

Dans la deuxième partie de ce chapitre, nous présentons un essai de définition pour le concept des systèmes multi-agents ainsi que quelques approches qui ont permis d'intégrer l'utilisateur comme élément à part entière dans l'interaction avec un système de réalité virtuelle.

## 2.1 Réalité virtuelle

### 2.1.1 Définition de la réalité virtuelle

On entend souvent par réalité virtuelle un dispositif technologique permettant à un utilisateur d'interagir avec un monde virtuel à l'aide de dispositifs hautement technologiques tels que le casque de visualisation ou encore des gants de données. Cette image ne reflète plus aujourd'hui de manière exhaustive toutes les applications liées à la notion de réalité virtuelle.

Plusieurs travaux se sont intéressés soit à des définitions plus ou moins formelles de cette notion de réalité virtuelle soit à la catégorisation de ses applications. L'objectif de cette section est de présenter les principales caractéristiques que doit avoir un système de réalité virtuelle afin de correspondre au mieux à notre cadre d'étude.

Fuchs (Fuchs, 1996) propose une définition selon trois axes : temps, lieu, et interaction (figure 2.1). Une situation de réalité virtuelle est envisagée lorsque un ou plusieurs axes sont modifiés. Une modification sur l'axe du temps traduit une action dans un environnement passé ou futur, une modification sur l'axe du lieu correspond à un environnement de synthèse simulant la réalité ou un monde imaginaire et de manière similaire, une modification d'interaction implique une simulation d'interaction avec un monde réel ou une interaction purement imaginaire.

Burdea et Coiffet (Burdea and Coiffet, 1993) décrivent les systèmes de réalité virtuelle comme une *"interface qui implique de la simulation en temps réel et des interactions via de multiples canaux sensoriels."* Les auteurs proposent une définition de la réalité virtuelle qui s'exprime autour des *"trois I"* :

**Interaction** Le système de réalité virtuelle doit offrir à l'utilisateur une interaction en temps réel selon ses propres canaux de communications.

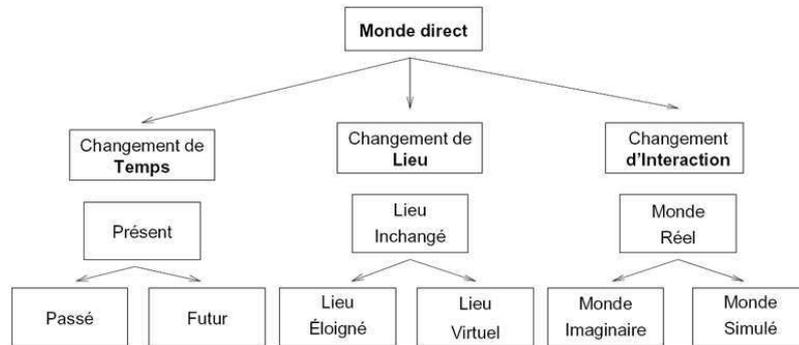


FIG. 2.1 – Définition de la réalité virtuelle d'après Fuchs.

**Immersion** La simulation doit être "réaliste" afin de produire chez l'utilisateur une sensation d'immersion dans l'environnement virtuel.

**Imagination** L'utilisateur participe activement à la simulation par son imagination. Cette composante est essentielle au bon fonctionnement de l'application.

Cette définition de Burdea et Coiffet est représentée par la figure 2.2 :

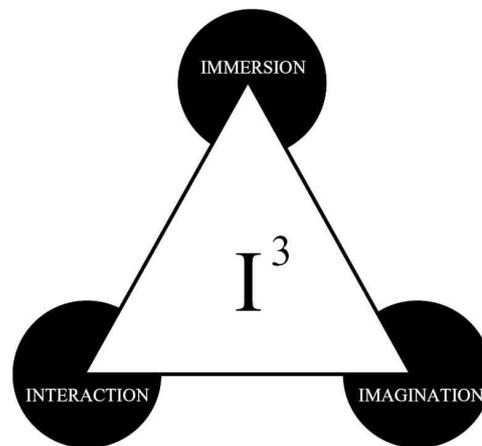


FIG. 2.2 – Définition de la réalité virtuelle d'après Burdea et Coiffet.

D'autres auteurs tentent de décrire la composante "immersive" de la réalité virtuelle comme une caractéristique essentielle du système. Pimentel et Teixeira décrivent la réalité virtuelle comme une expérience d'immersion de d'interaction générée par un ordinateur (Pimentel and Teixeira, 1994). Quéau reprend aussi cette idée et ajoute que la sensation d'immersion est produite par la conjonction d'une perception visuelle avec la sensation proprioceptive équivalente (Quéau, 1993).

De ces définitions, nous retenons les caractéristiques suivantes pour un système de réalité virtuelle :

**Univers virtuel :**

Qu'il s'agisse d'un monde imaginaire ou réaliste, la notion du monde virtuel est essentielle dans un système de réalité virtuelle. Ces mondes virtuels peuvent être catégorisés comme issus de la "réalité artificielle", "altérées" ou encore "réaliste". Le terme "réalité artificielle" a été introduit par (Krueger, 1991) pour décrire la capacité des systèmes de réalité virtuelle à créer des mondes totalement imaginaires sans aucun antécédent avec le monde réel. La "réalité altérée" est définie par (Vanderheiden and Mendenhall, 1994) comme une altération de certaines propriétés de l'univers virtuel afin d'assister des personnes présentant des déficiences mentales ou physiques (ex : ralentir le temps, agrandir des objets, etc.). Le terme "réalité réaliste" décrit une simulation aussi fidèle que possible du monde réel.

**Interaction :**

Il s'agit des possibilités offertes par le système pour faire interagir l'utilisateur avec le monde virtuel et les objets qui le constituent. Un utilisateur doit non seulement percevoir le monde virtuel mais aussi agir sur celui-ci et exercer des commandes de contrôle pour modifier son état.

**Immersion :**

Une forte interactivité entre l'utilisateur et le système doit conduire à une sensation d'immersion ou de "présence" dans l'univers virtuel. Différentes formes d'immersion peuvent exister lors d'une interaction avec un système de réalité virtuelle (physique, stratégique, tactique, sociale ou émotionnelle).

**Imagination :**

La sensation d'immersion d'un utilisateur n'est réellement efficace dans un système de réalité virtuelle que lorsque l'utilisateur participe activement à la simulation par son imagination. Cette imagination permet d'aider à "oublier" les imperfections du système et de se concentrer sur les aspects fonctionnels de l'interaction.

### **2.1.2 Intérêt de la réalité virtuelle pour la formation**

Plusieurs environnements d'apprentissage trouvent un support idéal dans la réalité virtuelle. Ces environnements de formation permettent à l'utilisateur (appelé également apprenant) de construire ses connaissances en agissant sur un univers virtuel.

De nombreux travaux ont souligné l'intérêt de ces environnements virtuels pour la formation et leurs avantages par rapport aux formations dans

des environnements réels (Lourdeaux, 2001).

Nous reprenons dans ce qui suit quelques intérêts qui justifient notre choix pour la réalité virtuelle comme outil pour la formation d'acteurs professionnels à la prise de décision.

**1. La mise en situation de l'utilisateur :**

Le premier intérêt de la réalité virtuelle est de mettre les apprenants dans des situations proches de celles rencontrées au cours de leurs vies professionnelles ou dans des situations qui représentent une particularité spécifique (domaines où la pratique est trop coûteuse à mettre en pratique ou très hasardeuse comme dans les domaines des secours, des combats ou du contrôle de processus de productions).

La réalité virtuelle permet ainsi assez naturellement de mettre en condition un apprenant dans de nouvelles situations en lui permettant d'apprendre en agissant en contexte.

**2. L'interaction avec les autres utilisateurs du système :**

L'intérêt d'utiliser une plate-forme multi-utilisateurs est de permettre non seulement des interactions apprenant-tuteur mais aussi des interactions entre apprenants eux-mêmes. Ils peuvent ainsi échanger des informations et dialoguer en langage naturel pour participer à la construction d'œuvres communes.

**3. L'aspect ludique :**

Un autre intérêt de la réalité virtuelle est aussi de pouvoir reproduire un scénario d'interaction en lui donnant un aspect ludique comme dans le cas des jeux vidéo ou des jeux de rôle. Les apprenants peuvent apprendre en jouant voire en 's'amusant'.

**4. La réutilisation de la simulation :**

Il s'agit des possibilités offertes par la réalité virtuelle de pouvoir reconfigurer les paramètres d'une simulation (en apportant des changements sur les environnements de formation, modifier les conditions climatiques, compliquer le déroulement d'une tâche, etc.), de réutiliser des objets de l'univers virtuel pour d'autres formations, de répéter la simulation pour revenir sur une action particulière ou même rejouer la session après l'entraînement pour visualiser les erreurs commises.

### **2.1.3 État de l'art des systèmes existants**

Plusieurs systèmes ont été présentés dans la littérature pour décrire l'utilisation de la réalité virtuelle dans des domaines plus ou moins divers comme l'éducation, la santé, la formation, etc.

Les outils que nous avons choisi de présenter dans cette section relèvent particulièrement du monde de la formation professionnelle.

Nous commencerons par décrire les principaux outils pédagogiques de formation dans un cadre général puis nous nous intéresserons particulièrement

aux travaux qui se sont intéressés à former des professionnels du métier à la gestion des secours et des situations de crises.

Nous présentons également les principaux outils commerciaux qui utilisent la réalité virtuelle pour la formation au travail collaboratif en situation de crises.

Parmi les outils pédagogiques les plus cités utilisant la simulation interactive pour la formation, nous pouvons identifier STEVE (Rickel and Johnson, 1999) un tuteur intelligent basé sur le moteur d'intelligence artificielle SOAR (Newell, 1990). STEVE est un agent virtuel interactif intégré à l'environnement virtuel VET (Virtual Environment for Training) qui a pour objectif d'aider à la formation de tâches procédurales. Il dispose de capacités cognitives permettant de créer et reconnaître la parole, démontrer des actions, répondre à des questions, se souvenir des actions passées de l'utilisateur, etc.

L'application FIACRE (David and Lourdeaux, 2001) est un environnement virtuel de formation des conducteurs de TGV à l'intervention sur les voies ferrées. Elle vise à permettre l'apprentissage des procédures d'actions pour des situations d'intervention sur les rails.

Dans (Lester et al., 1999), les auteurs décrivent un assistant pédagogique, HERMAN THE BUG, qui aide les utilisateurs à créer un environnement composé de plantes qui évoluent de manière autonome. L'utilisateur peut modifier et faire évoluer son environnement et ainsi étudier les règles de botanique et de physiologie.

D'autres outils pédagogiques proposent des simulations pour la formation à la gestion des secours et des situations de crises.

SécuRéVi est une application de réalité virtuelle développée pour la sécurité civile (Querrec et al., 2003; Querrec and Chevaillier, 2001). L'objectif de cette plateforme est de former des officiers sapeurs pompiers au commandement dans des situations de crises et de leur permettre d'acquérir une connaissance du site à risque où ils peuvent être amenés à intervenir.

Le système Mission Rehearsal Exercise (MRE) est un environnement de formation à la prise de décision et à la gestion des facteurs émotionnels (contrôle du stress)(figure 2.3). Il tire profit de l'environnement VET (Virtual Environment for Training) en concevant des agents virtuels basés sur le modèle de STEVE (Rickel et al., 2001, 2002; Hill et al., 2003). Ce système a été développé pour l'armée américaine afin de former des lieutenants à des situations d'interventions (mission de maintien de la paix). Ce système comporte des personnages virtuels dotés de capacités émotionnelles (Gratch and Marsella, 2001; Marsella and Gratch, 2002; Gratch et al., 2002). Une architecture

comportementale permet d'attribuer des expressions émotionnelles aux personnages virtuels en fonction des objectifs à atteindre (Gratch, 2000).



FIG. 2.3 – Aperçu d'un village Bosniaque dans le système MRE.

Dans (Julien and Shaw, 2003), l'environnement virtuel de formation au commandement permet à l'utilisateur de naviguer autour de la scène et observer une maison en feu sous différents angles. L'apprenant peut alors donner des ordres à ses co-équipiers et observer le déroulement de l'intervention. Ainsi les dommages causés à la maison en feu ainsi que les risques pris par l'équipe de secours sont d'autant plus minimes que les ordres pour organiser les secours sont correctement émis.

Cette application se base essentiellement sur un modèle réaliste de propagation du feu et de la fumée. L'interaction au sein de ce dispositif se déroule à travers une interface de commande où l'apprenant émet ses ordres pour le déroulement de l'intervention que le formateur transforme en commandes de contrôle des animations.

D'autres outils commerciaux utilisent la réalité virtuelle pour améliorer l'immersion de l'utilisateur dans le monde virtuel.

ADMS <sup>2</sup> (Advanced Disaster Management Simulator) est un environnement de réalité virtuelle interactif pour la formation au commandement, à la prise de décision, à la communication et à la coordination des secours (figure 2.4). Cet outil permet de simuler des situations d'urgences tels que des accidents d'avions, des actes terroristes, des catastrophes naturelles, etc.

Un autre outil de réalité virtuelle utilisé pour la formation est Vector-Command<sup>3</sup>. Le simulateur de VectorCommand crée des scénarios d'incendies réalistes et met à l'épreuve le commandement des opérations de secours en

---

<sup>2</sup> [www.admstraining.com](http://www.admstraining.com)

<sup>3</sup> [www.vectorcommand.com](http://www.vectorcommand.com)



FIG. 2.4 – Aperçu d’une simulation d’un accident de route dans ADMS.

reproduisant en temps réel les processus et les conditions rencontrées par les cadres des sapeurs-pompiers lorsqu’ils prennent le commandement des opérations de secours (figure 2.5).



FIG. 2.5 – Plusieurs vues d’une maison en feu sous VectorCommand.

La société E-Semble propose également un outil de formation DiaboloVR<sup>4</sup> capable de reproduire de façon réaliste des situations d’incidents. Les participants peuvent se déplacer dans l’environnement virtuel et prendre des décisions sur le déroulement de l’intervention et le déploiement des forces de secours.

<sup>4</sup>[www.e-semble.com](http://www.e-semble.com)

Cet outil propose également d'enregistrer les actions de l'apprenant afin de pouvoir les visualiser après la simulation (figure 2.6).



FIG. 2.6 – Aperçu d'une simulation d'un incident avec DiaboloVR.

L'ensemble de ces outils pédagogiques permettent d'assurer à l'utilisateur une immersion physique avec un bon réalisme au niveau de l'univers virtuel et une immersion stratégique au niveau de l'évolution des scénarios d'intervention mais n'offrent pas ou peu de moyens d'engagement social et émotionnel dans l'interaction avec d'autres utilisateurs (tableau 2.1). L'utilisateur n'est que rarement mis en contexte d'interactions avec d'autres opérateurs humains (formateurs, utilisateurs, etc) où l'engagement émotionnel et le contexte social sont déterminant lors de la prise de décision.

## 2.2 Les systèmes multi-agents

La recherche sur les systèmes multi-agents se focalise sur l'étude des comportements collectifs et sur la répartition de l'intelligence sur des agents plus ou moins autonomes, capables de s'organiser et d'interagir pour résoudre des problèmes.

Toute démarche de conception visant à décomposer un problème et à proposer une solution basée sur les interactions de telles entités est donc qualifiée par une approche multi-agents. A titre d'exemples, ces agents peuvent être des processus, des robots, des êtres humains ou encore des avatars.

	Système	Types Interaction	Intégration de la dimension émotionnelle	Intégration de la dimension sociale
Outils pédagogiques de formation	Steve	Apprenant - Agent	Non	Non
	Fiacre	Apprenant-Tuteur	Non	Non
	Herman The Bug	Apprenant - Agent	Non	Non
Outils de formation au secours	SécuRévi	Apprenant - Agent	Non	Peu
	MRE	Apprenant - Agent Apprenant - Tuteur	Oui	Peu
	Julien and Shaw	Apprenant - Tuteur	Non	Non
Outils commerciaux de formation au secours	ADMS	Apprenant - Tuteur	Non	Non
	Vector Comamnd	Apprenant - Tuteur	Peu	Peu
	DiaboloVR	Apprenant - Tuteur	Peu	Peu

TAB. 2.1 – Tableau comparatif des principaux travaux existant.

Dans la suite de cette section, nous nous intéressons aux caractéristiques d'un agent selon notre cadre d'étude ainsi que ses capacités de perception de son environnement et d'interaction avec les autres agents.

### 2.2.1 Notion d'agent

Cette notion d'agent, et malgré plusieurs efforts pour en donner une définition acceptée à l'unanimité, reste elle aussi floue et délicate. Le terme agent est souvent utilisé pour désigner des concepts différents.

Pour donner plus de précisions à cette notion d'agent, nous présentons dans ce qui suit quelques définitions importantes données dans la littérature :

- Un agent est une entité qui perçoit son environnement et agit sur celui-ci (Russell, 1997)
- Un agent est un système informatique, situé dans un environnement, et qui agit d'une façon autonome pour atteindre les objectifs (buts) pour lesquels il a été conçu (Wooldridge and Jennings, 1995) ;
- Les agents intelligents sont des entités logiciels qui réalisent des opérations à la place d'un utilisateur ou d'un autre programme, avec une sorte d'indépendance ou d'autonomie. A cette fin, ils utilisent une sorte de connaissance ou de représentation des buts ou des désirs de l'utilisateur ;
- Un agent est une entité qui fonctionne continuellement et de manière autonome dans un environnement où d'autres processus se déroulent et d'autres agents existent (Shoham, 1993) ;

- Un agent est une entité autonome, réelle ou abstraite, qui est capable d’agir sur elle-même et sur son environnement, et qui, dans un univers multi-agents, peut communiquer avec d’autres agents, dont le comportement est la conséquence de ses observations, de ses connaissances et des interactions avec les autres agents (Ferber, 1995).

Partant de ces définitions de la notion d’agent, nous pouvons distinguer plusieurs caractéristiques d’un agent selon divers points de vue (figure 2.7) :

- La nature de la relation entre un agent et éventuellement un opérateur extérieur désigne son autonomie ;
- la capacité d’un agent à entretenir des échanges avec les autres agents du système caractérise sa sociabilité ;
- le degré de couplage que peut avoir un agent avec son environnement désigne son caractère situé ;
- la capacité d’un agent à prendre des décisions comportementale en fonction de son environnement (agents et objets du monde) caractérise son intelligence.

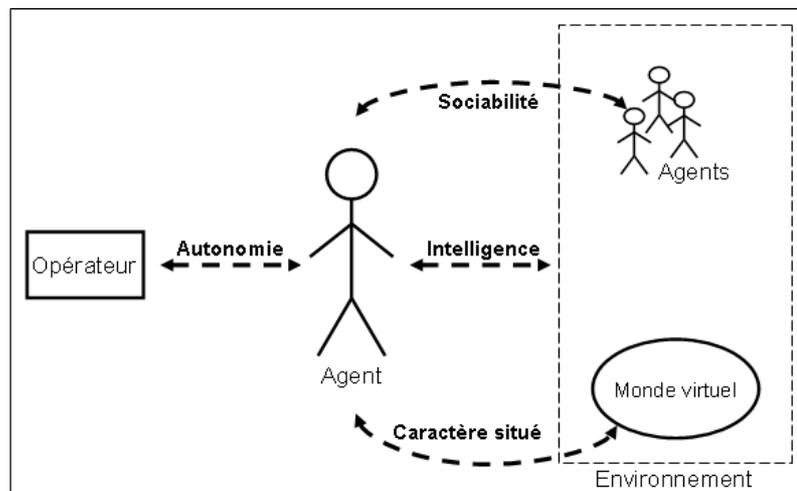


FIG. 2.7 – Principales caractéristiques d’un agent.

### 2.2.1.1 Autonomie

La notion d’autonomie décrit la capacité d’un agent à prendre des décisions et d’agir sur son environnement indépendamment de toute intervention extérieure.

Selon leur degré d’autonomie, nous distinguons trois classes d’agents :

- Les agents indépendants qui possèdent les moyens de prendre des décisions et d’agir selon leurs propres critères sans l’intervention d’un tiers (humain ou autre agent). Ils sont capables de percevoir l’état de leurs mondes et de gérer leurs propres états. Ces agents sont parfois qualifiés de ”pro-actifs”.
- Les agents semi-autonomes qui sont des agents contrôlés par un tiers qui est généralement un opérateur humain. Ils disposent cependant de capacités d’actions propres pour compléter les décisions prises par cet opérateur.
- Les agents contrôlés qui disposent de faibles capacités d’action voire pas du tout et qui sont complètement dépendant du contrôle de leurs opérateurs.

### **2.2.1.2 Caractère situé de l’interaction**

La relation d’un agent avec son environnement est une autre caractéristique que l’on peut associer à un agent. Nous considérons par le terme ’environnement’ l’ensemble des entités extérieure à l’agent avec lesquelles il peut interagir mais qui ne sont pas des agents. Nous distinguons ainsi la relation que peut avoir un agent avec les entités de son environnement (objets, paramètres, etc.) de celle qu’il peut avoir avec les autres agents (même si ces derniers font partie de l’environnement).

Lorsque l’agent est capable d’agir sur son environnement à partir des entrées sensorielles qu’il reçoit, il est dit situé. Ce couplage agent-environnement est plus ou moins fort selon la façon dont l’agent acquiert et traite l’information en provenance de son environnement (Guessoum, 1996). L’agent situé est donc un agent qui prend en compte sa situation pour agir.

### **2.2.1.3 Sociabilité**

La sociabilité d’un agent est souvent définie en Intelligence Artificielle comme sa relation interne avec les autres agents de son environnement. Un agent est dit sociable lorsqu’il est capable d’interagir avec les autres agents pour collaborer afin de poursuivre une œuvre commune. Il peut ainsi échanger des informations et partager des connaissances afin d’accomplir ses tâches ou aider d’autres agents à accomplir les leurs.

Nous soulignons dans cette caractéristique d’un agent le fait qu’une entité est décrite comme sociable (selon une terminologie Multi-Agents) dès lors qu’elle est capable de communiquer avec d’autres agents. Ceci ne veut pas dire que cette même entité peut accomplir des comportements ’sociaux’ ou peut respecter certaines pratiques sociales dans ses interactions.

#### 2.2.1.4 Intelligence

L'intelligence d'un agent est définie par rapport à la richesse de son comportement global dans son environnement (y compris avec les autres agents). D'après Steels, un comportement est considéré comme intelligent lorsqu'il permet au système de survivre au mieux dans son environnement (Steels, 1994).

Un agent peut paraître avoir un comportement intelligent pour l'utilisateur sans toutefois disposer de capacités de raisonnement ou de représentation symbolique sur son environnement. Pour distinguer ces deux types d'intelligences, on parle d'agents cognitifs et d'agents réactifs.

Les agents cognitifs (désignés aussi par délibératifs ou proactifs) sont habituellement dotés de capacités de représentation symbolique et de raisonnement. Ils peuvent décider d'agir et sont capables d'anticiper l'évolution de l'environnement et de planifier des actions.

Les agents réactifs ne font que subir l'environnement et réagir à sa modification. Ils ne disposent généralement pas de représentation symbolique ni d'historique et leurs comportements est souvent régi par un cycle de Perception/Action.

La distinction entre agents cognitifs et agents réactifs tend à disparaître avec les nombreux travaux qui tendent à proposer des approches hybrides afin de tirer profit des avantages des deux approches (Ferguson, 1992; Müller and Pischel, 1993; Wooldridge and Jennings, 1995).

La démarche adoptée est généralement l'utilisation d'une architecture réactive conjointement à un système symbolique (figure 2.8). La partie réactive permet à l'agent de réagir 'immédiatement' aux perceptions de son environnement et la partie cognitives se charge des opérations de planification, anticipation, mémorisation, apprentissage, etc.

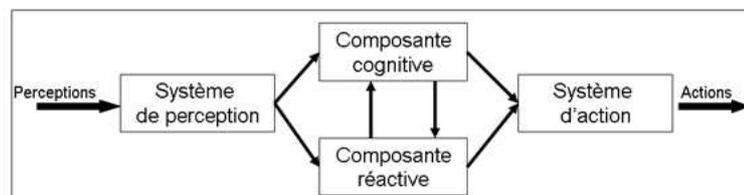


FIG. 2.8 – Architecture typique d'un agent hybride.

### 2.2.2 Action et communication

Une des principales propriétés d'un agent dans un système multi-agents est celle d'interagir avec les autres agents et son environnement. Ces interactions sont définies comme toute forme d'action permettant de modifier l'état du système.

Les moyens d'interactions employés dans les systèmes multi-agents sont généralement séparés en actions et communications. Cette taxonomie consiste à distinguer les actions effectuées sur l'environnement (déplacements, modifications apportées sur les objets, etc.) des actions utilisées pour communiquer avec d'autres agents de ce même environnement (envois de messages, etc.). D'après Ferber (Ferber, 1995), la communication entre les agents, bien qu'elle soit aussi une forme d'action, a pour but de modifier l'état mental d'un autre agent et de provoquer un comportement spécifique alors qu'une interaction avec l'environnement contribue simplement à modifier son état.

La figure 2.9 résume cette approche.

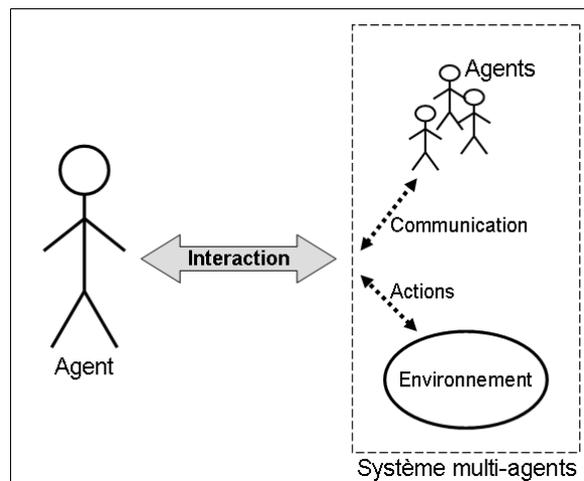


FIG. 2.9 – Interaction dans un système multi-agents.

#### 2.2.2.1 Actions et environnement

La distinction entre les actions effectuées sur l'environnement afin de changer son état et les actions effectuées sur les autres agents (communication) peut toutefois entraîner des confusions. En effet, cette distinction reste très réductrice puisque certaines actions peuvent avoir un but communicationnel comme par exemple les comportements ostensifs inférentiels (Sperber and Wilson, 1989).

La définition à donner à la notion d'action reste vague du fait de son interprétation multiple mais nous considérons toutefois que l'action est la plus petite unité permettant de décrire un comportement et que son résultat sur l'environnement peut être connu par la perception.

### 2.2.2.2 Communication

La communication est la forme d'action destinée à échanger de l'information avec les autres agents du système.

On distingue deux grands types de communications dans un système multi-agents : les communications directes qui correspondent à un envoi direct d'un message à un ou plusieurs agents et les communications indirectes qui utilisent l'environnement comme média de transmission des messages.

**Communication directe** Dans une communication directe dite également communication adressée, les agents utilisent des protocoles de communications évolués. Ces protocoles permettent de structurer l'échange de messages entre les agents. Plusieurs travaux de recherches se sont intéressés à ces protocoles et ont donné naissance à des langages de communications ou ACL (Agent Communication Languages) basés sur les actes de langages. Parmi ces langages de communications, nous pouvons citer KQML (Knowledge Query and Manipulation Language) ou le langage de la FIPA (Foundation of Intelligent Physical Agents).

**Communication indirecte** Dans une communication indirecte, un agent utilise l'environnement pour déposer un ou plusieurs messages dans des espaces communs. Pour atteindre ses destinataires, il faut que ces derniers se déplacent pour aller le chercher et le percevoir dans l'environnement.

### 2.2.3 Rôle de l'utilisateur

L'utilisateur joue un rôle plus ou moins important dans un système multi-agents. La plupart du temps, il intervient dans le système pour fixer les paramètres de la simulation, observer l'évolution du système ou analyser les résultats. On parle alors de simulation scientifique où toute l'interaction est basée sur les modèles mis en œuvre dans le système.

Lorsque l'utilisateur intervient dans le système pour faire évoluer l'environnement par ses actions, on parle alors de simulation interactive. L'utilisateur est considéré comme un modèle à part entière dans l'univers des modèles.

La représentation de l'utilisateur comme un agent à part entière dans un système multi-agents pose le problème de la modélisation de ses connaissances, l'analyse et l'interprétation de son comportement, la priorité de ses actions et

le protocole à mettre en œuvre pour la communication avec les autres agents logiciels.

Plusieurs travaux se sont intéressés à intégrer l'utilisateur dans l'interaction avec un système de réalité virtuelle à travers son avatar<sup>5</sup>.

Le rôle d'un avatar dans un système de réalité virtuelle peut se traduire par une interaction sur deux plans distincts de l'interface entre l'utilisateur et l'univers virtuel. D'une part, il interagit dans l'environnement virtuel à la place de l'utilisateur, c'est-à-dire qu'il perçoit l'environnement et effectue des actions. D'autre part, par l'intermédiaire du système d'interaction, il présente les perceptions sous une forme compréhensible par l'utilisateur et réceptionne les commandes à effectuer.

Cette notion d'avatar est parfois confondue avec des agents autonomes humanoïdes dits aussi acteurs virtuels ou humains virtuels.

Dans les travaux de Richard (Richard, 2001), l'utilisateur interagit dans un environnement constitué d'agents et formant un monde INVIWO (*Intuitive Virtual Worlds*). L'auteur propose un modèle d'avatar permettant d'intégrer complètement l'utilisateur dans le système multi-agents. L'avatar est représenté dans le système par un agent autonome permettant de contraindre ou d'assister l'utilisateur dans l'environnement.

Dans les travaux de Querrec (Querrec, 2002), l'environnement de réalité virtuelle est un système multi-agents hétérogène et ouvert où l'utilisateur peut participer via son avatar qui est un agent rationnel disposant de connaissances qui peuvent servir aux fonctions pédagogiques. Cet environnement virtuel de formation repose sur le modèle MASCARET (MultiAgent System for Collaborative and Adaptive Realistic Environment for Training) pour structurer les interactions entre les agents en interactions. Le modèle de l'avatar permet de prendre en compte les décisions de l'utilisateur et en même temps de suivre la procédure prescrite. Disposant de cette capacité, l'avatar de l'apprenant peut expliquer, conseiller ou montrer la réalisation d'une tâche à l'utilisateur.

Dans la suite de ces approches, nous nous proposons de concevoir un modèle de l'avatar permettant une intégration quasi complète de l'utilisateur dans l'interaction avec d'autres utilisateurs (représentés également par leurs avatars). Le rôle de l'avatar n'est plus de représenter l'utilisateur dans l'univers virtuel mais également d'accomplir les décisions prises par l'utilisateur et d'offrir des moyens de communications à la fois souples et intuitifs permettant l'émergence d'un couplage pertinent entre les actions engagées et les percep-

---

<sup>5</sup>Le terme avatar provient d'avatara, un mot 'sanskrit' qui désigne chacune des différentes incarnations du dieu hindou 'Visnu' sur terre.

tions obtenues.

La problématique de recherche devient ainsi le développement d'une architecture multi-agents capable de prendre en compte les décisions prises par l'utilisateur et de permettre à l'avatar de produire un comportement non intentionnel qui soit à la fois complémentaire aux décisions de l'utilisateur et en même temps indépendant de sa volonté.

Il s'agit donc de créer un couplage pertinent entre le contrôle utilisateur et le comportement non intentionnel de son avatar afin de produire une sensation de présence dans le dispositif de réalité virtuelle.

## 2.3 Synthèse

Dans ce chapitre, nous avons présenté quelques caractéristiques essentielles qui découlent des définitions données dans la littérature au concept de réalité virtuelle. Nous avons rappelé également l'intérêt que peut avoir un environnement de réalité virtuelle pour la formation et l'apprentissage de façon générale. Nous avons également souligné que les approches multi-agents fournissent un support conceptuel approprié à un environnement de réalité virtuelle.

Ce nouveau concept de système multi-agents repose sur la notion d'agent qui malgré plusieurs efforts pour en donner une définition unanime reste difficile à cerner. Nous avons présenté dans la section 2.2.1 les principales caractéristiques de la notion agent avant de discuter le rôle de l'utilisateur dans un système de réalité virtuelle. Ce rôle a souvent été limité à un ensemble d'actions prédéfinies à l'avance avec possibilité la plupart du temps de restreindre cette liberté d'action afin d'orienter l'utilisateur vers sa tâche.

Des approches comme celles de Richard ou de Querrec ont permis de développer des modèles d'avatars permettant d'intégrer l'utilisateur comme élément à part entière dans un système de réalité virtuelle.

Notre travail porte sur la conception et le développement d'un environnement de réalité virtuelle multi-utilisateurs composés d'agents en interactions. Cet environnement est constitué d'agents homogènes qui représentent des avatars contrôlés par des opérateurs humains. Chaque avatar dispose de capacités de perception et d'actions mais ses possibilités comportementales se limitent à l'accompagnement des décisions de l'utilisateur pour exprimer selon le contexte de l'interaction des expressions émotionnelles et sociales capables de produire une notion de présence dans l'univers virtuel. L'utilisateur reste ainsi libre de percevoir l'univers virtuel selon sa propre expertise et historique et de prendre librement des décisions. Cette approche permet à l'utilisateur

d'être cognitivement et culturellement situé dans l'interaction avec le système de réalité virtuelle. Dans le chapitre 3, nous essayons d'expliquer plus en détail ces notions d'interaction située et d'immersion émotionnelle et sociale.

## Chapitre 3

# Interaction virtuelle située

*C'est là le paradoxe suprême de la pensée que de  
vouloir découvrir quelque chose qu'elle-même ne  
puisse penser.  
(Sören Kierkegaard)*

La notion de l'interaction située est issue des théories de la 'cognition distribuée' qui sous-tend notre travail. Ces théories ont pour caractéristiques de considérer le groupe comme un seul système cognitif et d'accorder une place importante à l'environnement physique et social dans l'explication des phénomènes cognitifs.

Une place particulière est accordée par ces théories au contexte social et aux facteurs culturels dans l'analyse de l'interaction homme-machine et particulièrement dans l'analyse de l'apprentissage.

Dans la suite de ce manuscrit, nous allons considérer le terme 'interaction' dans un cadre d'étude homme-machine c'est à dire comme étant l'échange d'informations entre un utilisateur et un système de réalité virtuelle.

Ce chapitre est organisé autour de trois sections.

La section 3.1 présente la notion de l'interaction située que nous cherchons à produire dans nos simulations. Nous expliquons l'intérêt et le rôle de mécanismes sociocognitifs comme l'indexicalité et la réflexivité dans la communication en univers virtuel.

La section 3.2 de ce chapitre présente la notion du contexte qui sous-tend nos interactions en univers virtuel.

Finalement, nous donnons à la section 3.3 de ce chapitre, quelques éléments de définition de la notion d'immersion et nous présentons la nouvelle forme d'immersion sociale et émotionnelle que nous souhaitons reproduire dans un espace virtuel.

### **3.1 Interaction située en univers virtuel**

Dans un système de réalité virtuelle multi-utilisateurs, le sens donné par chaque utilisateur à son action qu'elle soit langagière, gestuelle ou corporelle, et aux actions des autres utilisateurs contribue à former son contexte de l'interaction. L'action ainsi engagée et interprétée ne correspond plus aux plans d'actions prédéfinis à l'avance mais devient fonction de la situation, de son interprétation et de la construction commune de sens entre les utilisateurs. Cette notion de l'action située contribue au réalisme des interactions et à l'immersion de l'utilisateur dans l'univers.

La crédibilité de ces interactions n'est assurée qu'avec la prise en compte de mécanismes sociocognitifs dans l'interaction tels que l'indexicalité et la réflexivité.

Ces deux concepts sont issus de l'ethnométhodologie, une discipline fondée par Garfinkel (Garfinkel, 1967) au cours des années 60. Il s'agit d'une approche sociale qui s'est construite en opposition à la sociologie et qui s'intéresse à la façon dont un ou plusieurs individus résolvent leurs problèmes concrets.

L'indexicalité concerne le problème du sens, de la communication, et donc de la création du sens en tant que facteur constitutif fondamental du fonctionnement social. Nous décrivons ci-après ces notions d'indexicalité dans la communication, de réflexivité des interactions humaines et de négociation du sens qui sous-tendent toute la réflexion autour de nos travaux.

#### **3.1.1 Indexicalité**

La notion d'indexicalité a été initialement formulée en linguistique par (Bar-Hillel, 1954) pour exprimer le fait qu'une expression langagière n'a de sens que lorsqu'elle est rapportée à son contexte.

L'ethnométhodologie a emprunté cette notion pour rendre compte de la nécessité qu'il y a, pour comprendre les échanges au sein d'une interaction, de les indexer sur les situations locales qui les ont produites.

Plus précisément, cette notion renvoie au fait que les expressions langagières et comportementales naturelles renvoient en permanence au contexte pour être comprises et interprétées. D'après Amiel (Amiel, 2004), elle correspond à l'idée qu'il n'y a pas de définition claire et générale d'un mot ou

d'un concept d'un langage parce qu'en situation naturelle, le sens vient par référence à d'autres mots, d'autres concepts et au contexte dans lesquels ils sont exprimés.

Cette notion d'indexicalité peut être définie d'après (Conein, 1984) par trois caractéristiques :

1. Ce sont des expressions dont la signification ne peut être donnée sans recours à des éléments liés au contexte pragmatique (espace, temps, sujets présents, objets présents, etc.).
2. Chaque fois que le contexte pragmatique change, la signification de l'expression change, car dans chaque contexte elle se réfère à des états différents.
3. Ces expressions comportent des indicateurs réflexifs sans valeur descriptive. Les déictiques et les index qui marquent la dépendance contextuelle.

En effet, le sens donné par un acteur à une expression langagière comme les expressions indexicales (ici, en haut, etc.), pronoms indéfinis, descriptions ambiguës (ça, là, etc.) est souvent fonction de plusieurs éléments contextuels tels que l'identité du locuteur, la forme de l'énonciation, l'intonation de la voix, l'auditoire cible, le lieu de son élocution, etc. Supprimer ces éléments contextuels limite grandement la compréhension de l'expression et l'attribution d'un sens précis à l'interaction.

Par exemple, une pancarte dans la rue affichant le terme "Maman" peut prêter à confusion à cause de l'absence d'un contexte, alors qu'un enfant dans la rue qui crie "Maman" nous renseigne qu'il appelle sa mère, et ce, sans que rien dans la phrase ne précise un contexte et alors qu'aucun possessif ne l'indique. D'après Theureau (Theureau, 1999), même une formulation 'correcte' risque d'induire des incompréhensions (si l'enfant crie "ma mère! ma mère!", on pensera qu'il arrive quelque chose à sa mère) et ne peut donc remédier aux traits indexicaux des phénomènes.

Cette indexicalité peut se retrouver également dans l'interprétation des expressions gestuelles des acteurs. Par exemple, les gestes déictiques, ces gestes de référence aux objets de l'environnement renvoient pour chacune de leur interprétation à un contexte unique. Dans une conversation, une expression gestuelle indiquant une direction donnée n'a de sens que lorsqu'elle est indexée sur le contexte de la situation (conversation en cours par exemple). Le sens donné au geste indexical ne sera plus le même lorsqu'il est reproduit dans un autre espace d'interaction.

### 3.1.2 La réflexivité des interactions

La notion de réflexivité renvoie à l'idée qu'un objet est en relation avec lui-même. Une relation réflexive est une sorte de miroir qui renvoie une image

de l'objet lui-même.

Ce caractère réflexif a d'abord été envisagé dans (Garfinkel, 1967) pour les pratiques et descriptions du langage naturel. Cela veut dire qu'une description faite par une personne renvoie automatiquement à cette même personne et à son contexte.

Les éléments utilisés par une personne pour décrire une situation jouent un rôle de miroir et renseignent sur une certaine image de cette même personne. Un enfant, par exemple, qui décrit une situation en employant des termes spécifiques reflète par sa description une certaine mentalité enfantine.

La réflexivité est donc essentiellement l'idée selon laquelle le sujet qui s'exprime apporte autant sinon davantage d'informations et de renseignements sur ce qu'il est lui-même que sur l'objet de son discours.

Par extension, il arrive que certains ethnométhodologues appellent 'réflexivité' le lien tout court qui met en correspondance une expression indexicale et son contexte.

Par une extension encore plus lointaine, on définit parfois la réflexivité comme une capacité subjective à gérer l'indexicalité, permettant d'effectuer sans hésiter des choix parmi les significations multiples qui se proposent à partir de plusieurs réseaux contextuels auxquels on appartient (Lecerf, 1985).

Pour donner un exemple de processus réflexifs qui apparaissent lors des dialogues chez la communauté des sapeurs pompiers en situation d'intervention, nous citons le dialogue suivant extrait de nos observations sur les lieux de formation (Darcy et al., 2003) :

- La personne qui se trouve de ce côté a besoin d'aide, tu t'en occupe !
- Celle qui est à l'entrée du bâtiment ?
- Oui oui, la personne là (et il désigne une direction) en pull bleu !
- Ok, chef je m'en occupe !

Cet exemple illustre comment l'indexicalité est à l'origine de la réflexivité. un geste indexical ou un comportement qui nécessite d'être indexé sur la situation pour être compris constitue l'origine de la création de la réflexivité.

Dans un dialogue à plusieurs, cette réflexivité se retrouve dans les interactions (qu'elles soient verbales, non verbales, gestuelles, etc.) qui provoquent des processus réflexifs pour la création du sens.

### **3.1.3 La négociation du sens**

Les communications humaines sont caractérisées par la mise en commun d'objets usuels et de concepts servant à désigner ceux-ci.

Dans les interactions langagières quotidiennes, cette mise en commun de

concepts suppose des déplacements sémantiques de ces concepts lorsqu'au départ leurs définitions diffèrent d'un individu à un autre ou que les conditions contextuelles ne permettent pas un partage 'correct' du sens.

Une négociation du sens s'avère ensuite nécessaire pour établir une interprétation de la situation. Cette négociation est arbitraire au niveau des résultats qu'elle peut produire.

Le sens négocié n'est jamais fixé et relève souvent des caractéristiques intersubjectives des interlocuteurs.

Ainsi, lors de toute conversation, le champ des interprétations possibles peut être multiple et imprévisible. Il devient donc essentiel, afin de bien prendre en compte ce mécanisme naturel de communication, de considérer en univers virtuel de nouveaux moyens d'interactions permettant aux utilisateurs d'accomplir des processus réflexifs lors de leurs communications et négocier le sens de leurs actions.

#### **3.1.4 Articulation indexicalité-réflexivité en univers virtuel**

D'après (Suchman, 1987), dire que l'action émerge des circonstances signifie d'une part, que l'action est dépendante des circonstances (c'est l'indexicalité), et que d'autre part, le fait d'agir définit le contexte de l'action (c'est la réflexivité).

La prise en compte de ces deux concepts dans une interaction en univers virtuel nous permet d'espérer atteindre une dimension nouvelle dans l'engagement social et émotionnel des acteurs en co-présence.

Permettre de 'naturellement' effectuer des expressions indexicales (usage du langage naturel, usage de gestes déictiques, exploration de l'univers par le regard, etc.) dans des interactions virtuelles contribue à renforcer les mécanismes mis en œuvre en situation réelle pour la construction commune du sens.

La notion de réflexivité dans une interaction virtuelle repose sur l'idée de dynamique et de couplage permanent entre les différents éléments du système (actions décidées par les utilisateurs, contexte partagé de la situation, etc.).

Les situations évoluent et rapidement leurs répercussions s'opèrent sur les acteurs sociaux (avatars en co-présence) et leurs manières de raisonner. Réciproquement, ceux-ci vont créer de nouvelles situations d'interaction par l'influence des actions qu'ils exercent sur l'univers virtuel. Une interaction entre plusieurs utilisateurs donne lieu à une compréhension toujours différente puisque chacun l'opère à partir de sa propre expérience. En effet, l'interprétation d'une action particulière est obtenue en fonction de la culture, des

croyances et des suppositions de chaque acteur social.

Par exemple, dans une situation de dialogue en univers virtuel entre plusieurs avatars, la perception commune de l'interaction langagière n'empêche pas que tous les acteurs en co-présence aient une image différente de celle-ci. Certains vont, par exemple, être satisfaits du déroulement de la conversation, d'autres peuvent avoir un avis négatif sur le sujet de la discussion, etc.

Au fur et à mesure que chaque personne présente dans cette interaction langagière apporte de nouvelles informations à l'échange, il contribue par son action à la construction commune de sens entre les différents interlocuteurs.

## **3.2 La notion de contexte**

L'importance de la notion de contexte pour la modélisation des activités humaines (raisonnement, perception, compréhension du langage, etc.) est un fait acquis aussi bien en sciences humaines qu'en informatique. Les approches traditionnelles de l'Intelligence Artificielle ont d'ailleurs très tôt mis en évidence les difficultés théoriques associées à la formalisation de cette notion. En effet, dans le cadre des théories de l'Intelligence Artificielle représentationnelle, le contexte peut difficilement être identifié par une liste finie de facteurs (Guha, 1991; Mc Carty, 1993).

Ce problème qualifié de 'problème de la qualification' se retrouve en sciences humaines et particulièrement dans les environnements complexes de travail : le choix par un opérateur humain des éléments qui vont constituer le contexte associé à sa prise de décision est de nature opportuniste et dans la plupart des cas imprévisible : l'opérateur 'choisit' les éléments du contexte en fonction du problème qu'il a à résoudre. Les références utilisées pour élaborer ce contexte sont à la fois internes (contraintes sur la mémoire, les connaissances, etc.) ou externes (environnement de travail, règles organisationnelles, etc.) (Salembier, 1996).

De nombreuses solutions formelles ont été avancées pour tenter de contourner cette difficulté (Giunchiglia and Bouquet, 1997; Mc Carty and Hayes, 1969) sans toutefois remédier définitivement au problème de la qualification qui semble bien être un 'avatar' encombrant lié à la nature représentationnelle des approches formelles (selon laquelle la pensée, peut toujours être représentée sous forme langagière et donc formalisée en termes de prédicats). En effet, cette difficulté théorique disparaît avec les approches constructivistes ou éthnométhodologiques qui s'appuient précisément sur le caractère 'ouvert' et 'émergent' des prises de décisions pour structurer leurs démarches de modélisation.

### 3.2.1 Éléments de définition du contexte

La notion de contexte est présente dans plusieurs disciplines et connaît de multiples définitions plus ou moins précises (Karsenty and Pavard, 1997). Le contenu vaste et vague du terme 'Contexte' fait que son utilisation est très rarement uniforme dans la bibliographie et fait de lui un sujet d'éternelles discussions (Dey and Abowd, 2000). Nous allons dans ce qui suit explorer certaines notions du contexte à travers certains champs disciplinaires et présenter notre cadre conceptuel sur la notion de 'contexte' dans un système de réalité virtuelle interactif.

– **En linguistique**

En linguistique, le contexte est traditionnellement défini comme l'environnement linguistique immédiat d'un élément ou unité linguistique. Cet environnement est composé de plusieurs éléments qui conditionnent la présence, la forme, la fonction ou le sens de cette unité linguistique et interviennent dans sa compréhension. Ces éléments peuvent être : l'environnement physique et social, l'identité des interlocuteurs, les événements qui ont précédé l'énonciation, leurs activités en cours, etc.

Par ailleurs, certains auteurs préfèrent utiliser le terme "situation" pour caractériser ces éléments et réservent le terme 'contexte' à l'environnement linguistique.

– **En ethnométhodologie**

En ethnométhodologie, le contexte est considéré comme un "accomplissement" des interlocuteurs, c'est-à-dire un élément qui se construit progressivement au fil de l'interaction, et non identifiable à priori.

Produit et modifié à chaque moment par les acteurs qu'ils utilisent comme ressource d'intelligibilité ou de compréhension mutuelle, il influence également en retour leurs comportements.

L'action produit le contexte auquel elle s'adapte et c'est ce processus généralisé qui, selon les ethnométhodologues, produit et maintient la consistance sociale. Cette approche met en avant le caractère dynamique et intersubjectif du contexte.

– **En psychologie**

En psychologie, la notion de contexte est définie de différentes façons.

Pour certains auteurs, le contexte se réfère aux représentations ou de façon très large aux états mentaux qui jouent un rôle dans ou influencent les processus cognitifs (perception, raisonnement, mémoire, compréhension, etc.).

Dans le cadre des études de psychologie expérimentale, le contexte renvoie aux conditions dans lesquelles apparaît un stimulus cible, pour s'intéresser aux effets de celle-ci sur son traitement.

– **En ergonomie**

Le contexte en ergonomie est généralement défini comme la situation ou l'ensemble des conditions ou encore l'environnement dans lequel s'exerce une activité et qui déterminent ou influencent celle-ci (Hollnagel, 1993; Leplat, 2001). Ces facteurs peuvent être : l'environnement physique (outils disponibles dans l'environnement, disposition spatiale des individus, etc.), l'environnement social et organisationnel, etc.

– **En informatique**

En informatique et dans le domaine des Interfaces Homme Machine (IHM), on définit une caractéristique contextuelle comme une information qui peut être utilisée pour définir et interpréter une situation dans laquelle interagissent des agents (Suchman, 1987; Winograd and Flores, 1986). En Intelligence Artificielle, le contexte est ce qui n'intervient pas explicitement dans la résolution d'un problème mais contraint cette résolution (Brézillon, 1999).

### 3.2.2 Structure du contexte en univers virtuel

Plusieurs travaux se sont intéressés à la notion de contexte et particulièrement à sa classification en vue de le modéliser (Kaenampornpan et al., 2004; Goker and Myrhaug, 2002; Tazari et al., 2003). Ces modèles présentés couvrent souvent des domaines d'application particuliers et la notion de contexte étudiée diffère d'une application à une autre. Nous allons nous inspirer de ces modèles de classifications pour proposer notre classification des différents niveaux de contextes qui interviennent dans la prise de décision dans un système de réalité virtuelle.

**Contexte de la tâche** Il représente l'ensemble des buts poursuivis, les étapes déjà réalisées et les contraintes à satisfaire. Il est important de faire la distinction entre la tâche à accomplir et l'activité mise en œuvre pour la réaliser.

**Contexte physique** Il représente tous les aspects de l'environnement externe à l'interaction : la configuration physique de l'environnement (nature des objets, positions et orientations, personnages dans le champ de vision, etc.) et les différents paramètres environnementaux (temps, bruit, luminosité, etc.).

**Contexte émotionnel** Il représente tous les états internes pouvant intervenir lors de la prise de décision d'un utilisateur. Parmi ces états, on peut citer l'ensemble des états émotionnels ressenties (tristesse, satisfaction, colère, peur, surprise, etc.), l'humeur (bonne, mauvaise, etc.), le stress ou encore la personnalité (ouvert, consciencieux, extraverti, etc.) qui peuvent avoir une influence sur le comportement de l'utilisateur.

**Contexte cognitif** Il regroupe l'ensemble des perceptions subjectives qu'un utilisateur obtient de son environnement. C'est l'ensemble des croyances, suppositions, objectifs, attitudes ou besoins que les processus interprétatifs d'un acteur utilisent pour l'analyse et la compréhension d'une situation.

**Contexte social** L'identité sociale des interlocuteurs (rôles, statuts, hiérarchies, etc.) influence également la compréhension de la situation. Le contexte social couvre le type de situation interactive et les rôles des participants dans cette situation.

**Contexte culturel** Le sens donné à une situation repose aussi sur un ensemble de croyances implicites qui forment la culture de chacun. Ces croyances implicites se distinguent des représentations abordées précédemment par le fait qu'elles sont partagées et mutuellement admises par les membres d'un collectif. Le contexte culturel d'un individu regroupe les éléments de l'environnement qui se réfèrent à une communauté d'individu et donnent un sens à la compréhension de la situation (ex : habitudes, coutumes, langues, accent, gestes spécifiques, etc.).

**Contexte linguistique** Il s'agit du contenu sémantique du dialogue, de la fonction communicative de ce contenu et la forme que prend l'énoncé.

Ces différents niveaux de contexte caractérisent la situation d'interaction virtuelle. Nous nous baserons sur cette classification pour proposer une solution de modélisation des situations d'interactions en univers virtuel (Chapitre 6).

### 3.3 Immersion

Le notion d'*immersion* désigne de façon métaphorique la sensation d'être entièrement plongé dans un environnement et de pouvoir interagir avec les objets qui s'y trouvent.

Un univers virtuel est capable de produire une sensation d'immersion sur un utilisateur lorsqu'il arrive à capter son attention et modifier l'ensemble de ses caractéristiques sensorimotrices.

Plusieurs formes d'immersions peuvent apparaître dans une interaction avec un système de réalité virtuelle. Parmi les formes d'immersions souvent citées, nous pouvons distinguer :

- l'immersion tactique,
- l'immersion stratégique,
- l'immersion narrative,
- l'immersion physique.

A ces formes d'immersions, nous ajouterons l'immersion sociale et émotionnelle.

Pour mieux expliquer cette notion d'immersion sociale et émotionnelle, nous allons dans ce qui suit présenter la notion d'immersion d'une façon générale puis décrire chacune de ces formes d'immersion tel que nous l'entendons avant de présenter cette nouvelle forme d'immersion que nous cherchons à produire dans un espace de réalité virtuelle.

### 3.3.1 la notion d'Immersion

Dans (Fuchs et al., 2003), les auteurs proposent deux niveaux de définition de la notion d'immersion :

- Les auteurs définissent le degré d'immersion comme le rapport de 'couverture sensorielle' du dispositif virtuel, c'est à dire que plus un dispositif est capable de couvrir un organe sensoriel (champ de vision par exemple) plus il est décrit comme immersif.  
Des dispositifs comme les visiocasques par exemple permettent de masquer complètement le champ de vision réel du sujet et symbolisent de ce fait l'immersion dans un univers nouveau.
- Le second niveau de définition est la richesse et l'adéquation entre les modes d'interfaçages retenus au plan sensori-moteur et la physiologie de l'utilisateur. Il s'agit d'évaluer modalité par modalité les propriétés du dispositif d'interaction (temps de latence, qualité du signal, etc.) afin que les réponses de l'environnement virtuel paraissent 'naturelles' à l'utilisateur.

Au delà de ces propositions de définitions, cette notion d'immersion demeure tout de même subjective. Un environnement virtuel peut être immersif même si au sens métrologique de couverture de l'espace sensoriel, il ne l'est pas.

Dans (Mellet-d'Huart, 2004), l'auteur discute de l'opérationnalité du concept d'immersion pris dans sa première acception. Il interroge l'opérationnalité de ce concept en évoquant l'exemple de CS WAVE (Mellet-d'Huart and Michel, 2005) : un environnement virtuel de formation au soudage. Cet environnement offre un affichage graphique de la simulation sur un écran de 19 pouces. Bien qu'au sens métrologique de couverture de l'espace sensoriel, cet environnement ne correspond pas à la définition donnée aux environnements immersifs, il offre tout de même l'ensemble des informations nécessaires au soudeur pour réaliser son activité.

D'autant plus qu'en situation réelle, un soudeur en activité porte une cagoule et effectue sa tâche à travers un écran en verre teinté relativement étroit (qui limite considérablement la lumière) et ne dispose ainsi pas plus de champ de vision que ce qui est offert par l'environnement virtuel.

Il s'avère donc que l'environnement virtuel reste tout de même immersif pour le soudeur puisqu'il correspond bien à son activité et peut ne pas l'être pour d'autres types d'activités.

### 3.3.2 Les différentes formes d'immersion

Nous présentons dans cette section les différentes formes d'immersion bien connues comme l'immersion tactique, stratégique, narrative et physique puis nous ajoutons la nouvelle forme d'immersion sociale et émotionnelle que nous cherchons à produire en univers virtuel.

#### 3.3.2.1 Immersion tactique

L'immersion tactique est une immersion provoquée par un couplage de type perception-action lors d'une interaction avec un système de réalité virtuelle. Ce type d'immersion est typiquement retrouvé dans les jeux vidéos où l'utilisateur est physiquement immergé dans l'action. Le processus décisionnel de l'utilisateur est réduit à un cycle rapide de perception/action.

Cette immersion est produite par de simples défis permettant à l'utilisateur de trouver rapidement (généralement de l'ordre de fractions de secondes) une solution appropriée. Ce genre d'immersion se retrouve souvent dans les jeux vidéos comme le jeux tetris, flippers, etc.

#### 3.3.2.2 Immersion stratégique

L'immersion stratégique requiert à la différence de l'immersion tactique un engagement basé sur la réflexion. Il s'agit par exemple d'une recherche de la meilleure solution pour optimiser une situation. Dans une immersion stratégique, l'utilisateur est complètement pris dans l'interaction pour observer, calculer, déduire, etc. Cette forme d'immersion est produite par des défis mentaux à résoudre lors de l'interaction avec le système de réalité virtuelle. Le meilleur exemple pour cette forme d'immersion est certainement le jeu d'échec.

#### 3.3.2.3 Immersion narrative

L'immersion narrative est une immersion du type de celle généralement vécue en lisant un livre ou bien en regardant un film. Dans un système de réalité virtuelle, cette forme d'immersion est obtenue lorsque l'utilisateur commence à s'intéresser aux personnages et/ou au scénario de l'interaction et s'attache à connaître la fin de l'histoire.

Dans cette forme d'immersion, l'utilisateur peut tolérer des imperfections dans l'engagement stratégiques ou tactique du scénario. Une immersion narrative est généralement créée par une bonne scénarisation de l'interaction.

#### 3.3.2.4 Immersion physique

L'immersion physique est la forme d'immersion où l'utilisateur est pris par le réalisme de l'environnement. Cette forme d'immersion est obtenue lorsque l'attention de l'utilisateur est retenue par le réalisme des objets de la scène, les personnages en interaction ou encore l'éclairage et la sonorisation du système

de réalité virtuelle. A cette forme d'immersion s'ajoute également et, depuis peu, celle du toucher, grâce à des périphériques à retour haptique.

### **3.3.3 Immersion émotionnelle et sociale**

L'immersion émotionnelle en univers virtuel multi-utilisateurs est la forme d'immersion où l'utilisateur est engagé par un phénomène émotionnel (empathie, haine, etc.) dans l'interaction avec d'autres utilisateurs. Elle caractérise la capacité de l'utilisateur à ressentir les émotions ressenties par les autres personnes et à réagir en fonction de cette empathie.

L'immersion sociale est la forme d'immersion où l'utilisateur est engagé dans l'interaction selon les règles de comportements sociaux utilisés dans la réalité sociale. Elle désigne le degré au bout duquel l'utilisateur est capable d'oublier les imperfections et les limites du système d'interaction pour communiquer de façon proche de la réalité en exploitant les modalités de communications verbales et non verbales offerts par le système.

## **3.4 Synthèse**

Nous venons de présenter dans ce chapitre un bref tour d'horizon des concepts théoriques qui sous-tendent notre travail. Face aux limites rencontrées par l'intelligence artificielle dite représentationnelle pour modéliser les connaissances en univers ouvert et dynamique, le recours aux théories de l'action située nous semble approprié pour la simulation d'interactions virtuelles impliquant des opérateurs humains où la dimension sociale et émotionnelle est primordiale pour l'accomplissement de la tâche. L'apprentissage et l'acquisition de compétences sont alors réalisés en contexte.

La notion de l'interaction située que nous entendons reproduire en univers virtuel ne peut être obtenue sans la prise en compte de deux mécanismes sociocognitifs importants lors des communications humaines qui sont l'indexicalité et la réflexivité.

L'indexicalité caractérise les références implicites ou explicites aux objets de l'environnement que peut faire un acteur en situation d'interaction.

La réflexivité est l'image que se fait chaque individu de l'interaction et qui construit son cadre d'action dans l'interaction.

La prise en compte en univers virtuel du langage naturel comme moyen de communication permet aux utilisateurs de s'exprimer naturellement en utilisant des références indexicales langagières et de véhiculer des émotions à travers la voix, des expressions corporelles et des expressions faciales.

La prise en compte des gestes déictiques (indiquer une direction en parlant, fixer du regard un interlocuteur, etc.) permet également d'espérer reproduire en univers virtuel une partie des mécanismes de communications naturels en

rajoutant des éléments contextuels supplémentaires pour la construction du sens entre les interlocuteurs.

L'utilisateur humain reste libre de considérer les éléments contextuels qu'il juge appropriés pour sa tâche et de prendre des décisions selon sa propre expérience, culture et histoire. Il est ainsi cognitivement et culturellement situé dans son interaction avec le dispositif de réalité virtuelle.

Cette interaction située est étroitement liée à la notion de contexte d'interaction et d'immersion émotionnelle et sociale (ou de 'présence' sociale) en univers virtuel.

Cette notion de contexte a souvent été évoquée en littérature et dans différents domaines. Nous avons tenté d'apporter notre contribution dans la définition de cette notion dans le cadre des interactions virtuelles.

L'immersion sociale et émotionnelle diffère également des autres formes d'immersion connues jusqu'à aujourd'hui. Il s'agit d'atteindre un degré d'immersion où l'utilisateur interagit en univers virtuel en reproduisant des mécanismes de comportements social et émotionnel proche de la réalité.

Dans la partie II de ce manuscrit, nous décrivons notre démarche de développement de cette plateforme ainsi que les modèles que nous utilisons pour implémenter ces concepts.



Deuxième partie

**Emotions & Interactions  
sociales**



## Chapitre 4

# Modèle émotionnel

*Si l'émotion est une ivresse, la passion est une maladie  
qui résiste à tous les moyens thérapeutiques...  
c'est un enchantement qui exclut l'amélioration morale.  
(Emmanuel Kant)*

Une condition nécessaire pour améliorer la sensation d'immersion dans le cadre d'une application de réalité virtuelle collaborative est de produire chez les agents virtuels des émotions en rapport avec la situation d'interaction.

Les interactions ainsi obtenues sont d'autant plus crédibles que les émotions véhiculées par les agents virtuels reflètent au mieux les émotions ressenties par l'utilisateur humain.

Pour atteindre cet objectif, deux alternatives sont possibles :

1. Mesurer les émotions ressenties par l'utilisateur pour les transposer chez son avatar.
2. Simuler les états émotionnels probables des agents virtuels de façon analytique en fonction de l'interaction de l'utilisateur dans l'univers virtuel.

Aucun système ne permet à ce jour d'identifier en temps réel les différentes émotions ressenties par un humain et mesurer leurs intensités de façon fiable .

Les travaux sur la physiologie de l'émotion fournissent des indices indirects sur l'état émotionnel d'un sujet. Par exemple, la conductivité dermique est connue comme étant une variable liée au stress (Meehan et al., 2002). Le diamètre pupillaire représente également un indice dynamique en rapport avec l'attention que peut porter le sujet à une situation. Cet indice est malheureusement sensible à bien d'autres facteurs comme la surprise, la peur, les sensations négatives, etc.

Il est donc bien difficile d'exploiter ces indices de façon fiable bien qu'ils peuvent fournir des informations directement en provenance de l'état émotionnel du sujet.

La seconde solution (utiliser un modèle analytique pour représenter l'état émotionnel d'un utilisateur) présente l'inconvénient de ne pouvoir modéliser de façon fidèle les émotions ressenties par l'utilisateur. En effet, cette solution présente des limites notamment à cause de notre faible connaissance des conditions et processus d'émergence des émotions mais également du degré de réduction que nous serons amenés à faire pour simuler les émotions.

Néanmoins, nous choisissons de considérer cette solution pour modéliser les états internes de nos agents malgré son réductionnisme. Nous pensons que malgré le décalage probable entre les émotions simulées et les émotions ressenties réellement par l'utilisateur, un modèle analytique peut engager émotionnellement un utilisateur dans ses interactions et produire une sensation de présence dans le dispositif. Ce modèle s'appuie sur des études psychologiques pour simuler au mieux les mécanismes d'émergence des émotions chez les avatars selon les interactions de leurs utilisateurs.

L'objectif de ce chapitre est de présenter le modèle analytique que nous mettons en œuvre pour doter nos agents de capacités émotionnelles et renforcer ainsi la sensation d'immersion dans le dispositif.

Nous débutons ce chapitre par une présentation générale des émotions et leurs intérêts pour les simulations interactives en univers virtuel (section 4.1). Nous présentons également dans cette section les principaux travaux qui se sont intéressés à la modélisation émotionnelle.

Dans la section 4.2, nous présentons le modèle émotionnel que nous implémentons dans notre plate-forme de réalité virtuelle pour simuler les états émotionnels de nos agents.

## **4.1 Emotions en univers virtuel**

### **4.1.1 Définition des émotions**

Les émotions jouent un rôle important dans la vie quotidienne et exercent une influence non négligeable sur les comportements humains. Elles œuvrent pour transmettre plus de sens aux échanges et renseignent sur les intentions et motivations.

Bien que la littérature offre un large éventail de définitions des émotions (Kleinginna and Kelinginna, 1981), deux aspects sont généralement acceptés :

1. Les émotions sont des réactions à des situations considérées appropriées aux besoins, objectifs et attitudes d'un individu.
2. Les émotions comportent des composantes physiologiques, affectives, comportementales et cognitives.

Dans les interactions sociales, les émotions sont des facteurs de régulation des comportements humains car elles permettent d'avoir un feed-back des actions produites (en percevant comment un comportement est perçu). Elles permettent également de coordonner les échanges sociaux en renseignant sur les états émotionnels de chaque acteur social.

Cet échange d'informations se réalise à travers des expressions émotionnelles comme les expressions faciales, gestuelles, paraverbales, la posture, etc.

Les émotions affectent également la dimension cognitive chez les individus en modifiant les mécanismes d'interprétation et en influençant les processus de prise de décision.

Modéliser les mécanismes d'émergence des émotions et d'influence sur les aptitudes comportementales et cognitives représente un défi constant notamment à cause de notre faible connaissance des facteurs qui guident cette émergence des émotions. Le rapprochement entre les recherches théoriques du domaine des sciences psychologiques, cognitives et informatiques permettent aujourd'hui de voir apparaître différents modèles simulant les processus émotionnels vivants dans des espaces virtuels.

### **4.1.2 Qu'est ce qu'un modèle émotionnel**

D'une façon générale et dans le cadre de la modélisation émotionnelle, nous considérons un modèle émotionnel comme : "un schéma de fonctionnement pour le phénomène d'émergence des émotions".

Une démarche de modélisation émotionnelle devient alors une démarche pour représenter en langage machine tout ou partie des processus émotionnels connus.

Le rôle des simulations émotionnelles et leurs impacts sur les systèmes artificiels a été sujet d'éternelles discussions (Hollnagel, 2003). Il est également important à ce stade de rappeler que notre objectif n'est pas de reproduire l'état émotionnel réel d'un utilisateur mais d'en proposer une représentation qui soit compatible avec la situation et avec la personnalité de ce même utilisateur.

### 4.1.3 Quelques modèles émotionnels

Plusieurs travaux de recherche se focalisent sur la création d'agents capables de communiquer et d'interagir comme les êtres humains.

Beaucoup de travaux se sont intéressés à intégrer la personnalité et les émotions dans les agents virtuels pour leur permettre d'interagir avec des humains en utilisant le langage naturel, les expressions faciales et corporelles de façon crédible.

Parmi les modèles les plus couramment cités, nous retrouvons celui d'Ortony, Clore et Collins (Ortony et al., 1988) appelé également modèle OCC du nom de ses auteurs. Ce modèle se base sur des théories d'évaluations qui permettent de spécifier des propriétés critiques aux événements qui peuvent causer des émotions particulières (Lazarus, 1991; Roseman et al., 1996; Scherer, 1988).

Des systèmes comme 'Émile' (Gratch, 2000; Marsella and Gratch, 2001) implémentent le modèle OCC pour intégrer la dimension émotionnelle dans leurs agents virtuels. Les travaux de Marsella et Gratch (Marsella and Gratch, 2002) se sont également intéressés aussi à l'étude de l'influence des émotions et de la personnalité sur le comportement d'agents autonomes.

Ces systèmes intègrent une forte composante Intelligence Artificielle : ils sont basés sur des systèmes de planification qui génèrent des émotions selon une évaluation des risques qui évolue en fonction d'un ensemble de buts prédéfinis.

Les modèles proposés dans ces travaux s'avèrent inappropriés dans le cadre de notre objectif puisqu'ils nécessitent un système cognitif sophistiqué au niveau des agents pour être implémentés et ne sont que 'faiblement' couplés à leur environnement.

D'autres modèles émotionnels utilisant la personnalité ont été développés pour modéliser les émotions.

Ball and Breese (Ball and Breese, 2000) ont expérimenté des modèles probabilistes avec les réseaux bayésiens.

Une approche similaire est aussi adoptée par Kshirsagar et Magnenat-Thalmann (Kshirsagar and Magnenat-Thalmann, 2002) qui introduisent une

notion supplémentaire dans leur modèle qui est l'humeur. Un modèle générique pour la personnalité et les émotions est également décrit dans (Egges et al., 2004).

D'autres approches utilisent la logique floue pour modéliser l'émergence des émotions suite à des événements (El-Nasr et al., 2003) ou encore l'usage de règles pour prendre en compte la personnalité (André et al., 1999).

Ces approches se basent essentiellement sur des modèles probabilistes pour l'émergence des émotions. Suite à une évaluation d'une situation particulière, un agent peut 'ressentir' une ou autre émotion selon le résultat de calculs probabilistes.

Nous pensons que les émotions émergent comme réactions à des situations considérées appropriées aux besoins, objectifs et attitudes d'un individu. Elles sont déterminées à l'avance en fonction de la situation et ce sont plutôt leurs intensités et conditions d'émergence qui sont sujettes aux calculs probabilistes.

Dans notre système, nous adoptons une approche différente pour modéliser les états émotionnels des agents. D'abord, nous essayons, à partir d'un état de l'art de travaux en psychologie sur les théories des émotions, d'intégrer l'ensemble des notions qui influencent l'émergence des émotions comme l'humeur et la personnalité.

Ensuite, nous essayons de proposer un modèle simple qui intègre entièrement toutes les notions identifiées et qui soit facilement implémentable chez des agents informatiques qu'ils soient automatiques ou dirigés.

Cette simplicité de mise en œuvre s'avère importante dans une simulation interactive multi-utilisateurs très exigeante au niveau du réalisme des personnages virtuels et des réponses du système d'interaction.

## 4.2 Description du modèle émotionnel

Nous présentons dans cette section notre modèle émotionnel pour l'identification des états internes des agents en interactions. Ce modèle intègre plusieurs aspects de la dimension émotionnelle des interactions humaines, comme :

- les émotions,
- la personnalité,
- les humeurs.

La figure 4.1 représente les principaux éléments de ce modèle.

Avant de présenter le détail de modélisation de ce modèle, nous allons justifier selon un point de vue psychologique le choix de ces éléments (émotions, personnalité et humeur) ainsi que l'intérêt de modéliser les relations d'influences entre eux.

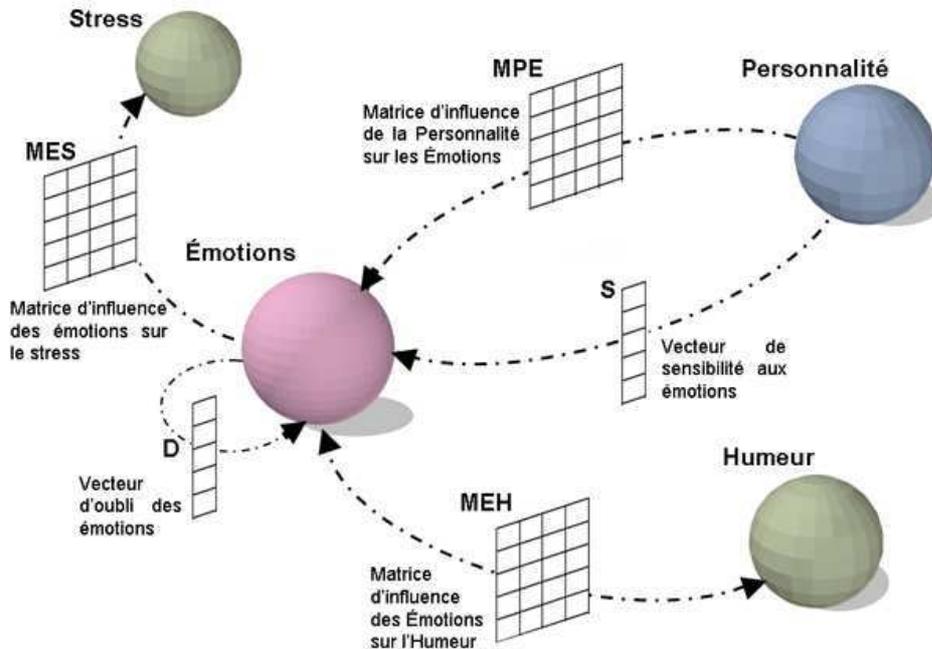


FIG. 4.1 – Schéma d'influences entre les éléments du modèle émotionnel.

#### 4.2.1 Quelques repères psychologiques

Nous distinguons dans notre modélisation la notion d'émotion de la notion de l'humeur. En effet, comme l'a expliqué (Frijda, 1994), les émotions sont 'intentionnelles'. Elles émergent d'une relation avec des objets particuliers. Par exemple, nous sommes content *de* quelque chose, en colère *contre* quelque chose, satisfait *à propos* de quelque chose, etc.

Les humeurs par contre sont 'non intentionnelles'. Elles ne sont pas dues directement à un objet en particulier et elles sont ressenties de façon plus diffuse, globale et générale. Une personne peut être déprimée (il s'agit ici d'une humeur) mais également triste à cause de quelque chose (il s'agit ici une émotion).

Une autre distinction entre les émotions et les humeurs peut être faites selon un point de vue fonctionnel. Comme réaction à différentes situations, les émotions influencent nos actions, elles préparent le corps et l'esprit à une réaction immédiate. Elles tendent également à être limitées dans la durée.

Les humeurs, par contre, tendent à influencer les processus cognitifs des individus et leurs stratégies d'actions pendant des durées plus importantes (Davidson, 1994).

L'interaction entre les émotions et les humeurs est également intéressante à modéliser. D'une part, les humeurs tendent à influencer l'émergence des émotions et d'autre part les émotions contribuent et sont à la cause de l'émergence des humeurs.

Dans notre modèle, nous représentons cette influence des émotions sur l'humeur par une matrice que nous appelons MEH (Matrice d'influence des Emotions sur l'Humeur). Cette relation d'influence sera détaillée plus loin dans ce chapitre.

Les principales causes des émotions sont :

- Les événements perçus : plusieurs travaux se sont intéressés à l'étude de l'impact des événements perçus sur les émotions. Les théories d'évaluations ("appraisal theories" en anglais) permettent de spécifier les caractéristiques des événements qui induisent des émotions particulières (Lazarus, 1991; Ortony et al., 1988; Scherer, 1988).
- La contagion par les émotions des autres : ce phénomène est bien décrit dans les travaux de (Hatfield et al., 1994). Sans toutefois comprendre la logique de cette contagion, une personne qui rit aux éclats amuse le plus souvent son entourage (voire même transmet complètement ses rires à son entourage).
- Les humeurs : comme nous venons de l'invoquer précédemment, agissent comme un filtre pour l'émergence des émotions et tendent à modérer les intensités des émotions ressenties.
- Les états émotionnels antérieurs : les anciens états émotionnels d'un individu influencent également l'émergence de nouvelles émotions. Cette influence apparaît à travers des mécanismes de transfert d'excitation ou d'habitude.

Les mécanismes de 'transfert d'excitation' sont décrits par (Zillmann, 1991) comme l'amplification de l'intensité d'une émotion nouvelle par l'intensité d'une émotion antérieure. Ce transfert d'intensité sera perçue comme faisant partie de la nouvelle émotion ressentie.

Ainsi, une émotion de colère va accentuer l'intensité d'une nouvelle émotion de peur ressentie.

L'émotion est également sujette au phénomène d'habituation. L'habituation est le fait que l'intensité d'une émotion diminue au cours du temps lorsque l'émotion est ressentie de façon répétitive (Frijda, 1988). En effet, lorsqu'une personne est soumise à un impact émotionnel répétitif, elle finit par s'y habituer et s'y attendre. Les émotions ressenties perdent

alors graduellement leurs intensités.

Nous essayons dans notre modèle émotionnel de prendre en compte tous ces facteurs dans l'émergence de nouvelles émotions.

Les émotions affectent également notre mémoire. Des études en psychologie ont permis de montrer que les événements négatifs sont généralement mieux mémorisés que les événements positifs (Newhagen and Reeves, 1991, 1992).

Dans notre modèle émotionnel nous représenterons par le vecteur (D) la capacité d'oubli de chaque émotion ressentie.

La personnalité d'un individu peut également affecter son état émotionnel. Son effet sur le comportement a longtemps été traité dans les recherches en psychologie. Par exemple, les individus considérés comme extravertis utilisent plus de phrases directes que les individus introvertis (Furnham, 1990). Ils parlent également plus rapidement et plus gravement (Scherer, 1979) en utilisant des gestes plus expansifs (Gallaher, 1992).

La relation entre les émotions ressenties et les traits de personnalité est également intéressante à modéliser. Cette relation a été examinée dans plusieurs travaux comme (Marsella and Gratch, 2002) pour étudier l'influence générale de la personnalité sur le comportement ou dans (Johns and Silverman, 2001) qui s'intéressent à l'influence des émotions et de la personnalité sur la prise de décision.

Dans notre modèle émotionnel, la relation d'influence des émotions sur la personnalité est représentée par la matrice MPE (Matrice d'influence de la Personnalité sur les Emotions).

Nous prenons également en compte la sensibilité d'un individu aux émotions. Cette sensibilité représente le seuil d'intensité à partir duquel une émotion peut être ressentie. Elle est représentée par un vecteur (S). Lorsque ce seuil est élevé, l'agent est décrit comme peu sensible à l'émotion. Cette relation d'influence sera détaillée plus loin dans ce chapitre.

Finalement, nous prenons en compte dans ce modèle le facteur stress afin de modéliser le couplage entre les émotions calculées par le modèle émotionnel et les réactions comportementales intentionnelles et non intentionnelles de l'avatar qui seront produites par le modèle d'interaction sociale. Ce facteur stress sera présenté au chapitre 5.

### 4.2.2 Émotion

Des émotions toujours différentes interviennent dans la compréhension des interactions humaines. Plusieurs travaux se sont intéressés à identifier ces émotions et les classer dans différentes catégories selon divers points de vues. Cette identification des émotions s'inscrit dans un débat éternel dans la théorie des émotions : est-ce que les émotions sont innées ou apprises ?

D'un côté, il existe des théories qui argumentent que les émotions sont innées et qu'à chaque émotion correspond une réponse cognitive et physiologique unique (Darwin, 1872/1998; Neese, 1990; Tooby and Cosmides, 1990).

D'un autre côté, il existe des théories qui font l'hypothèse que les émotions sont apprises dans l'environnement social et varient donc selon les cultures et les structures sociales (Averill, 1980; Ortony and Turner, 1990; Shweder, 1994).

Une position intermédiaire consiste à penser que des émotions de base sont partagées par tous les êtres humains (Ekman, 1992; Oatley and Johnson-Laird, 1987; Panksepp, 1992).

La réponse à la question : "quelles sont les émotions de base ?" est également sujet de discussion mais une liste typique comprend souvent des émotions comme la peur, la colère, la tristesse, la joie, le dégoût ou encore l'intérêt et la surprise. Les autres émotions sont souvent vues comme des combinaisons de ces émotions de base.

L'ensemble des émotions que nous considérons est inspiré du modèle OCC décrit par (Ortony et al., 1988). Ce modèle considère qu'une émotion est présente lorsqu'un contexte cognitif particulier existe. Il définit plusieurs types d'émotions regroupées en trois classes de contextes cognitifs généraux. Chaque classe est caractérisée par les conditions nécessaires à l'apparition d'une de ses émotions (Figure 4.2).

A chaque instant, un agent peut ressentir des émotions différentes causées par différents événements.

Il peut par exemple (dans le cadre des secours d'urgence) être satisfait du déroulement global d'une mission, être déçu du retard de l'arrivée du renfort et avoir peur que l'état d'une victime s'aggrave.

Nous représentons, dans notre modèle, les émotions d'un agent par un vecteur  $E_t$  de dimension  $m$ . Chaque émotion du vecteur  $E_t$  est représentée par une intensité variant entre 0 et 1.

$$E_t = \begin{pmatrix} e_1 \\ e_2 \\ \dots \\ e_m \end{pmatrix} \forall i \in [1, m] : e_i \in [0, 1] \quad (4.1)$$

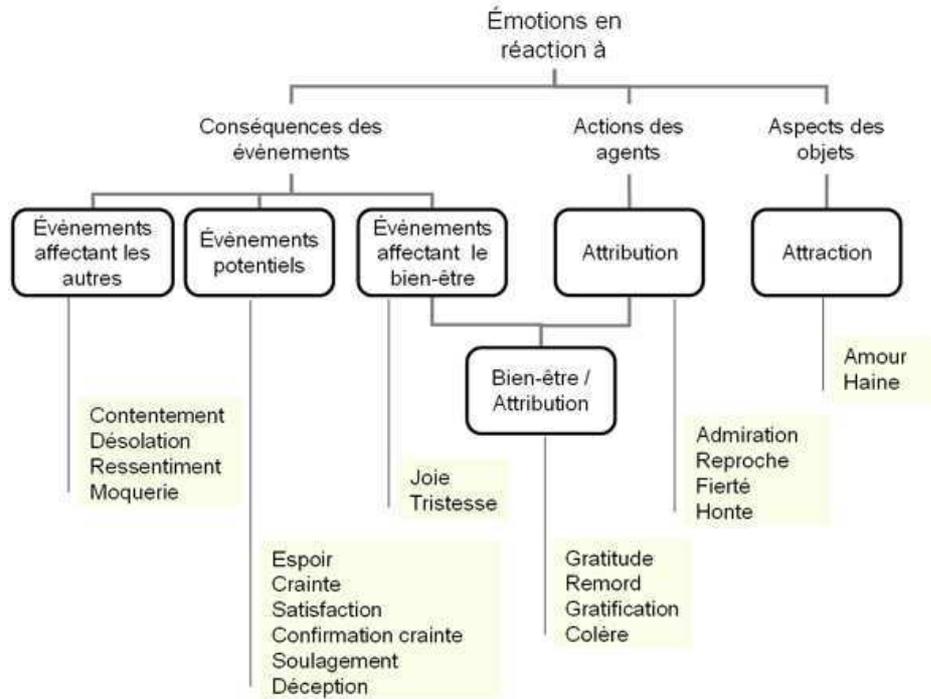


FIG. 4.2 – Classes d'émotions du modèle OCC.

Pour simuler les émotions dans notre système, nous adoptons un vecteur d'émotions à quatre dimensions : Satisfaction, Déception, Colère et Peur.

$$E_t = \begin{pmatrix} e_{satisfaction} \\ e_{deception} \\ e_{colere} \\ e_{peur} \end{pmatrix} \quad (4.2)$$

Ces émotions ont été choisies parce qu'elles reflètent les émotions les plus pertinentes dans le cadre de notre application.

Une autre raison pour cette considération est la réduction du temps de calcul pour l'évaluation des émotions.

Par exemple, soit les émotions d'un agent représentées à un instant  $t$  par le vecteur  $E_t$  suivant :

$$E_t = \begin{pmatrix} 0.4 \\ 0.6 \\ 0.2 \\ 0.1 \end{pmatrix} \quad (4.3)$$

Ce vecteur indique que cet agent ressent :

- une émotion de satisfaction d’une intensité de 0.4,
- une émotion de déception d’une intensité de 0.6,
- une émotion de colère d’une intensité de 0.2,
- une émotion de peur d’une intensité de 0.1.

Nous remarquons dans cet exemple que l’émotion de *déception* et de *satisfaction* malgré qu’elles sont représentées dans d’autres travaux comme des émotions opposées apparaissent dans notre modèle comme des émotions relatives à des événements différents. Un agent peut être satisfait de la réalisation d’un événement 1 et être en même temps déçu concernant le déroulement d’un événement 2.

### 4.2.3 Personnalité

La personnalité représente un thème de recherche important dans le domaine de la psychologie sans qu’il existe toutefois un accord dans la littérature sur sa définition. La plupart des descriptions de personnalités mettent l’accent sur les différences de traits de caractères entre les individus qui se retrouvent dans le comportement et les réponses émotionnelles.

Dans notre modèle, nous définissons la personnalité comme un vecteur à  $n$  dimensions.

$$P = \begin{pmatrix} p_1 \\ p_2 \\ \dots \\ p_n \end{pmatrix} \forall i \in [1, n] : p_i \in [0, 1] \quad (4.4)$$

$p_i$  représente une dimension de la personnalité et sa valeur est comprise entre 0 et 1.

Parmi les modèles de personnalités les plus complets dans la littérature en psychologie, nous considérons le modèle FFM ou Five Factor Model (McCrae and John, 1992). Ce modèle décrit 5 dimensions de la personnalité :

**Openess** décrit l’ouverture de l’esprit chez un individu et son intérêt pour la culture. Elle définit le caractère de curiosité, d’intérêt, de créativité, d’originalité et d’imagination chez un individu.

**Consciencious** décrit l’organisation et la persistance dans l’achèvement des objectifs chez un individu. Elle définit le caractère organisé, travailleur, discipliné, honnête et net d’un individu.

**Extravert** décrit le comportement d’un individu dans des situations sociales. Elle définit le caractère sociable, actif et optimiste d’un individu.

**Agreeable** décrit les interactions d’un individu avec les autres. Elle définit le caractère naturel, sincère, serviable et tolérant d’un individu.

**Neurotic** définit les tendances d'un individu à exprimer des pensées négatives. Elle décrit le caractère anxieux, nerveux, inquiet et émotif d'une personne.

La personnalité est représentée dans notre modèle par un vecteur à 5 dimensions :

$$P = \begin{pmatrix} p_{openess} \\ p_{conscientiousness} \\ p_{extravert} \\ p_{agreeableness} \\ p_{neurotic} \end{pmatrix} \quad (4.5)$$

La personnalité d'un agent est statique et intervient à 2 niveaux dans notre modèle émotionnel :

- Détermination de l'état émotionnel en intégrant le modèle OCC pour conditionner l'émergence des émotions.
- Influence la production d'expressions gestuelles chez les agents. Cette influence sera décrite au chapitre 5.

Par exemple, soit un agent représenté par la personnalité P suivante :

$$P = \begin{pmatrix} 0.1 \\ 0.9 \\ 0.9 \\ 0.5 \\ 0.1 \end{pmatrix} \quad (4.6)$$

Cette personnalité décrit cet agent comme :

- peu ouvert (valeur  $p_{openess} = 0.1$ ),
- très consciencieux ( $p_{conscientiousness}=0.9$ ),
- très extraverti ( $p_{extravert} = 0.9$ ),
- moyennement agréable ( $p_{agreeableness} = 0.5$ ),
- peu anxieux (valeur  $p_{neurotic} = 0.1$ ).

#### 4.2.4 Relation entre personnalité et émotions

La relation entre personnalité et les émotions reste problématique et aucun model unifié et implementable ne permet à ce jour de définir cette relation (André et al., 1999).

Plusieurs approches ont toutefois été développées comme (Bates, 1994) qui associe des comportements aux émotions selon la personnalité. La personnalité est aussi utilisée comme variable pour déterminer l'intensité des émotions dans (Allen, 2000). Dans (Egges et al., 2004), les auteurs définissent une relation entre les dimensions de la personnalité et le modèle OCC. Ce même modèle

OCC a été amélioré dans (Ortony, 2003) pour expliquer les relations entre des modèles de personnalités et les émotions.

Dans notre modèle, chaque émotion peut être influencée par une ou plusieurs dimensions de la personnalité.

Nous définissons une matrice MPE à  $(m \times n)$  dimensions qui représente la relation d'influence de chaque dimension de la personnalité sur les émotions.

L'intensité de cette influence est représentée par une valeur entre 0 et 1.

$$MPE = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ \alpha_{m1} & \alpha_{m2} & \cdots & \alpha_{mn} \end{pmatrix}$$

$$\forall i \in [1, m], j \in [1, n] : \alpha_{ij} \in [0, 1] \quad (4.7)$$

La matrice suivante donne un exemple d'une matrice MPE à  $(4 \times 5)$  dimensions. Chaque ligne représente la valeur d'influence d'une dimension de la personnalité (5 dimensions) sur une émotion particulière (4 émotions de notre modèle).

$$MPE = \begin{matrix} & O & C & E & A & N \\ \begin{matrix} Sat \\ Dec \\ Col \\ Peur \end{matrix} & \begin{pmatrix} 0 & 1 & 0 & 0.3 & 0 \\ 0 & 1 & 0 & 0.3 & 0 \\ 0 & 0 & 0.5 & 0 & 0 \\ 0 & 0 & 0.2 & 0 & 1 \end{pmatrix} \end{matrix}$$

Dans cet exemple, l'émotion de *satisfaction* affichée à la première ligne est influencée par la dimension *conscientious* (valeur vaut 1) et *agreeableness* (valeur vaut 0.3).

$$\begin{pmatrix} 0 & 1 & 0 & 0.3 & 0 \end{pmatrix} \quad (4.8)$$

Par la matrice MPE, nous supposons juste l'existence de liens d'influences entre chaque dimension de la personnalité d'un agent et ses émotions. Les intensités ou poids attribués à ces influences sont pour le moment arbitraire et ne sont pas justifié d'un point de vue psychologique.

#### 4.2.5 Humeur

Une autre caractéristique de la dimension émotionnelle des interactions humaines est l'humeur. Dans notre modèle, l'humeur agit comme un filtre pour l'émergence des émotions et pour l'interprétation des événements.

Une personne de bonne humeur tend à avoir une interprétation positive d'une situation donnée d'où une modération pour les émotions négatives qu'il peut ressentir. Inversement, une personne de mauvaise humeur aura tendance à interpréter négativement des événements positifs et accentuer les émotions négatives ressenties (Davidson, 1994).

Aucun modèle unifié n'existe dans la littérature pour la définition de l'humeur. Nous supposons donc l'existence de plusieurs types d'humeurs (comme par exemple être de bonne ou mauvaise humeur, être d'une humeur dépressive ou encore être d'une humeur amoureuse.).

Dans notre modèle, nous considérons l'humeur comme un vecteur à  $k$  dimensions où chaque dimension de l'humeur est représentée par une valeur comprise entre -1 et 1.

$$H_t = \begin{pmatrix} h_1 \\ h_2 \\ \dots \\ h_k \end{pmatrix} \forall i \in [1, k] : h_i \in [-1, 1] \quad (4.9)$$

Pour implémenter l'humeur dans notre système, nous considérons l'humeur comme un vecteur à 1 dimension où sa valeur varie entre -1 et 1. L'agent peut avoir une bonne ou mauvaise humeur.

$$H_t = h_t \quad (4.10)$$

#### 4.2.6 Relation entre humeur et émotions

Afin de modéliser l'influence de l'état émotionnel sur l'humeur, nous définissons une matrice MEH à  $(k \times m)$  dimensions qui associe un poids d'influence de chaque émotion sur une dimension de l'humeur. Nous supposons que chaque émotion affecte différemment une dimension de l'humeur et que l'intensité de l'humeur dépend de l'intensité des émotions ressenties.

$$MEH = \begin{pmatrix} \beta_{11} & \beta_{12} & \dots & \beta_{1m} \\ \beta_{21} & \beta_{22} & \dots & \beta_{2m} \\ \dots & \dots & \dots & \dots \\ \beta_{k1} & \beta_{k2} & \dots & \beta_{km} \end{pmatrix} \quad \forall i \in [1, k]; j \in [1, m] : \beta_{ij} \in [-1, 1]. \quad (4.11)$$

$\beta_{ij}$  représente la valeur de la dimension  $i$  de l'humeur quand l'intensité d'une émotion  $j$  ressentie est maximale. Par exemple, dans notre implémentation, nous utilisons la matrice MEH  $(1 \times 4)$  suivante :

Ces valeurs représentent le poids des émotions de notre modèle sur l'humeur. Chaque émotion ressentie à sa valeur maximale (valeur = 1) tendra à mettre la valeur de l'humeur à la valeur qui lui correspond dans la matrice

$$MEH = \begin{pmatrix} & Sat & Dec & Col & Peur \\ 0.8 & -0.5 & -1 & -0.6 \end{pmatrix}$$

MEH. Autrement, elle tendra à fixer la valeur de l'humeur proportionnellement à ce poids.

Par exemple, une émotion de satisfaction d'une intensité de 0.5 tentera de fixer l'humeur à la valeur 0.4 c'est à dire proportionnellement au poids 0.8 indiqué par la matrice MEH lorsque l'intensité de l'émotion de satisfaction vaut 1.

D'une façon générale, la valeur de l'humeur est calculée en prenant en compte les intensités non nulles de toutes les émotions ressenties :

$$h_t = \frac{\sum_{i=1}^m e_i \times \beta_i}{\sum_{i=1}^m \frac{e_i}{e_i}} \quad \forall e_i \neq 0 \text{ and } \forall i \in [1, m]. \quad (4.12)$$

Avec  $\sum_{i=1}^m \frac{e_i}{e_i}$  représente le nombre des émotions non nulles ressenties.

Soit, par exemple, un agent avec des émotions  $E_t$  et une matrice MEH comme suit :

$$E_t = \begin{pmatrix} 0.4 \\ 0.2 \\ 0 \\ 0.7 \end{pmatrix} \text{ et } MEH = \begin{pmatrix} 0.8 & -0.5 & -1 & -0.6 \end{pmatrix} \quad (4.13)$$

L'humeur de cet agent est calculée comme suit :

$$h_t = \frac{(0.4 \times 0.8) + (0.2 \times -0.5) + (0.7 \times -0.6)}{3} \quad (4.14)$$

### 4.3 Processus d'évaluation des émotions

Nous expliquons dans cette section le processus que nous mettons en œuvre pour évaluer les émotions ressenties par un agent.

Dans notre modélisation, les émotions ressenties sont provoquées par ses perceptions. Ces perceptions peuvent être des événements potentiels relatifs à la tâche (actions et décisions prises par l'utilisateur durant la simulation) mais également les émotions des autres agents en interaction (qui seront considérés dans notre modèle comme des événements).

Suite à la perception d'un évènement, nous estimons d'abord si l'évènement est capable de produire une émotion sur l'agent. Si oui, nous procédons à l'évaluation de la nouvelle émotion ressentie sinon nous considérons que l'évènement n'a aucun impact émotionnel sur l'agent.

L'évaluation des émotions se fait en deux temps, d'abord nous évaluons une émotion potentielle en considérant juste la personnalité de l'agent puis dans un second temps nous affinons cette émotion potentielle calculée en considérant l'humeur du moment et l'ancien état émotionnel.

Ce mécanisme de calcul nous permet de prendre en compte tous les facteurs d'émergence des émotions que nous venons de décrire comme les évènements perçus, les humeurs et les anciens états émotionnels.

### 4.3.1 La perception des évènements

Pour simuler les évènements perçus, nous définissons  $I_t$  comme un vecteur à  $m$  dimensions représentons l'impact émotionnel de l'évènement perçu sur l'agent. Les valeurs de ce vecteur représentent l'impact émotionnel de l'évènement sur chaque émotion.

Ces valeurs sont définies en fonction de la désirabilité et de l'importance des évènements par rapport aux objectifs de chaque agent (selon le modèle décrit par (Ortony et al., 1988)).

$$I_t = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \dots \\ \lambda_m \end{pmatrix} \forall i \in [1, m] : \lambda_i \in [0, 1] \cup \{-1\}. \quad (4.15)$$

Chaque évènement est porteur d'une ou plusieurs émotions selon les objectifs de chaque agent. Lorsque l'évènement n'a aucun impact émotionnel sur une émotion, la valeur de cet impact dans le vecteur  $I_t$  vaut -1 sinon elle sera comprise entre 0 et 1.

Par exemple, le vecteur  $I_t$  suivant correspond à un évènement porteur d'une émotion de colère probable d'une intensité de 0.8.

$$I_t = \begin{pmatrix} -1 \\ -1 \\ 0.8 \\ -1 \end{pmatrix} \quad (4.16)$$

Nous définissons pour chaque scénario d'interaction et pour chaque agent en interaction, une liste des évènements relatifs à la tâche pouvant survenir lors de simulation.

Cette liste des événements est recensée dans une 'table des événements' associée au scénario.

Le format de cette 'table des événements' est donné par le tableau 4.1 :

Évènement	Désirabilité	Vraisemblance	Emotion
-----------	--------------	---------------	---------

TAB. 4.1 – Format des données de la table des événements.

Dans le champ 'Évènement', nous recensons les principaux événements pouvant avoir lieu dans la simulation. Le système à simuler étant un système socio-technique complexe, la liste n'est bien évidemment pas exhaustive.

La 'Désirabilité' d'un événement est le degré d'importance qu'associe chaque avatar de la simulation à l'évènement (selon son rôle, ses objectifs, etc.). Elle prend une valeur entre -1 et 1.

Une valeur de 1 représente un événement très désirable et inversement.

La 'Vraisemblance' d'un événement est la probabilité qu'associe un personnage à l'apparition d'un événement. Elle prend une valeur entre 0 et 1.

L'évènement est peu probable lorsque son importance tend vers 0 et inversement.

Un exemple d'évènements associés à un scénario d'intervention pour un feu d'hôtel est donné par le tableau 4.2 :

Évènement	Désirabilité	Vraisemblance	Emotion
Arrivée secours	0.81	0.84	SATISFACTION
Explosion	-0.88	0.62	PEUR
Proximité feu	-0.72	0.75	PEUR
Placement échelle	0.65	0.66	DECEPTION

TAB. 4.2 – Exemple de données la table des événements associée à un scénario de Feu d'hôtel.

Les valeurs contenues dans cette table ainsi que les émotions associées aux événements sont modifiées dynamiquement en fonction de l'interaction de l'utilisateur dans l'univers virtuel.

Par exemple et dans le cadre d'une simulation d'un feu d'hôtel, un agent représentant un sapeur pompier aura pour tâche (d'après les procédures prescrites de formation) d'analyser la situation afin de demander les véhicules de secours appropriés à la gestion de l'incident.

Lorsque cet événement potentiel qu'on note dans notre table "arrivée secours" se produit pendant les trois premières minutes d'une intervention, il

est considéré dans notre système comme évènement porteur d'une émotion de *satisfaction*. La désirabilité de cet évènement dans les premières minutes de l'intervention augmente au cours du temps.

Lorsque ce même évènement ne se produit pas pendant les trois premières minutes prévues, nous considérons que l'évènement devient porteur d'une émotion probable de *déception*. Sa désirabilité continue tout de même à augmenter au cours du temps.

Si au bout d'un temps important de l'intervention, les secours ne sont toujours pas sur place, l'agent sapeur pompier commence à considérer qu'il y a un écart par rapport à la procédure et l'évènement devient porteur d'une émotion de *colère*. La vraisemblance de l'évènement commence alors à diminuer.

Un même évènement peut donc produire différentes émotions selon le contexte de l'interaction et son impact émotionnel diffère également selon son temps d'occurrence durant la simulation.

D'autres évènements (qui ne sont pas recensés dans notre table) sont aussi génériques (ils s'adaptent à tous les scénarios de simulations et à tous les agents) comme la contagion des émotions lorsque deux agents sont en proximité l'un de l'autre pendant un espace temps assez important.

### 4.3.2 Sensibilité aux émotions

Certains évènements bien qu'ils peuvent former un contexte cognitif requis à l'émergence d'une émotion, peuvent ne pas posséder certaines intensités suffisantes pour causer l'émergence d'une émotion qui soit effectivement ressentie par l'agent.

Par exemple, nous supposons dans notre modélisation qu'un évènement porteur d'une émotion de *peur* ne soit pas perçu de la même manière selon que l'agent est décrit comme ayant un caractère anxieux ou non.

Pour simuler ce mécanisme, nous définissons une nouvelle notion dans notre modèle qui est la sensibilité d'un agent aux émotions. Cette sensibilité représente le seuil minimum que l'intensité d'une émotion probable doit avoir pour être considérée.

Par exemple, soit un agent défini avec une sensibilité  $S$  qui perçoit deux évènements  $I1_t$  et  $I2_t$  porteurs respectivement d'une émotion de *satisfaction* probable et d'une émotion de *peur* comme suit :

$$S = \begin{pmatrix} 0.3 \\ 0.2 \\ 0.4 \\ 0.5 \end{pmatrix}; \text{ et } I1_t = \begin{pmatrix} 0.6 \\ -1 \\ -1 \\ -1 \end{pmatrix} \text{ et } I2_t = \begin{pmatrix} -1 \\ -1 \\ -1 \\ 0.4 \end{pmatrix} \quad (4.17)$$

L'émotion de *satisfaction* probable portée par l'évènement *I1* peut émerger car elle dépasse la sensibilité de l'agent à cette émotion ( $0.6 > 0.3$ ).

Par contre l'émotion de *peur* portée par l'évènement *I2* n'aura aucun impact sur l'agent puisqu'elle n'atteint pas le seuil d'intensité minimum nécessaire à son émergence ( $0.4 < 0.5$ ). L'agent ne réagira donc pas à cet évènement.

Pour calculer les seuils nécessaires à l'émergence de chaque émotion, nous supposons dans notre modèle que cette sensibilité dépend de la personnalité de l'agent. Nous la modéliserons donc à partir de la matrice MPE (matrice d'influence de la Personnalité sur les Emotions) par un vecteur  $S$  à  $m$  dimensions.

Chaque valeur de ce vecteur correspond à la sensibilité de l'agent à une émotion particulière. Cette sensibilité est représentée par une valeur entre 0 et 1.

$$S = \begin{pmatrix} \theta_1 \\ \theta_2 \\ \dots \\ \theta_m \end{pmatrix} \forall i \in [1, m] : \theta_i \in [0, 1]$$

$$\text{Avec } \theta_i = \frac{\sum_{j=1}^n \alpha_{ij} \times p_j}{\sum_{j=1}^n \alpha_{ij}}, \quad \forall j \in [1, n] \text{ and } \forall i \in [1, m]. \quad (4.18)$$

Comme nous l'avons souligné, une ou plusieurs dimensions de la personnalité influencent chaque émotion. A ce jour, cette sensibilité est calculée dans notre modèle émotionnel en fonction de la personnalité et reste statique au cours du temps mais nous pourrions envisager de faire varier cette sensibilité en fonction d'autres paramètres tels que la fatigue, le stress ou l'expérience déclarée de l'agent.

L'exemple qui suit illustre le processus de calcul de la sensibilité d'un agent avec une personnalité  $P$  et une matrice MPE à une émotion de satisfaction (qu'on note  $\theta_{sat}$ ).

$$P = \begin{pmatrix} 0.5 \\ 0.2 \\ 0.7 \\ 0.4 \\ 0.8 \end{pmatrix} \text{ et } MPE = \begin{pmatrix} 0 & 1 & 0 & 0.3 & 0 \\ 0 & 1 & 0 & 0.3 & 0 \\ 0 & 0 & 0.5 & 0 & 0 \\ 0 & 0 & 0.2 & 0 & 1 \end{pmatrix} \quad (4.19)$$

La ligne correspondant à l'émotion de satisfaction dans la matrice MPE affiche les valeurs suivantes :

$$\begin{pmatrix} 0 & 1 & 0 & 0.3 & 0 \end{pmatrix} \quad (4.20)$$

La sensibilité de cet agent à une émotion de satisfaction est calculée comme suit :

$$\theta_{sat} = \frac{(0.2 \times 1) + (0.4 \times 0.3)}{1 + 0.3} \quad (4.21)$$

Cette sensibilité représente l'intensité minimale qu'un évènement porteur d'une émotion de satisfaction doit avoir pour être considérée.

### 4.3.3 Évaluation des émotions

Les émotions ressenties sont causées par la perception de certains évènements ou la validation de certains objectifs. Le calcul des émotions ressenties  $E_{t+1}$  se fait en deux temps : d'abord nous évaluons la future émotion potentielle (que nous notons  $EP_{t+1}$ ) de l'agent selon la nature de l'évènement perçu et sa personnalité. Ensuite, cette émotion potentielle est modulée par l'humeur de l'agent et son ancien état émotionnel.

Chaque évènement perçu ( $I_{t+1}$ ) est comparé au vecteur  $S$  représentant la sensibilité de l'agent aux émotions pour décider si une émotion peut émerger ou non et ainsi évaluer la future émotion probable.

$$EP_t = \begin{pmatrix} ep_1 \\ ep_2 \\ \dots \\ ep_m \end{pmatrix} \forall i \in [1, m] : ep_i \in [0, 1] \quad (4.22)$$

Lorsque l'émotion portée par un évènement dépasse le seuil de sensibilité de l'agent à la même émotion ( $\lambda_i > \theta_i$ ), l'intensité d'une émotion potentielle ( $ep_i$ ) peut être calculée selon la formule [4.23]. Sinon nous considérons que l'évènement n'a aucun impact émotionnel sur l'émotion de l'agent et le module émotionnel peut continuer à maintenir cette émotion et gérer son oubli au bout d'un certain temps.

$$ep_i = \frac{\lambda_i - \theta_i}{1 - \theta_i} \quad (4.23)$$

Par exemple, un agent dans un état émotionnel  $E_t$  et une sensibilité aux émotions  $S$  perçoit un évènement  $I_{t+1}$  porteur d'une émotion de *peur*.

$$E_t = \begin{pmatrix} 0 \\ 0 \\ 0.5 \\ 0 \end{pmatrix}; S = \begin{pmatrix} 0.2 \\ 0.5 \\ 0.3 \\ \mathbf{0.4} \end{pmatrix} \text{ Et } I_{t+1} = \begin{pmatrix} -1 \\ -1 \\ -1 \\ \mathbf{0.8} \end{pmatrix} \quad (4.24)$$

L'impact de cette perception émotionnelle ( $\lambda_{peur}$ ) est comparé à la sensibilité de l'agent à l'émotion *peur* ( $\theta_{peur}$ ) :  $0.8 > 0.4$ . Une émotion de peur peut donc émerger chez l'agent et la valeur de son état émotionnel potentiel sera calculée par la formule :

$$EP_{t+1} = \begin{pmatrix} 0 \\ 0 \\ 0.5 \\ \frac{0.8-0.4}{1-0.4} \end{pmatrix} \quad (4.25)$$

Le calcul des émotions a jusqu'à maintenant été évalué en fonction de l'impact émotionnel de l'évènement ( $I_{t+1}$ ) sur l'agent et sa sensibilité (S) qui dépend de sa personnalité. La nouvelle émotion ressentie par un agent à un instant  $t+1$  sera évaluée en fonction de cet état émotionnel potentiel calculé ( $EP_{t+1}$ ) mais aussi en prenant en compte l'humeur de l'agent ( $H_t$ ) et son état émotionnel ( $E_t$ ) à l'instant  $t$ .

$$E_{t+1} = EP_{t+1} + \sigma(H_t, E_t, EP_{t+1}) \quad (4.26)$$

Avec  $\sigma$  = Facteur de modération.

La figure 4.3 décrit cette approche de calcul de l'intensité des nouvelles émotions ressenties.



FIG. 4.3 – Processus de calcul de l'intensité d'une nouvelle émotion.

L'humeur dans notre modèle agit comme un filtre 'affectif' qui tend à modérer ou accentuer les émotions ressenties. Un agent de bonne humeur aura tendance à accentuer les émotions positives et minimiser l'impact des émotions négatives.

Par exemple, un agent avec une humeur négative qui ressent une émotion de *peur* aura tendance à accentuer l'intensité de cette émotion ressentie. Lorsque ce même agent avec cette humeur négative ressent une émotion de *satisfaction* il aura plutôt tendance à modérer cette intensité ressentie.

Le tableau 4.3 donne les différentes valeurs possibles du facteur de modération que nous utilisons.

Le nouvel état émotionnel d'un agent est calculé par la formule suivante :

$$E_{t+1} = EP_{t+1} + \tau \times |h_t \cdot (EP_{t+1} - E_t)| \quad (4.27)$$

facteur de modération ( $\tau$ )	Humeur négative	Humeur positive
Émotion positive ( <i>ex : Satisfaction</i> )	-1	+1
Émotion Négative ( <i>ex : Déception, Peur, Colère</i> )	+1	-1

TAB. 4.3 – Valeurs possibles du facteur de modération.

Avec :

$\tau$  = Le facteur de modération :  $\tau \in \{-1, 1\}$

$E_{t+1}$  = L'état émotionnel à l'instant  $t$ .

$EP_{t+1}$  = L'état émotionnel potentiel à l'instant  $t+1$ .

$h_t$  = Humeur de l'agent à l'instant  $t$ .

Par exemple, un agent avec un état émotionnel  $E_t$  et une humeur  $h_t$  perçoit un évènement estimé provoquant un état émotionnel probable  $EP_{t+1}$

$$E_t = \begin{pmatrix} 0 \\ 0 \\ 0.8 \\ \mathbf{0.2} \end{pmatrix}; \quad h_t = -0.5 \quad \text{et} \quad EP_{t+1} = \begin{pmatrix} 0 \\ 0 \\ 0.8 \\ \mathbf{0.7} \end{pmatrix} \quad (4.28)$$

Le changement dans ce nouvel état émotionnel potentiel affecte seulement l'émotion de Peur qui est évaluée pour passer de la valeur 0.2 à 0.7. Selon notre modèle émotionnel, le nouvel état émotionnel de l'agent est calculé par :

$$E_{t+1} = \begin{pmatrix} 0 \\ 0 \\ 0.8 \\ 0.7 + 1 \times |-0.5 \times (0.7 - 0.2)| \end{pmatrix} \quad (4.29)$$

#### 4.3.4 Mise à jour de l'humeur

L'humeur affecte l'état émotionnel, il est donc essentiel de mettre à jour l'humeur dès que l'état émotionnel change. Supposons que le nouvel état émotionnel d'un agent est donné par  $E_t$  et que la matrice MEH (Matrice de l'influence des Emotions sur les Humeurs) affiche les valeurs suivantes :

$$E_t = \begin{pmatrix} 0 \\ 0 \\ \mathbf{0.5} \\ \mathbf{0.8} \end{pmatrix}; \quad MEH = \begin{pmatrix} 0.8 & -0.5 & -1 & \mathbf{-0.6} \end{pmatrix} \quad (4.30)$$

L'impact de chaque émotion ressentie sera calculé proportionnellement à son intensité selon les poids d'influences affichés par la matrice MEH. Soit

$h_{emotion}$  l'impact d'une émotion ressentie sur l'humeur. l'agent ressent deux émotions : une émotion de colère (0.5) et une émotion de peur (0.8).

La nouvelle valeur de l'humeur est donnée par :

$$h_t = \frac{h_{anger} + h_{fear}}{2} = \frac{(0.5 \times -1) + (0.8 \times -0.6)}{2} \quad (4.31)$$

#### 4.3.5 La mémoire des émotions

Dans notre modèle émotionnel, l'intensité d'une émotion est maintenue pendant un certain temps avant de commencer progressivement à diminuer pour tendre vers la valeur 0. L'oubli des émotions diffère d'une émotion à une autre. En effet, un individu tend en réalité à oublier les émotions positives plus rapidement que les émotions négatives (Reeves and Nass, 1996; Reeves et al., 1991).

Pour simuler ce mécanisme d'oubli des émotions, nous définissons le vecteur  $D$  à  $m$  dimensions comme vecteur d'oubli des émotions.

$$D = \begin{pmatrix} \delta_1 \\ \delta_2 \\ \dots \\ \delta_m \end{pmatrix} \forall i \in [1, m] : \delta_i \in [0, 1] \quad (4.32)$$

Chaque valeur contenue dans le vecteur  $D$  représente le facteur d'oubli d'une émotion au cours du temps.

Ce facteur sera utilisé par notre système pour maintenir l'émotion ressentie ( $e_i$ ) durant un certain temps proportionnel à ce facteur ( $\delta_i$ ). L'intensité de l'émotion ressentie commence ensuite à diminuer progressivement selon un rythme inversement proportionnel à ce facteur.

Par exemple, soit un agent avec un vecteur d'émotions  $E_t$  et un vecteur  $D$  comme suit :

$$E_t = \begin{pmatrix} 0.7 \\ 0.4 \\ 0.8 \\ 0.2 \end{pmatrix} \text{ et } D = \begin{pmatrix} 0.2 \\ 0.4 \\ 0.7 \\ 0.8 \end{pmatrix} \quad (4.33)$$

L'émotion de satisfaction va être ressentie pendant un temps proportionnel à son facteur d'oubli ( $\delta_{sat} = 0.2$ ) et donc sera maintenue par notre système moins longtemps qu'une émotion de peur dont le facteur d'oubli est plus important ( $\delta_{peur} = 0.8$ ).

Ensuite cette émotion de satisfaction sera oubliée plus rapidement que l'émotion de peur puisque les nouvelles intensités de chaque émotion sera calculée de façon inversement proportionnelle à son facteur d'oubli.

La figure 4.4 montre la variation dans notre modèle de l'intensité d'une émotion de *Satisfaction* et de *Peur* au cours du temps.

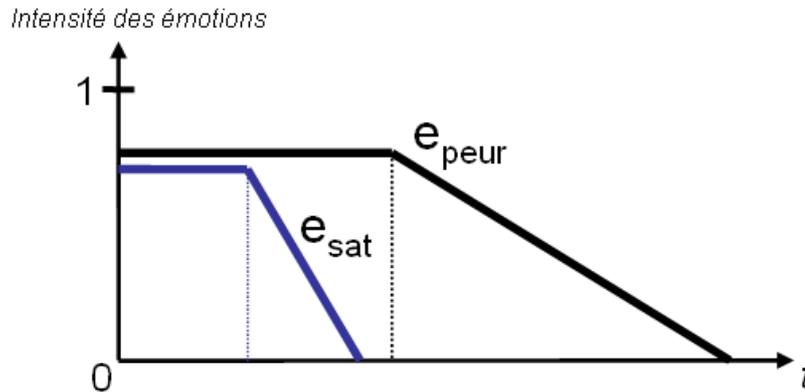


FIG. 4.4 – Processus d'oubli des émotions au cours du temps.

#### 4.4 Discussion

Dans ce chapitre, nous décrivons un modèle émotionnel pour produire des états émotionnels chez des agents virtuels.

Ce modèle décrit le processus d'évaluation, maintien et oubli des émotions chez les agents en interaction. Il prend en considération la personnalité, l'humeur et les anciens états émotionnels des agents pour évaluer les nouveaux états émotionnels.

Pour cela, nous analysons au cours de la simulation les actions et décisions prises par l'utilisateur afin de les transformer en événements pouvant avoir un impact émotionnel sur l'état interne de l'avatar (représentation de l'utilisateur dans l'univers).

Pour évaluer l'intensité des émotions ressenties par l'utilisateur, nous avons recours à un modèle analytique qui simule en fonction des interactions de l'avatar dans l'univers virtuel les états émotionnels.

Les évaluations de ce modèle qui sont discutées au chapitre 7 nous permettent de mesurer cet écart entre les émotions simulées par notre modèle et les émotions ressenties par l'utilisateur. Nous essayons également d'évaluer le degré d'immersion émotionnelle de l'utilisateur et jusqu'à quel point il peut oublier les imperfections du système pour réagir émotionnellement dans l'univers virtuel.

Nous décrivons dans le chapitre 5 le modèle que nous mettons en œuvre

pour produire des communications situées entre des agents en interactions dans un univers virtuel multi-utilisateurs.



## Chapitre 5

# Vers des communications situées

*Une des choses qui fait que l'on trouve  
si peu de gens qui paraissent raisonnables  
et agréables dans la conversation,  
c'est qu'il n'y a presque personne  
qui ne pense plutôt à ce qu'il veut dire  
qu'à répondre précisément à ce  
qu'on lui dit. (La Rochefoucauld)*

Dans ce chapitre, nous nous intéressons à la modélisation de communications situées dans des simulations virtuelles de gestion de crises où la dimension émotionnelle et sociale de l'interaction sont déterminantes dans la gestion de la situation.

Il s'agit d'offrir à des utilisateurs de système de réalité virtuelle de nouveaux moyens d'interactions capables de renforcer la notion de présence sociale.

Cet objectif ne peut être atteint qu'avec la prise en compte d'une part du contexte de la situation d'interaction, et d'autre part de ces mécanismes socio-cognitifs comme la réflexivité et l'indexicalité dans les interactions sociales.

Nous présentons dans ce chapitre, notre modèle d'interaction sociale (section 5.1 ) que nous avons implémenté dans notre plateforme de réalité virtuelle afin de prendre en compte cette dimension sociale et émotionnelle de l'interaction.

## 5.1 Les communications situées

### 5.1.1 Introduction

La réalité virtuelle offre un support idéal pour la simulation des comportements humains.

Plusieurs travaux se sont intéressés à la modélisation d'agents virtuels capables de produire des interactions conversationnelles avec des utilisateurs humains.

Ces agents sont apparus dans différents rôles selon les domaines d'applications comme :

- les agents pédagogiques pour la formation (Rickel and Johnson, 1999; Rety et al., 2003);
- les agents conversationnels pour le web (Lester et al., 2000; Pelé et al., 2003);
- les agents présentateurs de kiosques (Gustafson et al., 1999);
- les agents conseillers médicaux (Pelachaud et al., 2002).

Ces travaux adoptent des approches basées sur des traitements représentationnels de l'information où l'agent dispose généralement de modèles analytiques de perception, décision et action pour interagir avec l'utilisateur.

Nous pensons que ces approches atteignent vite des limites conceptuelles non par rapport à leurs méthodes de traitement mais essentiellement à cause de la nature représentationnelle des modèles de l'Intelligence Artificielles et des théories classiques des Sciences Cognitives. En effet, ces modèles s'inscrivent habituellement dans une logique déterministe qui associe un sens unique ou une signification prédéfinie aux éléments lors des processus de traitements. L'usage de tels modèles s'avère inefficace pour modéliser des systèmes complexes où le processus décisionnel est étroitement lié à l'interprétations des éléments contextuels de la situation.

Ces limites théoriques peuvent disparaître avec des approches issues de courants constructivistes, connexionnistes ou post-structuralistes. Ces approchent repensent la relation univoque qui existe entre un signifiant (élément ou objet d'un contexte particulier) et son signifié (sa signification pour l'action).

Ils proposent une vue plus dynamique et plus instable de cette relation où signifiants et signifiés sont en relation mutuelle permanente. Le sens d'un élément n'est donc pas prédéfini (spécifié selon un sens unique) mais déterminé en fonction de sa relation avec les autres éléments contextuels.

Un simple changement de l'univers (au niveau d'un signifiant) entraîne évidemment un changement global sur tous les autres éléments opérant ainsi

une transformation globale sur l'interprétation et le sens commun.

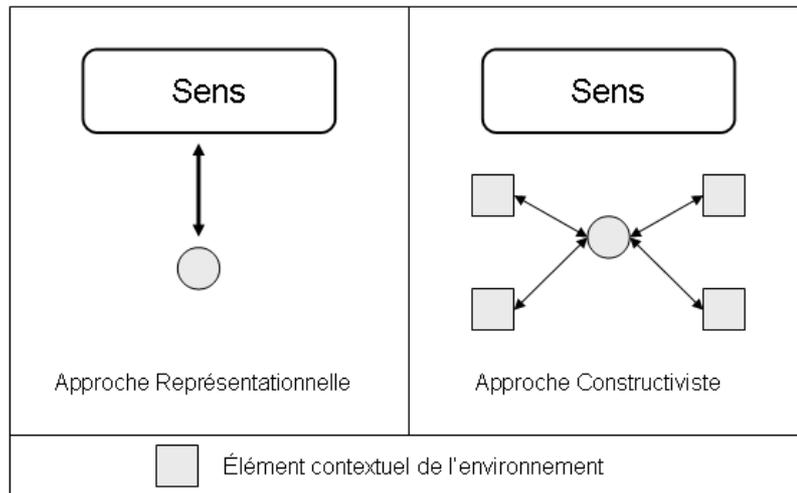


FIG. 5.1 – L'émergence du sens selon que l'on s'adresse aux approches représentationnelles ou constructivistes.

### 5.1.2 Approche pour des communications situées

Nous nous proposons dans nos travaux de développer un système de réalité virtuelle capable de prendre en compte une partie de ces mécanismes naturels des communications humaines et ainsi offrir un cadre permettant aux utilisateurs d'engager des communications 'situées' lors des interactions virtuelles.

Pour créer une interaction située en univers virtuel, nous avons développé un modèle d'interaction sociale et émotionnelle qui crée des couplages entre l'expression des différents utilisateurs.

Ce modèle permet d'une part à l'utilisateur de contrôler 'naturellement' son avatar, et d'autre part d'enrichir les actions décidées par l'utilisateur par des comportements non intentionnels compatibles avec les règles d'interactions sociales comme celles de la gestion du dialogue, de la proxémique, etc.

Il offre la possibilité de compléter les actions des utilisateurs par des comportements non intentionnels chez son avatar comme :

- la production d'expressions gestuelles lors du dialogue pour accompagner le discours de l'utilisateur ;
- la production d'expressions émotionnelles qui traduisent les états internes de l'avatar ;
- la production de hochements automatiques de la tête lors de l'audition ;

- la distribution automatique du regard entre les différents interlocuteurs lord de la locution ;
- l'orientation automatique du regard à la rencontre de nouveaux avatars ;
- le changement dynamique des postures et des animations en fonction du contexte de l'interaction.

Néanmoins, les utilisateurs peuvent intentionnellement 'prendre la main' sur ces comportements non intentionnels et par exemple autoritairement obliger l'avatar à fixer un interlocuteur particulier lors d'un dialogue.

Le modèle d'interaction sociale offre également des possibilités d'effectuer naturellement des actions intentionnelles capables de renforcer l'engagement émotionnel et social comme :

- l'usage du langage naturel pour communiquer avec les autres utilisateurs ;
- l'usage des gestes déictiques pour renforcer les mécanismes indexicaux des communications humaines ;
- la direction intentionnelle du regard pour faciliter le repérage dans l'univers virtuel et la désignation des directions, des personnes, etc. ;
- les actions volontaires comme les déplacements dans l'environnement qui permettent de reproduire les normes culturelles de proxémique ;
- la gestion des tours de paroles en groupe.

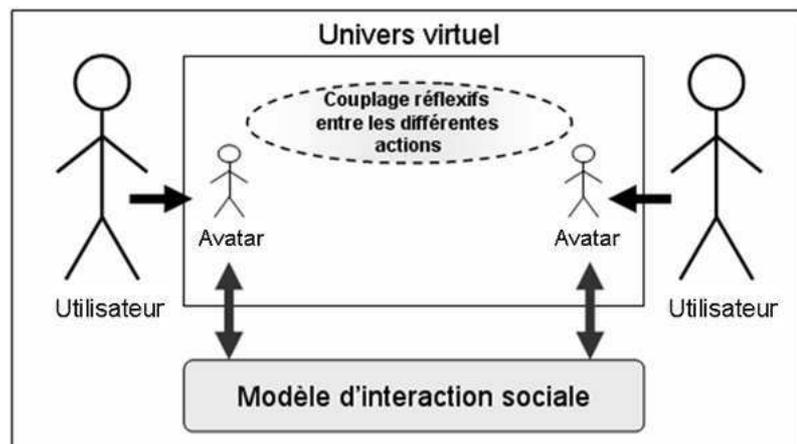


FIG. 5.2 – Cadre situé pour les interactions engagées en univers virtuel.

Cette architecture de contrôle des comportements a été conçue pour :

- D'une part, permettre de ne pas surcharger cognitivement l'utilisateur

en lui demandant en permanence de spécifier ses expressions gestuelles ou corporelles ;

- D'autre part, pour permettre à l'utilisateur de 'repren- dre la main' sur son avatar et par exemple imposer une direction de regard vers un interlocuteur particulier.

Un des points importants de cette recherche a été de trouver une articulation 'naturelle' entre ces deux types de contrôle.

Le détail des mécanismes d'interactions mis en œuvre dans ce modèle est décrit dans la suite de ce chapitre.

## **5.2 Les communications intentionnelles**

Pouvoir reproduire en univers virtuel les normes conversationnelles de la communication verbale constitue un préalable pour renforcer la sensation d'immersion sociale, émotionnelle et culturelle dans les interactions en univers virtuel.

Nous exposerons dans ce qui suit les solutions choisies pour renforcer les moyens d'interaction et le caractère situé des communications.

### **5.2.1 Usage du langage naturel**

L'expression para verbale a des incidences importantes sur l'impact conversationnel.

Les marqueurs comme la vitesse d'élocution, l'intensité vocale, la hauteur de la voix (" pitch "), les intonations (ou courbes mélodiques obtenues par variation de la hauteur de la voix) sont des indices fondamentaux pour l'interprétation des actes de communication (Cappella, 1985).

Ces informations sont par exemple utilisées pour communiquer ou interpréter nos états émotionnels ainsi que pour identifier les traits de personnalité de l'utilisateur (38% des émotions d'une communication sont véhiculées par les éléments vocaux) (Mehrabian, 1981).

Les références langagières indexicales ou déictiques ('je', 'tu', 'ici', 'là-bas', 'demain', etc.) sont également des éléments du discours qui, pour être correctement interprétés, doivent faire référence au contexte de l'énonciation. L'interprétation et la compréhension d'une situation sont étroitement liées au sens et aux références spatiales que nous associons à ces déictiques.

C'est avec l'objectif de reproduire ces modes d'expression que notre plateforme permet aux utilisateurs de dialoguer en réseau à travers des microphones et de communiquer en utilisant le langage naturel. L'expression verbale des avatars est synchronisée avec la production de sons par les acteurs humains. Nous verrons plus loin que les expressions émotionnelles sont ajoutées à ces expressions verbales en fonction du contexte. Ainsi chacun peut s'exprimer naturellement selon son histoire, sa culture et ses expériences passées.

### 5.2.2 Usage des gestes indexicaux

Les gestes déictiques permettent de désigner une personne, un objet, un niveau ou une direction. Ils peuvent, selon leurs modalités (amplitude, pointage, direction du regard) entraîner le déplacement de l'attention de l'utilisateur vers un focus d'intérêt.

Ces gestes sont parfois effectués de façon inconsciente pour accompagner un discours ou orienter vers une direction mais le plus souvent on a recours à ce type de gestes de façon intentionnelle pour insister sur un discours ou pour confirmer un sens.

L'utilisation des expressions gestuelles déictiques est un mécanisme naturel de communication non verbale que nous utilisons dans nos interactions quotidiennes pour nous orienter dans l'univers et construire du sens.

Afin de permettre aux utilisateurs de reproduire ces déictiques en univers virtuel, nous avons développé une interface de pointage intuitive basée sur l'usage d'une manette de contrôle 3D.

Cette interface permet d'orienter un geste déictique vers n'importe quelle direction dans l'espace virtuel. Notre modèle d'interaction sociale se charge d'associer une direction à ces déictiques orientée vers la direction du regard (figure 5.3).



FIG. 5.3 – Exemples d'expressions gestuelles indexicales.

On remarque que la direction du regard est cohérente avec la direction du déictique (comportement automatique généré par le modèle d'interaction sociale).

Ce couplage assez naturel et souple de l'orientation dans un espace virtuel

est très semblable aux situations de repérage en réalité où nous sommes amenés inconsciemment à orienter le regard vers la direction que nous cherchons à pointer par la main.

Notre modèle contrôle également la direction du regard des interlocuteurs en fonction du contexte (situation de dialogue, statut social des interlocuteurs, etc.).

Par exemple, dans une situation de dialogue professionnel, la direction des regards des interlocuteurs sera automatiquement orientée vers la direction indiquée par le bras du locuteur produisant le déictique.

On peut remarquer que si cette action socialement organisée n'est pas produite, le locuteur pourrait interpréter ce non respect des conventions sociales comme une marque de désintérêt ou de préoccupation de la part de son interlocuteur.

### **5.2.3 La direction du regard**

Le regard représente une forme de communication à la fois puissante et directe dans les échanges inter-humains (Leathers, 1997).

Dans les interactions sociales en univers virtuel, pouvoir contrôler la direction du regard permet aux utilisateurs de transmettre plus de sens dans leurs échanges. Ainsi, la direction du regard peut être utilisée pour indiquer des directions et éviter le langage naturel dans des environnements bruyants.

Dans une situation de dialogue le même regard peut être 'fuyant' ou 'insistant' et transmet des messages à chaque fois différents et nouveaux.

Sa durée et son orientation également obéissent à des règles précises totalement inconscientes (La France and Mayo, 1976).

Les travaux de (Kendon, 1967) ont également souligné l'importance de la direction du regard dans la régulation des tours de parole lors des conversations.

Cette gestion des tours de parole dans une communication entre un groupe de personnes relève des caractéristiques essentielles d'une conversation. Elle est caractérisée par plusieurs aspects qui relèvent de la culture de chacun (Sacks et al., 1974).

Ainsi le temps qui sépare les tours de parole (ou 'gap') peut être un indice contextuel important dans une conversation que les interlocuteurs exploitent pour prendre la parole.

Un autre phénomène de la gestion des tours de paroles est le chevauchement ('overlap') qui est plus ou moins toléré et diversement interprété selon les habitudes et traditions culturelles de chacun.

D'autres aspects du fonctionnement des tours de paroles comme l'ordre des tours, leurs longueurs ainsi que les techniques mises en œuvre pour la prise du

tour sont également des éléments contextuels importants dans le réalisme des interactions langagières.

Cette gestion des tours de paroles peut être utilisée par les interlocuteurs soit pour introduire le nouveau locuteur soit pour implicitement exprimer une intention, une marque d'intérêt ou de désintérêt, etc. Il est donc essentiel de permettre en univers virtuel à des utilisateurs de reproduire des comportements adaptés à leurs normes sociales et culturelles.

Dans notre modèle d'univers virtuel, la direction du regard est contrôlée par un double mécanisme :

- Elle est orientée par le modèle d'interaction sociale (en fonction du contexte de la situation). Un module spécialisé détecte la localisation des différents interlocuteurs, leur statut social et leurs états émotionnels afin de produire une alternance du regard entre les différents avatars en co-présence.
- Elle peut également être intentionnellement orientée par les acteurs humains (grâce à la rotation de la manette de commande). L'utilisateur reste libre de prendre le contrôle de son avatar afin d'orienter son regard pendant son discours vers une personne particulière (avatar) transmettant ainsi des indices contextuels supplémentaires pour l'interprétation de la situation (discours adressé à une personne en particulier par exemple)(figure 5.4).

#### 5.2.4 La proxémique

La façon de se positionner dans une interaction est également un élément très significatif lors d'un échange humain.

La 'bonne distance' à adopter, dans une situation donnée, vis-à-vis de son partenaire d'interaction est étroitement liée au contexte de la situation mais également de la conception culturelle que les interlocuteurs se font de cette distance (Schefflen and Ashcraft, 1976).

Les travaux de (Hall, 1966) distinguent quatre zones de communication :

- la zone 'intime' (entre 15 et 45 cm)<sup>1</sup> : les personnes qui se trouvent à cette distance sont souvent très liées et les conversations deviennent souvent plus faciles et spontanées.
- la zone 'personnelle' (entre 45 cm et 1,20 m) : indique en général une bonne entente entre les interlocuteurs. Il est important de noter que 1,20 mètres est, de façon approximative, la distance maximale à laquelle on

---

<sup>1</sup>Les distances associées à ces zones sont données à titre approximatif et varient selon les cultures.



FIG. 5.4 – Contrôle intentionnel de la direction de regard.

peut toucher une personne en allongeant un bras. On retrouve cette distance dans les relations professionnelles, voire amicales ;

- la zone 'sociale' (entre 1,20 m et 3,50 m) : elle marque le statut social de chacun. Elle implique un certain détachement et dans certains cas cette distance indique une supériorité sur le plan hiérarchique d'un des interlocuteurs ;

- la zone 'publique' (> 3,50 m) : elle n'est utilisée qu'en présence d'un grand nombre de personnes. La communication est souvent unidirectionnelle (ou presque).

En permettant en univers virtuel aux utilisateurs de placer eux-mêmes leurs avatars en situation de dialogue, nous contribuons ainsi à reproduire les mécanismes naturels de proxémique adoptés dans notre vie quotidienne et qui sont également étroitement liés à la culture de chacun.

### 5.3 Les communications non intentionnelles

Les comportements non intentionnels produits par notre modèle d'interaction sociale concernent les postures, les expressions gestuelles et les expressions faciales que les interlocuteurs utilisent pour donner de la consistance à leurs interactions émotionnelles et sociales.

Des études ont montré que 65% des informations échangées durant une interaction en face à face sont exprimées de façon non verbale (Argyle, 1975).

Les interactions non verbales, perçues inconsciemment, sont essentielles pour donner sens à la situation.

#### 5.3.1 Gestualité lors du dialogue

Les gestes sont très utilisés dans les communications humaines pour transmettre du sens et accompagner le discours.

De nombreux travaux de classification des gestes ont été présentés dans la littérature (McNeill, 1992; Kendon, 1988; Nespoulos and Lecours, 1986; Le Breton, 1998).

Notre objectif n'est pas de reproduire de façon réaliste l'ensemble de ces gestes mais d'en sélectionner une partie qui soit pertinente pour notre situation expérimentale et analyser dans quelle mesure les utilisateurs, en situation d'interaction sociale, peuvent en temps réel exploiter cette information et ainsi réaliser des 'couplages' dynamiques qui ont un sens par rapport à leurs objectifs.

Afin de produire chez nos avatars une gestualité lors du discours, nous considérons parmi les principales classes de gestes identifiés :

- les gestes 'expressifs' qui traduisent l'affectivité du sujet au fil de son écoute ou de sa parole ;
- les gestes 'rythmiques' qui scandent l'énoncé sans ajouter du sens au discours ;
- les gestes de 'régulation' qui contribuent à maintenir le contact entre les interlocuteurs ;

Le choix de ces classes de gestes est dicté par la difficulté technique liée à l'identification en temps réel du sens des phrases prononcées par l'utilisateur lors de l'interaction.

En effet, ces classes de gestes choisis sont moins dépendantes du discours prononcé (contrairement aux gestes descriptifs par exemple qui accompagnent le discours) et peuvent donc être modélisées dans notre univers virtuel par des animations génériques.

Dans notre modèle, nous prenons également en compte le fait que la production de ces gestes est influencée par les dimensions de la personnalité d'un individu.

Par exemple, une personne extravertie aura tendance à produire plus de gestes expansifs qu'une personne introvertie (Gallaher, 1992).

Le modèle ainsi mis en œuvre pour générer des expressions gestuelles chez nos avatars lors des situations de dialogue prend en compte la personnalité déclarée de l'avatar.

De ce fait, nous définissons une matrice MPG ( $g \times n$ ) qui associe à chaque catégorie de gestes ( $g$ ) et pour chaque dimension de la personnalité d'un individu une valeur d'influence comprise entre 0 et 1.

$$MPG = \begin{pmatrix} \gamma_{11} & \gamma_{12} & \cdots & \gamma_{1n} \\ \gamma_{21} & \gamma_{22} & \cdots & \gamma_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ \gamma_{g1} & \gamma_{g2} & \cdots & \gamma_{gn} \end{pmatrix}$$

$$\forall i \in [1, g], j \in [1, n] : \gamma_{ij} \in [0, 1]$$

La matrice MPG implémentée dans notre système est donnée par la matrice suivante où chaque ligne représente respectivement les gestes de régulation, rythmiques et expressifs :

$$MPG = \begin{matrix} & O & C & E & A & N \\ G_{regul} & \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix} \\ G_{rythm} & \\ G_{express} & \end{matrix}$$

Dans un dialogue, chaque personnage exprime ces types de gestes de manière non intentionnelle mais la fréquence de leurs apparitions varie selon sa personnalité.

Soit  $f_i$  la fréquence d'apparition d'une catégorie de gestes chez un avatar (avec  $i \in \{ \text{régulation, rythmiques, expressifs} \}$ ).

$$f_i = \sum_{j=1}^n \gamma_{ij} \cdot p_j, \quad \forall j \in [1, n] \text{ and } \forall i \in [1, g].$$

Le taux d'apparition  $\tau_{gestes}$  de chaque catégorie de gestes pendant le discours est donné par la formule :

$$\tau_j = \frac{f_j}{\sum_{i=1}^g \tau_i} \quad \forall i \in [1, g]$$

Ainsi, un avatar avec une personnalité 'consciencieuse' aura par exemple tendance à exprimer des gestes de régulation pour maintenir le contact avec ses interlocuteurs. Un autre avatar décrit comme 'agréable' et sociable aura plus tendance à exprimer des gestes expressifs lors de son discours.

La vitesse d'animation de ces gestes est également prise en compte par notre modèle. Cette vitesse est couplée à la valeur du facteur *stress* de l'avatar.

### 5.3.2 La notion de stress dans le comportement

Le stress caractérise la réaction de l'organisme à une sollicitation extérieure jugée 'négative'. Cette réaction de l'organisme a pour conséquence une mobilisation physiologique et/ou psychique qui influence la perception et l'interprétation d'une situation donnée (Franken, 1994).

L'objectif de cette prise en compte du facteur stress est de modéliser le couplage entre les émotions calculées par le modèle émotionnel et les réactions comportementales intentionnelles et non intentionnelles (produites par le modèle d'interaction sociale) de l'avatar.

A ce jour, aucun modèle unifié ne permet de donner une définition précise à la notion de stress et à ses causes (Selye, 1982). Par contre, beaucoup de travaux s'accordent pour considérer le stress comme un état généré par les émotions ressenties et qui influence la perception et les comportements.

Pour modéliser le stress d'un avatar, nous supposons qu'il existe une relation entre les émotions et le stress. Étant donné que le calcul des émotions est influencé par la personnalité et l'humeur, le facteur stress prend ainsi en compte l'ensemble des notions abordées dans notre modèle.

Nous modélisons cette relation par un vecteur MES (Matrice de l'influence des Emotions sur le Stress) à  $m$  dimensions qui contient les valeurs d'influences de chaque émotion sur le stress.

$$MES = \left( \begin{array}{cccc} \sigma_1 & \sigma_2 & \cdots & \sigma_m \end{array} \right) \\ \forall i \in [1, m] : \sigma_i \in [0, 1]. \quad (5.1)$$

Le vecteur suivant donne un exemple d'un vecteur MES à 4 dimensions. Chaque valeur représente le poids d'influence d'une émotion (4 émotions considérées dans notre modèle) sur le facteur stress.

$$MES = \begin{pmatrix} Sat & Dec & Col & Peur \\ 0.3 & 0.5 & 0.7 & 0.9 \end{pmatrix}$$

La valeur du stress d'un avatar est calculée par :

$$stress = \frac{\sum_{i=1}^m e_i \times \sigma_i}{\sum_{i=1}^m \sigma_i} \quad (5.2)$$

Ainsi, un avatar dont le modèle émotionnel identifie les émotions ressenties comme  $E_t$  et dont la matrice MES affiche les valeurs suivantes :

$$E_t = \begin{pmatrix} 0.75 \\ 0.62 \\ 0.54 \\ 0.89 \end{pmatrix}; \quad MES = \begin{pmatrix} 0.3 & 0.5 & 0.7 & 0.9 \end{pmatrix} \quad (5.3)$$

Cet avatar se retrouve avec un niveau de stress estimé à 71% par notre modèle d'interaction sociale comme :

$$stress = \frac{(0.75 \times 0.3) + \dots + (0.89 \times 0.9)}{0.3 + 0.5 + 0.7 + 0.9} = 0.71 \quad (5.4)$$

Ce niveau de stress influence les comportements automatiques ou non intentionnelles d'un avatar comme ses expressions gestuelles, ses animations (déplacements, postures, etc.) et ses hochements de tête.

Pour les expressions gestuelles, le modèle d'interaction sociale module la fréquence d'animation des gestes produits en fonction du niveau de stress estimé. .

Ainsi, plus un avatar est stressé et plus il aura tendance à produire des expressions gestuelles rapides lors de son discours. Inversement, plus un avatar est calme et plus ses expressions gestuelles seront animées à une vitesse modérée.

Cette relation d'influence entre la fréquence des expressions gestuelles et le niveau de stress est exprimée dans notre modèle par :

$$Freq\{Geste\} = \psi \times stress \quad (5.5)$$

avec :

- $Freq\{Geste\}$  : La fréquence d'animation du geste.
- $\psi$  : facteur utilisé pour adapter l'animation selon le taux d'affichage des frames dans le système de réalité virtuelle.

Ce couplage permanent entre la production d'expressions corporelles, le stress et la personnalité de l'avatar permet de penser que les interactions virtuelles engagées peuvent être 'situées'.

### 5.3.3 Postures et animations

La posture est une notion essentielle dans l'interprétation des comportements humains.

Comme les travaux de (Mehrabian, 1972) l'ont soulignés, la posture d'une personne renseigne sur son état psychologique.

Dans nos travaux, notre modèle d'interaction sociale module de façon automatique les postures et animations d'un avatar selon son niveau de stress calculé (lequel est évalué en permanence en fonction du contexte de la situation, de la personnalité de l'utilisateur et de ses états émotionnels.).

Par exemple, un avatar qui marche dans une direction dangereuse (source de flammes dans l'univers par exemple) peut voir son niveau de stress augmenter (en fonction de la personnalité, du statut social, etc.). Immédiatement, notre modèle d'interaction va modifier sa démarche pour la rendre compatible avec celle d'une personne stressée.

Le tableau 5.1 donne un aperçu de quelques variations dynamiquement apportées par le modèle d'interaction sociale à une animation de marche selon le niveau de stress estimé de l'avatar.

Description de l'animation de marche	Fréquence de l'animation
Démarche tranquille	$\psi_1 \times stress$
Démarche tranquille mais accélérée	$\psi_2 \times stress$
Démarche stressée	$\psi_1 \times stress$
Démarche très agitée	$\psi_2 \times stress$

TAB. 5.1 – Exemple d'une adaptation permanente de la démarche d'un avatar selon son niveau de stress. (Avec :  $\psi_2 > \psi_1$ )

L'adaptation permanente du comportement des avatars à leurs états émotionnels ainsi qu'à celui des autres acteurs en interaction nous permet de simuler des effets de propagation d'états émotionnels ainsi que l'émergence de conduites collectives.

C'est ce couplage réflexif permanent entre les acteurs de l'univers virtuel qui, comme l'a montré l'approche ethnométhodologique, peut assurer la consistance sociale du dispositif (présence sociale).

Ce couplage n'est évidemment que partiel par rapport à la réalité puisque les interactions sociales sont en partie sous le contrôle de notre modèle. L'objectif de notre expérimentation (chapitre 7) sera d'évaluer dans quelle mesure les acteurs seront capables d'intégrer de façon réflex la dimension émotionnelle des acteurs avec lesquels ils interagissent.

### 5.3.4 Expression faciale

Le visage est considéré comme l'espace le plus expressif de tout le corps humain. D'après Knapp (Knapp, 1978), il est le premier outil de communication des émotions lors des interactions humaines et d'après Mehrabian (Mehrabian, 1967), les expressions faciales fournissent 55% du sens lors des échanges inter-humains.

Nous avons, dans notre système de réalité virtuelle, un modèle qui calcule et maintient ces états émotionnels en fonction de l'interaction de l'utilisateur dans le dispositif, de sa personnalité déclarée et de son humeur estimée (El Jed et al., 2004).

Ce modèle permet de générer de façon automatique des expressions faciales en adaptant de façon permanente des textures optimisées pour exprimer les différents états émotionnels au visage des personnages virtuels (figure 5.5).



FIG. 5.5 – Exemples d'expressions émotionnelles chez un avatar : une émotion de déception et une émotion de satisfaction.

Le modèle émotionnel permet également d'exprimer des mouvements labiaux en accord avec la vitesse d'élocution ainsi que des mouvements de sourcils en fonction de la hauteur de la voix ('pitch') dans le discours (Chovil, 1992).

Ces expressions faciales participent à l'expression émotionnelle globale du personnage (figures 5.6 et 5.5).

Le module des expressions labiales ne permet à ce jour une parfaite synchronisation labiale (une parfaite co-articulation avec le discours) mais assure seulement des mouvements labiaux coordonnés par la cadence de l'énoncé.



FIG. 5.6 – Exemples d’expressions labiales lors d’une élocution.

### 5.3.5 Hochement de tête

Le hochement de tête peut signifier (en fonction des cultures) l’affirmation, la compréhension ou l’approbation.

Les hochements de tête sont des mouvements automatiques qui sont produits chez nos avatars essentiellement dans un contexte de dialogue.

Par exemple, lorsqu’un avatar joue un rôle d’auditeur dans une conversation à plusieurs, le modèle d’interaction sociale déclenche automatiquement des hochements de tête pour rajouter des éléments contextuels supplémentaires à l’interaction nous permettant d’espérer maintenir une forme de couplage réflexif entre l’auditeur et son interlocuteur.

Ce comportement automatique, lorsqu’il est perçu par l’interlocuteur permet à ce dernier de se savoir écouté et renforce ainsi son niveau d’immersion sociale.

Aucun modèle dans la littérature ne permet à ce jour de définir la relation entre les hochements de tête et les caractéristiques du contexte de la situation comme la personnalité d’un individu, son degré de stress, le statut social du locuteur, etc.

Notre faible connaissance de ces mécanismes des hochements de tête nous pousse à supposer l’existence d’une relation entre les propriétés des hochements de tête comme la vitesse du mouvement, son amplitude, son rythme, etc. et certaines propriétés du contexte de la situation comme le sens d’une discussion, le rôle social de l’interlocuteur, le niveau de stress de l’individu, etc.

Pour modéliser cette relation, nous supposons que la vitesse d’animation des hochements de tête chez un avatar dépend de son degré de stress et que

le rythme de ces hochements est fonction de sa personnalité déclarée.

$$Freq\{Hochement\} = \psi \times stress \quad (5.6)$$

$$Rythme\{Hochement\} = \varphi.P \quad (5.7)$$

avec :

- $Freq\{Hochement\}$  : La fréquence d’animation du mouvement de hochement de tête.
- $\varphi.P$  : Un rythme d’animation du hochement de tête qui diffère d’un avatar à un autre selon la personnalité déclarée.

Cette modélisation, bien qu’elle soit réductrice par rapport à la complexité des mécanismes de production du hochement, nous permet d’implémenter à minima certains couplages importants lors des dialogues tels que l’approbation du discours par l’auditeur ou encore la signification au locuteur de son écoute.

## 5.4 Conclusion

Dans ce chapitre, nous avons décrit un modèle pour l’interaction sociale en univers virtuel. Ce modèle permet de coupler les comportements non intentionnels exprimés par les avatars en accord avec les actions intentionnelles décidées par les utilisateurs qui les contrôlent et du contexte de leurs situations.

Notre objectif est d’essayer de reproduire en univers virtuel des interactions sociales, émotionnelles et culturelles proches de la réalité. Notre intention n’est pas de valider le modèle émotionnel et social mais de tester si le résultat de l’interaction obtenu produit chez les avatars des comportements suffisamment crédibles pour engager l’utilisateur dans une immersion sociale et culturelle à travers des modalités verbales et non verbales.

Cette modélisation émotionnelle et sociale en univers virtuel, du fait de son caractère analytique, risque évidemment d’induire une ‘mauvaise’ réflexivité dans l’interaction voire des incompréhensions entre les utilisateurs.

Néanmoins, nous ferons l’hypothèse que les utilisateurs peuvent en partie ”s’approprier” les comportements autonomes de leurs avatars et rétablir éventuellement le sens commun de la situation pour produire des interactions cohérentes et significatives.



**Troisième partie**

**Plate-forme de formation**



## Chapitre 6

# Architecture de la plate-forme

*Deviens ce que tu es.  
Fais ce que toi seul peut faire  
(Friedrich Nietzsche)*

Nous décrivons dans ce chapitre quelques aspects techniques de notre plate-forme de réalité virtuelle. Cette plate-forme s'appuie sur la technologie des systèmes multi-agents pour représenter les interactions entre les agents.

Dans un premier temps, nous décrivons à la section 6.1 l'architecture informatique de nos agents. Puis, nous présentons à la section 6.2 un aspect fonctionnel de cette architecture pour l'identification du contexte de l'interaction. A la section 6.3, nous donnons quelques exemples d'interactions sociales possibles dans notre environnement.

Finalement, nous présentons à la section 6.4 quelques aspects du développement de notre univers virtuel. Ce développement s'appuie sur des techniques d'animations, de modélisation et de rendu 3D que nous utilisons pour créer un univers suffisamment 'crédible' pour produire des interactions situées.

## 6.1 Architecture d'un agent

Selon la technologie multi-agents, plusieurs entités (appelées agents) réalisent chacune une tâche spécifique, interagissent et communiquent entre elles pour assurer la cohérence, la complétude et la correction d'une activité globale.

Notre environnement de réalité virtuelle représente donc un système multi-agents où chaque avatar en interaction est un agent poursuivant un but local : compléter les actions intentionnelles décidées par l'utilisateur par un comportement social et émotionnel adapté au contexte de la situation afin d'améliorer sa crédibilité et engager davantage l'utilisateur dans l'interaction.

Il doit pour cela percevoir son contexte de l'interaction, coordonner ses actions et communiquer avec les autres agents de l'univers. Son architecture doit permettre d'intégrer de manière cohérente et efficace un ensemble de capacités comme la perception, le processus décisionnel et finalement la production d'action qu'elle soit intentionnelle ou automatique.

L'architecture générale de chaque agent est donnée par la figure 6.1 :

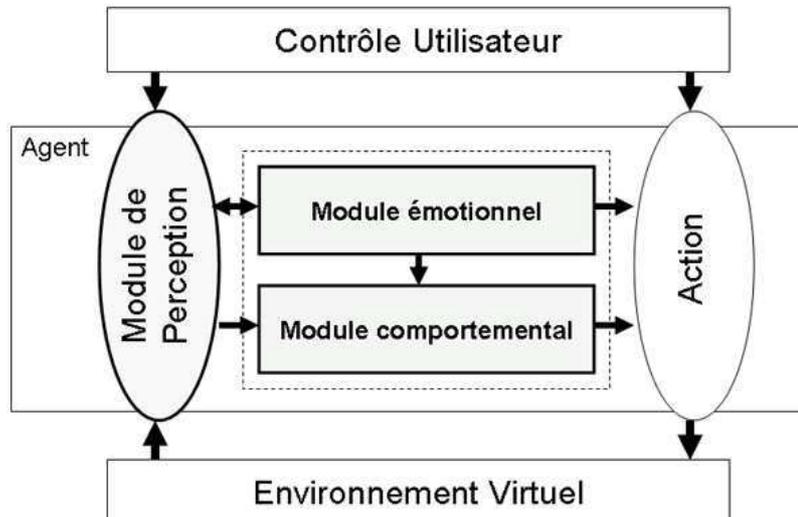


FIG. 6.1 – Architecture générale d'un agent représentant un avatar.

Cette architecture est composée de :

- **Un module pour la perception et l'identification du contexte d'interaction** qui permet à chaque agent de percevoir l'état du monde qui l'entoure et d'identifier le contexte de l'interaction à laquelle est situé l'utilisateur (ex : en dialogue, action, etc.).

- **Un module pour la perception des émotions.** Il maintient à jour les états émotionnels d'un agent en fonction de ses perceptions depuis l'environnement, ses anciens états émotionnels, sa personnalité et son humeur. Ces états émotionnels influencent le comportement de l'agent ainsi que sa perception du contexte de sa situation.
- **Un module pour la production d'un comportement** à la fois social et émotionnel en fonction des éléments contextuels perçus. Il définit la réaction de l'agent en fonction de ses perceptions et de l'action courante décidée par l'utilisateur.

Cette architecture est conçue de façon modulaire c'est à dire qu'elle favorise l'amélioration, l'ajout et la réutilisation des modules.

Nous allons dans ce qui suit nous intéresser aux principaux modules de cette architecture.

### 6.1.1 Le module de perception

La capacité de perception d'un agent lui permet d'acquérir des informations sur son environnement et sur lui-même. Dans le cadre des agents de notre dispositif de réalité virtuelle, ce module de perception assure à l'agent de détecter les objets à proximité et de reconnaître les agents (affiliation, rôle social, etc.) avec qui il entre en interaction. Le système perceptif d'un agent est continuellement excité par des stimuli provenant de l'environnement d'interaction et du contrôle de l'utilisateur.

La conception de l'architecture du module de perception s'effectue selon deux approches : une approche fonctionnelle dite également classique et une approche comportementale dite 'orientée comportement'.

Dans une approche classique, la perception de l'environnement est souvent vue comme un processus opérant une fusion des informations provenant de plusieurs systèmes perceptifs afin d'élaborer et mettre à jour une représentation symbolique de l'environnement (figure 6.2). Ce type d'architecture est aujourd'hui relativement désuet à cause de la complexité de modélisation des environnements (essentiellement ouverts et dynamiques) et des limites atteints par l'approche symbolique de l'Intelligence Artificielle dite classique.

L'architecture du module de perception est donc conçue selon une approche 'orientée comportement' (figure 6.3).

Cette approche trouve ses origines dans des théories sur l'intelligence (Minsky, 1988) et la robotique de la fin des années 80 (Brooks, 1986; Maes, 1989; Arkin, 1989). Elle propose la notion de perception modulaire ou perception 'orientée action'. Le système perceptif est ainsi composé de différents sous-systèmes perceptifs spécialisés, qui extraient l'information pertinente pour déclencher un

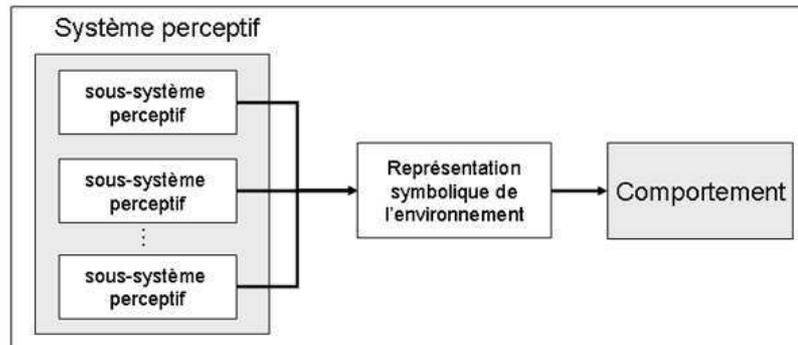


FIG. 6.2 – Approche classique pour la perception.

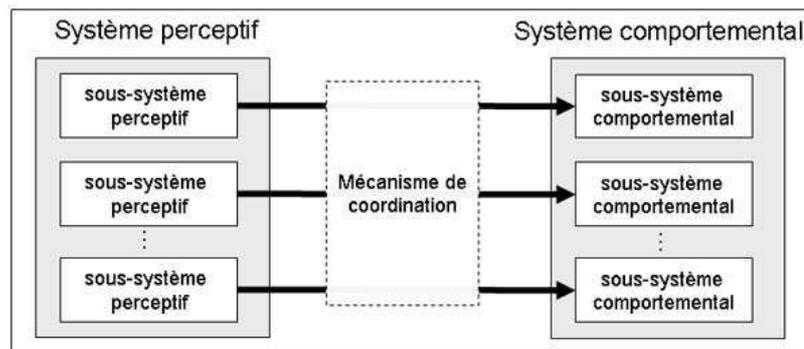


FIG. 6.3 – Approche 'orientée-comportement' pour la perception.

comportement spécifique.

L'agent ne dispose ainsi d'aucune représentation globale sur son environnement mais identifie juste les informations nécessaires à l'activation de comportements spécifiques. Le comportement global de l'agent résulte de l'interaction des comportements qui le composent.

Cette architecture permet de prendre en compte le fait qu'un avatar peut répondre à différents stimuli et mener plusieurs actions en parallèle (comme par exemple tourner la tête en continuant à marcher ou encore distribuer le regard, effectuer des expressions gestuelles, des mouvements labiaux, etc. lors d'un dialogue).

Cette approche de conception du système perceptif a l'avantage d'être réactive et beaucoup moins coûteuse en terme de traitement d'informations qu'une approche classique qui opère une fusion des informations pour élaborer

et mettre à jour une représentation unique de l'environnement.

Elle est de plus très modulaire et évolutive c'est à dire que l'ajout ou la suppression de nouveaux modules perceptifs ne perturbe en rien la fonction globale du système perceptif de l'agent.

### 6.1.2 Mécanisme de coordination

Le processus de décision d'un agent est architecturé autour de modules comportementaux s'exécutant en parallèle.

Le problème inhérent à ce type d'architecture est que les comportements souhaités par les différents modules de perception (ou sous-systèmes perceptifs) peuvent être conflictuels. Il est alors nécessaire d'ajouter un système de coordination capable de sélectionner le 'bon' comportement à accomplir.

Deux types de mécanismes de coordination d'actions sont utilisés (Pirjadian, 1999) : l'arbitrage et la fusion de commandes.

La fusion des commandes est également appelée la sélection coopérative. Elle consiste à choisir une action à accomplir en prenant en compte les propositions d'actions de plusieurs systèmes comportementaux. L'action ainsi choisie représente un compromis entre les différentes recommandations d'actions.

Cette stratégie est particulièrement adaptée aux systèmes à base de champs de potentiels. Par exemple, elle a été utilisée par (Arkin, 1998) pour la navigation de robots autonomes. Chaque système perceptif (*schème moteur*) propose un vecteur d'action qui correspond à la direction que doit prendre le robot. La direction finale prise par le robot correspond à un compromis entre les différentes directions proposées.

Un mécanisme de coordination basé sur une solution d'arbitrage (appelé également solution compétitive) sélectionne un comportement jugé pertinent parmi plusieurs propositions de comportements.

Ce principe est utilisé par l'architecture de subsomption (Brooks, 1986) qui sélectionne le comportement actif de plus haute priorité ou encore l'architecture de Maes (Maes, 1989) qui sélectionne le comportement jugé le plus pertinent en fonction du contexte.

Dans notre cadre d'application, une solution de coordination basée sur l'arbitrage apparaît comme plus appropriée puisqu'un avatar doit adopter un comportement unique en fonction de son contexte d'interaction (par exemple jouer un rôle d'auditeur, de locuteur, etc.).

Dans ses travaux de thèse (Tyrrell, 1993), Tyrrell a réalisé une comparaison des mécanismes de sélection de l'action pour différentes architectures basées sur les comportements. Il propose qu'une bonne architecture de sélection d'action doit posséder un certain nombre de caractéristiques dont les principales sont :

- **Motivé** : le système comportemental sélectionné doit correspondre à celui de plus forte motivation en fonction des objectifs.
- **Interruptible** : le système comportemental courant doit pouvoir être interrompu au profit d'un autre.
- **Persistant** : le système comportemental en cours doit pouvoir continuer à s'exécuter même si d'autres systèmes comportementaux peuvent être éligibles (on évite ainsi l'oscillation entre plusieurs comportements).
- **Opportuniste** : un système comportemental de moins forte priorité peut être sélectionné le temps de satisfaire un objectif secondaire.
- **Direct** : le système doit préférer les actions qui mènent directement à la satisfaction des objectifs par rapports aux autres.
- **Conciliant** : les actions sélectionnées doivent correspondre à un compromis entre plusieurs solutions au lieu de satisfaire un seul objectif de façon optimale.

Dans notre système, un arbitre est représenté par un module spécial 'Gestionnaire de Contexte' qui est chargé de réceptionner les propositions d'actions en provenance des différents modules de perception puis de choisir le comportement le plus approprié à activer parmi les modules comportementaux (figure 6.1).

Cette sélection effective du comportement à adopter se fait en associant des priorités aux différents modules comportementaux. Selon les caractéristiques de (Tyrrell, 1993), notre module de sélection des actions apparaît comme :

- **Motivé** : le 'Gestionnaire de Contexte' active le comportement le plus prioritaire par rapport à la situation. A titre d'exemple, nous fixons l'ordre suivant pour les priorités des comportements à adopter dans une situation de dialogue (tableau 6.1). Ces priorités permettent de respecter les règles du dialogue coopératif.
- **Interruptible** : chaque système comportementale peut être interrompu dans un délai 'raisonnable' par rapport au temps de l'animation. Nous avons donc dimensionné les temps d'exécution des comportements globaux de façon à proposer un temps de déclenchement acceptable.

Par exemple, un agent qui passe du statut d'auditeur à celui de locuteur, doit interrompre dans un délai raisonnable son comportement d'auditeur (regarder le locuteur, hocher la tête, etc.) afin de déclencher dans un temps acceptable un comportement de locution (distribuer le regard alternativement entre les différents auditeurs, effectuer des expressions gestuelles au cours de sa parole, etc.).

- **Persistant** : un système comportemental en cours continue à s'exécuter même si un comportement de même priorité est éligible. Par exemple, un agent qui parle continue à adopter un comportement de locuteur même si un autre événement lié à sa tâche de même priorité apparaît. Nous évitons ainsi une oscillation du système comportemental entre plusieurs

actions possibles.

- **Conciliant** : lorsque deux actions sont possibles, notre système de sélection d'action opère de façon probabiliste en sélectionnant au hasard un comportement à adopter. Cette stratégie d'action permet d'assurer un équilibre entre les différents comportements éligibles.

Priorité	Situation
1	Rencontre
2	Auditeur
3	Récepteur de déictique
4	Locuteur
5	Émetteur de déictique

TAB. 6.1 – Exemple d'ordre de priorités attribuées à certains comportements : (1) faible priorité, (5) forte priorité.

Par exemple, lorsqu'un agent A détecte une situation de dialogue (un autre agent B est entrain de lui parler), son module gestionnaire de contexte va tenter de lui attribuer le comportement d'auditeur (jugé approprié à la situation).

Dans le cas où l'agent A est déjà entrain de parler (c'est-à-dire que son état est identifié à celui d'un 'locuteur'), l'arbitre doit décider de l'état à activer vu que la situation est conflictuelle (l'agent est à la fois locuteur et auditeur dans ce dialogue). Il va vérifier l'ordre des priorités associés aux comportements de locuteur (4) et d'auditeur (2).

Étant donné que l'état de 'locuteur' est plus prioritaire, l'agent va continuer à adopter un comportement de locuteur même lorsqu'un autre agent est entrain de lui parler.

Ce système d'arbitrage nous permet également d'attribuer, en cas d'ambiguïté, une même priorité pour deux ou plusieurs comportements. Dans le cas où plusieurs sous-systèmes perceptifs tentent d'activer simultanément deux ou plusieurs sous-systèmes comportementaux, l'arbitre sélectionne alors aléatoirement un comportement unique à adopter. La réaction de l'avatar dans ce cas est imprévisible.

### 6.1.3 Les modules comportementaux

Le rôle du module comportemental est de contrôler le comportement d'un agent en sélectionnant les actions adéquates en fonction de ses perceptions (contexte de la situation) et des émotions qu'il ressent.

L'architecture que nous proposons respecte le principe énoncé par Brooks, Steels et Arkin (Brooks, 1999; Steels, 1994; Arkin, 1998) sur les architectures 'orientées-comportements' : le comportement global d'un agent doit émerger de

l'interaction entre des systèmes comportementaux concurrents et coordonnés, qui correspondent à des cycles perception-décision-action indépendants et qui sont chargés de réaliser les grandes fonctionnalités de l'agent.

Les systèmes comportementaux sont des modules capables d'assurer une fonctionnalité particulière pour l'agent. Par exemple, pour gérer l'animation de locuteur, un premier système comportemental se charge d'orienter la tête de l'agent vers les différents interlocuteurs, un second active les mouvements labiaux de l'agent en fonction du discours de l'utilisateur qui le contrôle et un dernier système comportemental se charge de produire la gestualité correspondante au discours.

Le modèle comportemental que nous décrivons permet de définir un comportement comme un ensemble de systèmes comportementaux s'exécutant en parallèle ou séquentiellement (figure 6.4).

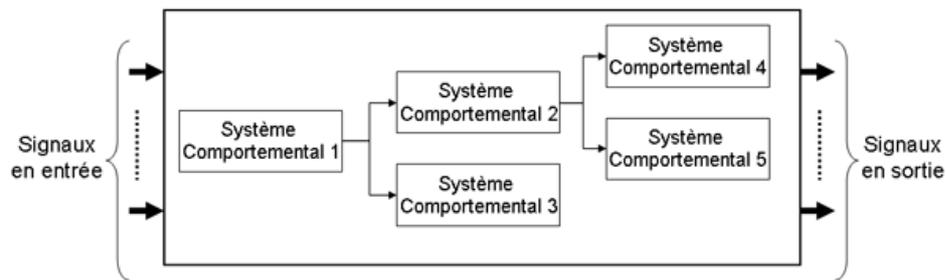


FIG. 6.4 – Exemple d'une architecture d'un module comportemental.

L'avantage de cette architecture est sa modularité et son évolutivité. Il est aisé de faire évoluer les comportements des agents en apportant des améliorations à leurs systèmes comportementaux sans perturber le comportement global de l'agent.

En faisant, par exemple évoluer les mécanismes comportementaux permettant d'activer les mouvements labiaux d'un agent, on peut améliorer la synchronisation labiale sans perturber le comportement global du locuteur.

Le lancement d'un comportement de l'agent déclenche une cascade de réactions des systèmes comportementaux. Ce déclenchement se fait selon une stratégie d'activation dès que l'ensemble des actions de l'instant courant a été réalisé. C'est la stratégie qui permet de réagir continuellement et le plus rapidement à l'environnement. Dans ce modèle de comportement, nous supposons que l'exécution d'un comportement prend un temps négligeable par rapport à nos besoins.

#### 6.1.4 Adressage et communication

Les communications, dans les systèmes multi-agents comme chez les humains, sont à la base des interactions et de l'organisation.

Une communication peut être définie comme une forme d'action locale d'un agent vers d'autres agents.

Dans notre dispositif de réalité virtuelle, les agents virtuels communiquent afin d'échanger des informations sur leurs situations courantes dans le but de se coordonner pour paraître crédibles et cohérents avec leurs contextes d'interactions.

Lorsque l'utilisateur qui contrôle son avatar décide de prendre la parole au sein d'un groupe, l'agent (représentant l'utilisateur) manifeste cette intention de communication verbale en informant les autres agents à proximité et en communiquant son état à son module 'Gestionnaire de Contexte' afin d'adopter un comportement adapté à ce nouveau contexte (un comportement de locuteur en l'occurrence).

Les agents recevant un message de notification vont à leurs tours activer leurs 'Gestionnaires de Contexte' afin d'adopter des comportements d'auditeurs. Les communications ainsi engagées se font suite aux commandes de l'utilisateur ou à une modification du contexte de l'interaction.

Selon le contexte de l'interaction, un agent peut adresser un message à l'ensemble des agents à proximité (le groupe des agents en interaction) ou bien envoyer un message unique vers un seul agent (notification d'un clone à travers le réseau par exemple).

De nombreux travaux ont porté sur la description de langages de communications entre agents (Barbuceanu and Lo, 1999; Cassell et al., 1998; Chicoisne and Pesty, 2000). Vu la simplicité du protocole de communication implémenté dans notre plate-forme de réalité virtuelle, nous avons opté pour un langage ad hoc pour assurer les communications entre les agents.

La figure 6.5 illustre le processus de communication entre trois agents dans un même environnement virtuel réparti sur deux postes (1 et 2). Au départ, nous supposons que le 'Gestionnaire de Contexte' de chaque agent indique un comportement par défaut (aucun comportement particulier n'est activé).

Suite à une commande de l'utilisateur sur le poste 1 (parler par exemple), l'agent A qui vient de percevoir cette commande informe tous les autres agents A distribués sur les autres postes (en particulier l'agent A du poste 2 dans notre exemple).

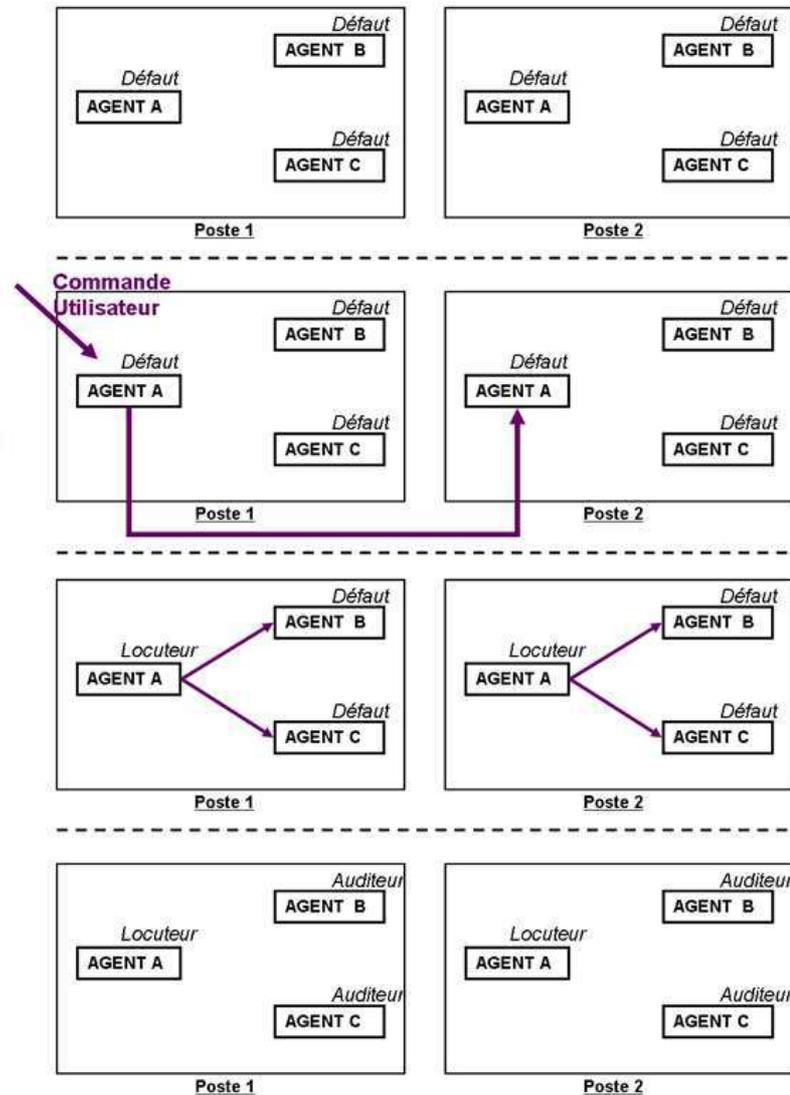


FIG. 6.5 – Processus de communication entre agents.

Tous les agents A (sur tous les postes) activent alors un comportement adapté à ce nouveau contexte (comportement de locuteur) et informent les agents dans leurs espaces interpersonnels (agents B et C) qu'ils sont entrain de produire un discours.

A la réception du message provenant de l'agent A, chaque agent B et C va mettre à jour son nouveau contexte d'interaction et activer un comportement adapté à la situation (comportement d'auditeur).

### 6.1.5 Architecture d'intégration

L'idée principale de notre architecture est que le comportement intelligent de l'agent émerge des interactions entre ses comportements plus simples.

Il s'agit d'une architecture hybride basée essentiellement sur des comportements réactifs où chaque comportement est régi par un cycle de perception/décision/action dans lequel l'agent effectue directement une action en fonction de sa perception sans réaliser beaucoup de réflexion.

La figure 6.6 présente un exemple d'une organisation logique des différents modules pour un agent.

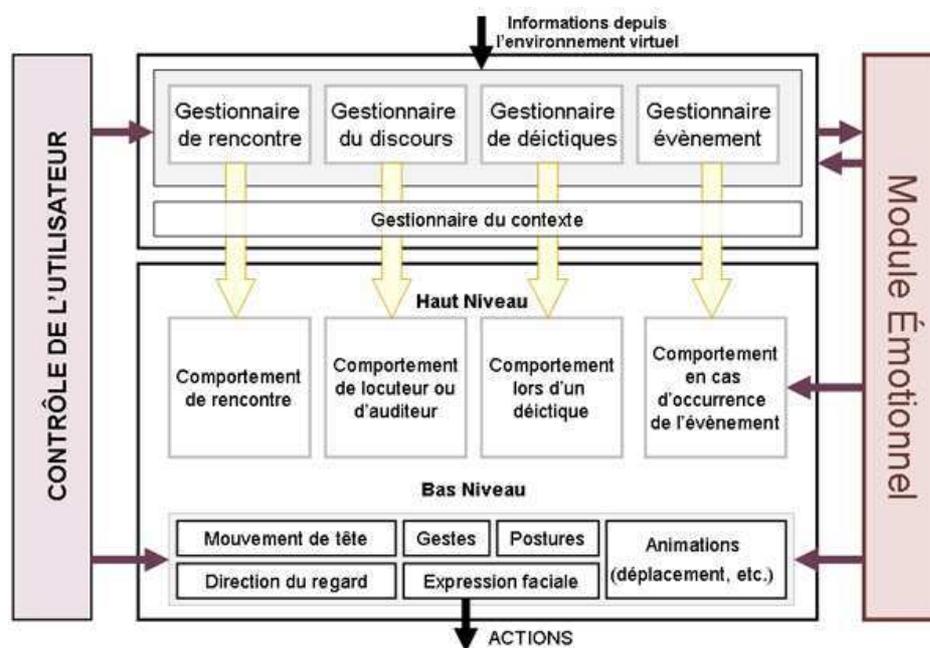


FIG. 6.6 – Exemple d'une architecture d'intégration pour un agent.

Le module de perception assure à l'agent l'identification du contexte de la situation. Il définit plusieurs sous-systèmes perceptifs qui se chargent chacun de la perception d'une situation particulière comme celle du dialogue, d'une rencontre, ou d'un évènement particulier selon le scénario.

Le gestionnaire du contexte permet d'activer un comportement unique jugé le plus approprié parmi les propositions de chaque sous-système perceptif. Il se charge de désactiver le comportement actif en cours et d'activer le nouveau comportement à adopter.

Le module émotionnel implante directement notre modèle émotionnel décrit au chapitre 4. Il fournit les états émotionnels simulés aux différents modules comportementaux et influence le module de perception. Il permet, par exemple, de réduire le champ visuel de l'utilisateur lorsque le niveau de stress est jugé important ou encore de réduire la distance minimale permettant à un agent de percevoir les autres agents.

Le module comportemental organise les comportements en deux catégories logiques :

- Les comportements de haut niveau qui spécifient les règles qui décrivent le comportement à adopter.
- Les comportements de bas niveau qui fournissent les comportements de base qui seront exécutés soit par les comportements de haut niveau lors de l'activation d'un comportement soit directement par l'utilisateur.

Un système de priorité permet ensuite d'attribuer une plus haute priorité aux commandes de l'utilisateur (pour l'accès aux comportements de bas niveau) qu'aux comportements de hauts niveaux favorisant ainsi le contrôle intentionnel de l'avatar au dépend des comportements automatiques produits par le modèle d'interaction sociale.

## 6.2 Aspect fonctionnel : le contexte de l'interaction

Dans cette section, nous décrivons notre approche pour identifier le contexte de la situation d'un avatar.

D'abord, il nous paraît important de distinguer dans ce qui suit la notion du contexte de l'utilisateur de celle de l'avatar.

### 6.2.1 Le contexte de l'avatar

Le contexte de l'utilisateur représente l'ensemble des éléments de la situation qui influencent le processus décisionnel de l'utilisateur et permettent à ce dernier de modifier l'environnement virtuel. Ces éléments peuvent être implicites (croyances, suppositions, pratiques, etc.) ou explicites (nature des objets, localisations, états émotionnels, etc.).

Le contexte de l'avatar représente l'ensemble des éléments accessibles dans l'environnement virtuel qu'un module informatique peut exploiter pour produire des comportements automatiques complémentaires aux actions intentionnelles décidées par l'utilisateur.

Malheureusement, ces deux types de contextes sont encore loin d'être complètement identifiés. Certains éléments contextuels (comme la reconnaissance d'intention par exemple) relèvent de mécanismes cognitifs dont l'humain

à jusqu'à aujourd'hui du mal à comprendre le fonctionnement d'où une difficulté de modélisation par un outil informatique.

Afin de limiter cette complexité de représentation du contexte, nous allons nous intéresser, pour représenter le contexte de l'avatar, uniquement aux éléments contextuels explicites (observables et accessibles) de l'environnement virtuel (figure 6.7).

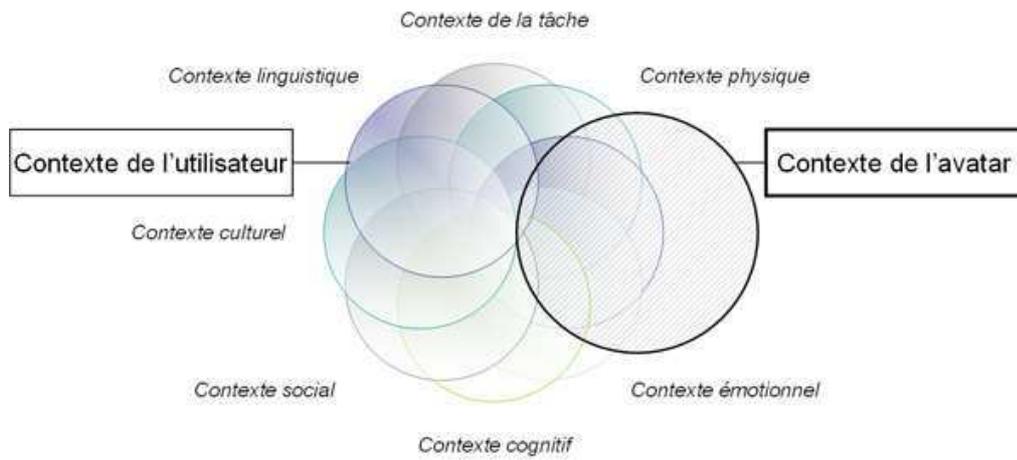


FIG. 6.7 – Contexte de l'avatar versus contexte de l'utilisateur.

Malgré le caractère réducteur (en terme d'éléments contextuels) de cette solution d'identification du contexte de l'avatar, nous pensons qu'elle reste assez riche pour identifier les éléments nécessaires à l'avatar pour produire des expressions corporelles (gestualité, expressions faciales, etc.) adaptées à la situation.

### 6.2.2 Approche de modélisation

Afin de produire des expressions corporelles adaptées à la situation d'interaction, notre modèle doit identifier le contexte de l'avatar (qui représente au mieux un sous ensemble du contexte réel de l'utilisateur).

Les éléments contextuels permettant de constituer ce contexte changent d'une situation à une autre. Pour produire le comportement le plus adéquat à un contexte donné, les utilisateurs considèrent des éléments contextuels toujours différents en fonction de leurs situations.

Par exemple, les éléments contextuels considérés lors d'une prise de décision dans une situation de dialogue entre un groupe d'individus sont différents de ceux considérés lors d'une prise en charge d'une victime dans une situation de secours.

De ce fait, la construction du contexte par notre modèle est étroitement dépendante du type de la situation d'interaction (que nous appelons dans notre modèle le contexte de la tâche).

Le contexte de l'avatar est donc constitué des différents éléments contextuels appartenant aux différents types de contextes identifiés à la section 3.2.2 et choisis en fonction de leur adéquation par rapport au contexte de la tâche.

Dans notre architecture d'agent, chaque sous-système perceptif se charge d'identifier un contexte particulier de la tâche (un type de situation prédéfinie) qui peut apparaître lors du scénario. Ensuite, il se charge de fournir au sous-système comportemental associé l'ensemble des éléments contextuels (appropriés à ce contexte de la tâche) pour être exploités dans la production d'un comportement approprié.

Par exemple, le contexte d'un avatar dans une situation de dialogue avec trois autres interlocuteurs peut être caractérisé par cet ensemble d'éléments (Table 6.2) :

Contexte de la tâche	Situation de dialogue
Contexte physique	Trois auditeurs : A, B et C. A.nom, A.position, A.orientation, etc. B.nom, B.position, B.orientation, etc. C.nom, C.position, C.orientation, etc.
Contexte social	A.statut-social B.statut-social C.statut-social
Contexte émotionnel	A.émotions, A.stress, etc. B.émotions, B.stress, etc. C.émotions, C.stress, etc.
Contexte culturel	Activation (si besoin) d'un ensemble de règles comportementales adaptées au contexte culturel du dialogue.

TAB. 6.2 – Exemple d'éléments contextuels dans une situation de dialogue pour un avatar particulier.

D'une façon générale, le processus d'identification du contexte d'interaction au niveau de chaque sous-système perceptif est donné par la figure 6.8.

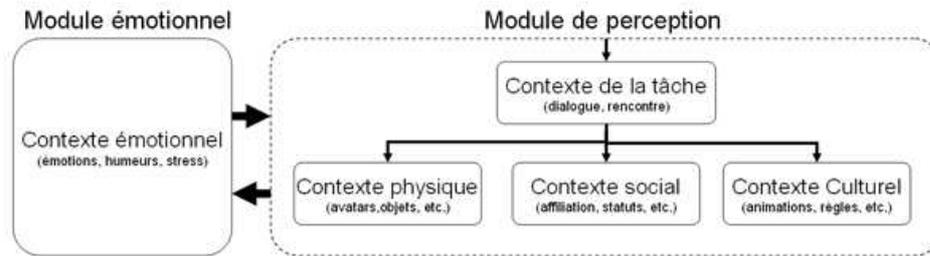


FIG. 6.8 – Processus de perception du contexte chez un agent.

### - Prise en compte du contexte de la tâche

Le contexte de la tâche regroupe l'ensemble des types de situations d'interaction dans l'univers. La prise en compte de toutes ces situations dans un environnement imprévisible et dynamique est une tâche complexe.

Pour simplifier, nous définissons un ensemble de situations d'interactions types (situations de dialogue, évènements spéciaux, objectifs atteints, etc.) que nous chercherons à identifier lors de la simulation.

En particulier, nous nous focalisons sur les types de situations qui permettent aux utilisateurs de s'exprimer socialement dans un univers virtuel à travers des modalités verbales (dialogue à travers un réseau de communication, etc.) et non verbale (déictiques pour désigner des objets, direction du regard, etc.).

Parmi ces types de situations, nous considérons particulièrement :

- **Situation de dialogue** : Une situation de dialogue est caractérisée par la forme de la communication (verbale ou non verbale) et le rôle joué par chaque acteur dans cette interaction. Le module de perception du contexte permet de caractériser chaque situation de dialogue en identifiant l'utilisateur qui initie une forme de communication (verbale ou non verbale) et associe à chaque avatar un rôle dans ce dialogue (auditeur, locuteur, etc.). Ce mécanisme nous permet de détecter et d'identifier à tout instant l'occurrence de situations de dialogue entre deux ou plusieurs utilisateurs.
- **Situation de rencontre** : Une situation de rencontre est caractérisée par la proximité dans l'univers virtuel d'un ou plusieurs avatars. Le module de perception du contexte permet de détecter cette proximité dans l'entourage d'un avatar et de signaler qu'il y a rencontre entre deux ou plusieurs avatars (lorsque cette proximité est détectée pour la première fois). Cette situation est identifiée ensuite comme dialogue dès lors qu'un utilisateur émet un acte de communication.

Ces situations permettent aux utilisateurs de réagir selon leurs propres

états émotionnels, expertises et historiques. Ils permettent ainsi aux utilisateurs d'être cognitivement et culturellement situés dans les interactions virtuelles engagées.

De ce point de vue, ils représentent un cadre d'étude intéressant pour l'expression de comportements émotionnels et sociaux.

### **- Prise en compte du contexte physique de la situation**

Il s'agit de la capacité d'un avatar à reconnaître, pour chaque type de situation rencontrée, les éléments contextuels environnementaux associés (les objets à proximité, leurs types, compositions, localisations, etc.). Connaissant l'identité d'un élément de l'univers virtuel (comme un avatar par exemple), le sous-système perceptif peut retrouver différentes informations associées (comme son nom, position, activité, etc.).

Dans une situation de dialogue, comme pour une situation de rencontre, un avatar peut identifier ses interlocuteurs, leurs positions et orientations afin d'adopter un comportement adapté.

### **- Prise en compte de l'environnement social**

La connaissance des affiliations sociales des différents avatars dans un contexte d'interaction permet de faire 'émerger' des comportements sociaux adaptés au contexte de la situation et en accord avec les identités des acteurs en co-présence.

Par exemple un avatar qui rencontre un autre avatar de même statut social tournera la tête dans sa direction pour le regarder quelques instants. Lorsque l'identité sociale de l'avatar rencontré est identifiée comme supérieur hiérarchique dans l'organisation des acteurs, un geste de salutation peut émerger.

### **- Prise en compte du contexte émotionnel**

Le contexte émotionnel de l'avatar (fournit par le module émotionnel) influence la production de comportements adaptés à la situation.

Par exemple, lors d'une situation de rencontre, un avatar dont le niveau de stress dépasse un certain seuil (prédéfini dans notre modèle), ne va pas regarder l'avatar rencontré (supprimant ainsi une réponse sociale naturelle).

Un avatar peut également percevoir les émotions des autres avatars et changer en conséquence son comportement par contagion des autres émotions. Par exemple, il peut devenir à son tour stressé lorsqu'il est en interaction avec d'autres avatars stressés.

### - Prise en compte du contexte culturel

Le sous-module du contexte culturel permet aux avatars d'exprimer des expressions corporelles adaptées aux règles d'interactions implicites partagées par le groupe d'appartenance.

Par exemple, certaines expressions gestuelles spécifiques sont partagées par les membres d'un collectif comme les sapeurs pompiers pour indiquer la fin de déroulement d'une intervention ou pour communiquer dans des environnements bruyants.

### 6.2.3 Le conflit d'intentionnalité entre l'utilisateur et son avatar

Nous adressons dans cette section le problème inhérent à ce type d'architecture qui est le conflit permanent entre les actions décidées par l'utilisateur et les actions automatiques proposées par le modèle d'interaction sociale pour s'adapter au nouveau contexte d'interaction.

Ce problème est résolu au niveau de notre architecture par :

- Un système d'arbitrage que nous définissons au niveau de la production des actions. Ce système considère les actions intentionnelles décidées par l'utilisateur comme prioritaires par rapport aux actions automatiques proposées par le modèle d'interaction sociale. De ce fait, nous permettons à l'utilisateur d'avoir un contrôle intentionnel total sur son avatar ;
- Une limitation du registre des actions automatiques. Nous limitons l'ensemble des comportements produits automatiquement par notre modèle à des séquences d'actions courtes (comme orienter le regard vers une direction, effectuer des expressions gestuelles lors du dialogue, saluer, etc.) qui peuvent être facilement assimilées par l'utilisateur. Ainsi, un utilisateur qui 'voit' son avatar effectuer des comportements automatiques peut s'approprier aisément ce comportement. Il peut ensuite décider de l'interrompre, si ce comportement ne correspond pas à ses intentions, ou l'intégrer dans son processus d'action (s'il considère qu'il correspond à ses intentions) ce qui garantit un cadre plus situé à l'interaction et une sensation de présence dans l'univers virtuel.

## 6.3 Exemples d'interactions sociales

A titre d'exemples d'interactions situées dans l'univers virtuel, nous analysons le comportement d'avatars représentant les membres d'une équipe de sapeurs-pompiers dans deux situations d'interactions différentes : une situation de dialogue et une situation de rencontre.

### 6.3.1 Cas des situations de dialogue

Le module de perception du contexte de chaque avatar permet de détecter les situations de dialogue entre deux ou plusieurs interlocuteurs.

Le modèle d'interaction sociale permet d'associer à chaque interlocuteur un rôle dans ce dialogue (auditeur, locuteur, etc.) selon la forme de la communication identifiée (communication verbale, ou non verbale comme les déictiques par exemple). Il permet ainsi d'activer un ensemble de règles comportementales adaptées au rôle de l'avatar dans le dialogue (figure 6.9).



FIG. 6.9 – Aperçu d'une situation de dialogue entre plusieurs avatars.

Par exemple, un avatar *Locuteur* distribuera son regard entre ses différents interlocuteurs selon le rythme de son discours (Chovil, 1992), il clignera les yeux à chaque silence et avant chaque mouvement de tête.

Selon leurs personnalités, les avatars associent certaines expressions gestuelles à leurs discours en fonction des taux calculés par le modèle émotionnel.

Le contenu du discours n'intervient pas dans le choix des expressions corporelles décidées par le modèle émotionnel et social. Il demeure bien évidemment possible pour l'utilisateur d'intervenir pour contrôler son avatar et fixer par exemple la direction de son regard vers un interlocuteur particulier ou indiquer une direction donnée. De telles interventions permettront d'ajouter d'autres indices contextuels que les autres utilisateurs pourront exploiter pour la construction de leurs contextes d'interactions. Ils peuvent ainsi tenir compte

de cette action dans la prise de leurs décisions.

Le modèle émotionnel intervient dans le dialogue pour moduler la fréquence des expressions gestuelles et la posture en fonction du stress supposé de l'avatar. Ce modèle fournit également l'ensemble des états émotionnels pour produire l'expression faciale adéquate.

### 6.3.2 Cas des situations de rencontre

Une situation de rencontre est caractérisée par la proximité dans l'univers virtuel d'un ou plusieurs avatars. Lorsqu'un avatar émet une forme de communication, cette situation sera identifiée comme situation de dialogue. Le module d'interaction sociale permet d'identifier chaque interlocuteur dans



FIG. 6.10 – Aperçu d'une situation de rencontre entre deux avatars.

une situation de rencontre et d'activer un ensemble de règles de comportement social.

Bien que l'utilisateur continue de contrôler son avatar pour marcher vers une direction donnée par exemple, son avatar peut de façon autonome regarder les personnes qu'il rencontre sur son chemin, les saluer s'il s'agit d'un collègue ayant un rôle hiérarchique plus élevé, etc.

Le module émotionnel intervient pour réguler ce mouvement de regard en fonction du stress et empêcher le cas échéant un regard vers d'autres interlocuteurs lorsque le niveau de stress est jugé important.

Les autres utilisateurs peuvent ainsi utiliser ce comportement atypique lors d'une rencontre comme un indice contextuel supplémentaire pour interpréter la situation.

D'une façon générale, le tableau 6.3 présente un exemple de quelques règles appliquées pour la production de comportements automatiques chez un avatar

selon son contexte de l'interaction.

	Comportement d'auditeur
Expression faciale	exprime l'état émotionnel de l'avatar
Regard	orienté vers le locuteur
Hochement de tête	aléatoire pendant le discours de l'interlocuteur
Posture	traduit le stress de l'avatar
	Comportement de locuteur
Expression faciale	exprime à la fois l'état émotionnel de l'avatar et les mouvements labiaux lors de l'élocution
Regard	orienté vers les différents auditeurs
Gestualité	exprime des gestes expressifs, rythmiques et de régulation selon la personnalité de l'avatar
Posture	traduit le stress de l'avatar
	Comportement lors d'une rencontre
Expression faciale	exprime l'état émotionnel de l'avatar
Regard	orienté vers l'avatar rencontré si le degré de stress ne dépasse pas un certain seuil
Gestualité	saluer l'avatar rencontré s'il est considéré comme supérieur hiérarchique
Animations	déplacement traduit le stress

TAB. 6.3 – Exemple de comportements selon le contexte d'interaction.

Ces comportements implémentés peuvent ensuite être facilement améliorés en rajoutant d'autres règles comportementales ou encore en modifiant certaines règles déjà existantes. On peut par exemple considérer davantage les états de sortie du modèle émotionnel (les émotions, la personnalité et l'humeur) ou encore les caractéristiques sociales de chaque agent (statut social, expérience, etc.) dans la spécification des règles comportementales pour rendre les avatars encore plus émotionnels et sociaux.

## 6.4 Modélisation de l'univers virtuel

La modélisation d'un environnement virtuel regroupe à la fois les aspects de rendu de l'image (qualité des personnages virtuels, objets 3D, etc.), les animations à produire et les moyens mis en œuvre pour réaliser des interactions intuitives.

### 6.4.1 Objets de l'environnement virtuel

Une interaction en univers virtuel est d'autant plus réaliste qu'elle est assimilée par l'utilisateur à une interaction en univers réel.

Il est donc essentiel de reproduire en univers virtuel les mêmes caractéristiques physiques de l'univers réel (bâtiments, véhicules, etc.) afin d'accroître la crédibilité du dispositif et augmenter la notion d'immersion physique chez l'utilisateur.

Dans notre environnement virtuel, un soin particulier a été apporté à la modélisation des objets 3D. L'élaboration d'une image commence par la constitution d'un modèle de l'objet, appelé maquette numérique, qui est la représentation informatique de cet objet à partir d'informations géométriques.

Une méthode classique consiste à raisonner en termes de surfaces. Chaque objet peut être décomposé en 'facettes' (ou polygones) qui, mis bout à bout, permettent de rendre compte de l'enveloppe extérieure d'un solide.

Plus une maquette comporte de polygones, plus l'image qui en résulte est précise.

Au moment de l'affichage, l'objet ainsi reproduit se présente sous la forme d'une juxtaposition de facettes, dite 'structure en fil de fer'. Il s'agit d'une représentation purement géométrique qui ne prend pas en compte les caractéristiques optiques de l'objet.

Chaque objet a un aspect et une texture qui permettent d'identifier s'il s'agit d'une pierre, d'un bois, d'un tissu, etc. Le texturage consiste à appliquer sur une surface un motif qui respecte les caractéristiques d'une matière, pour suggérer visuellement la nature de cette surface, comme par exemple le revêtement d'un mur ou une peau sur un visage (en l'espèce, le modèle géométrique).

Le texturage est l'une des principales composantes du rendu réaliste de l'image.

Une fois texturée, l'image doit faire l'objet du traitement des ombres et des intensités de lumière (lorsque le passage de l'ombre à la lumière se fait de manière continue : chaque point est affecté d'une luminosité différente, créant un dégradé de couleur qui permet de rendre compte des éclairages). Les propriétés de réflexion des objets rentrent aussi en jeu : chaque matériau, en effet absorbe ou renvoie la lumière. Cette propriété va jouer sur l'objet lui-même (les reflets, le scintillement) mais aussi sur les objets à proximité, puisque la lumière est envoyée sur les objets voisins.

La figure 6.11 présente un aperçu de quelques objets 3D intégrés dans nos simulations. Ces objets correspondent à des bâtiments existants et des



FIG. 6.11 – Modélisation de quelques objets 3D dans notre univers virtuel.

véhicules (utilisés par les sapeurs pompiers dans le cadre de leurs interventions) que nous avons photographiés. Les photos obtenues ont ensuite été plaquées comme textures sur ces objets 3D.

#### 6.4.2 Personnages virtuels et animations

Dans notre plate-forme de réalité virtuelle, les personnages virtuels représentent des avatars contrôlés par des opérateurs humains.

Différentes techniques existent pour la modélisation de personnages virtuels de forme humaine (ou humanoïdes) et de leurs actions. Les travaux de thèse de (Thomas, 1999) décrivent quelques solutions possibles.

Dans nos travaux, nous avons utilisé le logiciel POSER<sup>1</sup> pour créer les personnages virtuels puis nous avons utilisé le logiciel LightWave3D<sup>2</sup> pour réaliser les animations et les mouvements. L'animation interactive de ces personnages virtuels a été ensuite réalisée au sein de l'atelier d'animation VIRTOOLS<sup>3</sup>.

La figure 6.12 donne un aperçu de quelques personnages virtuels de notre univers virtuel.

<sup>1</sup><http://www.e-frontier.com/>

<sup>2</sup><http://www.newtek.com/>

<sup>3</sup><http://www.virttools.com/>



FIG. 6.12 – Quelques avatars de notre environnement virtuel.

### 6.4.3 Géolocalisation du son

Le son est un élément important du réalisme lors des interactions virtuelles. Il permet de les enrichir et de les rendre plus 'naturelles' en simulant nos environnements auditifs naturels.

Dans notre expérience quotidienne, nous percevons l'espace sonore en trois dimensions en analysant le son parvenant à nos oreilles. Au delà d'une simple 'audition' des bruits, cette 'perception spatiale' des sons complète les informations récoltées par nos autres sens.

Elle a d'abord un rôle informatif. Elle nous renseigne sur les positions des sources sonores dans l'espace environnant (perception de la direction et de la distance) : c'est la localisation auditive.

Les sons qui se propagent agissent aussi comme un 'révélateur' sur les lieux, par le jeu des réflexions sur les parois et l'effet de réverbération. Ainsi l'auditeur peut identifier le lieu d'où est parti un son : il a par exemple la sensation d'être dans une salle plus ou moins grande, une salle de bain ou une

cathédrale, ou encore dans un espace ouvert, une ruelle ou une forêt, etc.

Le son 3D est aussi le support de l'intelligibilité : nous avons cette aptitude à isoler une conversation dans un environnement bruyé ou parmi d'autres conversations.

Le fait d'être baigné dans un espace sonore participe au confort de l'écoute et à une certaine dimension de plaisir. Il s'agit d'une 'propriété immersive' supplémentaire que nous utilisons dans nos simulations pour augmenter la sensation de présence dans l'univers virtuel.

## 6.5 Conclusion

Dans ce chapitre, nous avons présenté une description générale de l'architecture de notre plateforme. Cette architecture se base sur un système multi-agents pour modéliser les interactions entre les agents qui la composent. Il s'agit d'une architecture hybride qui permet aux agents d'être réactifs aux interactions dans l'univers mais suffisamment flexible pour permettre d'intégrer également des comportements 'rationnels' comme les processus de planification, etc.

Nous avons ensuite fourni un exemple de fonctionnement de cette architecture notamment pour l'identification du contexte d'interaction afin de produire un comportement adapté à la situation. Notre approche différencie le contexte général perçu par l'utilisateur de celui identifié par son avatar.

Nous avons également présenté une solution basée sur la décomposition des types de contextes et une organisation des modules de perception chez l'agent pour identifier un ensemble exhaustif d'éléments contextuels jugés appropriés à la situation.

Cette identification du contexte et malgré son caractère réducteur (par rapport au contexte général de l'utilisateur) nous permet tout de même de produire des comportements automatiques chez les avatars qui soient compatibles avec la situation de l'interaction.

Nous avons présenté ensuite les moyens mis en œuvre pour la modélisation de cette plate-forme de réalité virtuelle. Les personnages virtuels modélisés sont dotés d'apparences humaines, d'animations et de textures inspirés de nos observations réelles. Nous intégrons également sur cette plate-forme des sons en trois dimensions capables de renforcer l'immersion de l'utilisateur dans l'interaction.

# Chapitre 7

## Expérimentations

*Nous aurions souvent honte  
de nos plus belles actions  
si le monde voyait  
tous les motifs qui les produisent.  
(La Rochefoucauld)*

Ce chapitre présente une étude pilote destinée à fournir des résultats préliminaires nécessaires à la conception d'expérimentations de notre plateforme de réalité virtuelle à de plus grandes échelles.

Nous commençons par présenter le contexte général des évaluations à conduire sur notre dispositif de réalité virtuelle (section 7.1), puis nous nous intéressons dans le cadre de cette thèse à deux aspects principaux de ces évaluations :

- l'évaluation de l'ergonomie des interactions virtuelles engagées dans l'univers virtuel (section 7.2) ;
- l'évaluation de la pertinence des modèles émotionnel et social implémentés chez les avatars (section 7.3) ;

Nous discutons à la section 7.4 l'intérêt des solutions mise en œuvre pour améliorer l'interactivité en univers virtuel.

## 7.1 Contexte d'évaluation

Le dispositif consiste à mettre en réseau plusieurs acteurs (participants humains) qui collaborent dans un espace virtuel via des avatars qu'ils animent. Les prises de décisions sont alors effectuées par les participants sur la base des indices conceptuels qu'ils perçoivent de l'environnement.

L'environnement de réalité virtuelle est composé d'une scène 3D avec une reproduction aussi fidèle que possible des caractéristiques environnementales de l'environnement naturel c'est à dire les bâtiments à proximité, les panneaux d'affichages, les routes, etc.

Notre objectif est de conduire une série d'expérimentations préliminaires qui visent à évaluer l'ergonomie des interactions engagées et la pertinence des modèles d'interactions mis en œuvre.

Cet objectif s'inscrit dans un objectif encore plus large qui consiste à évaluer notre plate-forme de simulation selon plusieurs axes (figure 7.1) :

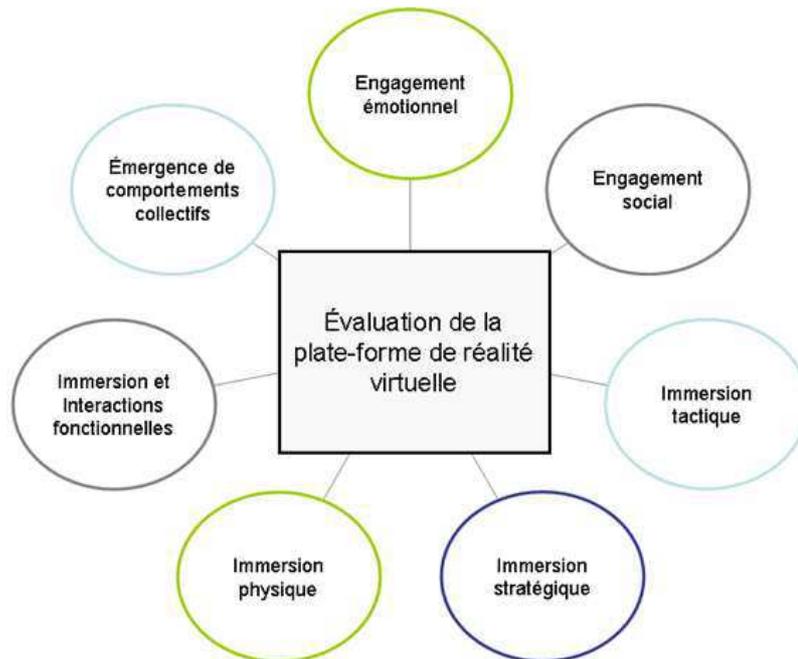


FIG. 7.1 – Extrait de la simulation d'un scénario de feu d'hôtel.

### **Engagement émotionnel de l'utilisateur :**

Il s'agit d'étudier la capacité du dispositif à produire une immersion émotionnelle lors des interactions.

Nous essayerons également d'étudier l'écart qui peut exister entre les

émotions simulées de l'avatar (par un modèle analytique) et les vraies émotions ressenties par l'utilisateur.

**Engagement social de l'utilisateur :**

Nous nous attachons à évaluer la pertinence des processus réflexifs engagés entre les utilisateurs et l'impact des comportements automatiques (générés par notre modèle d'interaction sociale) sur les processus décisionnels des utilisateurs.

Nous tenterons également d'étudier l'apport de la dimension indexicale (usage des gestes déictiques en particulier) dans le processus de négociation et de construction du sens entre les utilisateurs lors des interactions virtuelles.

**Immersion stratégique et tactique dans l'intervention :**

Il s'agit d'étudier la dimension individuelle dans la prise de décision.

Nous nous intéresserons à vérifier si la réalité virtuelle offre un support d'interaction suffisamment 'fiable' pour la reproduction des processus décisionnels qu'on obtient dans la réalité.

Nous essayerons d'analyser les scénarios de simulation virtuelle et les comparer à des simulations réelles pour identifier les indices contextuels utilisés par les acteurs pour prendre une décision en univers virtuel et jusqu'où l'utilisateur peut 'oublier' les imperfections du système pour une immersion stratégique dans le système.

**Immersion physique dans l'univers virtuel :**

Il s'agit d'évaluer les facteurs d'immersion physique que nous utilisons dans nos interactions comme le réalisme des personnages virtuels, de la scène 3D, la crédibilité des comportements (animations) et l'immersion sonore dans l'univers virtuel.

**Aspect ergonomique du dispositif :**

Il s'agit d'évaluer l'ergonomie du dispositif de contrôle. Nous nous intéresserons spécifiquement à l'usage de l'interface de contrôle et des moyens de communications mis en œuvre pour améliorer l'interactivité dans un univers virtuel.

**Aspect pédagogique de la formation :**

Nous essayerons dans cette évaluation d'étudier les possibilités de notre plate-forme de réalité virtuelle à répondre aux objectifs de formation (particulièrement pour la formation au commandement des sapeurs-pompier).

Nous nous intéressons également à étudier les facteurs d'immersion stratégique chez les utilisateurs pour la coordination des actions et la coopération de groupe (gestion des interventions d'urgences pour les sapeurs-pompier).

**Émergence de comportements collectifs :**

Il s'agit d'observer dans les simulations collaboratives l'émergence de phénomènes d'organisation sociale et de coordination de groupe.

Comme nous l'avons précisé dans l'introduction de ce chapitre, nous nous intéressons dans le cadre de cette thèse spécifiquement à deux aspects d'évaluations : l'ergonomie des interactions virtuelles et la contribution des modèles émotionnel et social au caractère 'situé' des interactions.

## 7.2 Évaluation de l'ergonomie d'interaction

Notre objectif dans cette section est d'évaluer l'ergonomie des interactions engagées en univers virtuel.

Nous nous intéressons d'abord aux moyens de contrôle mis à la disposition de l'utilisateur pour diriger son avatar et effectuer des actions intentionnelles (section 7.2.1), puis nous discuterons (à la section 7.2.2) du point de vue de l'utilisateur par rapport à son avatar.

### 7.2.1 Interface de contrôle

Le réalisme et l'immersion des acteurs du virtuel ne saurait se réaliser sans que l'interface de contrôle ne sache 'se faire oublier' sans pour autant réduire les possibilités d'actions de l'utilisateur.

Autant de contraintes difficiles à satisfaire par les interfaces de contrôles actuelles notamment en ce qui concerne les déictiques (les indications gestuelles destinées à orienter l'attention de l'interlocuteur) ou encore les mouvements d'exploration de l'univers par le regard.

Le casque de vision représente une interface 'intéressante' pour l'immersion et l'exploration dans un univers virtuel. Il permet la disposition d'écrans situés près des yeux (munis d'optiques spéciales afin d'éviter la fatigue visuelle) et d'agrandir l'image pour qu'elle remplisse tout le champ visuel de l'utilisateur. Cette disposition contribue à donner à l'utilisateur l'impression d'être à l'intérieur de la scène reconstituée en images de synthèse.

Une autre fonction du casque de réalité virtuelle est d'être un outil de commande de l'image, grâce à un dispositif de capteurs qui permettent de 'suivre' les mouvements de la tête. Ces capteurs permettent ensuite de déterminer l'orientation et la vitesse des mouvements de la tête de l'utilisateur.

L'image dans un tel casque occupe donc la totalité du champ de vision de l'utilisateur. De plus, la position de la tête de l'utilisateur étant transmise en temps réel au système, cette technique garantit l'immersion et la navigation dans l'image puisqu'elle donne l'impression de diriger l'image par le regard.

Malheureusement, en plus des difficultés techniques communes à l'ensemble des activités relatives aux images de synthèse, la fabrication de casques suppose des qualités de poids, d'ergonomie (aisance du port du casque), de design, et surtout de qualité d'optique encore complexe à mettre en œuvre pour une utilisation grand public d'outils de réalité virtuelle.

De plus, ces dispositifs sont reliés à l'ordinateur par un fil, ce qui restreint la liberté de mouvement de l'utilisateur.

Les 'data glove' ou encore les gants de capture représentent également une interface 'intéressante' pour le contrôle des personnages virtuels dans notre plate-forme. Ils disposent de nombreux capteurs, reliés par de la fibre optique, qui permettent au système de 'percevoir' et réagir aux mouvements de la main de l'utilisateur.

Les difficultés relatives à cette interface sont à la fois de nature technique (capter dynamiquement les gestes de la main) et cognitive (orientation, perception de l'espace). Dans le cas des déictiques, il convient de pouvoir visualiser simultanément la direction du bras et l'objet visé ce qui n'est pas toujours possible dans les univers virtuels.

Finalement, un système d'interaction représenté par une manette de jeu 3D et un casque doté d'un microphone nous semble adéquat pour le contrôle de l'avatar dans notre système de réalité virtuelle (figure 7.2).



FIG. 7.2 – Interface de contrôle.

Cette interface malgré sa simplicité nous semble offrir un contrôle souple et intuitif de l'avatar.

Le microphone permet de naturellement capturer les communications langagières engagées par l'utilisateur (dans un langage naturel) que le système de réalité virtuelle transmet ensuite aux différents utilisateurs en interaction.

La manette de contrôle 3D offre également des possibilités de contrôle souples des déplacements des avatars en orientant l'axe de la manette vers l'avant, en arrière, à gauche et à droite.

Cette manette offre également la possibilité d'orienter l'axe vertical selon un mouvement de rotation de poignée qui correspond intuitivement à des commandes d'exploration de l'univers virtuel par le regard ou pour effectuer des gestes déictiques vers des directions différentes de l'espace.

Ainsi l'utilisateur ne dispose que d'un ensemble réduit de commandes à maîtriser pour contrôler son avatar et interagir dans l'univers virtuel.

## 7.2.2 Champ visuel de l'utilisateur

Le point de vue de l'utilisateur lors des scénarios d'interaction en univers virtuel est problématique. D'après Büscher (Büscher et al., 2000), les environnements collaboratifs se centrent trop sur la représentation des 'personnes' (comme les avatars par exemple) au détriment de la représentation du 'matériel' c'est à dire de la tâche et des actions qui sont effectuées.

Plusieurs solutions sont envisageables pour offrir à l'utilisateur un angle de vue qui contribue à améliorer son interaction en univers virtuel.

Chacune de ces solutions présente des avantages et des inconvénients que nous discutons ci-après.

### 7.2.2.1 Quelques solutions envisageables

Une première solution consiste à placer la caméra (qui représente l'angle de vue de l'utilisateur) à la place de la tête de l'avatar (celui-ci ne voit donc pas le corps de son propre avatar).

Bien que cette solution contribue sensiblement à augmenter la sensation de présence dans le dispositif, nous pensons qu'elle présente tout de même quelques inconvénients.

Dans notre cas, les mouvements de tête sont soit générés indépendamment de l'utilisateur par notre modèle d'interaction sociale (maintien du contact visuel entre les interlocuteurs, regard vers un avatar rencontré, etc.) soit intentionnellement dirigé par l'utilisateur vers un objet ou un avatar de l'environnement.

Dans ces deux cas, l'utilisateur ne dispose d'aucune information 'proprioceptive' sur la position relative de son corps et de sa tête. Il ne perçoit que la direction de son regard d'où de réelles difficultés de représentation corporelle.

L'utilisateur ne peut également pas voir les expressions gestuelles de son propre avatar (gestes expressifs lors du dialogue, déictiques, etc), ni identifier ses propres mouvements (marcher, courir, etc.) d'où une gêne lors des interactions (figure 7.3).

Une autre solution (souvent utilisée dans les jeux vidéos) consiste à placer la caméra (qui représente la vue de l'utilisateur) au dessus de l'espace d'interaction selon un référentiel exocentrique (figure 7.3).



FIG. 7.3 – Angle de vue de l'utilisateur selon un référentiel égocentrique et exocentrique.

Cette orientation de la caméra s'avère plus confortable pour les interactions à caractère stratégique (déplacement de plusieurs agents, coordination de plusieurs tâches, etc.) mais risque de 'réduire' le sentiment d'immersion sociale que nous cherchons à produire à travers les comportements et émotions échangés entre les avatars.

Une autre approche consiste à fixer l'angle de vue des utilisateurs au dessus et en arrière de leurs avatars (l'utilisateur se voit ainsi de dos) mais à une distance que nous estimons 'raisonnable' pour permettre à la fois de voir les comportements produits par son propre avatar et les expressions émotionnelles et corporelles des autres avatars (figure 7.4).

Bien que cette solution ne corresponde pas totalement à la réalité (dans la réalité nous voyons plutôt la partie frontale de notre corps), elle présente un certain nombre d'avantages :

- Le sujet a une représentation visuelle distincte de la direction de son buste et de sa tête. Dans la réalité cette fonction est naturellement satisfaite puisque nos systèmes proprioceptifs, vestibulaires et visuels nous renseignent sur les orientations relatives de ces parties de notre corps.
- Il est possible de mieux se représenter la direction des déictiques par rapport aux objets de l'environnement (champ visuel artificiellement accru).
- Elle offre à l'utilisateur une meilleure vision globale de son corps en



FIG. 7.4 – Aperçu de l’angle de vue d’un utilisateur.

particulier des expressions gestuelles qui sont générées par le modèle émotionnel. L’utilisateur peut ainsi éventuellement intégrer cette information dans le flux de son activité intentionnelle.

- Elle offre également à l’utilisateur une vision claire de la direction de son corps et de sa tête ce qui lui permet de ne pas confondre ces deux directions lorsqu’il décide de se déplacer (il pourrait penser que la direction du regard reflète la direction du corps).

Le principal inconvénient de cette solution est que le corps de l’avatar masque une partie du champ visuel. Cet inconvénient est facilement compensé par le fait que le sujet peut déplacer son avatar et le positionner au mieux en fonction de ses intentions.

#### 7.2.2.2 Conclusion

Le point de vue subjectif de la ‘caméra reculée’ présente donc l’avantage de fournir à l’utilisateur une information visuelle qui lui permet de construire une représentation de son corps dans l’espace équivalente à la représentation proprioceptive qu’il aurait dans la situation naturelle.

Il peut ainsi produire des expressions déictiques (verbales et gestuelles) qui ont un sens par rapport à l’environnement.

Par exemple, le fait que le sujet puisse clairement percevoir la direction de son corps et de sa tête lui permettra de produire des expressions verbales égocentriques du type ‘à gauche’, ‘à droite’, etc. qui soient pertinentes pour son interlocuteur. Il peut également les accompagner d’expressions déictiques telles que des gestes dans la direction des objets référés.

Malgré le fait que l’avatar masque une partie du champ visuel de l’utilisateur, les utilisateurs s’approprient rapidement ce référentiel visuel et arrivent

même à 'intégrer ses limites' dans le sens qu'ils s'arrangent pour déplacer légèrement leurs avatars afin d'avoir leurs interlocuteurs complètement dans leurs champs visuels.

### **7.3 Évaluation de quelques aspects de l'engagement émotionnel et social**

Afin d'évaluer la pertinence du modèle d'interaction sociale dans une situation d'interaction en univers virtuel, nous avons mis en place plusieurs expériences exploratoires.

- La gestion du dialogue et la direction du regard. Cette expérience a pour objectif d'évaluer la pertinence du module de contrôle de la direction du regard en situation de dialogue.
- La prise en compte des expressions émotionnelles. Nous analyserons dans cette expérience la façon dont les utilisateurs prennent en compte les expressions émotionnelles en situation de dialogue.
- La gestion des déictiques gestuels et la dimension proxémique lors des interactions.

Nous décrivons ci-après ces expérimentations.

#### **7.3.1 Gestion du dialogue et direction du regard**

Le but de cette expérimentation est d'évaluer la pertinence du modèle d'interaction sociale dans la gestion du dialogue à plusieurs interlocuteurs.

Nous nous sommes intéressés à l'évaluation de la gestion des tours de paroles en univers virtuel, la pertinence des expressions gestuelles et des directions du regard en situation d'interaction multi-utilisateurs.

##### **7.3.1.1 Protocole expérimental**

Nous avons mis en situation de dialogue deux utilisateurs animant chacun son propre avatar (le participant et un instructeur) ainsi que deux autres personnages virtuels autonomes (figure 7.5).

L'instructeur s'adresse au participant pour l'informer de la situation (feu dans un hôtel).

Dans la perspective d'analyser la prise en compte de la direction du regard de l'interlocuteur, nous avons considéré deux situations : une situation où l'instructeur distribue ses regards sur l'ensemble du groupe avec une prédominance vers le participant et une situation où l'instructeur s'adresse ostensiblement à un avatar autonome, en évitant de regarder le participant.

Deux autres conditions nous ont permis d'évaluer le rôle synergique des déictiques gestuels lorsque le locuteur (l'instructeur) indique dans son discours une direction.

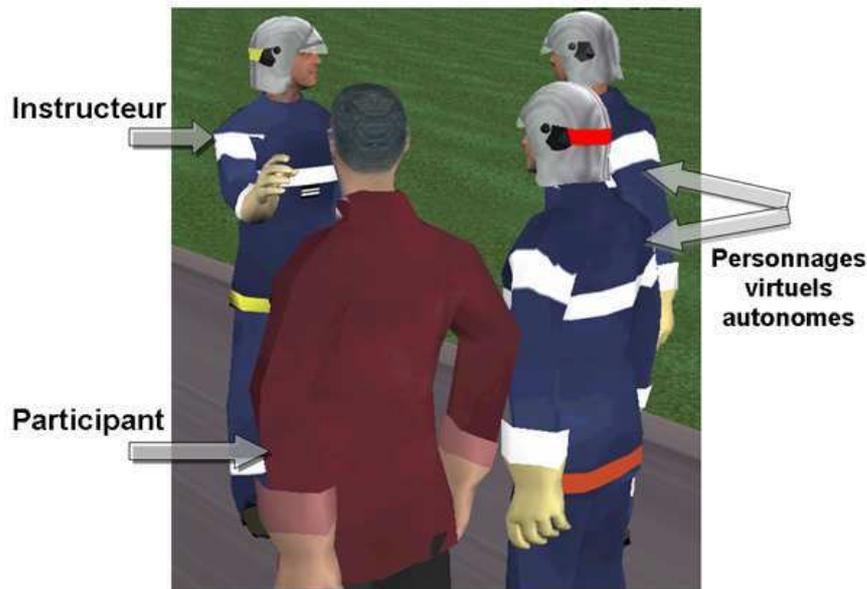


FIG. 7.5 – Disposition des avatars durant le dialogue selon l’angle de vue du participant.

### 7.3.1.2 Méthode

La population des participants se compose de 5 personnes représentant une tranche d’âge entre 20 et 40 ans.

Notre échantillon est constitué de participants familiers en informatique mais sans expertises particulières dans les jeux vidéo (étudiants IRIT).

Tous les acteurs et avatars autonomes ont été positionnés en respectant les distances interpersonnelles usuelles (1 m à 1,5m).

Nous avons filmé les différents utilisateurs pendant les simulations. Des interviews ont ensuite été réalisés à la fin de chaque simulation afin de recueillir les données spécifiques à chaque expérimentation et les impressions générales sur le déroulement de l’interaction.

Ces entretiens réalisés ont été libre afin de ne pas guider la réponse des utilisateurs.

### 7.3.1.3 Scénario d’interaction

L’instructeur (figure 7.5) qui contrôle son avatar (représenté par un sapeur pompier chef) s’approche du groupe où se trouve le participant.

Le participant immobile observe cette scène tout en ayant la possibilité de diriger son regard vers les différents acteurs de la scène. L’avatar de l’instructeur informe ensuite ses interlocuteurs de la situation :

*L'hôtel est en feu, la situation est assez grave, il faut agir très rapidement. Nous allons nous organiser de façon efficace. Il y a trois victimes qui sont bloquées par la fumée au niveau du deuxième étage. Un jeune homme s'est également défenestré du premier étage et est allongé au sol juste devant l'entrée du bâtiment. Il souffre du pied.*

Pendant le temps de ce discours, le module d'interaction sociale contrôle en temps réel :

1. la direction du regard de l'avatar de l'instructeur (le modèle distribue les regards vers les différents avatars de la scène) ;
2. les expressions gestuelles qui accompagnent le discours ;
3. les expressions faciales (mouvements de sourcils, expressions émotionnelles, etc.).

L'avatar de l'instructeur fixe ensuite du regard un personnage parmi ses 3 interlocuteurs et exprime une requête impérative :

*Suis-moi, s'il te plait !*

Le participant doit ensuite décider s'il suit l'instructeur ou non (en fonction de la direction du regard de l'instructeur car c'est le seul indice qui lui permet de savoir si l'instruction s'adresse à lui).

Suite à cette expérimentation, le participant doit donner ses impressions sur le déroulement du scénario et répondre à plusieurs questions portant sur :

1. l'alternance du regard lors de la locution de l'avatar en public ;
2. la pertinence de la désignation du regard lors de l'adressage ;
3. les expressions gestuelles qui accompagnent le discours ;
4. les expressions faciales lors de l'énoncé.

L'expérience est répétée 3 fois avec à chaque fois des variations concernant l'alternance du regard. Nous désactivons (une fois sur trois) le module d'alternance du regard afin que l'avatar de l'instructeur ne regarde pas le participant.

#### **7.3.1.4 Résultats et interprétations**

Concernant les impressions des utilisateurs sur la gestion du dialogue en groupe, il apparaît que :

**Désignation d'un interlocuteur par le regard** Dans la majorité des cas (14/15), nous avons constaté que les participants sont capables d'identifier correctement l'avatar qui est désigné par le regard.

Ceci est d'autant plus intéressant que le protocole expérimental ne donnait aucun renseignement explicite au participant sur la pertinence de

cet indice. Les participants prennent donc implicitement en compte la direction du regard du locuteur bien que celui-ci soit contrôlé la plupart du temps par le module d'interaction sociale.

Au moment où l'instructeur donne son ordre, le participant analyse la direction de son regard et s'attribue ou non le statut de destinataire.

Ce couplage réflexif pertinent qui s'établit entre l'action produite (orientation du regard) et l'interprétation de la situation contribue au caractère situé des interactions sociales obtenues.

**Distribution du regard du locuteur** Dans la condition où le module de gestion du regard s'attarde ostensiblement sur un personnage particulier alors que le contexte de la situation voudrait que celui-ci s'adresse à l'ensemble des interlocuteurs, les participants décrivent la situation comme n'étant pas naturelle.

Là encore, nous voyons que la direction du regard est bien décodée par les participants en fonction du contexte. Tout écart par rapport à la norme est interprété comme un événement devant avoir un sens particulier alors que rien dans notre expérimentation ne permettait d'inférer une interprétation particulière. Ce manque d'inférence constitue la principale cause de cette gêne ressentie par les participants.

**Les expressions gestuelles lors de la locution** La production des gestes automatiques qui accompagnent le discours de l'instructeur a été jugée globalement naturelle.

Aucun participant n'a remarqué une incohérence entre les gestes produits et le discours prononcé (malgré le caractère répétitif et non significatif des animations).

Par contre, dès lors que l'avatar s'adresse dans son discours à une personne particulière, les participants s'étonnent du fait qu'il n'y a pas eu de production de gestes déictiques associés.

Ceci est à mettre en rapport avec le fait que dans une conversation, une grande partie des gestes sont considérés comme non spécifiques (c'est à dire qu'ils peuvent être remplacés par une large palette d'autres gestes).

**Les expressions faciales** Les expressions faciales de l'avatar de l'instructeur ont été jugées naturelles et la seule critique a concerné la qualité de la synchronisation labiale qui effectivement manquait d'expressivité (pas de variabilité phonémique).

Dans une nouvelle version de notre système nous avons introduit une variabilité arbitraire et un bon synchronisme ce qui semble satisfaisant pour les prochaines expérimentations.

### 7.3.2 Prise en compte des expressions émotionnelles

Le but de cette expérimentation est d'évaluer le rôle des expressions émotionnelles et corporelles dans l'interprétation de la situation en fonction du

contexte.

### 7.3.2.1 Protocole expérimental

Dans cette expérimentation, les participants sont en situation de face à face avec un interlocuteur virtuel. Deux conditions ont été explorées :

- une situation où l'instructeur adopte un comportement émotionnel cohérent avec le contexte de la situation (condition 'contexte non stressant');
- une situation où l'expression émotionnelle du locuteur est 'décalée' par rapport au contexte.

La tâche du participant est d'observer l'avatar du locuteur puis de donner ses impressions sur la plausibilité de la situation (commentaires libres).

Nous nous attendons à ce que les participants prennent en compte l'état émotionnel du locuteur (qui s'exprime par des gestes plus ou moins agités et des expressions faciales exprimant le stress) et tentent de trouver un sens à toutes les situations.

#### - Evaluation de la cohérence émotionnelle dans un contexte non stressant (appartement)

La première situation situe les personnages dans un appartement (figure 7.6). L'instructeur dit :

*Effectivement, je me souviens bien de cette personne. Si je la voyais un jour !.*



FIG. 7.6 – Extrait de la simulation du personnage virtuel dans l'appartement.

Le ton de l'énonciation de ce discours est neutre.

Ce scénario, sera joué deux fois avec à chaque fois :

- Une expression émotionnelle et corporelle calme et détendue de l’avatar ;
- Une expression émotionnelle et corporelle d’agacement et de stress.

Nous avons également fait varier lors d’une simulation (parmi les deux qui sont jouées) la direction du regard de l’avatar lors de la locution afin d’étudier son impact sur l’interprétation du discours (regard orientée vers le participant et regard légèrement déviée par rapport au regard du participant).

Nous avons ensuite recueilli pour ces deux simulations, les impressions de chaque participant face à cette interaction.

### - Evaluation de la cohérence émotionnelle dans un contexte stressant (feu d’hôtel)

Dans cette seconde expérience, un personnage (un gérant d’hôtel) s’inquiète de la situation (figure 7.7) :

*Monsieur, Monsieur, mon hôtel est en feu! En attendant les pompiers, pouvez vous me donner un coup de main s’il vous plaît ?.*



FIG. 7.7 – Extrait de la simulation du personnage virtuel devant un hôtel en feu.

Le ton de l’énonciation de ce discours est neutre.

Ce scénario, comme le précédent, est joué deux fois avec à chaque fois :

- Une expression émotionnelle et corporelle calme et détendue de l’avatar ;
- Une expression émotionnelle et corporelle d’agacement et de stress.

Nous avons également fait varier lors de ces deux simulations la direction du regard de l’avatar avec :

- Une direction du regard orientée vers le participant au moment de la locution ;

- Une direction du regard légèrement déviée par rapport au regard du participant au moment de la locution.

Nous avons ensuite recueilli pour ces deux simulations, les impressions de chaque participant face à cette interaction.

### **7.3.2.2 Résultats et interprétations**

**Expressions émotionnelles en fonction du contexte** Dans la situation d'interaction dans l'appartement, tous les participants distinguent les deux états émotionnels (calme ou agité) et les deux situations sont perçues comme étant plausibles.

Par exemple, le personnage calme évoque souvent le souvenir d'une personne familière alors que le personnage agité induit une interprétation mettant en jeu des situations plutôt professionnelles.

Dans la situation du feu d'hôtel, tous les participants différencient également les deux états émotionnels (calme et agité). Par contre, la situation où le gérant d'hôtel est calme dans le contexte de feu d'hôtel n'apparaît ni naturelle ni crédible.

Nous constatons donc que dans un contexte environnemental avec contenu émotionnel neutre (comme l'intérieur d'un appartement), les différentes expressions émotionnelles et corporelles peuvent être jouées sans discréditer le réalisme de la situation.

Par contre, lorsque le contexte est très typé (comme dans une situation d'incendie), l'adéquation entre les expressions émotionnelles exprimées et le contexte de la situation est fondamental pour maintenir le réalisme de l'interaction.

Un modèle émotionnel capable de bien adapter les émotions exprimées du personnage virtuel en fonction de la situation s'avère donc d'autant plus nécessaire que le contexte de la simulation est plus typé.

**Direction du regard** Il apparaît également qu'en situation de face à face, la direction du regard a un impact très important pour la notion de présence sociale.

Nous constatons que la déviation du regard hors du champ des destinataires a produit une sensation de gêne chez tous les participants.

Trois participants ont pensé que le locuteur ne s'adressait plus à eux au moment de l'interaction.

**Interprétation de la situation** Les impressions recueillies concernant le passage d'une scène à une autre (d'une scène d'appartement vers un feu d'hôtel et vice versa) montrent que les participants tentent de trouver des explications cohérentes et à reconstruire un sens pour la situation.

Malgré le fait qu'aucun lien visible n'a été prévu entre ces deux scènes, un besoin permanent apparaît chez les utilisateurs pour rétablir la cohérence de la scène.

### 7.3.3 Proxémique et gestes déictiques

Le but de cette expérimentation est double :

- Évaluer si les utilisateurs vont reproduire en réalité virtuelle les mêmes normes culturelles qui régissent la proxémique en réalité ;
- Évaluer si les utilisateurs reproduisent de manière naturelle des gestes déictiques en univers virtuel.

#### 7.3.3.1 Protocole expérimental

Deux personnes (l'instructeur et le participant) contrôlent respectivement leurs avatars.

Le participant a été au préalable familiarisé avec l'usage du déictique.

Ce dernier est déclenché par une pression sur une des commandes de la manette. La direction du déictique correspond à la direction du regard du participant (figure 7.8).



FIG. 7.8 – Extrait d'une simulation de l'usage du déictique gestuel.

Le scénario d'interaction est composé de trois situations afin d'évaluer la prise en compte des règles de proxémique.

Les conditions expérimentales sont les mêmes que pour les précédentes expérimentations c'est à dire : même échantillon de participants et même méthode d'interview.

### **Situation 1 : L'instructeur enfreint les règles de proximité**

L'instructeur avance son avatar en direction de l'avatar du participant et se positionne en deçà de la distance inter-personnelle classique (0,5 m) et lui fait la requête suivante :

*Monsieur, avez-vous l'heure s'il vous plaît ?*

On observe ensuite la réaction du participant.

### **Situation 2 : Le participant positionne lui-même son avatar pour engager une discussion**

Il est demandé au participant de diriger son avatar vers celui de l'instructeur pour demander un renseignement : on note ensuite la position de l'avatar du participant par rapport à l'avatar de l'instructeur.

### **Situation 3 : Induction de l'usage d'un déictique gestuel**

L'instructeur s'approche du participant et lui demande :

*Excusez-moi Monsieur, je cherche l'hôtel Ariane. Savez-vous où il se trouve ?*

Le participant, en se promenant dans l'univers virtuel, a certainement pu voir l'hôtel Ariane et connaît son emplacement.

Nous nous attendons à ce que le participant, de lui même utilise un déictique gestuel (bien que la consigne ne lui demande pas explicitement de le faire).

## **7.3.3.2 Résultats et interprétations**

**Situation 1 : non respect des règles de proxémique** Dans cette expérimentation, tous les participants ont exprimé un sentiment de gêne forte face à l'invasion de l'espace personnel de leurs avatars.

La presque totalité des participants (4/5) a fait reculer spontanément son avatar pour accroître la distance interpersonnelle et la rendre socialement plus acceptable.

**Situation 2 : le participant positionne lui même son avatar** Dans cette expérimentation, tous les participants ont positionné leurs avatars à une distance moyenne d'environ 1,5m de l'avatar de l'instructeur.

Cette distance correspond à la fois à une distance socialement acceptable entre deux individus lors d'un dialogue mais également à une distance 'confortable' lors d'une interaction virtuelle pour observer les expressions faciales et émotionnelles des interlocuteurs.

**Situation 3 : induction de l'usage d'un déictique gestuel** Dans cette expérimentation, nous remarquons que la majorité des participants (4/5) réagissent spontanément à la requête en produisant un déictique gestuel. Le participant qui n'a pas réagi par un déictique gestuel à la question (il a préféré le déictique verbal), a effectué un geste déictique lorsqu'on lui

a posé la même question une seconde fois.

Il s'avère donc que les utilisateurs ont tendance à reproduire en univers virtuel les mêmes règles d'interaction sociale qu'en situation naturelle.

## 7.4 Discussions

Dans ce chapitre, nous avons exploré l'intérêt de l'usage d'un modèle émotionnel et social dans la gestion de l'interaction en univers virtuel.

Ces premiers résultats exploratoires nous apparaissent très prometteurs malgré le caractère arbitraire de tout modèle émotionnel.

Les principaux acquis de cette expérimentation pourraient être résumés de la façon suivante :

### 1 - Prise en compte des expressions émotionnelles des avatars :

Le premier résultat important de cette expérimentation est que les participants prennent systématiquement en compte les états émotionnels des avatars qui les entourent et tentent de leur donner un sens dans le contexte de l'interaction.

L'analyse des états émotionnels peut se faire soit à partir des expressions faciales soit en prenant en compte les expressions corporelles (gestes, postures, etc.).

Nous pensons que ce processus est facilité par le fait qu'une grande partie de nos activités gestuelles est 'non spécifique' (c'est à dire qu'un grand nombre d'expressions gestuelles ou posturales peut être associé à une situation sans la rendre incohérente).

L'activité interprétative du destinataire est néanmoins très présente pour analyser le sens de l'expression non verbale.

### 2 - Usage des indices indexicaux :

Nous avons mis en évidence la possibilité d'un usage intensif des indicateurs indexicaux spatiaux (gestes et direction de regard).

Contrairement aux résultats antérieurs relatant des difficultés à exploiter ce type d'indices en univers virtuel (Hindmarsh et al., 1998), nous avons montré qu'une interface adaptée permettait aux utilisateurs d'exploiter spontanément ce type d'indices.

La réduction du champ visuel inhérente à cette technologie peut être partiellement compensée par un positionnement en recul de la caméra et par une mobilité accrue des avatars.

Dans notre cas, nous avons implémenté un contrôle indépendant de la direction du corps et de la tête des avatars ce qui favorise une exploration active du champ visuel et compense la réduction de visibilité.

Nous avons remarqué que les utilisateurs peuvent utiliser spontanément des expressions indexicales (orientation du bras, direction de regard) pour structurer leurs discours.

Ils sont également très sensibles à la direction des regards et utilisent instinctivement ces derniers pour juger le statut des interlocuteurs (destinataires d'un message, transfert d'un tour de parole, etc.).

### **3 - Réflexivité de l'interaction et immersion sociale :**

L'ensemble de ces résultats milite en faveur du fait que les technologies de la réalité virtuelle sont aujourd'hui capables de produire de véritables couplages socio-cognitifs entre des acteurs en situation de communication.

Par exemple, les participants ont instinctivement repositionné leurs avatars dans le cas où leurs espaces interpersonnels n'était pas respecté. Des couplages liés à l'analyse de la direction du regard ont été également mis en évidence.

Très rapidement, les participants apprennent à se positionner dans l'espace pour contourner certaines difficultés représentationnelles.

Par exemple, il nous a été possible de constater qu'ils pouvaient exploiter des ombres pour inférer la présence d'autres acteurs.

La principale limite à cette réflexivité nous semble aujourd'hui être liée au manque de pertinence des expressions émotionnelles telles qu'elles sont évaluées par les modèles.

Nous avons vu que les participants prennent en compte naturellement ces expressions mais malheureusement ces modèles, malgré leur niveau de complexité, restent loin de pouvoir reproduire le véritable état émotionnel du participant.

Cette difficulté peut toutefois être partiellement contournée soit par des dispositifs capables de mieux évaluer en temps réel le véritable état émotionnel du participant (capteurs biologiques ou analyseur d'expressions émotionnelles) soit en donnant au participant la possibilité de jouer de la différence entre son propre état émotionnel et celui de son avatar.

L'ensemble de ces résultats nous permettent donc de penser qu'il est possible de produire en univers virtuel certains types d'interactions sociales situées.

Beaucoup reste à faire dans le domaine de l'expression émotionnelle mais nous pouvons penser que l'imperfection des modèles peut être partiellement compensée par les possibilités d'ajustement mutuel des comportements ainsi que par un surinvestissement du canal verbal porteur lui aussi d'informations émotionnelles.

L'impact social de ce type de technologie, s'il peut être mené à bout par une recherche pluridisciplinaire est très important puisqu'il ouvre la voix à de

nombreuses applications dans le domaine de la formation, la rééducation, de la conception et dans toutes les situations où la dimension émotionnelle dans l'interaction est essentielle.

Quatrième partie

**Conclusions & Perspectives**



## Chapitre 8

# Conclusions

### 8.1 Synthèse

Nos travaux portent sur les interactions situées en univers virtuel. Ces interactions posent des problèmes qui vont bien au delà des questions du réalisme de l'environnement virtuel et de l'immersion physique dans le dispositif d'interaction.

Nous pensons qu'une simulation interactive pour qu'elle soit crédible et susceptible de produire la notion de présence sociale, doit prendre en compte une partie des mécanismes naturels des communications humaines.

C'est seulement à cette condition que nous pouvons espérer produire des processus réflexifs en univers virtuel semblables à ceux que nous obtenons dans une situation réelle.

Reproduire ces processus réflexifs en univers virtuel passe obligatoirement par la prise en compte de mécanismes d'interaction sociale comme la réflexivité et l'indexicalité.

Dans cet objectif, nous avons développé une plate-forme d'interaction sociale multi-utilisateurs en la réalité virtuelle.

Chaque utilisateur collabore avec d'autres utilisateurs à travers son avatar et peut 'naturellement' interagir et prendre des décisions d'actions selon sa propre expertise, connaissance, historique, etc.

Nous nous sommes également intéressés à l'exploration des possibilités que peut offrir un modèle émotionnel et social dans l'amélioration des interactions virtuelles collaboratives.

Nous avons dans un premier temps développé un modèle émotionnel capable d'identifier les états internes des avatars en fonction de leurs interactions avec l'univers virtuel.

Ce modèle se base sur le modèle décrit par (Ortony et al., 1988) qui

calcule les états émotionnels selon une théorie d'évaluation qui associe des propriétés critiques aux événements qui peuvent naturellement produire des émotions particulières. Nous avons ensuite 'enrichi' ce modèle par d'autres notions comme la personnalité et l'humeur.

Au fur et à mesure de son interaction en univers virtuel, notre modèle émotionnel calcule les intensités probables des émotions ressenties (satisfaction, déception, peur et colère) en fonction des événements perçus qui dépendent du contexte de la tâche. Ce modèle gère également l'oubli des émotions au cours du temps.

Notre but n'étant pas de valider le modèle lui-même mais de l'intégrer dans nos simulations afin d'étudier son impact sur l'émergence de processus réflexifs lors des interactions virtuelles.

Nous avons ensuite intégré un modèle d'interaction sociale qui offre la possibilité de compléter les actions intentionnelles des utilisateurs par des comportements non intentionnels chez leurs avatars comme :

- la production d'expressions gestuelles lors du dialogue pour accompagner le discours de l'utilisateur ;
- la production d'expressions émotionnelles qui traduisent les états internes de chaque avatar ;
- la production de hochements automatiques de la tête lors de l'audition ;
- la distribution automatique du regard entre les différents interlocuteurs lors de la locution ;
- l'orientation automatique du regard à la rencontre de nouveaux avatars ;
- le changement dynamique des postures et des animations en fonction du contexte de l'interaction.

Nous avons intégré ces modèles au sein d'une même architecture d'agent capable de fusionner de façon 'réaliste' les actions intentionnelles décidées par les utilisateurs et les comportements non intentionnels produits par ces modèles.

Le conflit d'intentionnalité entre les actions intentionnelles et les comportements automatiques est résolu au niveau de notre architecture par un système d'arbitrage qui favorise la production des actions intentionnelles au dépend de celles qui sont automatiquement produites.

Nous veillons également à ce que les comportements non intentionnels que nous proposons pour enrichir le comportement global des avatars soient interruptibles et de courtes durées pour qu'un utilisateur puisse facilement les intégrer dans son processus de décision.

La dernière étape de nos travaux a consisté à conduire des expérimentations exploratoires destinées à concevoir de nouvelles expérimentations à plus grandes échelles et à fournir des résultats préliminaires destinés à orienter vers un futur protocole expérimental avec des acteurs du métier (comme les sapeurs

pompiers) qui permettra d'analyser d'autres aspects des interactions virtuelles comme les propriétés réflexives lors de la prise de décision et l'émergence de conduites collectives lors de la collaboration.

## 8.2 Conclusions

Nos résultats sont en faveur de l'hypothèse que certaines formes d'interactions sociales sont possibles dans des univers virtuels.

Les interactions humaines étant socialement et culturellement situées. Une expression donnée, qu'elle soit langagière ou corporelle, donne souvent lieu à un ensemble illimité d'interprétations possibles que les humains exploitent pour l'analyse de la situation.

Les références indexicales et la négociation du sens sont également des éléments majeurs et irréductibles dans ces interactions.

Les limites actuelles à la notion de présence sociale dans les univers virtuels nous semblent aujourd'hui essentiellement dues au caractère représentationnel des approches adoptées.

Les interactions virtuelles ont souvent été traitée selon des représentations formelles où les actions produites et le contexte général des interactions sont modélisés par des représentations symboliques qui limitent grandement leurs significations.

Cette réduction importante du champ des interprétations en situations d'interactions peut être compensée par une approche 'située' où l'utilisateur peut saisir par lui même les éléments contextuels de la situation pour leur donner du sens.

En proposant une architecture capable de gérer la production d'expressions indexicales pertinentes par rapport à la construction commune du sens, nous contribuons à rendre le contexte de l'interaction virtuelle suffisamment riche pour permettre la production de couplages socio-cognitifs pertinents par rapport aux objectifs à atteindre.

D'autres difficultés relatives aux interactions virtuelles ont été soulignées dans la littérature (Hindmarsh et al., 1998). Ces difficultés concernent à la fois la non adaptation de certaines interfaces de contrôle par rapport à certains types d'interactions ou encore les difficultés de maintien de la cohérence de l'univers sensible du fait de la réduction du champ visuel.

Nous proposons une solution simple pour compenser la réduction du champ visuel basée sur un positionnement en retrait et en hauteur de la caméra qui permet de voir son avatar de dos.

Malgré le fait que cette disposition ne correspond pas tout à fait à une vision réelle (un humain ne se voit jamais de dos) et qu'elle masque une partie

du champ visuel, nous remarquons que les utilisateurs intègrent rapidement ce nouveau référentiel visuel et arrivent à compenser ses limites par une mobilisation accrue au sein du dispositif. Ils déplacent ainsi leurs avatars pour percevoir les expressions faciales et corporelles de leurs interlocuteurs.

Nous constatons même qu'ils peuvent exploiter spontanément d'autres indices contextuels de l'univers virtuel pour compenser les limites représentationnelles du dispositif comme l'usage des ombres pour inférer la présence des objets dans leurs espaces d'interactions ou encore l'exploitation de la géolocalisation des sons pour identifier la localisation de certains objets.

Les difficultés liées à l'exploitation des expressions déictiques comme la direction du regard ou le pointage par la main dans l'univers virtuel ont également été résolues par un contrôle indépendant de la direction du corps et de la tête.

Ce contrôle indépendant des parties du corps associé au référentiel visuel situé en retrait permet une exploration souple et intuitive de l'espace virtuel en utilisant une simple interface de contrôle comme une manette 3D de jeu.

L'usage des expressions gestuelles déictiques lors des communications apparaît comme un facteur essentiel dans la production de couplages réflexifs entre les différentes actions des acteurs.

L'usage de l'orientation du regard joue également un rôle important dans les situations d'interactions. Une mauvaise distribution du regard entre les différents acteurs d'une conversation ou encore un décalage d'orientation par rapport au regard de son interlocuteur conduit rapidement à la 'destruction' de cette forme de couplage émotionnel et social qui se construit.

Toute la communication humaine étant basée sur la construction du sens, un usage spontané et intuitif de ces expressions déictiques qu'elles soient langagières ou gestuelles introduit donc une dimension supplémentaire aux interactions en univers virtuel et contribue sensiblement à leur caractère situé.

L'usage également du langage naturel comme moyen de communication en univers collaboratif représente une solution naturelle qui contribue à l'engagement social des utilisateurs lors des interactions virtuelles.

A travers l'intonation de leurs voix, les utilisateurs transmettent des indices contextuels supplémentaires (états émotionnels, statuts sociaux, etc.) nécessaires à l'interprétation de la situation et au rétablissement du sens commun.

Cette négociation du sens en univers virtuel peut être également facilitée par les expressions faciales et émotionnelles des avatars.

Les utilisateurs d'un système de réalité virtuelle prennent systématiquement en compte les expressions des personnages virtuels avec lesquels ils interagissent et tentent de donner une interprétation 'cohérente' à leurs interactions virtuelles.

Ces résultats militent en faveur du fait que les technologies de réalité vir-

tuelle sont aujourd'hui capable de produire en univers virtuel des couplages socio-cognitifs suffisamment pertinents pour engager les utilisateurs dans une nouvelle forme d'immersion sociale et émotionnelle.

### 8.3 Perspectives

Plusieurs perspectives nous paraissent intéressantes à explorer.

Une des premières perspectives est l'amélioration du modèle émotionnel. En effet, ce modèle s'appuie sur des théories issues du domaine de la psychologie et intègre plusieurs notions théoriques abordées dans la littérature qu'il est intéressant de formaliser avec l'aide d'experts du métier pour proposer un modèle analytique générique qui peut être implémenté chez des agents autonomes.

Les comportements des avatars peuvent également être améliorés. Nous souhaitons d'une part améliorer la qualité des animations et la richesse des bibliothèques de comportements. Il nous semble intéressant d'exploiter la possibilité de permettre à chaque utilisateur de sélectionner l'ensemble des animations comportementales que son avatar pourrait produire lors des interactions.

D'autre part, il nous semble également intéressant de doter nos personnages virtuels de plus de capacités d'autonomie.

En effet, le cadre d'étude que nous avons envisagé concerne uniquement des représentations des utilisateurs dans l'univers virtuel mais ceci ne représente pas le cadre d'application unique à nos travaux.

Nous pouvons envisager d'autres applications possibles pour notre plateforme de formation qui intègrent à la fois des avatars et des agents autonomes.

Dans ce cas, une extension de l'architecture des agents est nécessaire pour ajouter des possibilités de planification des actions chez les agents autonomes.

De plus, le mécanisme qui permet à un agent de percevoir les actions des autres agents en interaction est fondé, dans notre modèle sur l'envoi de messages.

Ce mécanisme peut être amélioré en fournissant aux agents la capacité de reconnaître les actions, voire les intentions des utilisateurs, ce qui permettrait à ces derniers d'agir plus naturellement.

Le modèle d'interaction sociale que nous proposons suffit pour son utilisation dans le cadre d'un environnement virtuel de formation destiné à la prise de décisions et non aux gestes techniques, mais l'amélioration de ce modèle permettrait une plus grande immersion de l'utilisateur dans son environnement.

Il s'agit donc d'offrir aux utilisateurs un moyen d'effectuer des gestes techniques correspondants à leurs tâches. Cela passe par l'utilisation de périphériques (de navigation ou de préhension) qui permettent à l'utilisateur d'effectuer

ses actions plus naturellement.

D'autres perspectives sont également en cours d'exploration. Comme nous l'avons précisé au début de ce chapitre, nos travaux s'inscrivent dans un objectif plus large : il s'agit de conduire des expérimentations à plus grande échelle pour explorer plusieurs pistes de recherche comme :

1. L'analyse des phénomènes réflexifs lors des interactions virtuelles. Il s'agit d'identifier par des expérimentations plus approfondies les propriétés de l'univers virtuel qui contribuent à l'émergence de couplages réflexifs lors des interactions virtuelles.
2. L'étude de l'émergence de phénomènes collectifs en univers virtuel comme la propagation des états émotionnels et la coordination des processus décisionnels. Il s'agit d'analyser les simulations interactives futures qui seront réalisées avec des acteurs professionnels (les sapeurs pompiers de l'Essonne à Paris en particulier) et les comparer avec des vidéos que nous avons filmés lors de précédentes sessions de formations réelles.

Ce travail est d'ores et déjà planifié dans le cadre d'une thèse de doctorat en cours de préparation dans l'équipe GRIC-IRIT.

Cinquième partie

**Annexes**



# Annexe A

## Application

### A.1 Implémentation

L'objectif de notre plate-forme de réalité virtuelle est de permettre à des utilisateurs distribués géographiquement d'interagir et de coopérer (en intégrant une dimension sociale et émotionnelle) dans un environnement de réalité virtuelle collaboratif.

Ce dispositif met facilement en relation et en interaction des personnes distantes et offre des moyens de collaboration à travers un réseau de communication.

Nous utilisons pour cette fin la suite logicielle Virtools qui permet la réalisation et le déploiement de contenus interactifs 3D (figure A.1). Les outils Virtools sont basés sur une interface de développement graphique qui permet d'assigner des comportements aux objets 3D et de les rendre interactifs.

Cette suite intègre un ensemble de technologies pour la visualisation et l'interaction 3D Temps réel :

**Application de création (An authoring application)** Il s'agit d'une application permettant de créer simplement des contenus 3D interactif. Il est possible d'intégrer la majorité des standards multimédia que ce soit des objets 3D, des animations, des images et des sons afin de les rendre vivant grâce à la technologie comportementale de Virtools.

Virtools ne permet pas de modéliser des objets 3D, par contre il permet facilement de créer des caméras, lumières, courbes, éléments d'interface ainsi que des frames 3D.

**Un moteur de comportement (A behavioral engine)** Un comportement est simplement une description d'une action d'un certain élément dans un environnement. Virtools offre une collection de comportements réutilisable qui permet de créer tout type de contenu, sans une seul ligne de code, à partir de l'interface graphique de l'éditeur de schéma.



FIG. A.1 – Aperçu de l'atelier graphique Virtools.

Ces comportements (ou 'Behaviors') sont représentés dans l'interface de Virtools par des modules graphiques ('Building Blocs') qui comportent des entrées et des sorties permettant de les interconnecter.

Ils sont organisés au sein de catégories clairement définies (3D Transformation, Cameras, Collisions, Controllers, Lights, Shaders, Sounds, etc.). Pour développer des comportements complexes, il existe en complément un script 'Virtools Scripting Language' (VSL) qui autorise les programmeurs à accéder au Virtools SDK.

**Un moteur de rendu (A render engine)** Virtools dispose également d'un moteur de rendu qui permet de dessiner les images visibles dans l'interface 3D.

**Un Web player** Les contenus Virtools peuvent être visualisé sur Internet et distribués sur le Web grâce à un Web player gratuit.

**Un kit de développement logiciel (Software Development kit)** Virtools inclu un kit de développement logiciel qui permet d'accéder à certaine partie des processus de comportement et de rendu.

Avec le SDK, on peut créer de nouveaux comportements (DLL), modifier les opérations des comportements existants, écrire de nouveau outils afin d'importer ou d'exporter de nouveau format, et enfin de modifier ou remplacer le moteur de rendu.

## A.2 Déploiement de l'application en réseau

Pour assurer la coordination de plusieurs utilisateurs distants, nous nous sommes basés sur un module 'multi-users' de Virtools qui apporte une di-

mension supplémentaire aux applications immersives en mettant facilement en relation et en interaction des personnes distantes dans des univers virtuels.

Notre dispositif est déployé sur un réseau local avec un serveur central pour assurer la cohérence des échanges entre les participants (figure A.2).

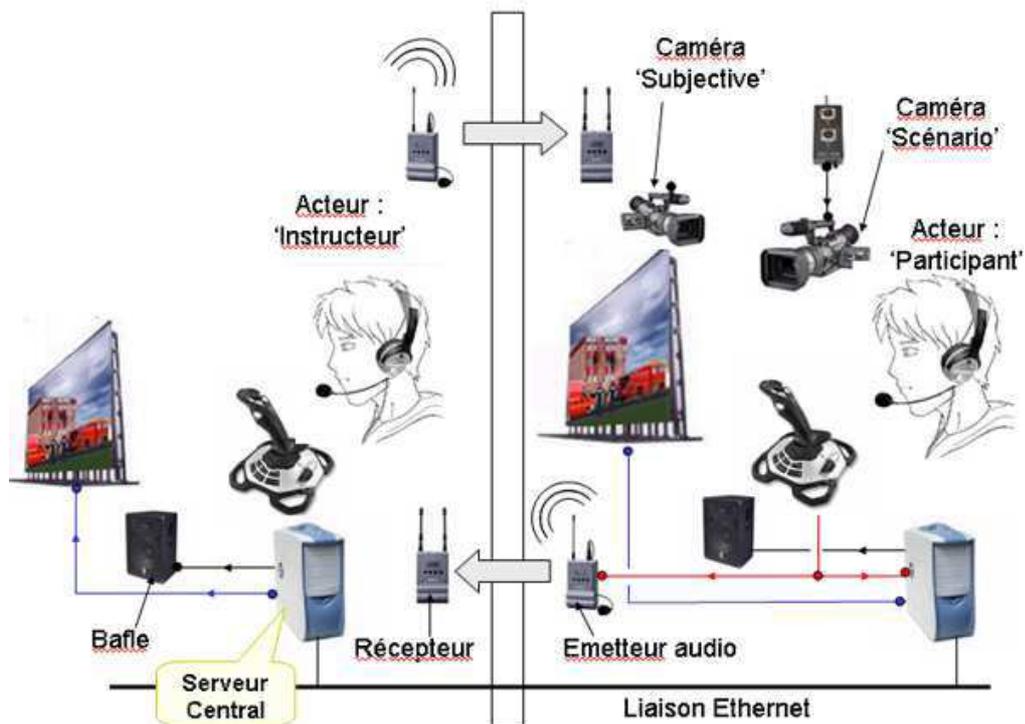


FIG. A.2 – Schéma expérimental pour une session collaborative en univers virtuel.

Le serveur définit une session collaborative permettant à plusieurs participants de la rejoindre. Chaque participant qui rejoint la session collaborative reçoit une mise à jour de l'environnement de réalité virtuelle avec les positions courantes des objets de la scène (positions des personnages, véhicules, bâtiments, etc.).

Les communications langagières entre les participants sont transmises à partir des microphones vers des émetteurs radio.

Chaque participant est filmé par deux caméras différentes :

- une caméra subjective permet de filmer les expressions émotionnelles et corporelles de chaque participant.
- Une caméra 'scénario' permet d'analyser tous les détails des interactions et l'ensemble des actions produites.



Sixième partie  
**Bibliographie**



# Bibliographie

- S. Allen. A concern-centric society-of-mind approach to mind design. *Proceedings of the AISB'00 Symposium on How To Design A Functioning Mind., Birmingham, England, 2000.*
- P. Amiel. *Ethnométhodologie Appliquée - éléments de sociologie praxéologique.* Les Presses du LEMA, 2004.
- E. André, M. Klesen, P. Gebhard, S. Allen, and T. Rist. Integrating models of personality and emotions into lifelike characters. *A. Paiva (Ed.) Affect in Interactions Towards a New Generation of Interfaces., 1999.*
- M. Argyle. *Bodily communication.* London : Methuen, 1975.
- R. C. Arkin. Motor schema-based mobile robot navigation. *The International Journal of Robotics Research*, 8(4) :92–112, 1989.
- R.C. Arkin. *Behavior-based robotics.* MIT Press, 1998.
- J. R. Averill. *A constructionist view of emotion*, volume 1. Emotion : Theory, research, and experience. NewYork : Academic Press, 1980.
- G. Ball and J. Breese. *Emotion and personality in a conversational character.* In : Cassell, J., Sullivan, J., Prevost, S., and Churchill, E. (eds.) : Embodied Conversational Agents, Cambridge, MA :MIT Press, 2000.
- Y. Bar-Hillel. *Indexical expressions*, volume 63 :359-379. *Mind*, 1954.
- M. Barbuceanu and W.K Lo. Conversation oriented programming in cool : current state and future directions. *Agents'99 : Working Notes of the Workshop on Specifying and Implementing Conversation Policies. Seattle (USA), 1999.*
- J. Bates. The role of emotions in believable agents. *Communications of the ACM*, 37(7) :122–125, 1994.
- R. A. Brooks. A robust layered control system for a mobile robot. *IEEE Journal of Robotics and Automation*, RA-2/1, pages 14–23, 1986.

- R. A. Brooks. *Cambrian Intelligence : the early history of the new AI*. MIT Press, 1999.
- P. Brézillon. Context in problem solving : A survey. *The Knowledge Engineering Review*, 14(1) :1–34, 1999.
- M. Büscher, M. Christensen, K. Gronbak, P. Krogh, P. Morgensen, D. Shapiro, and P. P. Orbak. Collaborative augmented reality environments : Integrating vr, working materials and distributed work. *Proceedings of CVE : Collaborative Virtual Environments, San Francisco, CA : ACM*, pages 47–56, 2000.
- G. Burdea and P. Coiffet. *La réalité virtuelle*. Hermès, Paris, 1993.
- J. N. Cappella. *The Management of Conversations*, volume pp. 393-435. 1985.
- J. Cassell, T. Bickmore, M. Billingham, L. Campbell, K. Chang, H. Vilhjálmsón, and H. Yan. An architecture for embodied conversational characters. In *First Workshop on Embodied Conversational Characters - Tahoe City (USA)*., 1998.
- G. Chicoisne and S. Pesty. The puppeteer behind the avatar. *SIGGRAPH 2000 - Sketches and Applications. ACM Press. Nouvelle-Orléans (USA)*., 2000.
- N. Chovil. Discourse-oriented facial displays in conversation. *Research on Language and Social Interaction*, 25 :163–194, 1992.
- B. Conein. Arguments ethnométhodologiques, problèmes d'épistémologie en sciences sociales. CEMS. EHESS. CNRS. Paris, 1984.
- S. Darcy, J. Dugdale, M. El Jed, N. Pallamin, and B. Pavard. Virtual reality - story building - story telling. *International Conference on Virtual Storytelling, Toulouse, France.*, November 20-21 2003.
- C. Darwin. *The expression of the emotions in man and animals*. London : HarperCollins, 1872/1998.
- P. David and D. Lourdeaux. A simulator using virtual reality techniques for training driver to manual interventions on the tracks. *WCRR'01 World conference on railway research. Poster presentation.*, 2001.
- R. J. Davidson. *On emotion, mood, and related affective constructs*. New York : Oxford University Press, 1994.
- A. K. Dey and G. D. Abowd. Towards a better understanding of context and context-awareness. *CHI 2000 Workshop on The What, Who, Where, When, Why and How of Context-Awareness, April 1-6*, 2000.

- A. Egges, S. Kshirsagar, and N. Magnenat-Thalmann. Generic personality and emotion simulation for conversational agents. *Computer Animation and Virtual Worlds.*, 2004.
- P. Ekman. An argument for basic emotions. *Cognition and Emotion*, 6 (3/4) : 169–200, 1992.
- M. El Jed, N. Pallamin, J. Dugdale, and B. Pavard. Modelling character emotion in an interactive virtual environment. In *AISB 2004 Convention : Motion, Emotion and Cognition* , Leeds, UK, page 9. The society for the study of Artificial Intelligence and the Simulation of Behaviour, 29 mars-01 avril 2004.
- M. S. El-Nasr, J. Yen, and T. R. Ioerger. Flame - fuzzy logic adaptive model of emotions. *International Journal of Autonomous Agents and Multi Agents Systems*, 3(3) :1–39, 2003.
- J. Ferber. *Les systèmes multi-agents : Vers une intelligence collective*. Inter-Editions, 1995.
- I. Ferguson. *TouringMachine : An Architecture for dynamic, Rational, Mobile Agents*. PhD thesis, Clare Hall, University of Cambridge, Great Britain, 1992.
- R. E. Franken. *Human Motivation*. Belmont, CA : Brooks/Cole Publishing Company.3rd ed., 1994.
- N. H. Frijda. *The laws of emotion*, volume 43(5), 349-358. American Psychologist, 1988.
- N. H. Frijda. *Varieties of affect : Emotions and episodes, moods, and sentiments*. New York : Oxford University Press, 1994.
- P. Fuchs. Les interfaces de réalité virtuelle. *Association des Journées Internationales de l'Informatique de Montpellier-District, Montpellier*, 1996.
- P. Fuchs, G. Moreau, B. Arnaldi, J. M. Burkhardt, A. Chauffaut, S. Coquillart, S. Donikian, T. Duval, J. Grosjean, F. Harrouet, E. Klinger, D. Lourdeaux, D. Mellet d'Huart, A. Paljic, J. P. Papin, P. Stergiopoulos, J. Tisseau, and I. Viaud-Delmon. *Le Traité de la réalité virtuelle*, volume 2 of ISBN 2-911762-47-9 et 2-911762-48-7. Les Presses de l'Ecole des Mines de Paris, deuxième édition edition, septembre 2003.
- A. Furnham. *Language and Personality*, volume 73-95. Handbook of Language and Social Psychology. Chichester, England UK :John Wiley & Sons., 1990.
- P. E. Gallaher. Individual differences in nonverbal behavior : Dimensions of style. *Journal of Personality and Social Psychology*, 63(1)(133-145), 1992.

- H. Garfinkel. *Studies in Ethnomethodology*. Prentice-Hall, New York, 1967.
- F. Giunchiglia and P. Bouquet. *Introduction to contextual reasoning. An Artificial Intelligence Perspective*, volume 3. NBU Press, Sofia (Bulgaria), 1997.
- A. Goker and H. I. Myrhaug. User context and personalization. *ECCBR Workshop on Case Based Reasoning and Personalisation, invited paper, Aberdeen, UK*, 2002.
- J. Gratch. Emile : marshalling passions in training and education. *Proceedings of the Fourth International Conference on Intelligent Agents, Barcelona, Spain.*, 2000.
- J. Gratch and S. Marsella. Tears and fears : Modeling emotions and emotional behaviors in synthetic agents. *Proceedings of the 5th International Conference on Autonomous Agents, Montreal, Canada.*, 2001.
- J. Gratch, J. Rickel, E. André, J. Cassel, E. Petajan, and N. Badler. Creating interactive virtual humans : Some assembly required. *IEEE July/August 2002*, 2002.
- Z. Guessoum. *Un environnement opérationnel de conception et de réalisation de systèmes multi-agents*. PhD thesis, LAFORIA, Université Paris VI, Mai 1996.
- R. V. Guha. *Contexts : A Formalization and Some Applications*. PhD thesis, Stanford, 1991.
- J. Gustafson, N. Lindberg, and M. Lundeberg. The august sopken dialog system. *Proceedings of Eurospeech'99, Budapest, Hungary*, 1999.
- E. T. Hall. *The Hidden Dimension*. Garden City, N.Y. : Doubleday, 1966.
- E. Hatfield, J. T. Cacioppo, and R. L. Rapson. *Emotional contagion*. Paris/Cambridge : Editions de la Maison des Sciences de l'Homme and Cambridge University Press (jointly published), 1994.
- R. Hill, J. Gratch, S. Marsella, J. Rickel, W. Swartout, and D. Traum. Virtual humans in the mission rehearsal exercise system. *Special issue on embodied conversational agents. KI*, 2003.
- J. Hindmarsh, M. Fraser, C. Heath, S. Benford, and C. Greenhalgh. Fragmented interaction : establishing mutual orientation in virtual environments. In *ACM conference on Computer-Supported Cooperative Work (CSCW'98)*, pages 217–226. ACM Press, November 1998.
- E. Hollnagel. Is affective computing next term an oxymoron? *International Journal of Human-Computer Studies*, 59 Issues 1-2 :65–70, July 2003.

- E. Hollnagel. *Human reliability analysis : Context and control*. London : Academic Press, 1993.
- M. Johns and B.G. Silverman. How emotions and personality effect the utility of alternative decisions : a terrorist target selection case study. *Tenth Conference on Computer Generated Forces and Behavioral Representation.*, 2001.
- T. U. St Julien and C. D. Shaw. Firefighter command training virtual environment, 2003.
- M. Kaenampornpan, E. O'Neill, V. Kostakos, and A. Warr. Classifying context classifications : an activity theory perspective. *2nd UK-UbiNet Workshop, 5-7th May 2004, University of Cambridge, UK*, 2004.
- L. Karsenty and B. Pavard. Différents niveaux d'analyse du contexte dans l'étude ergonomique du travail collectif. In *Réseaux*, number 85, pages 73–99, 1997.
- A. Kendon. Some functions of gaze direction in social interaction. *Acta Psychologica*, 26 :22–63, 1967.
- A. Kendon. *How gestures can become like words*. Crosscultural perspectives in nonverbal communication. Potyatos, F. (ed), 1988.
- P. R. Kleinginna and A. M. Kelinginna. A categorized list of emotion definitions, with suggestions for a consensual definition. *Motivation and Emotion*, 5(4) :345–379, 1981.
- M. L. Knapp. *Nonverbal Communication in Human Interaction (2nd ed.)*. New York, NY : Holt, 1978.
- M. Krueger. *Artificial Reality II*. Addison-Wesley, 1991.
- S. Kshirsagar and N. Magnenat-Thalmann. A multilayer personality model. *Proceedings of 2nd International Symposium on Smart Graphics*, (1) :107–115, 2002.
- M. La France and C. Mayo. Racial difference in gaze direction in social interaction. *Journal of Personality and Social Psychology*, 33 :547–552, 1976.
- R. S. Lazarus. *Emotion and adaptation*. New York : Oxford University Press, 1991.
- D. Le Breton. *Les Passions ordinaires : Anthropologie des émotions*. Armand Colin, 1998.
- D. G. Leathers. *Successful Nonverbal Behavior - Principles & Applications (3rd ed.)*. Needham Heights, MA : Allyn & Bacon, 1997.

- Y. Lecerf. Pratiques de formation (analyses), ethnométhodologies. In *Pratiques de formation*, volume 11-12, 1985.
- J. Leplat. *L'analyse du travail en psychologie ergonomique (Tome I)*. Toulouse : Octarès Éditions, 2e éd, 2001.
- J. C. Lester, B. A. Stone, and G. D. Stelling. Lifelike Pedagogical Agents for Mixed-Initiative Problem Solving in Constructivist Learning Environments. *User Modeling and User-Adapted Interaction*, 9(1-2) :1–44, 1999.
- J.C. Lester, S.G. Stuart, C.B. Callaway, J.L. Voerman, and P.J. Fitzgerald. Deictic and emotive communication in animated pedagogical agents. In *S.Prevest J.Cassell, J.Sullivan and E.Churchill, editors, Embodied Conversational Characters*. MITpress, Cambridge, MA, 2000.
- D. Lourdeaux. *Réalité Virtuelle et Formation : Conception d'Environnements Virtuels Pédagogiques*. PhD thesis, École des Mines de Paris, 2001.
- P. Maes. The dynamics of action selection. *Proceedings of the International Joint conference on Artificial Intelligence, IJCAI-89*, 1989.
- S. Marsella and J. Gratch. Modeling the interplay of emotions and plans in multi-agent simulations. *Proceedings of the 23rd Annual Conference of the Cognitive Science Society. Edinburgh, Scotland.*, 2001.
- S. Marsella and J. Gratch. A step towards irrationality : using emotions to change belief. *Proceedings of the Third International Joint Conference on Autonomous Agents and Multi-Agent Systems. Bologna, Italy.*, 2002.
- J. Mc Carty. Notes on formalizing context. *Proceedings of the thirteenth international joint conference on artificial intelligence*, 1993.
- J. Mc Carty and P. J. Hayes. *Some philosophical problems from the standpoint of artificial intelligence*. D. Michie (ed), Machine Intelligence 4, American Elsevier, New York, NY, 1969.
- R. R. McCrae and O. P. John. An introduction to the five-factor model and its applications. *Special Issue : The five-factor model : Issues and applications. Journal of Personality :60*, pages 175–215, 1992.
- D. McNeill. *Hand and Mind : What Gestures Reveal about Thought*. Chicago, IL/London, UK : The University of Chicago Press, 1992.
- M. Meehan, B. Insko, M. Whitton, and F. P. Brooks Jr. Physiological measures of presence in stressful virtual environments. *Proceedings of SIGGRAPH 2002, San Antonio, Texas. also in ACM Transactions on Graphics. (In Press)*, 2002.

- A. Mehrabian. Orientation behaviors and nonverbal attitude communication. *Journal Of Communication*, 16 :324–332, 1967.
- A. Mehrabian. *Non verbal communication*. Aldine-Atherton, 1972.
- A. Mehrabian. *Silent Messages*. Belmont CA : Wadsworth (2nd Edition), 1981.
- D. Mellet-d’Huart. *De l’intention à l’attention. Contributions à une démarche de conception d’environnements virtuels pour apprendre à partir d’un modèle de l’(én)action*. PhD thesis, Université du Maine, Décembre 2004.
- D. Mellet-d’Huart and G. Michel. *Réalité virtuelle et apprentissage*. Hermes - Collection 'Traité IC2 Information Commande Communication', 2005.
- M. Minsky. *La société de l’esprit*. InterEdition, Paris, 1988.
- J. Müller and M. Pischel. The agent architecture interrapp : Concept and application. Technical Report RR-93-26, DFKI Saarbrücken, 1993.
- R. M. Neese. Evolutionary explanations of emotions. *Human Nature*, 1990.
- J. L. Nespoulos and A. R. Lecours. Gestures : nature and function. *The biological foundations of gestures : motor and semiotic aspects*, pages 49–62, 1986.
- A. Newell. *Unified Theories of Cognition*. Harvard University Press, 1990.
- J. Newhagen and B. Reeves. *Emotion and memory responses to negative political advertising*. Televisions and political advertising : Psychological processes. Hillsdale, NJ :Lawrence Erlbaum., 1991.
- J. Newhagen and B. Reeves. This evening’s bad news : Effects of compelling negative television news images on memory. *Journal of Communication*, 42 :25–41, 1992.
- K. Oatley and P. N. Johnson-Laird. Towards a cognitive theory of emotions. *Cognition and Emotion*, 1 (1) :29–50, 1987.
- A. Ortony. *On Making Believable Emotional Agents Believable*. Trappl R., et al, 2003.
- A. Ortony and T. J. Turner. What’s basic about emotions. *Psychological Review*, 97(3) :315–331, 1990.
- A. Ortony, G. L. Clore, and A. Collins. *The Cognitive Structure of Emotions*. Cambridge University Press., 1988.
- J. Panksepp. A critical role for 'affective neuroscience' in resolving what is basic about basic emotions. *Psychological Review*, 99 (3) :554–560, 1992.

- B. Pavard and J. Dugdale. From representational intelligence to contextual intelligence in the simulation of complex social system. *CASOS conference. Pittsburg.*, 2002.
- D. Pelé, G. Breton, F. Panaget, and S. Loyson. Let's find a restaurant with nestor a 3d embodied conversational agent on the web! AAMAS2003 Conference, 2003.
- C. Pelachaud, V. Carofiglio, B. De Carolis, and F. De Rosis. Embodied contextual agent in information delivering application. *First International Joint Conference on Autonomous Agents et Multi-Agent Systems (AAMAS), Bologna, Italy*, July 2002.
- K. Pimentel and K. Teixeira. *La Réalité Virtuelle ... de l'autre coté du miroir*. Addison-Wesley, 1994.
- P. Pirjanian. Behavior coordination mechanisms - state-of-the-art. Technical report, Research Report Robotics Research Laboratory, University of Southern California, October 1999.
- P. Quéau. *Le Virtuel*. Champ Vallon, Seyssel, 1993.
- Q. Querrec. *Les systèmes multi-agents pour les environnements virtuels de formation. Application à la sécurité civile*. PhD thesis, Université de Bretagne Occidentale, Octobre 2002.
- R. Querrec and P. Chevaillier. Virtual storytelling for training : An application to fire fighting in industrial environment. *BALLET O. & SUBSOL G. & TORQUET P. Eds*, 2001.
- R. Querrec, C. Buche, E. Maffre, and P. Chevaillier. Securevi : Virtual environment for fire fighting training. *RICHIR S., RICHARD P. & TARAVEL B. Eds. Laval Virtual - 5th Virtual reality international conference - Proceedings 2003.*, 2003.
- B. Reeves and C. Nass. *The media equation : How people treat computers, television and new media like real people and places*. New York : Cambridge University Press, 1996.
- B. Reeves, J. Newhagen, E. Maibach, M. D. Basil, and K. Kurz. Negative and positive television messages : Effects of message type and message content on attention and memory. *American Behavioral Scientist*, 34 :679–694, 1991.
- J. H. Rety, J. C. Martin, C. Pelachaud, and N. Bensimon. Coopération entre un hypermédia adaptatif éducatif et un agent pédagogique. *Actes de H2PTM'03, créer du sens à l'heure numérique, Saint-Denis*, 2003.

- N. Richard. *Description de comportements d'agents autonomes évoluant dans des mondes virtuels*. PhD thesis, Ecole Nationale Supérieure des Télécommunications, Paris, 2001.
- J. Rickel and W. Johnson. Animated agents for procedural training in virtual reality : Perceptionperception, cognition, and motor control. *Applied Artificial Intelligence.*, 1999.
- J. Rickel, J. Gratch, R. Hill, S. Marsella, and W. Swartout. Steve goes to bosnia : Towards a new generation of virtual humans for interactive experiences. *AAAI Spring Symposium on Artificial Intelligence and Interactive Entertainment Stanford University*, March 2001.
- J. Rickel, S. Marsella, J. Gratch, R. Hill, D. Traum, and W. Swartout. Toward a new generation of virtual humans for interactive experiences. *IEEE Intelligent Systems July/August 2002*, 2002.
- I. J. Roseman, A. A. Antoniou, and P. E. Jose. Appraisal determinants of emotions :constructing a more accurate and comprehensive theory. *Cognition and Emotion*, 10(3) :241–277, 1996.
- S. J. Russell. *Rationality and intelligence.*, volume 94. p.57-77. Artificial Intelligence, 1997.
- H. Sacks, E. Schegloff, and G. Jefferson. *A simplest systematics for the organisation of turn-taking in conversation*, volume 50 : 696-735. Language, 1974.
- P. Salembier. Cognition(s) : Située, distribuée, socialement partagée, etc. Technical report, ENS, 1996.
- A. E. Schefflen and N. Ashcraft. *Human Territories. Hew we Behave in Space-Time*. Englewood Cliffs (N.J.) : Prentice Hall, 1976.
- K. Scherer. *Criteria for emotion-antecedent appraisal : A review*, volume Cognitive perspectives on emotion and motivation (pp. 89-126). Dordrecht : Kluver Academic Publishers., 1988.
- K. R. Scherer. *Personality Markers in Speech*. Social Markers in Speech, Cambridge University Press, 1979.
- H. Selye. *History and present status of the stress concept*. In L. Goldberger and S. Breznitz, eds. Handbook of Stress : Theoretical and Clinical Aspects. New York : The Free Press., 1982.
- Y. Shoham. *Agent-oriented programming*, volume 60, p.51-92. Artificial Intelligence, 1993.

- R. A. Shweder. *You're not sick, you're just in love : Emotions as an interpretive system*. New York : Oxford University Press, 1994.
- D. Sperber and D. Wilson. *La pertinence, Communication et cognition*. Paris : Minit. (A. Gerschendorf & D. Sperber, Trad.) (Edition originale 1986)., 1989.
- L. Steels. The artificial life roots of artificial intelligence. In MIT Press, editor, *Artificial Life Journal*, volume 1, page 75 110, 1994.
- L. A. Suchman. *Plans and situated actions : The problem of human-machine communications*. Cambridge, UK : Cambridge University Press, 1987.
- M. R. Tazari, M. Grimm, and M. Finke. *Modelling User Context*. Jacko, Julie A. (Ed.), 2003.
- J. Theureau. Cours des uv sc 23 (théories et méthodes d'analyse de l'action & ingénierie) et sh 12 (anthropologie cognitive & ingénierie). UTC/SHT, Compiègne (346 p.) (nouvelle édition remaniée), 1999.
- G. Thomas. *Environnements virtuels urbains : modélisation des informations nécessaires à la simulation de piétons*. PhD thesis, Université de Rennes I., 1999.
- J. Tooby and L. Cosmides. The past explains the present : Emotional adaptations and the structure of ancestral environments. *Ethology and Sociobiology*, 11 :407–424, 1990.
- T. Tyrrell. *Computational mechanisms for action selection*. PhD thesis, University of Edinburgh, 1993.
- G. Vanderheiden and J. Mendenhall. Use of a two-class model to analyse applications and barriers to the use of virtual reality by people with disabilities. *Presence - MIT Press*, 1994.
- T. Winograd and F. Flores. *Understanding Computers and Cognition : A New Foundation for Design*. Ablex Publishing Corporation, 335 Chesnut Street, Norwood, New Jersey, 1986.
- M. Wooldridge and N. R. Jennings. *Agent theories, architectures, and languages*, volume p.1-22. Intelligent Agents, Springer Verlag, 1995.
- D. Zillmann. *Television viewing and Physiological arousal*, volume pp 103-133. Responding to the screen : Reception and reaction processes. Hillsdale, NJ : Lawrence Erlbaum Associates., 1991.

