



UNIVERSITE KASDI MERBAH OUARGLA

**FACULTE DES SCIENCES ET SCIENCES DE
L'INGENIEUR**

**DEPARTEMENT DES MATHEMATIQUES ET
D'INFORMATIQUE**

MAGISTER

Spécialité : INFORMATIQUE

Option : Informatique et Communication Electronique

Par : BENKHEROUROU Chafika

Thème

**Le langage HumanMl comme support de
communication entre agents**

Soutenu publiquement le : 25 Mai 2009

Devant le jury composé de :

M^{me}	F.Z. LAALLAM	M.C	Université de Ouargla	Président
M	A.BILAMI	M.C	Université de Batna	Examineur
M	O.KEZZAR	M.C	Université de Biskra	Examineur
M	B. BELATTAR	M.C	Université de Batna	Rapporteur

Remerciements

Je tiens à remercier toutes les personnes qui ont contribué de manière directe ou indirecte à l'aboutissement de ce travail :

En premier lieu, je remercie très sincèrement Monsieur Brahim BELATTAR sans qui ce sujet passionnant ne m'aurait pas été proposé. Je le remercie pour son encadrement et surtout pour l'aide précieuse qu'il m'a apportée tout au long du travail.

Je remercie Madame Fatima Zohra LAALLAM, Maître de conférences à l'université de Ouargla d'avoir bien voulu accepter la présidence de ce Jury. Je remercie également Monsieur Azzedine BILAMI Maître de Conférence à l'Université de Batna et Monsieur Okba KAZZAR Maître de Conférence à l'Université de Biskra pour l'honneur qu'ils me font en acceptant de juger mon travail.

Je n'oublie pas avant de conclure ma page de remerciements à adresser un merci tout particulier à Mr HARROUZ Abdelhakim qui a contribué à ce que nos études en post-graduation se déroulent dans une atmosphère des plus agréables et qui nous a aidé dans la réalisation de démarches administratives fastidieuses.

Un remerciement particulier à mon mari Hakim pour sa patience et son soutien inconditionnel sans lequel je n'aurais jamais pu terminer mon travail, et à mes enfants Rym et Mahdi qui ont su me supporter pendant mes études.

Je termine par un grand remerciement à mes parents, à mes frères et soeurs pour leurs encouragements et pour leur soutien moral et surtout ma soeur Sonya qui m'a soutenue à sa manière.

Résumé :

Avec la prolifération d'internet, la communication a connu une véritable progression. Ainsi, un humain peut interagir avec un autre humain qui peut être d'une culture ou d'une langue différente par (email, Chat, environnement virtuel...) ou avec un agent artificiel dans le cadre des applications de la réalité virtuelle, l'IHM, l'e-learning, l'e-commerce...

Cependant, plusieurs informations liées aux caractéristiques humaines tel que la représentation des émotions, des gestes, de la posture du corps, du regard et des expressions faciales ne sont pas pris en considération lors de la transmission des messages. Dans ce cas, le risque d'avoir des malentendus ou des conflits est majeur.

Pour éviter tout malentendu, l'utilisation du langage HumanML permet d'inclure d'importantes informations liées aux caractéristiques humaines. De telles informations permettent une transmission plus précise et plus individuelle des messages.

Dans notre travail, nous proposons l'utilisation du langage humanML dans un système multi-agent dans le but d'humaniser le web en permettant aux humains de communiquer avec les machines d'une manière plus facile et moins ambiguë. Nous proposons les différents cas d'utilisation sur le web.

MOTS-CLES :

HumanML, agent conversationnel, SMA, communication, IHM, langage non-verbal, XML schéma.

ABSTRACT :

With the proliferation of Internet, the communication knew a true progression. Thus, human can interact with another human which can be of a different culture or a language by (email, Chat, virtual environment) or with an artificial agent within the applications of virtual reality, HCI, e-learning, e-business.

However, several information related to the human characteristics such as the representation of the emotions, the gestures, the posture of the body, the gaze and the facial expressions are not taken into account during the transmission of the messages. In this case, the risk to have misunderstandings or conflicts is major.

To avoid any misunderstanding, the use of the HumanML language makes it possible to include significant information related to the human characteristics. Such information allows a more precise and more individual transmission of the messages.

In this report, we propose the use of the language humanML in a multi-agent system with an aim of humanizing the Web while allowing human to communicate with the machines in an easier and less ambiguous way. We propose the various cases of use on the Web.

KEY WORDS :

HumanML, conversational agent, SMA, communication, HCI, non verbal language , XML schema.

ملخص:

مع انتشار شبكة الإنترنت شهدت الاتصالات تطورا مذهلا . حيث أصبح بإمكان أي شخص ان يتفاعل مع شخص آخر و الذي قد يكون ذا ثقافة أو لغة مختلفة عن طريق (البريد الإلكتروني ، والدرشة ، بيئة افتراضية) أومع وكيل برمجي في إطار تطبيقات الواقع الافتراضي، ال HCI ، التعلم الإلكتروني، التجارة الإلكترونية...

إلا أن العديد من المعلومات التي تخص الإنسان كالعواطف ، الحركات ، هيئة الجسم ، النظرات وتعابير الوجه لا تؤخذ بعين الاعتبار خلال الإتصال بواسطة الرسائل الالكترونية. وفي هذه الحالة ، احتمال حدوث سوء الفهم أو النزاعات وارد.

لتجنب أي سوء فهم ، يمكن استخدام لغة HumanML التي لها القدرة على تمثيل المميزات الخاصة بالأفراد. هذه المعلومات يمكنها إعطاء خصوصية ووضوح أكبر للرسائل المرسلة .

في عملنا هذا، نقترح استخدام لغة HumanML كسند لنظام متعدد الوكلاء من أجل إضفاء الطابع الإنساني على الانترنت عن طريق السماح للأشخاص بالتواصل مع الآلات بطريقة أسهل وأقل غموضا. كما نقترح عدة تطبيقات تعتمد على هذه اللغة على شبكة الإنترنت.

مفاتيح

HumanML، وكيل مُتحدِّث ، SMA ، الاتصالات ، HCI ، لغة غير لفظية، XML-Schema

Table des matières

REMERCIEMENTS	2
TABLE DES MATIERES	5
LISTE DE FIGURES.....	7
LISTE DE TABLEAUX	8
SIGLES ET ACRONYMES.....	9
CHAPITRE 1 : INTRODUCTION	10
1.1 PROBLEMATIQUE :	10
1.2 OBJECTIFS :	13
1.3 ORGANISATION DU MEMOIRE :	13
CHAPITRE 2 : LES SYSTEMES MULTI-AGENTS	16
2.1 DEFINITION D'UN AGENT :	16
2.2 PROPRIETES D'UN AGENT :	18
2.3 LES SYSTEMES MULTI-AGENTS (SMA):	20
2.4 LA TYPOLOGIE DES AGENTS :	21
2.5 LES INTERACTIONS ENTRE AGENTS :	24
2.6 ARCHITECTURE D'AGENTS :	25
2.6.1 <i>Structure interne</i> :	25
2.6.2 <i>Fonctionnement</i> :	27
2.7 L'HETEROGENEITE DES AGENTS :	28
CHAPITRE 3 : LA COMMUNICATION ENTRE AGENTS	31
3.1 DEFINITION DE LA COMMUNICATION :	31
3.2 LE PROCESSUS DE COMMUNICATION:	32
3.3 LA COMMUNICATION ENTRE AGENTS LOGICIELS :	34
3.3.1 <i>La théorie des actes de langages</i> :	34
3.3.2 <i>Les langages de communication entre agents</i> :	37
3.3 LA COMMUNICATION HUMAIN-AGENT :	40
3.4 ENVIRONNEMENTS DES RASSEMBLEMENTS D'AGENTS HUMAINS ET ARTIFICIELS :	41
3.5 LA COMMUNICATION CHEZ LES HUMAINS :	42
3.6 LES LIMITES DES ACL :	48
CHAPITRE 4 : LES AGENTS CONVERSATIONNELS.....	51
4.1 NOTIONS SUR LES AGENTS CONVERSATIONNELS :	51
4.2 AGENTS CONVERSATIONNELS ET AVATARS :	54
4.3 HISTORIQUE ET EVOLUTION DES ACA:	55
4.4 DOMAINES D'APPLICATION ET DE RECHERCHE:	56
4.5 LES ACA DANS LA COMMUNICATION AVEC LES HUMAINS :	57
4.6 ETAT DE L'ART DES AGENTS CONVERSATIONNELS :	58
4.7 LANGAGES DE REPRESENTATION POUR ECAS :	68
CHAPITRE 5 : LE LANGAGE HUMANML	72
5.1 PRESENTATION DU LANGAGE HUMANML :	72
5.2 BREF HISTORIQUE DE HUMANML :	73
5.3 LES DOMAINES D'APPLICATION DE HUMANML :	74
5.4 LES OBJECTIFS DU LANGAGE HUMANML :	75
5.5 COMMUNICATION HUMAIN-HUMAIN SUR LE WEB:	76
5.6 HUMANML ET LE WEB SEMANTIQUE :	78

5.7 LE LANGAGE XML ET LE XML SCHEMA:	81
5.8 RDF ET RDF SCHEMA :	84
5.9 LES STRUCTURES FONCTIONNELLES DU LANGAGE HUMANML :	85
5.10 LES CARACTERISTIQUES DE HUMANML :	86
5.11 LES COMPOSANTS DE HUMANML:	87
5.11.1 L'élément racine :	89
5.11.2 Les types complexes:	89
5.11.3 Les types simples :	91
5.11.4 Les groupes d'attributs :	92
5.12 LE LANGAGE HPCDML:	93
CHAPITRE 6: L'UTILISATION DU LANGAGE HUMANML DANS LES SMA	95
6.1 HUMANML DANS UN SYSTEME DE CHAT :	96
6.1.1 Description du système :	101
6.1.2 L'architecture du système de Chat :	102
6.1.3 Le fonctionnement du système:	104
6.2 L'UTILISATION DU LANGAGE HUMANML DANS LES IHM :	105
6.2.1 Agents présentateurs d'informations sur internet :	106
6.2.2 Agents conversationnels dans l'e-commerce:	109
CHAPITRE 7 : CONCLUSION ET PERSPECTIVES	119
7.1 CONCLUSION :	119
7.2 PERSPECTIVES :	121
BIBLIOGRAPHIE.....	123

Liste de figures

FIGURE 1.	REPRESENTATION D'UN AGENT SELON RUSSEL ET NORVIG	16
FIGURE 2.	PRINCIPALES CARACTERISTIQUES D'UN AGENT.....	19
FIGURE 3.	ARCHITECTURE TYPIQUE D'UN AGENT HYBRIDE.	23
FIGURE 4.	STRUCTURE D'UN AGENT	26
FIGURE 5.	FONCTIONNEMENT D'UN AGENT	27
FIGURE 6.	TYPLOGIE SELON LA NATURE DES AGENTS	29
FIGURE 7.	LE MODELE CLASSIQUE DE LA THEORIE DE LA COMMUNICATION	32
FIGURE 8.	SCHEMA DU MODELE DE SHANNON (1948)	33
FIGURE 9.	UN MESSAGE KQML	39
FIGURE 10.	UN MESSAGE FIPA-ACL	40
FIGURE 11.	UN EXEMPLE DU FONCTIONNEMENT DU CHATTERBOT CYBELLE SUR AGENTLAND.COM	53
FIGURE 12.	L'AVATAR CLASSIQUE	54
FIGURE 13.	EXPRESSIONS FACIALES SUR AVATARS	55
FIGURE 14.	AUGUST, UN AGENT ANIME	58
FIGURE 15.	EXEMPLE D'ACTION DE BODY CHAT	59
FIGURE 16.	L'AGENT VIRTUEL ANANOVA	60
FIGURE 17.	QUELQUES ATTITUDES DE MERLIN (MS-AGENT).....	61
FIGURE 18.	MERLIN FAIT UN GESTE DE DESIGNATION	61
FIGURE 19.	L'AGENT IMMOBILIER REA	62
FIGURE 20.	L'AGENT COSMO SYMPATHIQUE.....	64
FIGURE 21.	L'AGENT HERMAN THE BUG	65
FIGURE 22.	DEUX VUES DE L'ENVIRONNEMENT DE STEVE	66
FIGURE 23.	ADELE AIDE L'APPRENANT PENDANT UN EXERCICE	67
FIGURE 24.	L'AVATAR DE TESSA.....	68
FIGURE 25.	DEPAUL UNIVERSITY AMERICAN SIGN LANGUAGE.....	68
FIGURE 26.	LE « CAKE » DE TIM BERNERS LEE.....	80
FIGURE 27.	LA STRUCTURE DE HUMANML.....	86
FIGURE 28.	SMILEYS.....	99
FIGURE 29.	LES SIX EMOTIONS DE BASE	102
FIGURE 30.	ARCHITECTURE DU SYSTEME DE CHAT	103
FIGURE 31.	AGENT E-COMMERCE	110
FIGURE 32.	L'AGENT OZ DU SITE VIRTUOS.COM.....	116
FIGURE 33.	ARCHITECTURE DU L'AGENT POUR LE E-COMMERCE.....	117

Liste de tableaux

TABLEAU 1. :	DIFFERENTS OBJECTIFS VISES PAR LA COMMUNICATION	31
TABLEAU 2. :	LES CATEGORIES D'ACTES DE LANGAGE SELON SEARLE	36
TABLEAU 3. :	FORME D'UN ACTE DE LANGAGE	36
TABLEAU 4. :	LES PARAMETRES D'UN MESSAGE KQML.....	38
TABLEAU 5. :	FONCTIONS INTERACTIONNELLES ET LEURS REALISATIONS COMPORTEMENTALES.....	47
TABLEAU 6. :	ELEMENTS HUMANML	88
TABLEAU 7. :	QUELQUES EMOTICONES LES PLUS UTILISEES	99

Sigles et acronymes

ACA :	Agent Conversationnel Animé
ACL :	Agent Communication Language
AML:	Avatar Markup Language
APML:	Affective Presentation Markup Language
BAML:	Body Animation Markup Language
BDI :	Belief, Intention, Desire
BVH :	BioVision's Hierarchical
CML:	Character Markup Language
DMML :	Dialogue Management Markup Language
EML:	Emotion Markup Language
FAML :	Facial Animation Markup Language
FIPA :	Foundation for Intelligent Physical Agents
GML :	Gesture Markup Language,
MPML:	Multimodal Presentation Markup Language
HPCDML:	Human Physical Characteristics Description Markup Language
HTML:	HyperText Markup Language
HumanML:	Human Markup Language
KQML:	Knowledge Query Manipulation Language
MathML :	Mathematical Markup Language
OASIS:	Organization for the Advancement of Structured Information
OFX :	Open Financial eXchange
OWL:	OntologyWeb Language
RDF:	Resource Description Framework
RDFs:	RDF schema
RRL:	Rich Representation language
SGML:	Destiné à la modélisation de gros fonds documentaires
SMA :	Système multi-agent
SML:	Speech Markup Language
SMIL:	Synchronized Multimedia Integration Language
VHML:	Virtual Human Markup Language
VRML:	Virtual Reality Modeling Language
XML:	XML eXtensible Markup Language
W3C:	World Wide Web Consortium

Chapitre 1 : Introduction

1.1 Problématique :

La communication est une compétence fondamentale pour les êtres humains. Ils ont besoin d'interagir avec leurs semblables dans le but de satisfaire un besoin social, partager leurs idées, passions et savoir.

La communication a suivi une voie parallèle aux développements technologiques. Son évolution a connu plusieurs étapes : en commençant par la communication de face-à-face, ensuite avec la communication à distance avec support simple: les télégrammes, le téléphone, le courrier postal... et enfin avec le développement progressif d'Internet par l'usage du Chat, e-mail, téléconférence, système de réalité virtuelle, ...

Dans nos dialogues, il est commun d'omettre certains détails surtout s'ils nous paraissent inutiles ou triviaux. Parfois même les messages transmis peuvent être incomplets ou ambigus. Les conséquences sont parfois graves et conduisent à un rejet ou à une interprétation erronée du message. Il suffit de voir autour de nous pour constater les désastres commis par les mauvaises interprétations. Le résultat conduit à la haine, la violence, les conflits et encore plus grave les guerres entre les humains.

Tout cela est la conséquence de l'incapacité de communiquer certains aspects de la nature humaine. Les différences culturelles, religieuses, comportementales qui existent entre les hommes peuvent également mener aux mauvaises interprétations, aux conflits et malentendus.

En effet, la richesse de la communication humaine ne se limite pas à la parole, mais elle dépasse ce stade par l'intégration d'autres éléments non verbaux très porteurs d'informations dans le dialogue comme l'utilisation des expressions faciales et gestuelles pour transmettre des messages implicites ou encore l'utilisation du regard et des émotions pour mieux véhiculer le sens dans le discours.

Avec l'arrivée du net, la communication humain-humain a connu une autre progression très rapide. Mais, la situation n'est pas toute à fait idéale, puisque avec les outils actuels de communication sur le net, une grande partie des caractéristiques humaines nécessaires pour la

communication ne sont pas pris en considérations. Les émotions, les postures corporelles, les expressions faciales, les intonations de la voix, le caractère de l'utilisateur et sa culture, sont presque ignorés, ce qui nous ramène une autre fois à la problématique du départ. On perd donc une grande partie de l'information qui permet particulièrement de se rendre compte de l'état de notre interlocuteur, relativement au message qu'il est en train d'émettre et relativement au contexte.

Les internautes ont trouvé des solutions pour tenter de palier au mieux au manque d'informations et à la difficulté de les transmettre dans le message. Les plus connues sont l'emploi des émoticônes qui permettent de décrire l'état émotionnel de l'émetteur. Mais ces solutions ne sont pas officielles et manquent de standardisation.

L'acte naturel pour parer à ce problème est de raffiner la communication. Une bonne communication est celle qui permet d'appréhender une situation sans risque d'ambiguïté et autorise une transmission efficace de l'information. Enrichir le message initial lorsque qu'il est possible, permet de lever certains malentendus.

XML et les techniques du web sémantique peuvent aider à améliorer la communication entre humains. Pour cette raison, des membres du consortium OASIS ont pensé à créer un comité technique pour promouvoir la spécification d'un langage permettant d'éliminer tout conflits dans la communication humaine. Ce langage est HumanMI, son intention est d'humaniser le web et créer une voie standard pour représenter les émotions, le langage du corps et les caractéristiques humaines à travers XML. Il permet donc une amélioration de la communication en la rendant plus compréhensible et moins ambiguë.

D'une autre voie, les interactions entre les utilisateurs humains et les composants logiciels se multiplient. L'arrivée du net et de la technologie mobile a permis aux humains de communiquer avec des composants artificiels dans plusieurs types d'applications. De plus en plus, les utilisateurs de ces services sont des gens ordinaires qui ont du mal à manipuler correctement les outils et services auxquels ils ont affaire. Ils posent alors des questions qui leur semblent directes dans leur monde cognitif mais qui sont parfois décalées sémantiquement pour les applications et le dialogue tourne alors à la frustration.

Pour rendre l'interaction Homme-Machine plus réaliste, les agents intelligents sont apparus. Leur but dans les nouvelles IHM est de faciliter la tâche de la représentation du

comportement humain par la machine et de perfectionner les échanges entre l'utilisateur et l'ordinateur de manière à ce que ces échanges soient des plus naturels possible.

L'agent doit comprendre ce que l'utilisateur dit, par sa parole, par ses comportements non verbaux (gestes, mimiques faciales...) et réciproquement doit répondre à l'utilisateur par la parole et des comportements non verbaux. Pour répondre, la machine doit donc prendre l'apparence d'un être humain. Un nouveau type d'agents a été proposé: l'agent conversationnel.

Les agents ne sont plus vus par le grand public seulement pour leur côté ludique : leur utilité pour certaines applications a été vraiment affirmée... En effet, pour développer un agent conversationnel capable de montrer un comportement proche de celui des humains lorsqu'il communique avec un autre agent artificiel ou avec un humain, il faut en premier lieu enrichir le dialogue avec des informations non-verbales et des expressions émotionnelles.

Ces types d'applications nécessitent l'utilisation d'un langage de représentation capable d'inclure d'importantes informations liées aux caractéristiques humaines. Cependant, les langages de communication entre agents utilisés actuellement dans les SMA, comme KQML et FIPA-ACL ne sont pas adaptés à la communication entre agents et humains car ils ont été conçus pour l'échange d'informations. De plus, ces langages de communication sont destinés à des systèmes composés uniquement d'agents logiciels. Ils ne prennent pas en compte des paramètres de la complexité de la nature humaine.

Par contre, l'utilisation de certains langages de représentation pour agents conversationnels permet de manipuler les différents aspects de la communication humaine tel que la représentation des émotions, des gestes, de la posture du corps, du regard et des expressions faciales.

Dans les dernières années, plusieurs langages de représentation de très large envergure ont été développés : VHML, HumanML, AML, CML... certains d'entre eux sont destinés à des applications spécifiques : SiGML, MURML, MPML, RRL, APML, ... Dans notre travail, nous nous intéresserons au langage HumanML.

La principale contribution de ce mémoire est de présenter les principes du langage de représentation HumanML et son utilisation comme support de communication en SMA particulièrement dans le cas des agents conversationnels utilisés dans les interfaces Homme-

Machine. Ce langage permet d'encoder des informations liées aux comportements communicatifs des hommes. En effet, suivant la culture, l'âge, le travail... de l'envoyeur du message, le contenu du message peut varier. HumanML est basé sur les communications Homme-Homme et Homme-Machine dans les systèmes d'information digitale.

1.2 Objectifs :

Nous nous intéressons dans ce mémoire à l'utilisation du langage de représentation HumanML comme support de communication entre agents.

De part leur nature, les agents que nous considérons (humains et agents artificiels) n'ont pas les mêmes capacités, n'utilisent pas les mêmes langages de communication et ne disposent pas des mêmes connaissances. Mettre en relation des agents conversationnels avec des humains pour les faire interagir devient alors un véritable défi.

Notre objectif est d'améliorer la communication entre les utilisateurs et les agents artificiels dans les applications interactives sur le web. Nous proposons pour cela d'utiliser le langage HumanML qui a comme but d'améliorer la communication par l'ajout des aspects humains nécessaires pour la communication du face à face. Nous considérons la communication comme pertinente, lorsqu'elle n'est pas ambiguë et fournit suffisamment d'informations pour la compréhension du message. Dans ce mémoire, nous essayons de répondre aux questions suivantes afin d'atteindre nos objectifs :

- *Comment faciliter les échanges entre des intervenants différents les uns des autres ;*
- *Quel est le rôle du langage non-verbal et des émotions dans les nouvelles IHM;*
- *L'importance des agents conversationnels dans les applications en ligne ;*
- *L'apport du langage HumanML à la communication humain-humain et à la communication humain-agent.*

Pour répondre à ces questions, nous avons organisé le mémoire selon le plan suivant :

1.3 Organisation du mémoire :

L'organisation de ce mémoire en chapitre suit une progression régulière et ordonnée. Ce document se compose de six chapitres.

Chapitre 1: Introduction

Ce chapitre est une introduction générale où nous présentons le contexte général de nos travaux. Nous présentons ensuite les objectifs que nous voulons atteindre. Nous y détaillerons également le plan général du mémoire.

Chapitre 2 : Les systèmes multi-agents

Ce chapitre correspond à un ensemble de définitions et de concepts concernant les agents et les SMA, dont les notions apparaissent tout au long de ce mémoire. La notion de système multi-agents met l'accent sur l'interaction permettant le passage des comportements individuels à un comportement collectif. Nous expliquons que nous considérerons les humains comme des agents dans un SMA en soulignant l'intérêt, mais aussi les difficultés que cela apporte. Nous présentons ensuite notre problématique de recherche : notre but est de faire communiquer des agents de nature très différente alors qu'ils n'ont pas a priori de structure, ni de langage commun.

Ce chapitre permet de nous positionner par rapport à notre problématique et de spécifier nos orientations.

Chapitre 3 : La communication entre agents

Ce chapitre présente la communication d'une manière assez générale en introduisant divers concepts pouvant nous aider dans le cadre de notre étude. En premier lieu, nous parlerons de la communication Agent-Agent en présentant les généralités du domaine. Nous y introduisons la théorie des actes de langages sur laquelle se basent les langages de communication entre agents. Ensuite nous présentons un ensemble d'ACLs dont les plus connus les standards KQML et Fipa-ACL. Nous aborderons également dans ce chapitre la communication entre humain et agents en insistant sur les limitations des ACLs actuels pour ce type de communication d'où la nécessité de trouver d'autres langages de communication.

Chapitre 4 : Les agents conversationnels

Nous abordons dans ce chapitre le cas des agents conversationnels qui sont utilisés principalement dans le cadre des Interactions Homme-Machine. Nous présentons les principaux travaux existants relatifs aux agents conversationnels. Nous aborderons notamment les différents langages de représentation des ACAs, ils sont particulièrement nombreux et nous ne citerons que

les principaux (VHML, APML, CML, AML). Pour terminer ce chapitre, nous essayons de justifier les raisons de notre choix pour ce langage. Les principes du langage seront présentés dans le chapitre qui suit.

Chapitre 5 : Les fondements du langage HumanML

Pour commencer ce chapitre, nous revenons sur les bases du langage HumanML. Nous commençons par un aperçu sur le langage XML sur lequel se base un nombre important de langages de représentation des ACAs et en particulier HumanML. Ensuite, nous exposons la relation de HumanML avec le web sémantique. Les domaines d'application de ce langage seront également détaillés.

Chapitre 6 : Le langage HumanML comme support de communication entre agents

Dans ce chapitre, nous présentons la contribution principale de notre mémoire qui est l'utilisation du langage HumanML comme support de communication entre agents. Nous commençons d'abord par montrer les avantages que ce langage a apportés pour la communication entre humains, ensuite nous appliquons cette utilité dans la communication humain-agent en insistant sur la prise en compte des caractéristiques humaines. Pour démontrer l'importance du langage HumanML, nous proposons de l'exploiter dans 3 Applications sur le web qui mettent Humain et Agent en interaction.

Chapitre 7 : Conclusion et perspectives

Enfin, le chapitre 7 est la conclusion de ce mémoire dans laquelle nous faisons le bilan de nos contributions et où nous dégageons les perspectives de recherche.

Ainsi, de chapitre en chapitre, le sujet abordé se précise, partant du contexte vaste des agents et des SMA, pour proposer ensuite l'interaction entre les différents types d'agents en particulier lorsque cette interaction concerne un agent artificiel et un humain. De cette relation entre ces deux types d'agents, nous aborderons le point crucial de ce mémoire qui est la communication entre agents. Ensuite, nous nous intéresserons aux agents conversationnels. Enfin, dans notre dernière étape de focalisation, nous présentons le langage de représentation permettant de balayer les conflits existants entre les deux types d'agents

Chapitre 2 : Les systèmes multi-agents

Dans ce chapitre, nous allons présenter le cadre théorique dans lequel s'inscrivent nos travaux. Au début, nous essayons de donner une définition de l'agent. Cela n'est pas aussi simple qu'il y paraît au premier abord car chaque auteur a sa propre définition.

Dans la suite du chapitre nous passons à la définition de la notion des systèmes multi-agents. En fin de chapitre nous examinons les problèmes liés à l'interaction entre agents et en particulier lorsque ces agents sont de différents types à savoir humain et agents artificiels.

2.1 Définition d'un agent :

Le terme *Agent* est très répandu en informatique et il est utilisé dans plusieurs contextes. Mais, malgré plusieurs efforts pour en donner une définition totalement admise il reste difficile à cerner. Pour donner plus de précisions à cette notion d'agent, nous présentons dans ce qui suit quelques définitions importantes :

En premier lieu nous allons se baser sur une première définition simple, telle que celle donnée par Russell et Norvig [23] . Ils présentent *l'agent comme celui qui peut percevoir et agir sur son environnement à travers des capteurs et des effecteurs*.

Cette première définition, très généraliste, permet certes, de pouvoir identifier des agents dans plusieurs domaines d'applications, mais elle ne fournit aucune indication sur le comportement de l'agent et ses capacités à communiquer avec le monde extérieur [23].

La définition de Russel et Norvig reste très générale et constitue une bonne entrée en matière. La figure 1. permet d'illustrer ces propos :



Figure 1. Représentation d'un agent selon Russel et Norvig

Une autre définition [23] a été donnée par Wooldridge et Jennings en 1995. Elle est plus détaillée que la précédente :

« Un agent est un système informatique, situé dans un environnement, et qui agit d'une façon autonome pour atteindre les objectifs pour lesquels il a été conçu »

Et d'après Jacques Ferber [19], on appelle agent une entité physique ou virtuelle :

- a. qui est capable d'agir dans un environnement,
- b. qui peut communiquer directement avec d'autres agents,
- c. qui est mue par un ensemble de tendances (sous la forme d'objectifs individuels ou d'une fonction de satisfaction qu'elle cherche à optimiser),
- d. qui possède des ressources propres,
- e. qui est capable de percevoir (mais de manière limitée) son environnement,
- f. qui ne dispose que d'une représentation partielle de cet environnement,
- g. qui possède des compétences et offre des services,
- h. qui peut éventuellement se reproduire,
- i. dont le comportement tend à satisfaire ses objectifs, en tenant compte des ressources et des compétences dont elle dispose, et en fonction de sa perception, de ses représentations et des communications qu'elle reçoit.

Cette définition [18] met l'accent sur les capacités de l'agent à interagir avec les autres agents et à s'adapter en fonction de chaque changement de situation. Une entité devient un agent aussitôt qu'elle est capable d'exercer un contrôle local sur ses processus de perception, de communication, d'acquisition de connaissances, de raisonnement, de prise de décision ou d'exécution [16].

Toutes ces définitions se basent sur des notions semblables que sont : ***un agent est autonome, plongé dans un environnement et qui a des objectifs à atteindre.***

Cette définition peut être appliquée, entre autres, aux exemples suivants [07] :

- A l'utilisateur humain d'un système informatique. En effet les êtres humains correspondent à cette définition dans le sens où ils disposent de moyens d'acquérir et d'utiliser des connaissances, où ils perçoivent ce qui les entoure par les organes sensoriels, et où ils agissent sur cet environnement par action musculaire.

- A un robot qui explore une surface. Celui-ci peut aussi être considéré comme un agent, car il doit gérer l'énergie dont il dispose, il peut percevoir son environnement au moyen de caméras et d'autres capteurs ou appareils de mesure.
- A certains programmes informatiques que l'on appellera des agents logiciels. Par exemple les agents capables de trier les messages électroniques, de sélectionner des articles sur les forums ou encore de planifier des réunions. Les capteurs et les actionneurs des agents logiciels représentent alors les parties de programme qui gèrent les entrées et les sorties. Ainsi, un agent qui filtre des e-mails analyse les messages reçus (il perçoit), puis les classe ou les supprime (il agit) selon leur forme ou leur contenu (il décide).

2.2 Propriétés d'un agent :

Partant des définitions de la notion d'agent, nous pouvons distinguer plusieurs caractéristiques d'un agent selon divers points de vue :

- **Autonomie** : Ce point est le plus discuté dans les définitions d'agents. 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 (d'un autre agent ou d'un opérateur humain). L'agent a la possibilité de répondre par l'affirmative ou le refus à des requêtes provenant des autres agents. Il dispose donc d'une certaine liberté de manoeuvre, ce qui le différencie de tous les concepts semblables, qu'ils s'appellent "objets", "modules logiciels" ou "processus" [19]. Sans cette propriété d'autonomie, l'agent se restreint à un simple objet informatique qui possède des données et propose des méthodes [23]. Cette autonomie implique évidemment que l'agent doit être capable de gérer les conflits entre ses objectifs et son état interne ou l'état de son environnement [34].

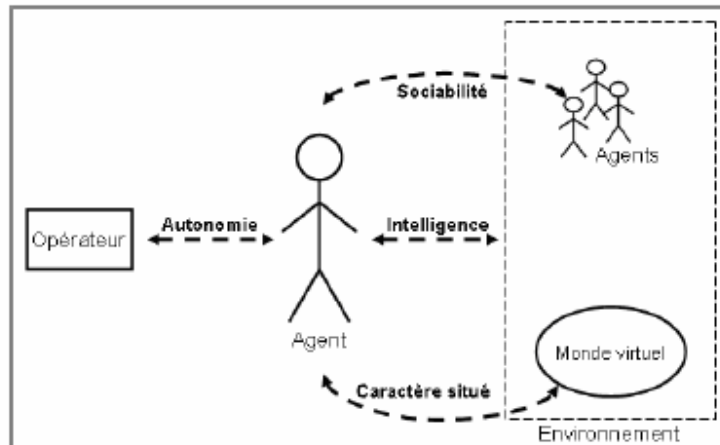


Figure 2. Principales caractéristiques d'un agent

- **L'adaptativité** : Un agent est capable de modifier son comportement et ses objectifs en fonction de ses interactions avec son environnement et avec les autres agents [34]. En particulier, il peut réagir différemment selon l'agent avec lequel il interagit. On peut aussi considérer dans ce point le fait qu'avec son expérience, l'agent acquiert une personnalité [08], qui le rend différent des autres agents. L'adaptativité signifie que l'agent est capable : (1) de réagir de façon flexible aux changements de son environnement ; (2) de prendre, lorsque nécessaire, des initiatives orientées vers un but ; et (3) d'apprendre, à partir de sa propre expérience, de son environnement et de ses interactions avec d'autres agents [29].
- **Conversation** : Il pose que pour être considérée comme un agent, une entité informatique doit posséder des capacités d'interaction avancées, permettant une rétroaction de l'utilisateur sur l'agent et de l'agent sur l'utilisateur. Il aborde explicitement la notion d'enchaînement conversationnel et d'évolution des conversations entre les deux partenaires au cours des différentes interactions [08].
- **Caractère situé de l'interaction** : Un agent autonome évolue de façon continue dans un environnement, dans lequel peuvent exister d'autres agents. Puisqu'un agent agit sur son environnement et que l'environnement agit sur l'agent, il existe une réelle interaction entre ces deux entités [34]. Les agents perçoivent leur environnement (qui peut être le monde physique, un utilisateur au travers d'une interface graphique, un ensemble

d'autres agents, l'Internet, ou peut-être une combinaison de tout cela), et répondent de manière opportune à ses changements [13].

- **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 [18]. Les agents sont considérés également de sociables [24] s'ils peuvent communiquer les uns avec les autres, et possiblement avec des humains, grâce à un langage de communication pour réaliser les tâches pour lesquelles ils ont été conçus.
- **L'intelligence** : Un agent est habituellement qualifié d'intelligent lorsqu'il possède des capacités de représentation symbolique et de raisonnement s'il est capable de raisonner sur son propre comportement et ses propres objectifs [34].

2.3 Les systèmes multi-agents (SMA):

Dans la partie précédente nous avons présenté le concept d'agent isolé comme une entité située dans un environnement qu'elle perçoit et sur lequel elle peut agir. Cette partie se concentre sur des systèmes composés de plusieurs agents.

Il est plutôt rare que les concepteurs d'agents n'aient besoin que d'un seul agent dans l'environnement qu'ils construisent. Lorsque plusieurs agents se retrouvent dans un même environnement et que ces agents ont besoin d'interagir entre eux, on parle alors de système multi-agent (ou SMA). Ainsi, D'après J.Ferber [19], on appelle système multi-agent, un système composé des éléments suivants:

1. Un environnement **E**, c'est-à-dire un espace disposant généralement d'une métrique.
2. Un ensemble d'objets **O**. Ces objets sont situés, c'est-à-dire que, pour tout objet, il est possible, à un moment donné, d'associer une position dans **E**. Ces objets sont passifs, c'est-à-dire qu'ils peuvent être perçus, créés, détruits et modifiés par les agents.
3. Un ensemble **A** d'agents, qui sont des objets particuliers (**A** _ **O**), lesquels représentent les entités actives du système.

4. Un ensemble de relations **R** qui unissent des objets (et donc des agents) entre eux.
5. Un ensemble d'opérations **Op** permettant aux agents de **A** de percevoir, produire, consommer, transformer et manipuler des objets de **O**.
6. Des opérateurs chargés de représenter l'application de ces opérations et la réaction du monde à cette tentative de modification, que l'on appellera les lois de l'univers.

Les SMA s'appuient sur la métaphore de l'organisation collective, par opposition à l'intelligence artificielle classique qui s'appuie sur la métaphore du penseur isolé. Ils proposent de modéliser des systèmes complexes à l'aide d'une société d'agents. Chaque agent a ses propres compétences, mais il a besoin d'interagir avec les autres pour résoudre les problèmes qui dépendent de son domaine et éviter les conflits. L'objectif de ces systèmes est donc de trouver une solution à des problèmes globaux ou de simuler des comportements complexes [21]. 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 [18].

L'étude des systèmes multi-agents a débuté vers le début des années 80 avec l'étude de l'intelligence artificielle distribuée (IAD). Les systèmes multi-agents sont devenus un champ d'étude très important du monde de la recherche particulièrement pour les applications Internet.

Un SMA est caractérisé par [06] :

- Chaque agent a des informations partielles, donc, une vue limitée du problème ;
- Il n'y a pas de contrôle global du système, c'est-à-dire de système central ;
- Les données sont décentralisées ;
- Les opérations sont asynchrones.

2.4 La typologie des agents :

Les agents ont une représentation partielle de leur environnement et d'eux-mêmes. On distingue deux types d'agents en fonction de leur architecture interne :

- Les agents réactifs
- Les agents cognitifs

- ✓ **Les agents réactifs** : Un agent est considéré comme réactif lorsqu'il agit uniquement en fonction de stimuli (internes et externes). Il ne possède ni représentation (ou une représentation minimale) de l'environnement dans lequel il évolue, ni mécanisme d'envoi de messages. Un SMA constitué d'agents réactifs, possède généralement un grand nombre d'agents et présente un comportement global intelligent [29]. L'exemple le plus manifeste d'organisation émergente est celle de la fourmilière que toutes les fourmis se situent sur un plan d'égalité et qu'aucune d'entre elles ne possède de pouvoir d'autorité stricte sur les autres, les actions des fourmis se coordonnent de manière que la colonie survive et affronte des problèmes complexes tels que ceux posés par la recherche de nourriture, les soins à donner aux oeufs et aux larves, la construction de nids, la reproduction, etc. [19]. Ainsi, ce n'est pas au niveau de l'individu que les agents réactifs sont intéressants, mais au niveau du groupe.

- ✓ **Les agents cognitifs** : ils sont qualifiés d'"intelligents". Ils possèdent une représentation partielle de l'environnement, des buts explicites et sont capables de planifier leur comportement, de mémoriser leurs actions passées, de communiquer, de négocier, etc. Les architectures **BDI** (Belief Desire Intention) [39] constituent un type d'architecture d'agents cognitifs. Cette architecture est basée sur les notions d'attitudes mentales que sont la Croyance (Belief), le Désir (Desire) et l'Intention (Intention) :
 - Les croyances correspondent aux informations (éventuellement incomplètes et incorrectes) qu'a l'agent de son environnement ;
 - Les désirs correspondent aux états de l'environnement que l'agent souhaiterait voir réalisés ;
 - Les intentions correspondent aux projets de l'agent pour satisfaire ses désirs.

Les prises de décisions d'un agent BDI sont alors effectuées à partir de la manipulation les états mentaux de l'agent et de la révision de ceux-ci au cours du temps.

Un SMA constitué d'agents cognitifs possède peu d'agents, mais l'architecture de ces agents est plus compliquée à mettre en œuvre que celle des agents réactifs [29]. Les agents cognitifs, peuvent travailler de manière relativement indépendante. Les tâches qu'ils

accomplissent sont complexes. Ils peuvent ainsi résoudre des problèmes compliqués de manière relativement individuelle [19].

Cependant, La distinction entre agents cognitifs et agents réactifs, tend à disparaître, grâce aux nombreux travaux qui proposent des approches hybrides afin de tirer profit des avantages des deux approches [18]. De façon habituelle, un agent n'est pas complètement de type réactif ni de type délibératif. C'est plutôt une combinaison de ces deux approches. Dans certaines circonstances, un agent doit agir très rapidement alors que dans d'autres, il aura la possibilité de prendre plus de temps pour avoir un processus délibératif de meilleure qualité [24].

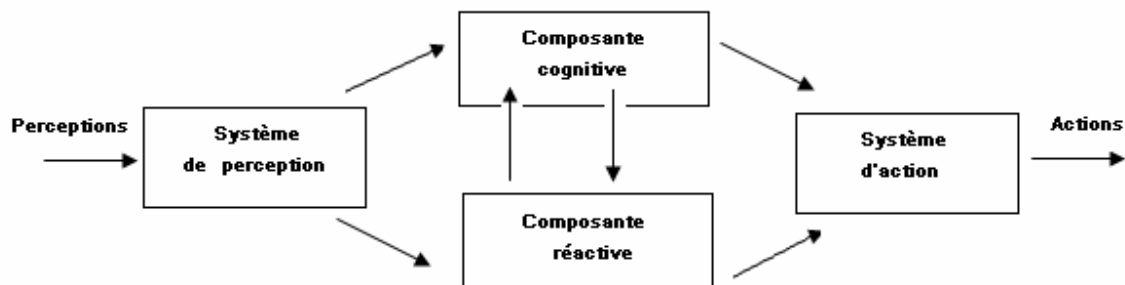


Figure 3. Architecture typique d'un agent hybride.

Dans ce cas, la partie réactive permet à l'agent de réagir de façon réflexe à son environnement et la partie cognitive se charge des opérations de planification, anticipation, mémorisation, apprentissage, de communication directe avec d'autres agents, etc

Une autre classification des agents est faite suivant leur utilisation, on y trouve :

- ✓ **Les agents collaboratifs** : Ces agents ont des habiletés de coopération. Un regroupement de ces agents permet, entre autres, de réduire un problème complexe en sous problèmes moins complexes.
- ✓ **Les agents d'interface** : Ces agents collaborent avec l'utilisateur pour effectuer certaines tâches.
- ✓ **Les agents pour la recherche d'informations** : Ces agents effectuent, en premier lieu, une recherche d'informations parmi une collection de données et, en second lieu, procèdent à une analyse des informations utiles trouvées afin de découvrir de nouvelles connaissances.
- ✓ **Les agents pour le commerce électronique** : La montée de l'Internet a bien entendu créé de nouvelles nécessités. Les agents issus de cette tendance permettent la promotion, la vente ainsi que l'achat de produits et de services par l'entremise des réseaux informatiques etc. [24].

On peut également les classer selon la technologie employée :

- ◆ **Agents stationnaires** : Il s'agit du cas où l'agent s'exécute toujours sur la même machine.
- ◆ **Agents mobiles** : Ces agents s'exécutent sur différentes machines en se promenant d'un hôte à l'autre. Ils réduisent la charge des réseaux en se rendant directement sur la machine "utile" pour y effectuer les interactions de manière locale [24].

2.5 Les interactions entre agents :

La notion d'interaction est au centre de la problématique des SMA. Une interaction est une mise en relation dynamique de deux ou plusieurs agents par le biais d'un ensemble d'actions réciproques [19]. Les interactions permettent à plusieurs agents intelligents de combiner leurs efforts. Un agent sans interaction avec d'autres agents n'est plus qu'un corps isolé [19].

Les interactions ont une influence sur le comportement futur d'un agent [19]. En fait, il doit s'adapter aux changements qui se produisent dans l'environnement et qui ont des répercussions sur son comportement.

Il existe différents types d'interactions que les agents peuvent utiliser comme la coordination, la coopération et la communication [24].

Les situations d'interactions sont nombreuses et diverses [19]:

- L'aide d'un robot par un autre,
- L'échange de données entre serveurs informatiques,
- La mise en commun des compétences de plusieurs spécialistes pour porter un diagnostic,
- L'utilisation d'une imprimante par deux programmes simultanément,
- La collision de deux véhicules ...

Un agent devra interagir avec les membres de son environnement pour quatre raisons principales [24]:

- Partager de l'information;
- Atteindre ses buts;
- Eviter, autant que possible, les conflits;
- S'assurer de l'interopérabilité avec des agents hétérogènes.

2.6 Architecture d'agents :

Dans ce paragraphe, nous présentons une architecture générale [35] d'agent cognitif et nous décrivons les processus mis en oeuvre lors de son fonctionnement.

2.6.1 Structure interne :

La figure 5. décrit l'architecture globale d'un agent cognitif. C'est une synthèse des architectures d'agents décrites dans la littérature [35]. On distingue essentiellement les modules suivants :

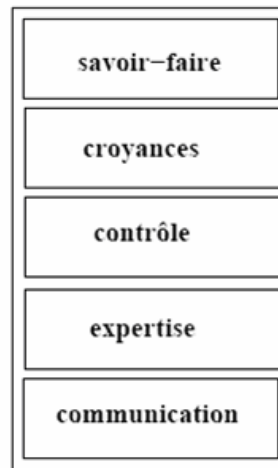


Figure 4. Structure d'un agent

- **Savoir-faire** : est une interface permettant la déclaration des connaissances et des compétences de l'agent. Il n'est pas nécessaire mais il est très utile pour améliorer les performances du système, quel que soit le mode de coopération utilisé.
- **Croyances** : Dans un univers multi-agents, chaque agent possède des connaissances sur lui-même et sur les autres. Ces connaissances ne sont pas nécessairement objectives, on parle alors de croyances d'un agent.
- **Contrôle** : La connaissance de contrôle dans un agent est représentée par les buts, les intentions, les plans et les tâches qu'il possède.
- **Expertise** : C'est la connaissance sur la résolution de problème.
- **Communication** : L'agent doit posséder un protocole de communication lui permettant d'interagir avec les autres agents pour une bonne coopération et une bonne coordination d'actions.

2.6.2 Fonctionnement :

Les agents sont immergés dans un environnement dans lequel et avec lequel ils interagissent. D'où leur structure autour de trois fonctions principales :

Percevoir, décider et agir. Parmi les sous fonctions importantes d'un agent on peut citer : la détection de conflits, la révision des croyances, la coopération (négociation, coordination), l'apprentissage, etc.

Un agent a la possibilité d'acquérir des connaissances sur l'environnement externe (perception). Il a aussi des capacités d'interaction avec les autres agents (communication, négociation). En fonction des connaissances et croyances dont il dispose et des buts qu'il se fixe suite à une perception ou à une interaction avec le monde extérieur, l'agent doit élaborer un plan d'action. Pour cela, il doit décider quel serait le but à retenir et à satisfaire en premier, ensuite planifier en fonction de ce but et passer à l'exécution.

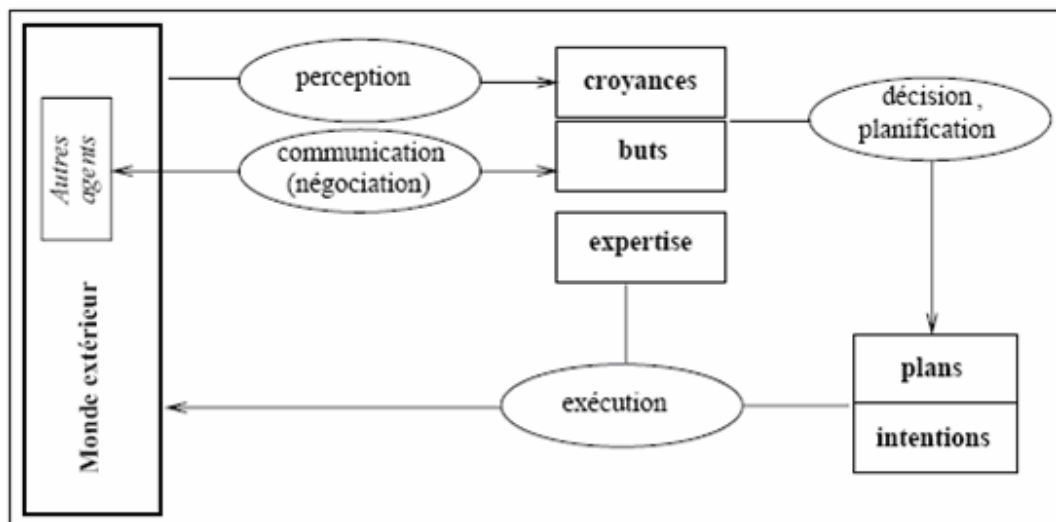


Figure 5. Fonctionnement d'un agent

❖ **Perception** : Les connaissances d'un agent ont plusieurs origines :

- Le savoir initial de l'agent ;
- La perception de soi et du monde ;
- La communication avec les autres agents.

Généralement, les informations issues de la perception et du savoir initial de l'agent sont considérées comme des connaissances certaines puisqu'elles n'ont subi aucune mise à jour, alors

que les connaissances provenant des autres agents sont considérées incertaines puisqu'elles évoluent sans que l'agent en soit forcément informé. On peut pour cela associer à chaque connaissance son origine afin d'en évaluer la crédibilité et d'en permettre la vérification.

- ❖ **Prise de décision** : Durant son exécution, un agent se fixe un certain nombre de buts, suite à ses observations et à ses interactions avec le monde (perception, communication, négociation). Il se trouve donc confronté au problème de la sélection du but à satisfaire en premier et pour chaque but, de l'action qui permet de l'atteindre.

Face à de telles situations, l'agent analyse les différentes alternatives en termes d'utilité (quel avantage l'agent pourrait en retirer) et d'incertitude (quelle chance a l'action d'être effectuée en fournissant le résultat attendu).

Le but à retenir sera celui qui présente le plus d'influences positives parmi les buts les plus interdépendants.

La prise de décision est une des caractéristiques des agents rationnels, l'agent tiendra compte de ses croyances pour faire son choix.

- ❖ **Planification** : La planification dans les SMA est une planification distribuée : il n'existe pas de plan global et chaque agent construit son propre plan en coordonnant avec les autres (cas d'agents coopératifs).

2.7 L'hétérogénéité des agents :

La notion de systèmes multi-agents est parfois utilisée pour décrire des systèmes qui ne font pas appel qu'à des agents logiciels, comme dans le cadre du travail collaboratif dans lequel agents et humains interagissent. Les communautés en ligne sont déjà le lieu de rencontre entre agents logiciels et humains. On cite l'exemple des forums de discussion. Plus généralement sur Internet, les robots logiciels (Chatterbots) interagissent avec les utilisateurs, et ces derniers se trompent parfois en les prenant pour des humains [22].

Dans un SMA, il est très rare que deux agents soient identiques en tous points. L'hétérogénéité dans un système multi-agent est définie par la présence de différences entre les agents. Dans [07], la typologie des Franklin et Graesser, permet de se retrouver avec un système comprenant des agents biologiques et des agents logiciels... Ensuite, même si on prend deux agents logiciels, ils peuvent être conçus de façon très dissemblable. Le premier pourra être

construit sur une architecture simple, alors que le second pourra être un agent BDI. Enfin, même s'ils sont du même type, ils n'auront pas forcément les mêmes connaissances et les mêmes capacités. Tout ceci montre que si on veut les faire interagir ensemble dans le même SMA, il faut remédier à toutes sortes d'incompatibilités plus ou moins perceptibles [07].

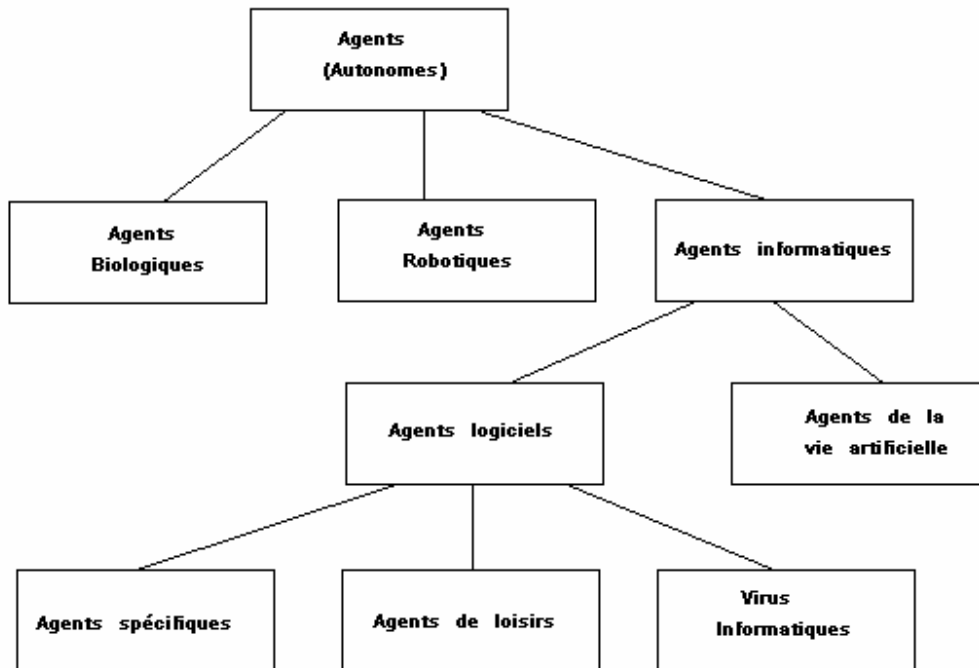


Figure 6. Typologie selon la nature des agents

Conclusion :

La progression des paradigmes de programmation a conduit aux agents et systèmes multi-agents. Dans notre mémoire, nous nous intéresserons spécifiquement à des problèmes faisant intervenir un ou plusieurs agents humains en interaction avec les agents artificiels. La tâche de modélisation est très difficile, car elle nécessite de prévoir au maximum les comportements qu'un humain peut adopter. La communication entre ces différents types d'agents devient alors une tâche difficile, car les protocoles de communication entre les humains et les agents artificiels ne sont pas les mêmes. On voit en effet davantage de problèmes se soulever lorsque l'un des agents est un humain parce qu'il y a un fossé de langage. En considérant cela, on conçoit que prendre en compte des humains dans un SMA est une tâche très difficile [07].

Le chapitre suivant va présenter la communication entre agents tout en considérant l'introduction de l'humain dans ce type de communication ainsi que les problèmes qui en découlent.

Chapitre 3 : La communication entre agents

Dans ce chapitre, nous allons aborder d'un point de vue général, la notion de la communication dans le but de mieux la comprendre afin d'appliquer certains de ses aspects aux communications entre agents artificiels ensuite entre agents artificiels et humains.

3.1 Définition de la communication :

Le terme communication provient du latin *communicare* qui signifie "mettre en commun" et "entrer en relations" [24]. Communiquer signifie établir des relations avec quelqu'un dans le but d'échanger de l'information. Elle se produit chaque fois qu'une source tente d'influencer un destinataire, au moyen de signaux transmis par le canal qui relie ces deux entités.

Les communications s'effectuent par l'intermédiaire donc de signaux, chaque signal caractérisant un indice ou un signe. Un signal est l'élément le plus primitif. On peut le définir comme une marque, une trace qui porte une information pour celui qui est capable de percevoir ce signal. Lorsque le signal ne produit qu'un comportement, on dit qu'il s'agit d'un stimulus. C'est le cas pour la plupart des animaux et, pour ce qui nous concerne, pour l'ensemble des agents réactifs. Lorsque le signal ne produit pas seulement un comportement, mais s'insère dans un système cognitif, on dit alors que le signal est porteur de signification [19].

On utilise régulièrement la communication dans un but bien précis. A titre d'exemple, le tableau 1. présente différents objectifs que la communication cherche à accomplir.

Emetteur	Recepteur
Informer	Comprendre
Enseigner	Apprendre
Plaire	Apprécier
Proposer, persuader	Ecarter un problème, prendre une décision

Tableau 1. : Différents objectifs visés par la communication [24]

L'essor des technologies informatiques a permis aux moyens de la communication de se diversifier. Dans le domaine de la recherche, la communication est un objet d'étude multiforme exploré par maintes disciplines comme la psychologie, la biologie, les sciences cognitives ou encore l'informatique et, en ce qui nous concerne, les systèmes multi-agents (SMA). Chaque discipline aborde la communication selon un axe privilégié, et suivant sa problématique propre. Communiquer est une activité que nous pratiquons de façon régulière, sans nous rendre totalement compte de la nature complexe de ce phénomène qui a évolué depuis déjà plusieurs siècles.

3.2 Le processus de communication:

Un schéma simpliste de communication est celui qui met en action un émetteur, un destinataire et un médium sur lequel transite le message. Il existe de nombreuses théories de la communication mais elles reposent essentiellement sur des variantes de la théorie de la communication issue des recherches en télécommunications des années 40 développées par Shannon et Weaver [19].

La théorie de la communication de Shannon sert souvent de référence tant dans les travaux sur la communication humain-machine que ceux sur la communication Humain-Humain.

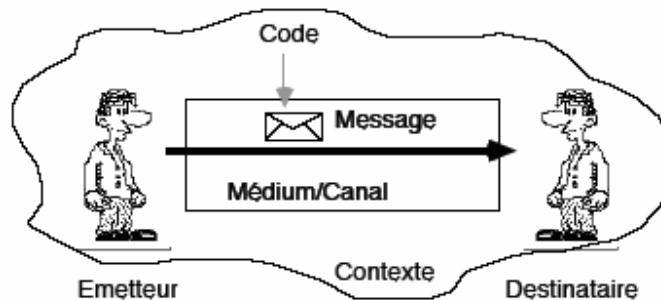


Figure 7. Le modèle classique de la théorie de la communication

Ainsi, la communication serait essentiellement composée de différentes étapes associées à la transmission. Ce modèle est schématisé dans la figure 9.

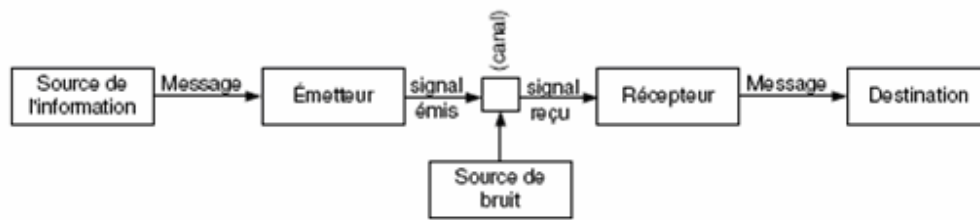


Figure 8. Schéma du modèle de Shannon (1948)

Les étapes du modèle de Shannon sont définies comme suit :

1. La source d'information produit un message devant être transmis.
2. L'émetteur traduit ce message en un signal émis sur le canal de communication vers le récepteur.
3. Le récepteur reçoit le signal et le transforme afin de reconstituer le message original et le transmettre à destination.
4. Il est possible que le signal ayant été transmis sur le canal soit altéré par du bruit.

Le modèle a été développé afin de résoudre certains problèmes techniques liés à la télécommunication [24]. Toutefois, dans certaines de ces situations de communication avec l'humain, ce modèle peut se montrer inadapté [08].

La notion de communication s'est depuis structurée et, de la simple transmission d'information initiale, on est passé à des formes plus élaborées, telles que les actes de langage et les structures conversationnelles qui mettent l'accent sur la notion d'interaction dans les communications [19].

Mais la communication va au-delà de l'échange verbal [19]. La communication est alors considérée comme un processus social permanent intégrant de multiples modes de comportement: la parole, le geste, le regard, la mimique, l'espace interindividuel, etc. Dans ce cas, les messages n'ont plus de signification intrinsèque: seul le contexte est capable d'apporter une lumière sur la signification des modes d'interaction.

Nous allons présenter dans les points suivants la communication entre les différentes catégories d'agents :

- ✓ La communication Agent-Agent
- ✓ La communication Humain-Agent
- ✓ La communication Humain-Humain

3.3 La communication entre agents logiciels :

La communication, dans les SMA comme chez les humains, est à la base des interactions et de l'organisation sociale [19]. C'est parce que les agents communiquent qu'ils peuvent coopérer, coordonner leurs actions, réaliser des tâches en commun et devenir ainsi de véritables êtres sociaux.

Les SMA utilisent principalement deux méthodes différentes pour modéliser les communications entre agents : soit ils communiquent indirectement par le biais d'une mémoire partagée, c'est le cas du modèle du tableau noir, soit d'une manière directe par envoi des messages les uns aux autres.

- **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 [18].
- **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.

3.3.1 La théorie des actes de langages :

Pour être en mesure de dialoguer, les participants d'une conversation doivent bien évidemment échanger des énoncés. A cet égard, il y a une théorie très importante qui a été développée par les philosophes et que les chercheurs ont utilisée pour gérer des unités de conversation entre agents. Il s'agit de **la théorie des actes de langage**.

La parution, en 1962, du livre d'Austin *How To Do Things With Words* (Quand dire, c'est faire) a donné une vue différente aux théories précédentes, avant les chercheurs s'intéressaient essentiellement à savoir quel était le degré de vérité d'une phrase constative, c'est-à-dire qui affirme quelque chose du monde [19]. L'idée maîtresse de cette théorie est qu'une instance d'utilisation de la langue est une action comme les autres : « dire c'est faire » [28]. Par cette formule, il montre que toute communication est action. Maintenant avec la théorie des actes de langage un énoncé n'est plus Vrai ou Faux mais il réussit ou il échoue [36].

Pour Austin et ses successeurs, John Searle [19], et Vanderveken, l'énonciation, c'est-à-dire le fait de produire un énoncé, est un acte qui sert avant tout à produire des effets sur son destinataire. Austin parle alors d'actes de langage [36].

Principalement, il identifie une catégorie de verbes qu'il baptise performatifs (de l'anglais *to perform*, accomplir, exécuter, effectuer, . . .) [08].

Les actes de langage sont définis comme des structures complexes formées de trois composantes, considérées comme des actes élémentaires. Austin en a identifié trois. Il s'agit de l'acte locutoire, de l'acte illocutoire et de l'acte perlocutoire :

- **Acte locutoire** : Il s'agit de l'action de dire quelque chose. Il se rapporte à l'énonciation en elle-même.
- **Acte illocutoire** : qui se rapporte à l'acte effectué par l'énoncé. Le locuteur doit interpréter le sens de l'énoncé en termes de son contenu propositionnel. Les actes illocutoires, les plus étudiés en pragmatique du langage, sont caractérisés par une force illocutoire (exemple: affirmer, questionner, demander de faire, promettre, ordonner, prévenir, etc.) et par un contenu propositionnel qui est l'objet de la force illocutoire [19].
- **Acte perlocutoire** : Il s'agit des effets que le message illocutoire a sur le destinataire : sur ses actions, ses croyances [36]. Cet effet n'est pas sous le contrôle du locuteur.

Par exemple, la phrase "Je t'affirme que la terre tourne autour du Soleil" est un acte illocutoire d'affirmation, qui peut être pris, dans un certain contexte d'énonciation, comme une tentative de convaincre le destinataire que la terre tourne et donc de modifier l'ensemble de ses

croyances. De même lorsqu'on demande à quelqu'un de faire quelque chose, le fait qu'il ait compris qu'il s'agisse d'un exercitif montre qu'il a reconnu la composante illocutoire de l'acte, et s'il accomplit effectivement ce qui lui a été demandé, alors on peut dire que la composante perlocutoire a été satisfaite [19].

Le but illocutoire relie la proposition énoncée au monde réel. Pour Searle [28], il existe cinq utilisations possibles du langage qui sont caractérisées par les cinq buts illocutoires (on parle aussi des cinq types d'actes de langage), le tableau 2. Illustre ces actes :

Catégorie	But	Exemple
Assertifs	Donner une information sur le monde en affirmant quelque chose	Il fait beau, le rectangle a quatre angles droits
Directifs	Donne un ordre représenté par le contenu propositionnel au destinataire	Ouvrez la porte, Donne moi le sel
Promissifs	Engager le locuteur à accomplir certains actes dans l'avenir	Je viendrai à la réunion de 5 heures. Je te promets de t'envoyer des cartes postales
Expressifs	Exprimer les sentiments de celui qui parle	Je suis heureux, je m'excuse pour hier
Déclaratifs	Provoquer un changement par une déclaration	Je déclare la séance ouverte

Tableau 2. : Les catégories d'actes de langage selon Searle

A la suite des travaux de formalisation de Searle et Vanderveken, un acte de langage se compose d'une force illocutoire et d'un contenu propositionnel, respectivement baptisés F et P [08].

A	=	F	(P)
Un acte de Langage	est égal	une force illocutoire	appliquée à un contenu propositionnel

Tableau 3. : Forme d'un acte de langage

Par exemple, avec le même contenu propositionnel “Il pleut” nous pouvons former les illocutions suivantes :

- Affirmer (Il pleut) «Tiens, il pleut ! »
- Questionner (Il pleut) «Sais-tu si il pleut ? »
- Prévenir (Il pleut) «Il risque de pleuvoir » etc.

La notion de performatif est utile pour les SMA, puisque ce sont les performatifs qui indiqueront les différents types d’actes de langage que peuvent émettre et interpréter les agents.

La théorie des actes de langage, Initialement pensée pour le langage naturel, sa nature formelle la rend utilisable pour les modèles computationnels [28]. Elle est reconnue depuis longtemps dans la communauté scientifique pour sa capacité à bien modéliser les communications [29].

3.3.2 Les langages de communication entre agents :

Depuis plusieurs années, la communication dans les SMA a cherché à s’inspirer des phénomènes de communication humains. En se basant sur le formalisme des actes de langage, des langages d’interaction ont été développés, parmi lesquels KQML et ACL-FIPA.

Les langages de communication d’agents se distinguent des approches plus traditionnelles de communication au niveau des programmes informatiques telles que UDP, TCP/IP, RPC, JAVA-RMI ou CORBA du fait qu’ils cherchent à échanger des propositions et des tâches ayant une sémantique de haut niveau comparativement à de simples données ou objets. Dès lors, les agents sont généralement impliqués dans une conversation ou un dialogue semblable à ce que font les humains, et non dans un simple processus d’échange d’informations [24].

Nous présenterons dans cette partie KQML et l’ACL de la FIPA (Foundation for Intelligent Physical Agents) qui font office de standards parmi les langages d’agents, puis nous expliquerons en quoi nous estimons les langages d’agents actuels inadaptés à une communication impliquant un agent humain.

3.3.2.1 Le langage KQML:

Premier apparu parmi les ACLs, le langage KQML (Knowledge Query Manipulation Language) est considéré comme un standard [29]. Il est issu d’un projet de la Darpa KSE

initialement prévu comme moyen d'échange d'informations entre programmes à base de connaissance. Sa structure orientée message et la généralité de ses primitives lui permettent d'être utilisé comme ACL [29].

KQML est à la fois un langage et un protocole permettant de structurer des messages afin de partager de l'information entre agents [24].

KQML est basé sur la théorie des actes de langage. Il propose une encapsulation des messages dans une performative qui définit l'acte illocutoire. Il est constitué d'un important nombre de performatifs qui sont en quelque sorte les opérations permises entre les agents.

Le tableau 4. représente l'ensemble des paramètres que l'on peut retrouver à l'intérieur d'un message Kqml ainsi que la signification de chacun d'entre eux [24].

Mot-clé	Signification
:sender	L'expéditeur du message.
:receiver	Le récepteur du message.
:from	Indique l'origine du performatif se trouvant dans l'étiquette :content lorsque forward est utilisé.
:to	La destination finale du performatif se trouvant dans l'étiquette :content lorsque forward est utilisé.
:content	L'information que le performatif exprime.
:in-reply-to	L'étiquette de réponse à laquelle se réfère ce message.
:language	Le nom du langage utilisé pour représenter le contenu.
:ontology	L' (Les) ontologie(s) utilisée(s) qui donne(nt) la signification aux éléments du contenu.
:reply-with	L'étiquette qui doit être utilisée pour les réponses si l'expéditeur en désire une.

Tableau 4. : Les paramètres d'un message KQML

Un message KQML ressemblera donc à ce qui est présenté dans la figure 10.

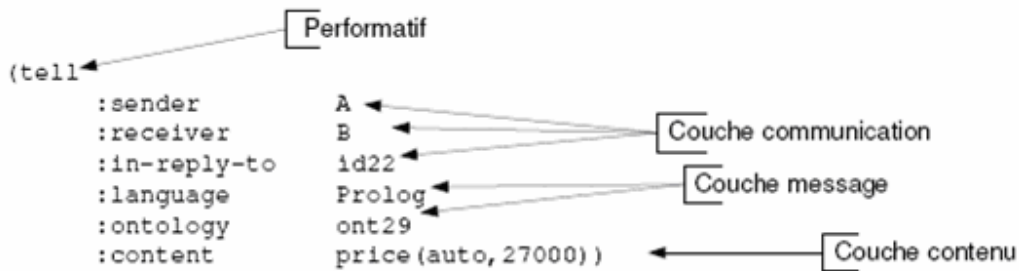


Figure 9. Un message KQML

Kqml du fait qu'il est le premier standard des ACLs présente certains avantages, ce qui faisait de lui, un langage fort utilisé [24]. Toutefois, KQML présente également certains inconvénients. Parmi lesquels, on identifie l'Ambiguïté dans certains performatifs qui ne sont pas clairs ou des performatifs manquants : Il s'agit sans doute de la critique la plus importante. Une classe indispensable d'action est manquante, celle des promissifs. Il n'est donc pas possible qu'un agent s'engage par un acte tel que la promesse (promise), et les expressifs (montrer des émotions ou des gestes nécessaire dans la communication du face à face) [24].

3.3.2.2 FIPA-ACL :

La fondation pour les agents intelligents physiques (FIPA pour "Foundation for Physical Intelligent Agents") est un organisme international à but non lucratif regroupant diverses compagnies dont l'objectif est de fournir des standards pour faciliter l'interopérabilité entre les systèmes à base d'agents. Le Langage qui a résulté de ces efforts est FIPA-ACL [24]. Comme KQML, ce dernier est basé sur la théorie des actes de langage : les messages sont vus comme des actes communicatifs [29].

FIPA-ACL est très proche de KQML: leur syntaxe est similaire sauf pour un petit nombre de performatives dont le nom diffère [29].

Par exemple, dans la figure 11, l'agent A informe l'agent B qu'il pleut aujourd'hui à l'aide d'un message FIPA-ACL.

```
(inform
:sender      A
:receiver    B
:content     " weather (today,raining) "
:language    Prolog
:ontology    weather01)
```

Figure 10. Un message FIPA-ACL

FIPA-ACL contient des actes de communication permettant les demandes et transmissions d'information, les négociations, les demandes d'action et la gestion des erreurs. Ces actes de communication peuvent s'utiliser comme primitives ou bien être combinés [29].

3.3 La communication Humain-Agent :

Avec la prolifération des applications de commerce électronique, d'e-learning, de propositions de services, les humains se sont mis en interaction avec les agents artificiels. Aujourd'hui, dans ces environnements, l'humain est tout simplement considéré comme un agent et intégré avec des agents artificiels au sein de systèmes multi-agents parfois qualifiés d'hétérogènes.

Cependant, lorsque un humain et une machine communiquent ensemble, le risque d'avoir des malentendus ou des conflits est majeur. Un agent logiciel et un utilisateur humain ne partagent ni le même vocabulaire, ni les mêmes expériences, ni les mêmes perceptions, ni les mêmes valeurs. Dans ce cas, il est très important de trouver un point commun entre les deux participants à la communication. Pour parvenir à une situation permettant d'exploiter au mieux les capacités spécifiques de chacun, il est nécessaire que chaque participant (humain et agent artificiel) s'adapte en partie à l'autre.

Cependant, beaucoup des mécanismes d'interaction employés entre agents artificiels sont inadaptés à l'humain, et les spécificités de l'humain commencent à peine à être prises en compte dans le domaine des SMA.

Pour franchir cet obstacle, il faut arriver à définir un mode de communication exploitable au mieux par chacune des deux parties.

Les modèles couramment utilisés en informatique sont très spécifiquement adaptés à la communication entre entités artificielles. Aucun des langages spécifiques aux agents artificiels tel que KQML et Fipa-ACL n'est fortement adapté aux interactions mixtes, car les informations traitées par les deux parties sont de type différent. Cependant, si l'on souhaite se baser sur un langage existant (la langue naturelle, ou plus précisément, la théorie des actes de langages, dans notre cas) pour développer un langage mixte, il faudra adapter et enrichir ce langage pour que ni agents artificiels ni humains ne soient bridés par les capacités expressives du langage.

Donc, pour cette raison, nous cherchons un point commun entre les capacités et les nécessités de chaque type d'agents. Un agent doué de réelles capacités conversationnelles doit permettre l'emploi d'un langage plus expressif et plus adapté à l'être humain. Symétriquement, un humain cherchant à exploiter au mieux les capacités des agents avec lesquels il interagit devra adapter son langage, qui est, lui, spécifiquement adapté à la communication entre humains [08].

Donner à des agents des capacités de communication aussi proches que possible de celles des humains permettra facilement aux agents et humains de communiquer sur des tâches pour lesquelles les méthodes de communications employées sont adaptées, c'est-à-dire, typiquement, des problèmes destinés à être résolus par des humains.

Bien sûr, pour permettre l'exploitation de la liberté accordée au niveau de la communication avec l'utilisateur, l'agent devra être doté de capacités de communication suffisamment riches et adaptées à cet usage [08].

3.4 Environnements des rassemblements d'agents humains et artificiels :

Les humains et les agents se rassemblent dans diverses applications, parmi les combinaisons d'environnements où ils se rencontrent, on peut citer :

- **Environnement réel** : C'est le cas des travaux sur les interactions de groupes d'humains et de groupes de robots, où des robots vivent dans les couloirs du laboratoire [08].

- **Environnement virtuel** : les humains et les agents se rencontrent au sein d'un espace virtuel. Très souvent, dans cette situation, l'espace est représenté en trois dimensions, et les agents (naturels ou artificiels) communiquent principalement par le biais du texte [08].

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 [18] comme :

- Les agents pédagogiques pour la formation ;
- Les agents conversationnels pour le web ;
- Les agents présentateurs d'information ;
- Les agents conseillers médicaux...

3.5 La communication chez les humains :

Lorsque les humains communiquent entre eux, ils utilisent plusieurs modalités comme la parole, les gestes, les postures du corps, le regard, les expressions faciales. Cette multimodalité est caractéristique de la communication Homme-Homme. Dans un dialogue Homme-machine, l'ordinateur peut aussi être perçu comme un partenaire dans la communication : il doit savoir s'adapter aux caractéristiques de l'utilisateur, à son langage et à sa manière d'interagir.

Malheureusement, les Interfaces Homme-Machine (IHM) actuelles utilisent souvent un ensemble restreint de médias (souris, clavier, écran) et restent très limitées vis à vis de la représentation des modalités comme la parole, les gestes, les expressions faciales...

Dans le cadre des agents, il est facile de comprendre que c'est un comportement assez difficile à reproduire étant donné la subtilité dont fait preuve ce mécanisme chez l'être humain [24]. Cette problématique pour les agents est toutefois très intéressante et même très importante puisque les agents seront constamment face à ce problème.

Il existe bien d'autres problèmes avec le dialogue, notamment au niveau du contenu qui y est véhiculé. Des difficultés comme les termes de remplissage, les pauses, les ellipses et les

phénomènes déictiques (anaphore, cataphore, ...) sont également présents dans les dialogues entre humains.

Pour les êtres humains, la façon de communiquer peut sembler évidente, puisqu'on peut directement déceler le sens, l'humeur et l'intention de l'émetteur. Pourtant, ce mécanisme est loin d'être évident dans les SMA.

L'observation des conversations humaines est la base de l'élaboration de protocoles sophistiqués. Il faut aussi garder à l'esprit que le langage naturel a le plus grand pouvoir d'expression. Pouvoir garantir que le langage développé, à l'instar du langage naturel, permet de tout dire et sous certaines conditions d'être compris de tous est un aspect majeur.

Pour concevoir des interfaces autorisant des modalités d'entrée plus naturelles, comme la parole et le geste, il nous faut préalablement connaître la manière dont l'humain utilise ces modalités pour communiquer avec un ordinateur. Dans ce cas, la métaphore de la communication Homme-Homme est appliquée pour créer des agents capables d'interagir avec l'humain.

Lors d'une conversation, les interlocuteurs utilisent une grande variété de comportements à la fois verbaux (paroles) et non verbaux. La communication Humaine non verbale est extrêmement riche et étendue ; elle comporte les expressions faciales, mouvements de la tête, regards, mouvements des mains, postures, distances interpersonnelles, apparences physiques... A ces facteurs humains, il faut ajouter les parfums, les vêtements, les accessoires...

Tous ces facteurs influencent ce que nous disons et comment nous nous comportons. Le non-verbal naît du mimétisme que chacun d'entre nous acquiert dès le plus jeune âge. Quant à son décodage, il est avant tout culturel.

Dans le paragraphe suivant, nous présentons les comportements non-verbaux par ordre décroissant d'importance des parties du corps humain :

- ◆ **Affichages faciaux** : Les expressions du visage sont un moyen important de communication. Le message verbal est souvent enrichi par des actions telles que le sourire, le mouvement des yeux, le soulèvement des sourcils, et hochement de tête... De telles expressions sont synchronisées avec nos propres paroles ou celles des autres. Elles sont utilisées pour

contrôler le flux de la conversation. Elles sont également le canal le plus important pour démontrer l'émotion [30].

D'un point de vue anatomique, les affichages faciaux font intervenir les parties suivantes du corps : le front, la position des sourcils, le mouvement des yeux, les expressions de la bouche et le mouvement de la tête.

Les affichages faciaux n'ont pas la même durée suivant la fonction qu'ils réalisent :

- Ceux liés à la personnalité sont constants, pendant toute la vie (regard avec les yeux grand ouverts)
- Ceux liés à l'état émotionnel peuvent durer aussi longtemps qu'une émotion est ressentie (les yeux baissés lors d'une dépression)
- Ceux synchronisés avec les unités de la conversation durent très peu de temps.

Le visage réalise 4 fonctions communicatives :

- a) Phonologique :** forme des lèvres (visèmes) qui change en fonction du phonème prononcé. Ce sont les affichages faciaux phonogènes, c'est-à-dire nécessaires à la production de parole.
- b) Syntaxique, grammaticale :** accompagne le flot de la conversation, co-occurrence avec les syllabes accentuées, les pauses, les questions. Cette fonction est réalisée par les sourcils levés, le mouvement de la tête ou le clignement des yeux.
- c) Sémantique :** accentue le discours, remplace un mot (mouvement de la tête à la place de dire "oui").
- d) Interactionnelle, liée au dialogue :** Ces fonctions sont modulées par les caractéristiques de l'orateur et de son interlocuteur (personnalité, émotion, le degré de formalité de la relation sociale)

◆ **Les gestes des mains :**

Les fonctions gestuelles sont nombreuses:

- **Fonction épistémique** : la main est un organe de réception par le sens tactilo-kinesthétique (informations sur la forme, l'orientation, la distance, la grandeur par l'utilisation du toucher et de mouvements exploratoires) et la perception proprioceptive (informations sur le poids, les trajectoires, les mouvements des objets par l'utilisation de récepteurs placés dans les articulations et les oreilles)

- **Fonction ergotique** : la main est un organe moteur qui agit sur le monde physique pour le transformer (application de forces grâce à la structure osseuse et les muscles).

- **Fonction sémiotique** : la main est un émetteur d'information. Elle s'adresse à la perception visuelle des interlocuteurs.

Il existe plusieurs catégories de gestes classés suivant leur puissance d'expression décroissante :

- **Les gestes des langues des signes** : ils forment une véritable langue dotée d'une syntaxe, permettent même la création dynamique de signes non standard en fonction des besoins.
- **Les gestes co-verbaux** : ils s'effectuent simultanément avec la parole et permettent d'illustrer ou de compléter le message verbal.
- **Les gestes propositionnels** : gestes conscients, présents en petit pourcentage, considérés comme un constituant grammatical de la parole. Par exemple, dire "Une caisse de cette taille-là" et utiliser les mains pour mesurer la taille d'un espace symbolique.
- **Les gestes spontanés** : ils sont produits de façon inconsciente. Ils représentent la majorité des gestes (représentent un élément du sujet qui a une forme physique, mimes d'actions, rouler les mains pour dire qu'un processus est en cours...)
- **Les gestes quasi-linguistiques** : ce sont des gestes conscients qui rendent la communication possible sans usage de la parole ; ils ne bénéficient pas d'un enrichissement du message par combinaison des gestes entre eux. Ils sont de deux types :

1. Naturels : on les appelle les emblèmes. Ils représentent 10% à 20% des gestes et sont culturellement dépendants (comme le “V” de la victoire).

2. Artificiels : comme le langage des aviateurs, des plongeurs ou des courtiers en bourse

La distribution des unités gestuelles est similaire à la distribution des unités d’intonation : le début et la fin sont synchronisés, le temps fort du geste et l’accent tonique de l’intonation sont sur la partie rhématique de la phrase [02].

◆ **Régulation des tours de parole :**

L’information interactionnelle a pour rôle principal de négocier la prise de parole et de réguler la conversation. Ce protocole se fait à l’aide de signaux appelés régulateurs. On peut identifier trois catégories de fonctions interactionnelles :

1. l’amorce et la fin : réagir à la venue d’une nouvelle personne, se retirer de la conversation, “dire” au revoir.

2. prise de parole : passer la parole, vouloir la parole, prendre la parole (établissement de l’orateur).

3. réaction (ou retour) : demander une réaction, donner une réaction.

Lorsqu’un individu parle, il émet le signal de l’orateur en cours de parole : il veut conserver la parole et s’assurer que son interlocuteur le suit. Ce signal est souvent suivi par une réaction de son interlocuteur, qui peut être suivi d’un signal de continuation de la part de l’orateur si ce dernier veut conserver le tour de parole.

La synchronisation entre les intervenants est primordiale : un retard dans une réponse peut être interprété par l’autre interlocuteur comme la volonté de ne pas coopérer ou comme un fort désaccord. Le tableau 5 montre comment sont réalisées les fonctions interactionnelles :

Fonction interactionnelle	Comportement interactionnel
<i>Amorce et Fin</i>	
Réagir à une nouvelle personne	Regarder rapidement vers l'autre
Se retirer de la conversation	Regarder ailleurs
Adieu	Regarder l'autre, mouvement de la tête, dire au revoir de la main
<i>Prise de parole</i>	
Donner la parole	Regarder, lever les sourcils (suivi d'un silence)
Vouloir la parole	Mains levées dans l'espace des gestes
Prendre la parole	Regarder ailleurs, commencer à parler
<i>Réaction</i>	
Demande de réaction	Regarder l'autre, sourcils levés
Donner une réaction	Regarder l'autre, mouvement de la tête

Tableau 5. : Fonctions interactionnelles et leurs réalisations comportementales

Un autre point important doit être pris en considération dans la communication humaine, celui des émotions.

◆ **Emotion :**

Une émotion est un ensemble de variations à court terme (quelques secondes) de l'état mental [02]. Cela inclut des réponses largement physiques comme la peur et des réponses cognitives comme la jalousie.

Elle peut être provoquée aussi bien par des stimuli internes que externes. Il peut aussi arriver que des émotions soient mélangées lorsqu'un événement provoque plusieurs sentiments.

Le visage est l'expression des émotions. C'est la partie du corps que nous regardons le plus lorsque nous parlons avec quelqu'un. On distingue 6 émotions fondamentales qui sont reconnues dans le monde entier, quels que soient les visages, les cultures, les races:

Pelachaud s'appuie sur les travaux d'Ekman et Friesen [02] qui ont trouvé que 6 émotions avaient des expressions faciales universelles :

- La joie (Bouche ouverte, joues en arrière, yeux rieurs, sourcils relevés)
- La surprise (bouche très légèrement ouverte, visage figé)
- La peur (bouche plus ouverte, yeux exorbités, visage tremblant)
- La colère (dents serrées, nez replié, sourcils froncés)

- La tristesse (bouche fermée, yeux légèrement fermés, front marqué)
- Le dégoût (bouche avec lèvres vers le bas, sourcils froncés)

Pelachaud distingue une émotion d'un réflexe ou d'une humeur par des considérations temporelles:

- un réflexe est bref (inférieur à 0.5 seconde) et ne peut pas être complètement inhibé. Il est impossible de le simuler parfaitement.
- une émotion a une durée limitée (entre 0.5 et 4 secondes). Les expressions faciales qui traduisent une émotion ne peuvent pas durer plusieurs minutes ou heures sans montrer de crampes.
- une humeur dure plus longtemps qu'une émotion (une journée entière) ou apparaît plus fréquemment qu'une émotion (dix fois par jour). Elle se réfère à la tendance qu'une personne peut avoir à ressentir une émotion.

◆ Règles d'affichage :

Certaines émotions sont universelles (indépendantes de la culture) et que l'intensité et le moment d'apparition de leur expression diffèrent cependant suivant les cultures.

Elle introduit la notion de règles d'affichage des émotions qui définissent "quelle personne peut montrer quelle émotion à qui et quand". Elles traduisent l'environnement social et les dépendances culturelles.

Ces règles n'induisent pas la neutralisation de l'émotion ressentie mais influencent son intensité et sa durée. Certaines situations peuvent demander de masquer une émotion ressentie par une autre émotion, ou de simuler une émotion. Le sourire est l'expression la plus fréquente dans les cas de simulation ou de masquage.

3.6 Les limites des ACL :

Bien que ces langages soient majoritairement utilisés dans le domaine des SMA, plusieurs études critiquent l'approche par protocoles en considérant que les protocoles sont trop limités et figés pour modéliser la dynamique de conversations entre agents et humains. En outre, dans le cadre des interactions entre agents logiciels et agents humains, les primitives de langage accessibles par un langage comme l'ACL/FIPA et KQML sont trop peu nombreuses et ne

permettent pas le développement d'agents conversationnels capables de prendre en compte les besoins de l'utilisateur humain notamment la communication non verbale.

Afin d'illustrer ces limitations, considérons les 2 messages suivants en ACL/FIPA :

```
(request (inform
  :sender I :sender J
  :receiver J :receiver I
  :content "weather(today,X)" :content "weather(today,raining)"
  :language Prolog :language Prolog
  :ontology weather42 :ontology weather42
) )
```

Dans cet exemple, le premier message correspond à une requête de l'agent I à l'agent J dont le contenu est écrit en Prolog et correspond à demander quel sera le temps aujourd'hui. L'ontologie (ontology) correspond au domaine de connaissances dans lequel s'inscrit le message. Le destinataire J informe en retour l'agent I qu'il va pleuvoir.

Considérer une conversation entre agents logiciels et utilisateurs humains à partir de ces seuls types d'échanges reste de façon évidente trop limitative, notamment si l'on considère des données de nature non verbales (telles que l'orientation du regard, les gestes, la posture du corps) qui sont nécessaire pour la communication du face à face.

Par ailleurs, les langages tels que KQML et ACL/FIPA ne permettent pas encore d'encapsuler des messages en langue naturelle et ne possède donc pas encore la richesse verbale nécessaire à des interactions conversationnelles Homme-Machine évoluées. Malgré que ces langages s'appuient sur la théorie des actes de langage, ils ne considèrent parmi ces actes que les assertifs et les directifs.

Nous considérons en revanche que, dans le cadre du développement d'agents conversationnels susceptibles d'interagir avec des utilisateurs humains, les langages de communication entre agents doivent permettre de prendre en compte l'ensemble des catégories d'actes définies par la théorie des actes de langage. Il s'agit alors d'inclure dans ces langages les promissifs, les déclarations et, surtout dans le cadre de la modélisation des émotions, les expressifs qui fournissent le principal vecteur de la dimension émotionnelle du langage [32].

Conclusion :

Dans ce chapitre, nous avons abordé différents aspects de la communication tant du point de vue de la structure, du contenu que des mécanismes sous-jacents. Bien que l'objet de notre étude ne réside pas dans la communication Humain-Humain ou Machine-Machine mais plutôt Humain-Machine, il est important de se laisser guider par les mécanismes étudiés depuis déjà plusieurs décennies dans ces deux champs de recherche. Cela permet un meilleur développement des bases d'un modèle de communication pour agents, particulièrement lorsque l'on désire construire des mécanismes de communication à la fois flexibles et hautement utiles comme c'est le cas avec la communication humaine.

Chapitre 4 : Les agents conversationnels

Au début de l'informatique, les ordinateurs étaient réservés à des utilisateurs élitistes qui étaient des ingénieurs et des informaticiens experts. La complexité des systèmes ne facilitait pas l'utilisation de l'outil informatique. Aujourd'hui, les ordinateurs sont partout, ils tiennent de plus en plus de place dans notre vie, ceci a permis à une grande variété de gens d'utiliser l'informatique dans diverses applications.

Cependant, malgré l'accroissement sans cesse des capacités des ordinateurs en terme de vitesse et de stockage, l'interaction de l'utilisateur avec la machine n'a pas beaucoup évoluée. La machine idéale est sans doute celle qui s'adapte à l'homme et non l'inverse. Pourtant, dans la courte histoire de l'informatique, c'est plutôt l'utilisateur qui a du s'adapter à l'ordinateur.

L'amélioration des IHM est devenue la préoccupation majeure de plusieurs chercheurs. Le but de ces recherches est de perfectionner les échanges entre l'utilisateur et l'ordinateur de manière à ce que ces échanges soient des plus naturels possible.

Or la communication que l'on considère naturelle est celle qui se fait de l'homme à l'homme, d'où est apparue l'intelligence artificielle qui a pour but de reproduire le comportement rationnelle de l'homme par la machine.

Pour rendre l'interaction Homme-Machine plus réaliste, les agents intelligents sont apparus dont le but est de faciliter la tâche de la représentation du comportement humain par la machine car l'utilisateur s'exprime le mieux et le plus spontanément par le langage naturel.

L'agent doit comprendre ce que l'utilisateur dit, par sa parole, par ses comportements non verbaux (gestes, mimiques faciales...) et réciproquement doit répondre à l'utilisateur par la parole et des comportements non verbaux. Pour répondre, la machine doit donc prendre l'apparence d'un être humain. Un nouveau type d'agents a été proposé: l'**agent conversationnel**.

4.1 Notions sur les agents conversationnels :

Un agent conversationnel est un composant logiciel capable de raisonner et de dialoguer de manière autonome avec un utilisateur. Les agents conversationnels tels qu'ils existent

actuellement, englobent l'ensemble des paramètres de communication et les signes non verbaux qui font partie intégrante des processus communicatifs. Ils associent dans leurs programmes le geste, l'intonation, l'animation multimodale de l'expression faciale.

Une autre définition considère l'Agent Conversationnel Animé comme une interface homme-machine utilisant un personnage animé affiché à l'écran doté de capacités communicatives inspirées de la communication humaine.

On trouve plusieurs types d'agents conversationnels pouvant revêtir différentes apparences :

- **Textuelle uniquement** : c'est le cas des rubriques d'aide avec recherche par questions libres.
- **Graphique** : avec une représentation statique de l'agent permettant de lui donner une visibilité supérieure.
- **Les ACA** : cela veut dire « Agent Conversationnel Animé ». Cela correspond au terme anglo-saxon ECA pour « Embodied Conversational Agent » qui aurait pu être traduit par « Agent Conversationnel Incarné » (ACI), c'est-à-dire ayant un 'corps' [36]. Ils ont une représentation visuelle (humanoïde ou non) capable de s'animer et qui facilite la communication [03]. Les ACA sont le type d'agents conversationnels le plus proche à la métaphore humaine.

Sur le web, on trouve un type d'agent conversationnel, appelé généralement Chatterbot. C'est un programme informatique qui simule une conversation entre l'Homme et la machine. Nous entrons un type de question ou d'affirmation quel qu'il soit, et le chatterbot répond, tout comme le ferait une personne. Les chatterbots tentent de créer l'illusion qu'un véritable échange a lieu entre deux entités pensantes vivantes. Il faut parfois nous pincer pour nous convaincre que nous ne parlons pas à une personne réelle.

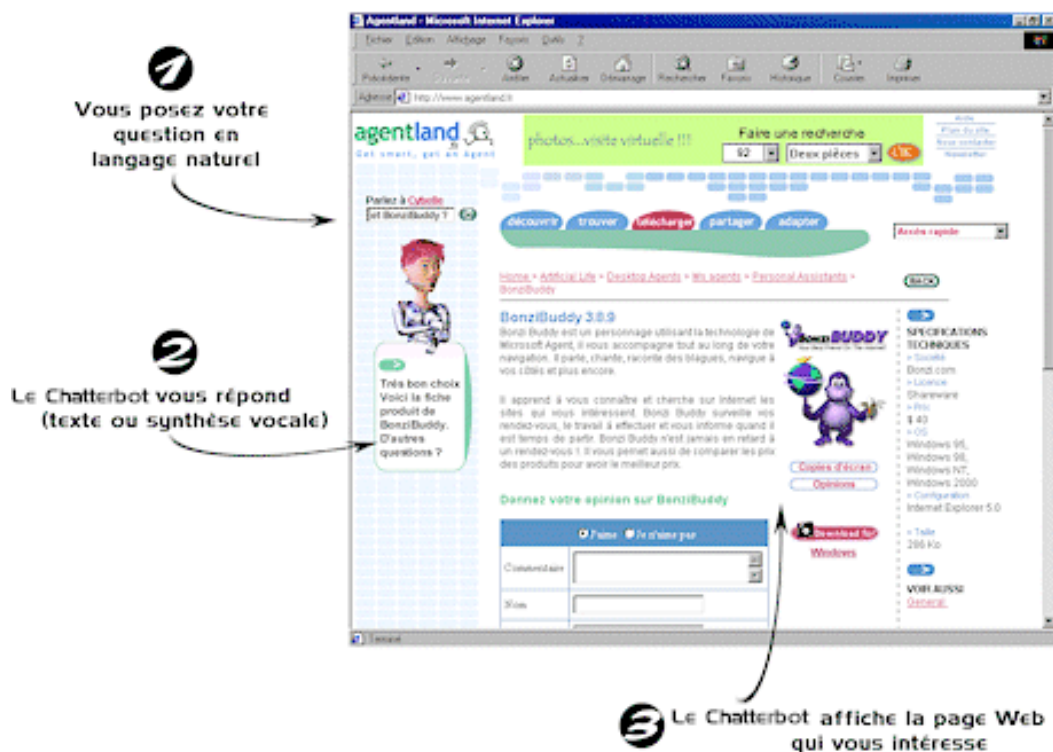


Figure 11. Un exemple du fonctionnement du chatterbot Cybelle sur AgentLand.com

La conception de tels Agents est par nature pluridisciplinaire [05] : les domaines concernés incluent le graphisme, l'animation, l'intelligence artificielle, le traitement automatique de la langue, l'architecture logicielle, le design d'interface, la synthèse vocale, l'ergonomie, la psychologie, la sociologie, les arts du spectacle... Il semble évident que peu d'équipes regroupent une telle diversité de compétences. Les recherches sur les Agents Conversationnels se focalisent donc généralement sur un ou plusieurs points de leur chaîne de conception, faisant plus ou moins abstraction des autres.

L'Agent Conversationnel idéal [05] devrait être intelligent, capable de comportements sociaux, et tirer parti de sa représentation visuelle pour renforcer sa crédibilité notamment par des comportements non verbaux sophistiqués et pertinents, par l'expression d'émotions...

Pour être utiles, les Agents Conversationnels doivent être compétents et rendre des services; pour être utilisables, ils doivent communiquer de manière efficace et robuste ; et pour être confortables, ils doivent satisfaire nos attentes sur la manière dont une conversation se déroule.

4.2 Agents conversationnels et avatars :

Le domaine des Agents Conversationnels rejoint par certains points celui des Avatars, qui ont cette spécificité de ne pas être autonomes [05]. Le terme avatar provient d'*avatāra*, le mot sanskrit désignant chacune des différentes incarnations du dieu hindou Visnu sur terre [34]. Ainsi, un avatar est le représentant d'un utilisateur dans un monde virtuel.

Contrairement à un ACA qui est contrôlé par un agent artificiel, un avatar est complètement contrôlé par un humain. Les avatars sont principalement utilisés dans les jeux vidéo (le joueur guide les actions du personnage grâce à une manette de jeu, par exemple), ainsi que dans les espaces de discussions virtuels, où, dans le but d'augmenter la convivialité, ils permettent de donner une représentation graphique aux utilisateurs participant à la communication [05].

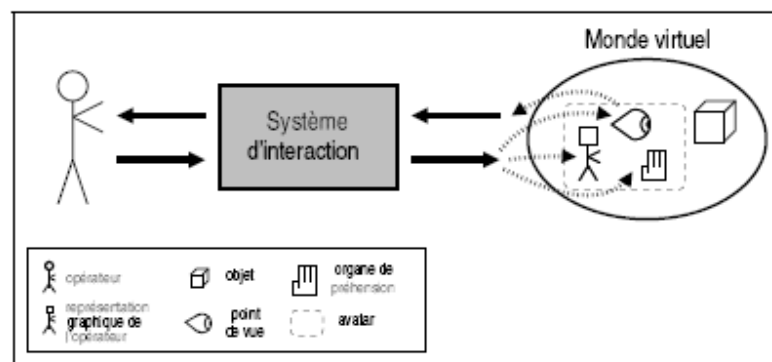


Figure 12. L'avatar classique

Dans les mondes virtuels, on peut distinguer deux rôles importants pour l'avatar [34] :

1. La représentation visuelle (texte, 2D ou 3D) de l'opérateur dans un environnement multi-utilisateurs; Dans ce cas, l'avatar est le support de communication entre les différents utilisateurs présents dans le monde virtuel.

2. le contrôle de l'ensemble des moyens d'interaction à la disposition de l'utilisateur : ici, l'avatar perçoit et agit pour l'opérateur dans le monde virtuel.



Figure 13. Expressions faciales sur avatars

4.3 Historique et évolution des ACA:

Le concept d'agent conversationnel est relativement ancien puisqu'il remonte à la fin des années 60, lorsque fut développé l'un des plus illustrés en la personne d'Eliza de Joseph Weizenbaum [03], qui jouait un rôle de psychanalyste virtuel. Lorsqu'une personne dialoguait avec Eliza, elle reformulait la phrase sous forme de question, invitant l'utilisateur à donner de plus amples explications et à l'infini. La conversation n'était pas tout à fait de haute qualité, mais néanmoins suffisamment " intelligente " pour créer une confusion à l'époque où il n'était pas commun d'interagir avec des ordinateurs [40].

Ces modèles ont rapidement montré leurs limites à cause de leur incapacité à gérer les non attendus (mots inconnus, phrases incomplètes) et les erreurs de manière positive comme dans les conversations courantes.

Dans les années 90, les agents conversationnels sont devenus majoritairement des agents conversationnels animés avec l'adjonction d'un avatar représentant l'agent, doté au minimum d'une représentation graphique statique, mais aussi souvent d'expressions animées et d'une voix de synthèse leur permettant de vocaliser leurs réponses.

La théorie des actes de langage est utilisée, surtout la distinction entre les forces illocutoires : assertion sur le monde, ordre ou interdiction, promesse, déclaration. De plus en plus, on tente également de leur donner une véritable incarnation en les dotant d'une personnalité propre, rendue perceptible notamment par le biais de leurs émotions [03]. Tous ces éléments contribuent à renforcer la présence de l'agent du point de vue de l'utilisateur grâce à l'utilisation de plusieurs modalités en entrée (reconnaissance vocale, des émotions, des mouvements de

l'utilisateur dans l'espace, des gestes déictiques. . .) comme en sortie (synthèse vocale, les expressions du visage et les gestes, représentation des émotions. . .).

4.4 Domaines d'application et de recherche:

Parallèlement, les agents conversationnels sont sortis des laboratoires de recherche universitaires et des solutions commerciales à base d'agents conversationnels commencent à être offertes par des entreprises spécialisées.

Les domaines d'utilisation de ceux-ci sont multiples et diversifiés : [03]

- **Le commerce électronique :** Sur internet, un agent agit de manière similaire à un vendeur en magasin. Il peut chercher un produit correspondant aux critères définis par l'utilisateur, il renseigne et rassure, augmentant ainsi le taux d'achats. Il apporte également un aspect vivant aux sites Internet, aide à fidéliser les clients et à recueillir plus naturellement des informations à leur sujet.
- **Les services d'assistance à distance :** Il est potentiellement très avantageux pour une entreprise de substituer un agent conversationnel (capable de gérer plusieurs sessions en parallèle) aux opérateurs humains pour répondre aux questions les plus fréquemment posées, avec éventuellement une possibilité de basculer de manière transparente sur un opérateur humain si les questions s'avèrent trop pointues par rapport au champ de connaissances de l'agent. Le client peut pour sa part ainsi espérer voir son temps d'attente réduit.
- **L'apprentissage et la formation à distance :** l'agent peut pallier le manque de contact humain qui est le principal travers des solutions classiques d'e-learning dans lesquels l'élève se sent souvent seul face à des documents statiques ou des présentations animées peu interactives.
- **Les jeux vidéos :** transformer les personnages contrôlés par l'intelligence artificielle en véritables agents conversationnels offre très certainement de nouvelles possibilités créatives et ludiques renforçant l'aspect vivant et réaliste des univers virtuels. C'est par exemple le cas des univers virtuels dotés d'agents discutant entre eux et sur lesquels le joueur peut intervenir pour modifier le cours de l'histoire.
- **Les assistants personnels :** offrir un agent personnel adapté aux goûts et aux habitudes de son utilisateur, omniprésent par son intégration dans un périphérique mobile de type PDA

ou téléphone portable, et qui l'assiste pour toutes ses démarches quotidiennes est un des objectifs les plus ambitieux promis par les agents conversationnels.

Depuis le début des années 2000, des projets de grande envergure généralement financés par les pouvoirs publics visent à aller au-delà des prototypes déjà réalisés en développant des agents mieux modélisés et aux capacités de dialogue plus poussées pour pouvoir élaborer des applications orientées vers le grand public.

Le premier axe de recherche concerne l'utilisation d'autres modalités que la langue naturelle pour que l'agent puisse à la fois percevoir davantage d'informations sur l'environnement réel dans lequel il se situe et mieux transmettre son message dans ses interventions en utilisant des signes de communication extra-verbaux (gestes et expressions du corps et du visage). De nombreux travaux tournent donc autour de la détection et la restitution d'émotions, qui sont jugés indispensables à la réussite d'une interaction Homme-Machine de qualité.

4.5 Les ACA dans la communication avec les humains :

L'interaction des usagers avec les ordinateurs est seulement une interaction par signes. Elle n'est pas symétrique ni du point de vue des modalités, ni du point de vue des rôles. La plupart du temps, la machine présente des pages sur un écran comportant des textes et des images et envoie des messages auditifs (mode visuel et auditif). L'utilisateur agit principalement par l'écriture (mode textuel) et avec la souris (mode gestuel). Cette interaction ne met en jeu que le monde de l'ordinateur, son monde interne et les flux arrivant par ses capteurs, pas celui de son interlocuteur.

Pour améliorer les IHM, certaines équipes de recherche proposent de présenter dans les systèmes de dialogue, un visage de synthèse animé, articulant de manière coordonnée la synthèse vocale prononcée par la machine.

C'est le cas, par exemple, de l'agent August [62] présenté sur la Figure 15, Il est capable de reproduire des mimiques telles que l'interrogation, le doute, l'étonnement, etc.

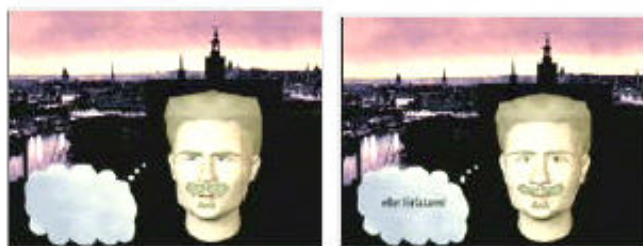


Figure 14. August, un agent animé

August a été intégré dans une borne interactive, en gare de stockholm. Sa tâche consistait à dialoguer avec les usagers, en langue naturelle orale, pour répondre à leurs questions et fournir des renseignements touristiques.

D'autres études montrent l'intérêt d'intégrer, non plus seulement des faces ou des visages, mais des représentations du corps entier.

Avec l'introduction des agents conversationnels dans les interactions Homme-Machine, l'utilisateur semble être plus captivé par un Dialogue dans lequel il perçoit son interlocuteur. Cela devrait être particulièrement vrai avec un public jeune (d'enfants et d'adolescents) et/ou novice.

Les " temps morts " comme lors des appels à des ressources distantes via le réseau, où d'ordinaire l'utilisateur s'impatiente rapidement, semblent être mieux acceptés lorsque l'agent animé prend le relais.

4.6 Etat de l'art des agents conversationnels :

Les ACAs ont été créés dans le cadre de la personnalisation d'interfaces utilisateur ainsi que dans le cadre de tâches pédagogiques, commerciales, d'assistance... ayant la capacité d'exprimer des comportements non-verbaux tels que la gestuelle, le regard, et les expressions faciales communicatives [30].

Les agents présentés dans le paragraphe suivant ont des caractéristiques très variées, dans leur apparence, leurs fonctions et leurs compétences. Nous les avons classés suivant leur degré d'autonomie.

4.6.1 BodyChat :

C'est un système qui analyse les phrases tapées par l'utilisateur lors de sa discussion avec son interlocuteur et qui se sert de cette analyse pour générer le comportement de l'avatar

matérialisant l'utilisateur [08]. BodyChat permet à l'utilisateur de déléguer le contrôle de l'avatar : les utilisateurs communiquent en utilisant le Chat, et les avatars calquent leur comportement sur les informations transmises par ce biais. Les gestes de l'avatar sont des gestes d'accompagnement de l'information (matérialisations d'emphase sur certains mots) ou de la conversation (prise ou don de parole).

La figure 17 montre un exemple d'animation du visage de l'avatar : le mot "very" est accentué par un hochement de tête, et le point d'exclamation retranscrit par un haussement de sourcils [08].

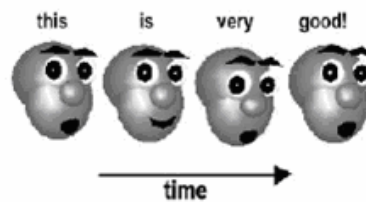


Figure 15. Exemple d'action de body chat

Le centre d'intérêt est donc l'avatar, auquel l'utilisateur peut déléguer une partie des tâches de communication, ce qui améliore en retour la qualité d'interaction entre utilisateurs, à la fois par l'enrichissement que pourra apporter l'avatar sur certaines modalités, et par l'allègement de la charge de l'utilisateur, qui n'a pas à contrôler systématiquement l'avatar.

L'autonomie donnée à l'avatar joue sur trois types d'action :

- L'envoi spontané de messages (lorsque l'avatar dispose d'une grande autonomie, il peut prendre l'initiative d'envoyer des messages),
- La manière d'accomplir les actions (l'agent peut influencer sur la forme d'une action requise par son utilisateur)
- Les messages complémentaires aux messages de l'utilisateur (expression faciale adaptée aux messages de l'utilisateur, par exemple).

Si on laisse à l'avatar le maximum d'autonomie, il peut alors être considéré comme un agent autonome [08].

4.6.2 *Le monde INWIVO : de Nadine Richard*

Dans les travaux de Nadine Richard [34], l'auteur s'est intéressé aux mondes virtuels habités, c'est-à-dire aux environnements virtuels dans lesquels interagissent des agents virtuels et des utilisateurs. Elle a cité Le Diamond Park qui est un exemple particulièrement intéressant de monde virtuel habité. Il s'agit d'un parc de loisirs virtuel dans lequel divers agents assistent les utilisateurs.

4.6.3 *Ananova :*

Ananova est un agent qui diffuse Vingt-quatre heures sur vingt -quatre, des informations accessibles sur le Web. Ananova démontre qu'un agent virtuel est parfaitement capable de remplacer un humain dans ce genre de tâches bien spécifique, qui consiste à diffuser de l'information aux internautes, ou aux personnes connectées grâce à un téléphone portable, par exemple. Les différents tons, gestes, mimiques et attitudes qu'elle prend imitent remarquablement un comportement humain. Toutefois il n'y a là aucune interaction entre l'utilisateur et cet agent virtuel. On ne peut pas l'interrompre, lui demander de répéter, etc. On doit se contenter, pour l'instant, de l'écouter [62].



Figure 16. L'agent virtuel Ananova

4.6.4 *Ppp persona :*

PPP (Personalized Plan-based Presenter) Persona est un Agent destiné pour la présentation des informations, c'est-à-dire capable de faire des exposés générés dynamiquement [05]. Il a été créé pour la personnalisation d'interfaces utilisateurs [30]. Il a été appliqué à des

contextes d'assistance technique, ou encore au domaine touristique (présentation d'informations sur Internet).

PPP Persona présente verbalement des informations et désigne des éléments graphiques à l'aide d'un pointeur (baguette virtuelle). Il produit également des comportements non-verbaux tels que des gestes déictiques et des expressions faciales communicatives.

4.6.5 MS-Agent

Microsoft propose, en effet, des agents virtuels animés dits MSAgents que l'on intègre aux pages Web (voir Figure 18). Ceci est censé améliorer l'interaction entre l'homme et la machine, puisque l'agent peut adopter des attitudes explicites, comme la surprise, l'explication, l'incompréhension ; se déplacer sur l'écran, désigner des parties de l'écran (à gauche, à droite, etc.). Ces MS-Agents disposent de logiciels de synthèse et de reconnaissance vocale, qui s'installent sur la machine cliente, lors du chargement de la page Web.



Figure 17. Quelques attitudes de Merlin (MS-Agent)

La Figure 19 ci-après présente l'agent Merlin, faisant un geste de désignation, pour signifier à l'utilisateur où regarder sur l'écran, en même temps qu'il donne la réponse à la question posée [62].



Figure 18. Merlin fait un geste de désignation

L'agent attire parfois l'attention de son interlocuteur de manière ludique, en cognant au carreau (l'écran) de l'intérieur. Des effets sonores accompagnent les gestes, et les mouvements labiaux sont assez réalistes. Et pour combler les temps morts, Merlin par exemple, sort une boule de cristal de sa poche et la manipule quelques instants pour signifier que la recherche est en cours...

4.6.6 Rea (*Real Estate Agent*) :

Rea, est un agent autonome chargée de vendre de l'immobilier [30]. Elle est capable de maintenir une conversation multi-modale avec un client: elle peut comprendre et répondre en temps réel. Elle bouge ses bras afin de désigner un élément de l'image. Elle utilise le regard, les mouvements de la tête, et les expressions faciales pour des fonctions telles que l'échange de prise de parole, l'emphase, les salutations aussi bien que pour fournir des feedbacks à l'utilisateur.

REA a une apparence humanoïde (Figure 20) lui permettant d'avoir des comportements verbaux et non verbaux (regard, expressions faciales, posture et gestes des mains) [05].

Pour renforcer le réalisme de l'interaction, Rea apparaît sur un grand écran de projection la mettant ainsi à taille humaine. Elle dispose en entrée de caméras pour juger de la présence ou de l'absence d'utilisateurs et d'un micro pour la reconnaissance vocale [08].

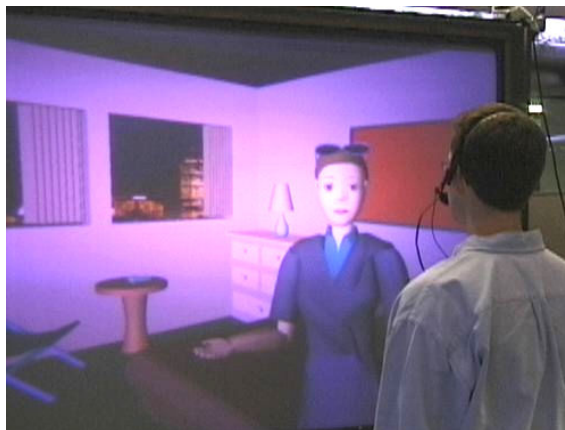


Figure 19. L'agent immobilier Rea

4.6.7 Agents conversationnels particuliers :

Parmi les agents conversationnels qui sont très utilisés, il existe deux types d'agents particuliers, ce sont les agents pédagogiques et les agents signeurs.

4.6.7.1 Agents pédagogiques :

L'enseignement est un champ d'application intéressant pour les Agents Conversationnels, qui sont susceptibles d'introduire un aspect ludique dans l'apprentissage.

Un Agent Pédagogique Animé est un ACA utilisé dans un EIAH. Le personnage est, par exemple, affiché à côté d'un support de formation ou intégré dans un environnement virtuel pédagogique en 3D. Le personnage vise à fournir des signes de communication verbale et non verbale (regard, expressions faciales, gestes de la main, postures) tout en cherchant à motiver les apprenants par des aspects communicationnels et émotionnels. Chaque mouvement, chaque posture corporelle, chaque expression faciale, chaque geste de l'agent pédagogique a ainsi une fonction communicative qui est interprétée par l'apprenant.

En effet, un agent pédagogue est plus dynamique et plus interactif que les personnages dessinés dans les manuels scolaires ou les CD éducatifs [30]. Il possède plusieurs avantages potentiels par rapport à des cours statiques en ligne. Il permet de motiver l'apprenant en lui posant des questions pour qu'il puisse réagir au cours, en l'encourageant et lui donnant des retours sur ses actions. Il doit présenter des informations pertinentes, donner des exemples faciles à retenir, interpréter les réponses de l'apprenant, et éventuellement faire preuve d'humour [30]. La présence d'un agent améliorerait la mémorisation [30]. Les élèves qui apprenaient avec un agent obtenaient de meilleurs résultats sur des tests de transfert et étaient plus intéressés par l'agent. Des travaux récents ont étudié l'impact du réalisme de l'agent, de son sexe, de son apparence culturelle et de son rôle pédagogique sur l'apprentissage [30]. Les élèves montraient un transfert d'apprentissage plus fort lorsque l'agent pédagogique était représenté de manière plus réaliste et non traditionnelle [30].

- **L'agent Cosmo**

Cosmo est un agent émotionnel pédagogique développé par le laboratoire IntelliMedia [30]. Il explique à l'utilisateur comment les ordinateurs d'un réseau sont connectés, le fonctionnement des routeurs,... Tout le corps de Cosmo est animé. Mais les principales parties du

corps porteuses d'informations émotives sont les yeux, les sourcils, le visage, la bouche, la position de la tête, la posture et les gestes des bras et des mains et particulièrement pour Cosmo : la position de ses antennes [05].

Par exemple, la figure 21 représente une image d'un Cosmo sympathique avec un air interrogateur : ses sourcils sont levés, sa tête est légèrement penchée, sa bouche est souriante, et ses mains sont ouvertes et en l'air.

Il est également capable de montrer des réactions émotionnelles, du type « interrogation » quand l'utilisateur sollicite de l'aide, « concentration » quand Cosmo donne une explication, « félicitation » quand l'utilisateur fait un choix correct ou « déception » quand il fait une erreur ou quand le temps de résolution d'un problème est lent [05]. Le ton et la forme de la phrase dépendent alors de l'émotion que Cosmo cherche à exprimer. Ainsi quand l'étudiant réussit à résoudre un problème, Cosmo félicite l'apprenant oralement, exprime sa joie en applaudissant par exemple, et suivant la difficulté du problème il modère ou exagère son comportement. Le but recherché est alors d'encourager l'apprenant [30].

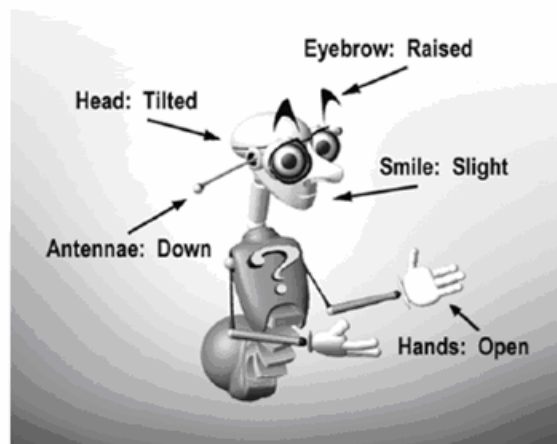


Figure 20. L'agent Cosmo sympathique

- **Herman the Bug :**

Le logiciel Design-A-Plant enseigne à des élèves le développement des plantes, et en particulier les mécanismes de la photosynthèse [05]. La figure 22 présente un personnage 2D qui est présent visuellement : il suit la progression de l'utilisateur. Les interventions de l'Agent visent à fournir des conseils appropriés au contexte, en rappelant des principes biologiques (ex : des feuilles plus larges permettent de capter plus de lumière) ou en donnant des recommandations

concrètes (ex : dans un environnement sombre, il faut que la plante ait de larges feuilles). Il donne aussi des feedbacks réguliers sur la progression de l'utilisateur.

Enfin, sa crédibilité est renforcée par des comportements annexes à sa mission pédagogique : par exemple des changements de posture (ex : debout, assis, couché), des déplacements à l'écran (ex : en marchant, sautant, volant, en faisant des acrobaties), des mouvements autocentrés (ex : se gratter la tête ou le dos, nettoyer ses lunettes), etc.

Cette application a été testée avec succès par des élèves d'une douzaine d'années [05].



Figure 21. L'agent Herman the Bug

- **Steve :**

Steve (Soar Training Expert for Virtual Environments) est un autre Agent pédagogique [05], techniquement sophistiqué: il habite un environnement 3D dans lequel des étudiants sont immergés par l'intermédiaire d'un casque de réalité virtuelle.

L'objectif de Steve présenté sur la figure 23 est de leur apporter des connaissances procédurales, par exemple comment faire fonctionner ou comment réparer des équipements complexes. Steve est capable de réaliser ces actions lui-même, de répondre aux questions des étudiants et de contrôler leur exécution de ces tâches (les étudiants peuvent agir grâce à une souris 3D et des datagloves).

Steve peut aussi exprimer des réactions émotionnelles. Par exemple, un des buts de Steve est de donner des explications sur le domaine considéré. Il manifeste donc un certain enthousiasme lorsqu'il a l'opportunité de discuter un détail intéressant avec un étudiant. Un autre

de ses buts est de maintenir l'attention de l'étudiant : il est donc parfois amené à rappeler à l'ordre de manière ferme un étudiant qui ne regarde pas ce que Steve est en train de montrer. Enfin, un autre but de Steve est que les étudiants retiennent leurs leçons : il manifeste donc de la joie s'il constate que l'étudiant a retenu un point important ; ou au contraire il apparaît préoccupé s'il constate l'inverse. [05]

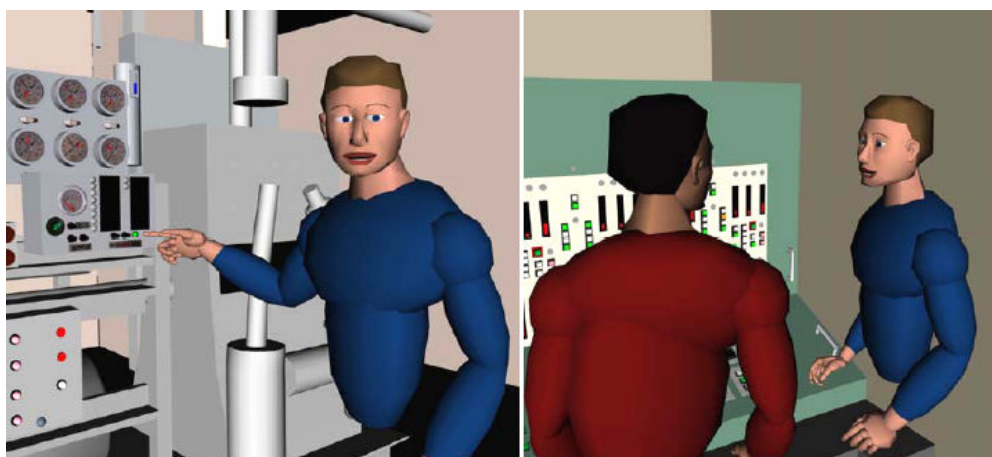


Figure 22. Deux vues de l'environnement de Steve

- **Adele :**

L'Agent féminin Adele (Agent for Distance Learning – light edition) ne requiert qu'un navigateur web [05], et peut donc être utilisée pour l'enseignement à distance, par exemple. Adele enseigne la médecine (voir figure 24), et propose des exercices aux étudiants, dont elle contrôle et évalue la réalisation. Elle suit du regard le pointeur de la souris, fournit des feedbacks verbaux et non verbaux (hoche la tête ou sourit pour les bonnes réponses, montre de la perplexité quand l'étudiant fait une erreur), commente et conseille [05].

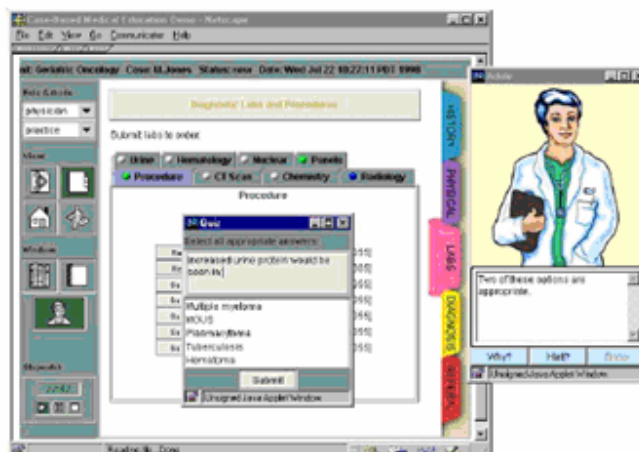


Figure 23. ADELE aide l'apprenant pendant un exercice

4.6.7.2 Les agents signeurs :

Dans le monde, il y a une catégorie de gens qui sont privé de la faculté d'entendre et de parler. Ces personnes souffrent d'un manque d'intégration sociale lié à une stérilité dans la communication avec leur entourage. Pourtant ces personnes peuvent s'exprimer dans une langue qui leur est propre : la langue des signes. Mais, lorsqu'ils sont face à une machine, sur le web ou en relation avec une application d'apprentissage ou de formation par exemple, ils sont totalement étrangers. Leur handicap les oblige à abandonner toutes tentatives de compréhension.

Heureusement des solutions tirées du domaine des SMA permettent aujourd'hui de faire intégrer cette catégorie à profiter des avantages de la technologie informatique et de mener leur vie normalement. Pour arriver à résoudre ce problème, des agents conversationnels signeurs ont été développés dans divers domaines et utilisant plusieurs langues de signes. On cite parmi les plus connus :

- **Tessa :**

Le projet TESSA (Text and Sign Support Assistant) [30] est une application développée par l'Université East Anglia en partenariat avec la poste anglaise. Il allie un système de reconnaissance vocale à un humain virtuel animé pour constituer un intermédiaire entre l'employé de la poste parlant anglais et le client sourd connaissant la langue des signes britannique (BSL – British Sign Language).

Les signes réalisés par l'avatar sont issus d'un enregistrement préalable des mouvements réellement effectués par une personne signant en BSL. De tels enregistrements sont fait au moyen d'un système de capture de mouvements placé sur la personne [30].



Figure 24. L'avatar de Tessa

- **eSign**

eSign (Essential Sign Language Information on Government Networks) est un autre projet européen [30]. L'objectif principal de ce projet est d'améliorer l'accès aux informations du gouvernement pour les personnes sourdes. Plus précisément, il s'agit d'intégrer un avatar signant sur des sites web, comme substitution à des vidéos en langue des signes [30].



Figure 25. DePaul University American Sign Language

4.7 Langages de Représentation pour ECAs :

Les agents peuvent être décomposés essentiellement en deux parties principales :

- **Le corps** : il joue le rôle de l'animation. Il peut être 2D, 3D, de type dessin animé, réaliste... Il peut se baser sur le standard Mpeg-4 ; il peut suivre les spécifications de H-

ANIM ; son animation peut venir de la méthode de capture de mouvements, ou de la dynamique de mouvements, ou encore de key-frames [30].

- **L'esprit de l'agent** : L'esprit est responsable du raisonnement, de la planification, de la perception du monde dans lequel l'agent est placé ainsi que de ou des utilisateur(s) et/ou de ou des agent(s) avec lesquels l'agent discute et interagit. L'esprit doit percevoir et comprendre ce qui est dit et quels sont les événements qui ocurrent dans le monde. Il doit générer ce que l'agent doit dire et comment il doit le dire ; il doit aussi savoir comment réagir aux événements, déclencher des émotions ou bien des actions à entreprendre... De plus l'esprit de l'agent reflète la personnalité et les facteurs d'identité de l'agent.

Cependant, dans la démarche de conception et de développement des ACA, il ne s'agit pas d'ajouter simplement un corps à des agents conversationnels pour qu'il puisse communiquer sur des dimensions non-verbales mais de reconsidérer cette notion de corps qui vient "matérialiser" le rapport des agents à leur environnement pour mieux comprendre et approfondir ce qu'apporte effectivement un corps, aussi virtuel soit-il, du point de vue des performances et des compétences langagières des agents en situation d'interaction avec, notamment des êtres humains et, éventuellement, d'autres agents.

Dans le domaine du développement des ACA, l'"embodiment" ou "l'incarnation" est traditionnellement considéré comme une couche dans l'architecture de l'agent chargée de traduire en comportements verbaux et non verbaux des représentations symboliques produits par l'application. La couche d'"embodiment" est présentée comme une couche supplémentaire permettant de traiter les aspects multimodaux de la communication. [33]

Pour arriver à faire le lien entre le corps et l'esprit de l'agent, nous avons besoin d'un langage de représentation qui spécifie les comportements de l'agent. Un tel langage sert d'interface entre les différents modules de l'architecture d'un système d'agent. De plus, il peut être utilisé pour contrôler l'animation des agents, et pour synchroniser les diverses modalités (visage, corps, voix...). Il peut contenir plusieurs niveaux d'abstraction : allant de la description des signaux (sourire, hochement de tête), aux informations sémantiques, aux fonctions communicatives (performatif, émotion) [30].

Plusieurs langages de représentation des ACAs ont été développés récemment :

➤ **Virtual Human Markup Language : VHML**

Il inclut plusieurs sous-langages pour chaque modalité (parole, visage, geste...) [30]; les éléments du langage peuvent faire référence à des informations de bas niveau (soulèvement sourcil droit) ou à des informations de haut niveau (émotion 'colère').

Les différents sous-langages sont spécialisés le long d'une dimension : l'organisation du dialogue, les émotions, l'animation faciale, l'animation corporelle, l'hypertexte, et la parole. On cite par exemple :

1. Dialogue Management Markup Language, **DMML**
2. Speech Markup Language, **SML**
3. Emotion Markup Language, **EML**,
4. Gesture Markup Language, **GML**, sont définis avec des éléments de trois sous-langages : le Facial Animation Markup Language, **FAML**, le Body Animation Markup Language, **BAML**, et le Speech Markup Language, **SML**.

➤ **Multimodal Presentation Markup Language (MPML)** son objectif est de permettre de créer des agents animés pour des applications de présentation interactive. Jusqu'à présent, pour la plupart des applications de présentation interactive, les agents web informent généralement de manière séquentielle. MPML permet la génération dynamique du contenu de la présentation au fur à mesure que la conversation évolue entre l'agent web et l'utilisateur.

➤ **Character Markup Language, CML et Avatar Markup Language, AML** sont deux langages pour animer des avatars à partir respectivement, d'une approche top-down (CML) et d'une approche bottom-up (AML). CML a été développé pour combler le vide existant entre la méthode de modélisation des émotions et le calcul des comportements à montrer. A l'opposé, AML souhaite développer un cadre multimédia compatible avec MPEG-4, en particulier pour les Chat-rooms pluri-utilisateurs et les agents 3D autonomes [01].

- **RRL : Le Rich Representation**, a été développé pour contrôler l'interaction entre deux ou plusieurs agents virtuels.
- **APML : Affective Presentation Markup Language** L'objectif d'APML est de spécifier le comportement de l'agent au niveau de sa signification. Par exemple, la classe liée aux informations sur le monde inclue les gestes déictiques, les gestes indiquant une direction (doigt pointé pour dire 'ce livre') ainsi que les gestes métaphoriques (ouverture des bras pour indiquer 'ce grand homme') et les gestes iconiques (gestes mimant la propriété saillante d'un objet, 'la table ronde'). Une autre classe rassemble les gestes exprimant un but de l'agent, tels que les performatifs (doigt levé et menaçant pour indiquer la menace). Les émotions sont plus souvent marquées par les expressions faciales mais aussi par les gestes (poing levé pour la colère).

Conclusion :

Dans ce chapitre, nous avons présenté les agents conversationnels, leur état de l'art, les domaines où ils sont utilisés. Dans ce parcours, nous avons remarqué que l'humain est le partenaire principal des agents conversationnels. Mais réunir humain et agents dans une interaction commune nécessite des recherches considérables dans le domaine de la communication. C'est la raison pour laquelle nous avons exploré les aspects théoriques de la conversation humaine dans le chapitre précédent.

Les langages de représentation des ACAs sont nombreux. Chacun d'eux est spécifique pour un type particulier de modalité humaine (émotion, geste, expressions faciale, ...). Parmi ces langages, nous avons choisi HumanML Dans le chapitre suivant, nous allons présenter le langage HumanML en montrant son intérêt dans la communication humaine ainsi que dans la communication dans les SMA.

Chapitre 5 : Le langage HumanML

Ce chapitre présente les fondements du langage HumanML développé par le consortium OASIS. Nous commençons le chapitre par un bref historique déterminant les conditions dans lesquelles a été créé ce langage et les motivations qui lui ont donné naissance. Ensuite, les différents composants de ce langage seront présentés. Et en fin, le langage HPCDML un langage dérivé du HumanML sera également présenté.

5.1 Présentation du langage HumanML :

Le langage de représentation HumanML (Human Markup Language) est un nouveau langage basé sur XML schéma et les schémas RDF. Il a été développé par l'organisation OASIS (Organization for the Advancement of Structured Information Standards), située à Massachusetts au Etats Unis. L'objectif de cette organisation est de créer un ensemble de normes internationales permettant l'évolution du Web à des niveaux supérieurs [09].

HumanML a été déterminé dans le but de s'assurer que toutes les caractéristiques naturelles humaines – que ce soit de nature comportementale, mentale, et physiques - sont correctement transmises pendant la communication. Le langage (HumanML) permet aux programmeurs d'écrire l'abstrait, d'intégrer la communication humaine non-verbale dans le code des logiciels, et donnant aux utilisateurs les moyens de communiquer leurs émotions aux autres à travers le Web.

Le but est d'améliorer la qualité et de réduire les ambiguïtés dans la communication humaine. Le nouveau langage de représentation doit jouer un rôle principal en supprimant les divers obstacles culturels rencontrés actuellement dans la communication sur le Web [09].

Actuellement, avoir un niveau supérieur de sémantique dans le Web est un sujet de grande envergure. Les chercheurs espèrent que HumanML pourrait être employé pour décrire tout qui fait partie d'une interaction humaine en commençant par les émotions et le comportement jusqu'à arriver à l'apparence physique et les significations culturelles complexes [10].

Évolué à partir de la pratique simple de l'utilisation des émoticônes pour améliorer la communication dans les email, des news groups et le Chat, HumanML vise vers une représentation plus officielle des caractéristiques humaines (par exemple, culturel, physique,

psychologique, etc.) [31]. Ainsi, l'objectif principal du langage HumanML est l'intention d'humaniser le Web et de créer une piste standard pour transmettre les émotions et le langage du corps ainsi que les caractéristiques générales de l'humain.

5.2 Bref historique de HumanML :

Les efforts pour la création du langage HumanML ont commencé avec les préoccupations de Ranjeeth Kumar Thunga pendant ses études en psychologie [43].

Il constatait que le développement en sciences humaines comme la psychologie, la sociologie, la psychologie, etc.... était très lent par rapport au progrès connu dans les domaines de la technologie, des sciences de la physique, chimie et biologie. Ce manque continuait à occuper Ranjeeth Kumar Thunga surtout durant son travail sur le développement du Web pendant les années 90. A partir de ces idées, il a conclu que la seule solution pour intégrer les caractéristiques et les intérêts humains dans la technologie de l'information était par le biais de XML.

En février 2001, Ranjeeth invitait tous les intéressés à se joindre au groupe de discussion sur YahooGroup dans l'intérêt de discuter la possibilité de créer un comité technique pour Oasis. Cette liste existera pour seulement 90 jours, après quoi le comité technique peut être formé ou pas. L'objectif de ce groupe de discussion est d'attirer les personnes qui ne sont pas uniquement issues du milieu des sciences de la technologie de l'information mais qui ont une grande pratique dans les sciences humaines spécialement les langues, la sociologie, la psychologie... L'objectif de la discussion est d'apporter leurs contributions pour mettre les bases du langage HumanML et aider à définir ses normes [12].

En septembre 2001, le comité technique du langage HumanML d'OASIS consortium a été fondé [43]. Les fondateurs de la liste sont [12]:

- (1) Ranjeeth KUMAR Thunga,
- (2) Emmanuel Batsis,
- (3) Joseph Norris,

Les efforts de ce comité ont donné naissance à une première version d'un document de besoins qui est apparue au 31 avril 2002, et au 12 Décembre 2002 est apparu le premier document de spécification « HumanML primary Base XML Specification 1.0 ». Ce document a

été légèrement modifié [43] par le public lors d'une période d'essai de 30 jours pour enfin l'adopter le 12 janvier 2003.

HumanML sera proposé en deux versions. La première reposera sur XML Schema , la recommandation du W3C spécialisée dans la présentation des données. La deuxième sur le langage RDF de Tim Berners Lee, dont l'objectif est de présenter des faits.

Le langage HumanML est actuellement en cours de développement [11]. Les fondateurs du langage espèrent travailler avec des groupes et organisations pour s'assurer que HumanML sera maintenu comme standard international. Ils comptent sur l'appui les individus ou les organismes ayant un intérêt pour XML, le web sémantique, et des communications humaines pour fonder ce langage [44].

5.3 Les domaines d'application de HumanML :

Les champs d'application du langage HumanML sont vastes et multiples, ils concernent la communication Humain-Humain et la communication Humain-Agent sur le Web.

- **En messagerie :** dans une plate-forme de messagerie, HumanML permettrait de dissiper les malentendus, grâce aux émoticônes. Les internautes ont déjà développé un système simple mais non standard pour enrichir la communication sur le net. HumanML peut aider à éliminer les malentendus dans les correspondances par e-mail entre les individus. Le langage aurait un but semblable aux émoticônes existantes actuellement comme :) pour un sourire, ou :(pour montrer la tristesse.
- **En réalité virtuelle :** le langage transmettrait directement la colère ou la joie. Un avatar peut aider à visualiser le contenu d'une page web à travers les gestes, interpréter les intentions et les émotions, ainsi que les aspects du comportement humain dans la communication non verbale [57].
- **Les applications commerciales :** les bases de données comportementales auront la possibilité de générer automatiquement des profils de consommateurs, indiquant leur provenance, coutumes, religions ou profils psychologiques [11].

- **En psychologie:** HumanML peut être utilisé pour trier les fichiers des patients sur la base de la culture, personnalité, types de désordre psychologiques qui peut être appropriée dans des diagnostics psychologiques [46].
- **Communications:** les fondateurs d'HumanMl estiment que tout types de communication doit profiter de ce nouveau langage. Cette nouvelle technologie permettra au gens de se communiquer entre eux en utilisant leurs propres termes.

D'autres domaine bénéficieront des avantages de ce langage, on peut citer : anthropologie, archéologie, l'aide à la décision dans le business, résolution du conflit, études culturelles, marketing, médecine... [04]

5.4 Les objectifs du langage HumanML :

Les initiatives du langage HumanMl permettent d'arriver à :

- ✓ Enrichir les communications humaines et réduire les malentendus par des mécanismes explicites pour représenter les dispositifs paralinguistiques de la communication humaine par l'utilisation d'un cadre standard pour les référents aux descriptions des états émotifs [43].
- ✓ Coder les émotions ou les intentions d'un utilisateur [11].
- ✓ Augmenter la fidélité dans la communication humaine [12]
- ✓ Permettre l'échange des informations exactes entre les personnes de cultures différentes à travers le Web. Le nouveau langage doit jouer un rôle important dans l'élimination des barrières culturelles qui se trouvent actuellement dans les communications sur le Web.
- ✓ Permettre à quelqu'un qui parle une langue complètement différente de comprendre la signification du message [45].
- ✓ Permet à des personnes de s'exprimer plus profondément [48] notamment par l'ajout du langage non verbal dans les messages transmis.
- ✓ Réduire les conflits interpersonnels liés à la transmission incomplète des caractéristiques de l'humain [48].
- ✓ HumanML prolonge l'utilisation de XML dans les domaines totalement nouveaux [49].

- ✓ Fournir une interface Homme-Machine plus attractive grâce à l'introduction de plusieurs modalités [43].
- ✓ Permet aux agents intelligents de posséder un état émotionnel interne dans le but d'humaniser le web [60]
- ✓ L'objectif principal de HumanML est de s'assurer que toutes les caractéristiques de l'humain sont correctement transmises.

5.5 Communication Humain-Humain sur le web:

Internet a connu une croissance très rapide du côté des services offerts ainsi que du côté du nombre d'utilisateurs qui augmente chaque jour notamment dans les pays non- occidentaux. Une étude [04] d'analyse estime que la population globale d'internet dans ces pays dépasse les 75 pc en 2007. Pendant les quinze années à venir on s'attend à ce que l'échange de l'information exacte entre les personnes de cultures et d'origines différentes soit un souci toujours croissant en ce qui concerne l'Internet et d'autres moyens de communication [43]. En fait, La capacité d'interpréter correctement les informations transmises réduira considérablement les malentendus et l'ambiguïté, ce qui permet d'avoir un effet positif sur les rapports personnels entre les personnes.

Les conséquences des malentendus dans la communication entre les humains sont nombreuses. Le résultat peut conduire aux conflits, à la haine, à la violence et encore plus grave à la guerre entre les humains. Tout ceci provient de l'incapacité de communiquer la nature humaine au niveau racine. Cette nature concerne les éléments tels que 'l'état émotionnel', 'le comportement', 'motivation', 'culture', 'les gestes du corps', et 'l'intention' qui ne sont pas transmis avec le message.

La mauvaise communication abonde dans notre société. Plusieurs exemples montrent cette situation :

- Dans le milieu du travail, un avertissement sévère à un subordonné peut être considéré comme une sous-estimation.
- Un nouvel immigrant peut être considéré grossier pour n'avoir pas dit "Merci".
- Une plaisanterie avec un ami peut être interprétée comme ironie.

- Les faits présentés au cours d'une discussion peuvent être obscurcis par les plaisanteries des autres.

- L'armée d'un pays peut interpréter une exposition de `de force 'par son adversaire comme appel `à la guerre '.

Ainsi, les situations de doutes et les mauvaises interprétations pendant la communication laissant chaque message susceptible de préjudices énormes.

Une autre source de conflit provient à partir des différentes interprétations du même événement, et le manque d'outils pour mieux détecter ces différences. Les différentes interprétations sont donc le résultat des points de vues différents des communicants. Le même message est interprété de façons complètement différentes.

Ceci se produit parce que les aspects humains nécessaires pour la communication ne sont pas explicitement exprimés [51].

En conséquence, il y a un besoin clair de technologies naissantes qui amélioreront la communication globale pour adresser le niveau plus élevé de l'interconnectivité globale [43]. Cette évolution exige l'adoption d'un standard permettant d'offrir une meilleure communication parmi la population d'Internet et aider à améliorer les échanges entre les utilisateurs de différentes langues et cultures.

HumanML peut être adopté comme un standard de communication sur le web. La fonction de HumanML est d'augmenter la fidélité dans le processus de communication, par l'utilisation de XML, en vue d'éliminer les malentendus dans la communication entre humains [04].

HumanML vise précisément à introduire les aspects humains nécessaire pour la communication dans le but de véhiculer l'information d'une manière correcte pour pouvoir la partager avec les utilisateurs sur le net.

Un autre objectif de HumanML est également fournir aux documents une interface Homme-Machine plus attractive grâce à l'introduction de plusieurs modalités [43] notamment le non-verbal.

Jusqu'ici, des applications ont commencé à inclure des aspects humains élémentaires dans la communication, à travers l'utilisation des émoticônes et des acronymes par exemple :), lol, :(, ☺, etc. Ceux-ci ont certainement aidé à donner l'expression humaine dans les messages. Cependant, leur utilisation manque de la standardisation, et leur usage est tout à fait limité.

HumanML intervient dans le cas de résolution de conflits en fournissant un vocabulaire précis pour l'annotation et incluant des caractéristiques et l'interprétation exacte de communication dans un document. On peut dans ce cas introduire les émotions, le langage du corps, et les intentions perçues. Ceci peut prendre beaucoup de temps pour arriver à estomper entièrement les mauvaises interprétations [42].

Avec HumanML, des pages Web pourraient être employés pour exprimer la signification à quelqu'un qui parle une langue entièrement différente. L'humeur est souvent perdue pendant la traduction, mais pourra être codée dans HumanML. Ceci pourrait alors être représenté dans un browser HumanML comme images, sons ou mots [52].

5.6 HumanML et le web sémantique :

Le Web Sémantique, qui est une extension du Web actuel, a comme objectif majeur d'apporter de la sémantique à l'information manipulée afin de faciliter la communication entre agents humains et artificiels [38]. Il rend possible la compréhension de l'information aux ordinateurs. Le partage des connaissances se fait dans des langages essentiellement humains. Le Web sémantique permet aux individus de s'exprimer dans des termes que les ordinateurs pourront également interpréter et échanger, et ce afin de résoudre des problèmes qui paraissent fastidieux. A la différence du web actuel, le web sémantique décharge les utilisateurs d'une bonne partie de leurs tâches de recherche, de construction et de combinaison des résultats, grâce aux capacités accrues des machines à accéder aux contenus des ressources et à effectuer des raisonnements sur ceux-ci.

En effet, le Web actuel n'est compréhensible que par les humains, l'objectif du Web sémantique est de fournir aux machines un meilleur accès à l'information grâce à des métadonnées qui décrivent les informations disponibles et ceci sans l'intervention des humains.

Les métadonnées sont fournies par le composant pivot du Web sémantique, les ontologies. Ces dernières représentent une technologie dont le but est d'améliorer l'organisation, la gestion et

la compréhension de l'information électronique. Elles servent de vocabulaire standardisé pour le partage de connaissances.

Pour cela, le système doit avoir accès non seulement aux termes utilisés par l'être humain mais également à la sémantique qui leur est associée, afin qu'une communication efficace soit possible. Les ontologies visent à représenter cette connaissance en étant à la fois interprétable par l'homme et par la machine.

Ainsi, le web sémantique doit permettre d'abord de localiser, d'identifier et de transformer des ressources de manière robuste et saine tout en renforçant l'esprit d'ouverture du Web avec sa diversité d'utilisateurs quelque soit leur niveau, leur expérience ou leur culture. Il peut également aider à réaliser une meilleure communication entre les humains [38].

C'est dans cette optique que les membres de L'OASIS Consortium ont pensé à créer un comité technique pour mettre les bases d'un langage qui permet de coder les caractéristiques humaines et de les introduire dans la communication entre humains ou entre humain et machine. Le nouveau langage HumanMI sera basé sur le langage XML qui connaît une utilisation croissante parmi des programmeurs du Web.

Grâce à HumanMI, l'émetteur d'un document XML pourra finalement intégrer des aspects humains beaucoup plus profonds dans le code du message émit.

- **Les composants du web sémantique:**

Afin de représenter l'information sur le Web sémantique et de la rendre simultanément inter opérable aussi bien syntaxiquement que sémantiquement à travers les applications, il est nécessaire d'utiliser des langages spécifiques. Certains d'entre eux sont dits de « haut niveau», d'autres de « bas niveau »

La vision d'un Web sémantique, dans lequel l'information serait accessible et manipulable automatiquement par la machine, s'appuie sur une pile de langages jouant chacun un rôle particulier.

- Les protocoles et langages standards du Web : le protocole HTTP et les URI (Uniform Resource Identifier ou Identificateur Uniforme de Ressource).

A ces standards s'ajoutent ceux qui sont propres au Web sémantique :

- XML fournit une manière de représenter des documents structurés, mais il n'impose aucune contrainte sémantique sur les documents produits ;
- XML Schéma permet de contraindre la structure des documents XML;
- RDF est un modèle de données simple, fondé sur des ressources et des relations entre ces ressources, équipé d'une sémantique et qui peut se représenter en XML; Il permet de décrire toute donnée ;
- RDF Schema permet de créer des vocabulaires, ensembles de termes utilisés pour décrire des choses ;
- OWL fournit d'avantage de primitives de modélisation pour décrire des ontologies plus riches sur le web. Il est destiné à être utilisé lorsque les informations contenues dans les documents doivent être traitées par des applications logicielles, c'est-à-dire lorsqu'elles ne sont pas simplement montrées à l'utilisateur. [15]

Ces trois standards sont ouverts et issus du W3C2. Ils forment l'ossature du Web sémantique [37].

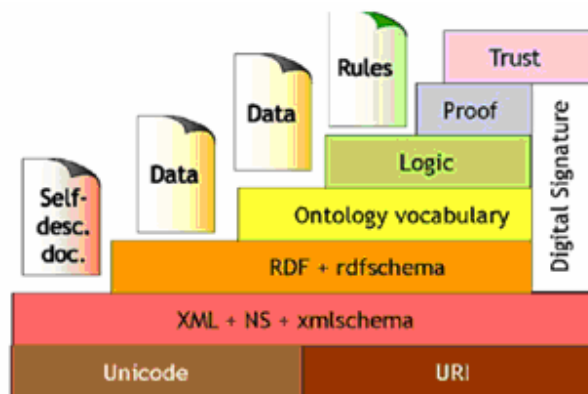


Figure 26. Le « cake » de Tim Berners Lee

La construction de HumanMI autour du XML schéma et du RDF schema lui a donné l'avantage d'être une extension du web sémantique [45]. Xml schema est utilisé pour définir l'utilisation des termes. Tandis que RDF Schema est utilisé pour associer ces termes avec les ressources standard dans le but d'assurer l'interopérabilité.

Afin de mieux comprendre XML Schema et le RDF schema et leurs implications, il est tout d'abord nécessaire de présenter brièvement le langage XML et le RDF schema ainsi que les différentes fonctionnalités qu'ils proposent.

5.7 Le langage XML et le XML schéma:

Crée en 1998 par le W3C, XML ou eXtensible Markup Language est un méta-langage qui permet la description et la structuration des données. Il est utilisé pour représenter les informations que les applications peuvent utiliser, gérer et afficher plus facilement.

Dérivé du SGML, ce langage est particulièrement performant pour l'échange et le stockage de tous type de données et de leurs structures [17]. Un document XML transporte les données ainsi que les informations sur les données transportées. Il est capable également d'intégrer des données provenant de différentes sources.

Contrairement à HTML, qui est à considérer comme un langage possédant un nombre de balises limité et prédéfini, XML peut être considéré comme un métalangage permettant de définir d'autres langages [41].

XML se révèle donc comme un langage à la fois puissant et flexible et permet de plus d'effectuer un certain contrôle sur la validité des données et de la structure qui composent un document XML. L'Extensibilité est une caractéristique qui permet à un document XML de pouvoir être utilisable dans tous les domaines d'applications. Ainsi, c'est pour cet atout que de nombreux langages de représentation de données issus de XML apparaissent (il en existe plus d'une centaine) :

- **OFX : Open Financial eXchange** pour les échanges d'informations dans le monde financier.
- **MathML : Mathematical Markup Language** permet de représenter des formules mathématiques.
- **SMIL : Synchronized Multimedia Integration Language** permet de créer des présentations multimédia en synchronisant diverses sources : audio, vidéo, texte,...
- HumanML, VHML, APML, FAML, BAML, CML, AML,

Un document XML est associé à un schéma. D'un point de vue global, les schémas XML développés par le W3C, permettent la description de la structure d'un document XML, de son

contenu et du type de données qui le constitue. Ils facilitent l'échange des données et permettent une meilleure homogénéité des documents.

Il existe un grand nombre de langages de schémas, parmi lesquels la DTD (Document Type Definition) et le XML Schema, ou XSD (XML Schema Definition Language).

Jusqu'à présent, la DTD est le langage le plus fréquemment utilisé pour définir la structure d'une classe de document XML mais, en raison de certaines limites, le XML Schema est en passe de devenir la norme de référence [14]. XML Schema offre donc un grand nombre de fonctionnalités supplémentaires comparé aux DTDs.

Le langage XML Schema doit donc répondre à un certain nombre d'exigences. Il doit, entre autre, être plus performant que les DTDs, auto-descriptif, utiliser la syntaxe XML et être compatible avec un grand nombre d'applications qui utilisent elles-mêmes XML.

L'un des atouts majeurs de XML Schéma repose sur sa syntaxe XML, ce qui permet à de nombreuses applications développées pour ce langage de pouvoir également gérer les schémas XML [14].

Le typage des données est introduit, ce qui permet la gestion de booléens, d'entiers, d'intervalles de temps... Il est même possible de créer de nouveaux types à partir de types existants.

Pour bien montrer l'écriture d'un schéma XML, il est préférable de commencer par un exemple pratique. Soit une instance de document XML décrivant un fichier client :

```
<?xml version = "1.0" ?>
  <Customer>
    <FirstName>Raymond</FirstName>
    <MiddleInitial>G</MiddleInitial>
    <LastName>Bayliss</LastName>
  </Customer>
```

Le schéma qui lui correspond et donc qui décrit sa structure est formalisé de la façon suivante [14] :

```
<?xml version = "1.0" ?>
<schema xmlns = "http://www.w3.org/2001/XMLSchema">
  <element name = "Customer">
    <complexType>
      <sequence>
```

```
<element name = "FirstName" type = "string" />
<element name = "MiddleInitial" type = "string"/>
<element name = "LastName" type = "string" />
</sequence>
</complexType>
</element>
```

Un schéma XML est composé principalement d'éléments qui sont représentés par les balises. Ces éléments peuvent comporter des attributs et sont associés à un type de données qui peut être défini comme « type simple » ou « type complexe ».

Un type simple est un élément ne contenant que du texte, tandis qu'un type complexe est un élément qui peut contenir d'autres éléments ou des attributs. Ces derniers servent à qualifier un élément et ne peuvent être que de type simple.

XML Schéma offre également la possibilité de constituer des « groupes d'attributs » afin de faciliter à la fois la lisibilité et la maintenance d'un schéma [59].

Les applications qui s'appuient sur les schémas pour connaître les structures des données XML peuvent alors avoir des informations importantes pour optimiser les évaluations de requêtes, obtenir des informations générales sur le contenu, faciliter l'intégration de données issues de différentes sources d'information, améliorer le stockage, faciliter la mise en place d'index ou de vues et aider à la classification de documents.

L'intérêt de disposer d'un format commun d'échange d'information dépend du contexte professionnel dans lequel les utilisateurs interviennent.

Par conséquent, XML est un bon moyen de décrire des données et de les stocker mais ne permet pas d'interpréter ces mêmes données. Il ne peut donc être utilisé en tant que tel pour décrire une ontologie.

Une des différences fondamentales entre XML et les ontologies RDF [15] réside dans le fait qu'XML ne traduit finalement que des grammaires tandis que les ontologies s'attachent à représenter la sémantique des objets d'un domaine en déterminant les concepts et les relations qui le peuplent.

5.8 RDF et RDF schema :

RDF (Resource Description Framework) est un standard du W3C dont l'objectif principal était d'associer aux documents du Web des données sémantiques exploitables par machine. Il utilise la syntaxe d'XML. Ces représentations comportent des métadonnées sur les ressources du Web comme les auteurs de pages Web, leur date de création . . .

Les ressources du Web sont l'élément de base de RDF. Chaque ressource est pourvue d'un identifiant uniforme de ressource (Uniform Resource Identifier, URI). Initialement recommandé par le W3C dans le but de standardiser les définitions et les usages des métadonnées, RDF est également utile à la représentation de données en elles-mêmes.

Les instances RDF sont décrites sous forme d'un ensemble de triplets constituant une structure sémantique. Un triplet RDF est constitué de trois éléments : <Sujet, Prédicat, Objet> noté communément (S, P, O) :

- Le Sujet est une URI (Uniform Resource Identifier) qui est un identifiant d'un objet appelé ressource. Une ressource peut être soit un document accessible sur le Web, soit une partie d'une page Web, soit une collection de pages.

- Le Prédicat est une propriété qui permet de caractériser la ressource. Chaque propriété possède une signification spécifique, définit ses valeurs permises, les types de ressources qu'elle peut décrire, et les relations qu'elle entretient avec les autres propriétés.

- L'Objet est la valeur de la propriété qui peut être une URI (ex : une ressource) ou une valeur littérale (ex :de type de données primitif défini par XML).

Le langage de balisage XML est utilisé pour le spécifier. On parle parfois de RDF/XML car d'autres formalismes existent.

RDF définit lui-même très peu de prédicats, ceux-ci pouvant être définis librement. Il a donc été très rapidement complété par RDFs qui fournit les prédicats essentiels pour représenter une ontologie en RDF.

L'utilisation conjointe de RDF et RDFs permet donc à la fois de représenter (en RDFs) une ontologie et (en RDF) des instances définies en termes de cette ontologie [15].

On peut donc décomposer le RDF en trois choses :

- un modèle de description de ressources Web : métadonnées.
- un langage de représentation d'ontologies (RDFs).
- une syntaxe XML pour l'échange de métadonnées et d'ontologies.

Les ontologies RDF ont été utilisées dans HumanMI pour bien interpréter les énoncés mais aussi comme aide au raisonnement.

5.9 Les structures fonctionnelles du langage HumanMI :

Le processus de la création du langage HumanMI a été décomposé en structures fonctionnelles. Pour un premier temps, il sera proposé en deux versions. La première reposera sur XMLSchema, la recommandation du W3C spécialisée dans la présentation des données. C'est le schéma primaire du langage HumanMI. La deuxième sur le langage RDF, dont l'objectif est de présenter des faits [60].

- **Le schéma primaire du langage HumanMI** : c'est le schéma principal du langage HumanMI. Il contient les éléments et les attributs pour décrire les caractéristiques élémentaires de l'humain et ses activités dans le but de les introduire dans des documents numériques [47]. Le schéma primaire du langage humanMI sera basé sur:

- XML 1.0, <http://www.w3c.org/TR/2000/REC-xml-20001006>
- XML Namespaces, <http://www.w3c.org/TR/1999/REC-xml-names-19990114/>
- XML Schema, <http://www.w3.org/TR/xmlschema-1/>, <http://www.w3.org/TR/xmlschema-2/>

Le schéma primaire de HumanMI est localisé à <http://www.oasis-open.org/committees/humanmarkup/schema/huml-primary-base-1.0.xsd>. C'est le document normatif pour valider les applications basées sur XML.

- **Le schéma secondaire de HumanMI**: à partir du schéma primaire, un ensemble de termes a été rassemblé en extensions pour former le schéma secondaire. Chaque extension désigne un langage pour représenter un champ d'application particulier. Le schéma secondaire contient les termes définis dans le schéma primaire.

Les schémas secondaires spécifiques qui sont planifiés pour être développés sont:

- La culture;
- La réalité virtuelle et l'intelligence artificielle.

- Human Physical Characteristics Description Markup Language (HPCDML) Schema;
- La résolution des conflits, et la communication diplomatique.
- Des schéma liés à la psychologie de l’homme.

Cette liste est partielle et elle peut être enrichie [47].

Les deux schémas ont été établis dans le but de distinguer entre les éléments fondamentaux qui représentent l’homme et les domaines dérivés.

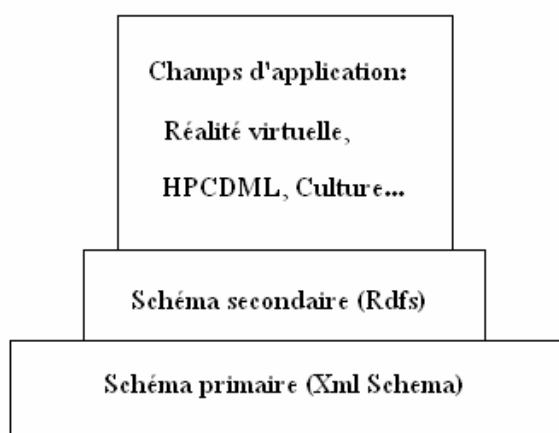


Figure 27. La structure de HumanMI

5.10 Les caractéristiques de HumanMI :

- **Compatibilité** : HumanML doit être conforme avec les règles et la syntaxe de XML.
- **Extensibilité** : Le schéma primaire et le schéma secondaire doivent être extensibles.
- **Modularité** : HumanMI est organisé en schéma primaire et schéma secondaire. Les deux schémas sont considérés comme modules. Tout les modules d’HumanMI doivent être intéropérables entre eux, ainsi qu’avec les standards utilisant XML [47].

Pour introduire du code HumanMI dans un document, nous allons voir, comment déclarer éléments et attributs à l'aide d'un schéma

Un élément, dans un document HumanMI est introduit grâce à la balise <Huml:element>, Par exemple, l’élément simple “personality” est défini avec [56] :

```
<huml:personality>
```

L'élément « address » représente l'adresse postale qui peut être attribuée à une personne. Il se compose de code postal, email, résidence,...Le format de l'élément nécessite une harmonisation suivant les conventions internationales pour décrire une adresse.

```
<xs:complexType name="Address">
  <xs:annotation>
    <xs:documentation>
      Address Type An address in a named address system, such as
      street, city, state, etc.
    </xs:documentation>
    <xs:appinfo>NONE</xs:appinfo>
  </xs:annotation>
  <xs:sequence>
    <xs:element name="postal" type="xs:string"/>
    <xs:element name="residential" type="xs:string" minOccurs="0"/>
    <xs:element name="email" type="xs:string" minOccurs="0"/>
    <xs:element name="previous" type="xs:string" minOccurs="0"/>
    <xs:element name="current" type="xs:string" minOccurs="0"/>
  </xs:sequence>
  <xs:attributeGroup ref="humlIdentifierAtts"/>
</xs:complexType>
```

Pour introduire l'élément address dans un document HumanML, le processus s'effectue de la manière suivante :

```
<huml :address>
  <postal> 30000</postal>
  <residential> Ouargla </residential/>
  <email>contact@ser.dz</email/>
</huml:address/>
```

5.11 Les composants de HumanML:

Cette partie présente les termes et les définitions du vocabulaire du langage HumanML. C'est un ensemble de concepts fondamentaux à partir desquels des extensions peuvent être établies. Le schéma secondaire va se baser sur ce vocabulaire pour définir un ensemble de spécifications à partir desquelles d'autres vocabulaires concernant des domaines spécifiques seront construits.

Le vocabulaire présenté dans le schéma primaire représente le vocabulaire de base [54]. Des changements et des améliorations dans le vocabulaire peuvent survenir suivant le besoin des domaines d'applications. Le langage HumanML fournit un vocabulaire qui permettra à une grande variété d'applications concernant l'humain d'être établie.

Le tableau 6 récapitule les éléments, attributs et types inclus dans le schéma primaire [54]:

Element	Type complexe	Type simple	Groupe
Huml	Address	Locator	Age
	Artifact	Range	gender
	Belief		humlCommAtts
	BodyLocation		humlIdentifierAtts
	Channel		humlTemporalAtts
	Chronemic		physicalDescriptors
	Community		
	Culture		
	Emotion		
	GeoLocator		
	Haptic		
	Human		
	HumanGroup		
	HumanNameElements		
	Intent		
	Kinesic		
	MeasurementUnit		
	Personality		
	Proxemic		
	Semiosis		
	Semiote		
	Sign		
	Signal		
	Symbol		
	Thought		

Tableau 6. : Eléments HumanML

5.11.1 L'élément racine :

- huml : est le préfixe qui sera associé aux spécifications du langage HumanMI.

```
<xs:element name="Huml">
  <xs:annotation>
    <xs:documentation>
      huml is the prefix that will be associated with the Human Markup
      language Specifications. It indicates the first Human Markup
      language schema and which forms the basis for succeeding schemata
    </xs:documentation>
  </xs:annotation>
</xs:element>
```

5.11.2 Les types complexes:

Nous présentons quelques uns:

- **Body Location** : le terme BodyLocation est utilisé pour localiser les parties du corps humains (bras, main, pied,...).

```
<xs:complexType name="BodyLocation">
  <xs:annotation>
    <xs:documentation>
      Body Location This is a location on a body ...
    </xs:documentation>
  <xs:appinfo>NONE</xs:appinfo>
</xs:annotation>
  <xs:attribute ref="bodyPart"/>
  <xs:attribute name="location" type="Locator"/>
  <xs:attributeGroup ref="humlIdentifierAtts"/>
</xs:complexType>
```

- **Human Culture** : les caractéristiques culturelles sont définis par l'ensemble des traits distinctifs caractérisant le mode de vie d'un peuple ou d'une société. Cela inclus les bases sémiotiques qui participent dans les processus de la communication dans les cultures différentes. Le terme huml est : *Culture*.

```
<xs:complexType name="Culture" abstract="true">
  <xs:annotation>
    <xs:documentation>
      Cultural characteristics emerge as the defining set of
      elements which a community shares and which distinguish
      the community as a culture per se. This partakes of the
      semiotic basis for communication processes as a
      contextual element.
    </xs:documentation>
  <xs:appinfo>NONE</xs:appinfo>
</xs:annotation>
  <xs:attributeGroup ref="humlTemporalAtts"/>
  <xs:attributeGroup ref="humlIdentifierAtts"/>
</xs:complexType>
```

- **L'émotion:** désigne l'ensemble des émotions humaines de base. Les six émotions de base seront fournies. Ces six expressions sont : la tristesse, la colère, la joie, la crainte, le dégoût, et la surprise. Le terme huml est *Emotion*.

```
<xs:complexType name="Emotion" abstract="true">
  <xs:annotation>
    <xs:documentation>
      A basic set of primitive human emotions. The six basic
      emotions that have well-documented facial expressions
      associated with them are provided.
    </xs:documentation>
    <xs:sequence>
      <xs:element name="sadness" type="boolean" default="true"/>
      <xs:element name="joy" type="boolean" default="true"/>
      <xs:element name="anger" type="boolean" default="true"/>
      <xs:element name="fear" type="boolean" default="true"/>
      <xs:element name="disgust" type="boolean" default="true"/>
      <xs:element name="surprise" type="boolean" default="true"/>
    </xs:sequence>
  </xs:annotation>
  <xs:appinfo>NONE</xs:appinfo>
</xs:complexType>
```

- **Personality:** Ce type complexe est utilisé pour établir l'usage d'un système de types de personnalités.

```
<xs:complexType name="Personality" abstract="true">
  <xs:annotation>
    <xs:documentation>
      Human Personality Type
    </xs:documentation>
  </xs:annotation>
  <xs:appinfo>NONE</xs:appinfo>
  <xs:attributeGroup ref="humlIdentifierAtts"/>
</xs:complexType>
```

- **Semiosis :** est un échange significatif des signes, des signaux et des symboles à travers des agents cognitifs. Ce processus est le mode de la communication humaine sur lequel le langage HumanMI a été fondé.

```
<xs:complexType name="Semiosis" abstract="true">
  <xs:annotation>
    <xs:documentation>
      Semiosis is a meaningful exchange of signs, signals and
      Symbols among cognitive agents.
    </xs:documentation>
  </xs:annotation>
  <xs:appinfo>NONE</xs:appinfo>
  <xs:attributeGroup ref="humlTemporalAtts"/>
  <xs:attributeGroup ref="humlCommAtts"/>
</xs:complexType>
```

- **Semiote** : est l'agent cognitive qui participe à l'échange significatif de signal. Cet élément est l'acteur dans le modèle sémiotique de la communication. Il est composé d'un ensemble de processeurs capables d'émission, de réception et de répondre aux signaux.

```
<xs:complexType name="Semiote" abstract="true">
  <xs:annotation>
    <xs:documentation>
      Cognitive Agent
      A Semiote is a cognitive agent who participates in
      meaningful signal exchange among cognitive agents.
    </xs:documentation>
    <xs:appinfo>NONE</xs:appinfo>
  </xs:annotation>
  <xs:attributeGroup ref="humlCommAtts"/>
</xs:complexType>
```

- **Kinesic**: La kinésique concerne l'étude des mouvements corporels utilisés dans la communication.

```
<xs:complexType name="Kinesic" abstract="true">
  <xs:annotation>
    <xs:documentation>
      Kinesic: Human Movements
      Communicational Kinesics constitute a vocabulary of body
      Language used to portray moods and emotions and to add
      emphasis to verbal communication.
    </xs:documentation>
    <xs:appinfo>NONE</xs:appinfo>
  </xs:annotation>
  <xs:attributeGroup ref="humlCommAtts"/>
  <xs:attributeGroup ref="humlTemporalAtts"/>
  <xs:attribute ref="intensity"/>
</xs:complexType>
```

5.11.3 Les types simples :

Cette spécification contient deux types simples.

- **Locator** :C'est un ensemble de noms des localisations qui peuvent désigner les positions où se trouve un objet.

```
<xs:simpleType name="Locator">
  <xs:annotation>
    <xs:documentation>
      A simple set of names of locations on an object.
    </xs:documentation>
    <xs:appinfo>NONE</xs:appinfo>
  </xs:annotation>
```

```

        <xs:restriction base="xs:string">
            <xs:enumeration value="upper"/>
            <xs:enumeration value="lower"/>
            <xs:enumeration value="back"/>
            <xs:enumeration value="front"/>
            <xs:enumeration value="inner"/>
            <xs:enumeration value="outer"/>
            <xs:enumeration value="left"/>
            <xs:enumeration value="right"/>
            <xs:enumeration value="top"/>
            <xs:enumeration value="bottom"/>
        </xs:restriction>
    </xs:simpleType>

```

- **Range:** l'attribut range est utilisé pour indiquer les valeurs de l'intensité des éléments par exemple l'émotion.

```

<xs:simpleType name="range">
    <xs:restriction base="xs:decimal">
        <xs:minInclusive value="0"/>
        <xs:maxInclusive value="1"/>
    </xs:restriction>
</xs:simpleType>

```

5.11.4 Les groupes d'attributs :

Cet ensemble de groupes d'attributs est utilisé par plusieurs types complexes.

- **Age :** est un ensemble d'attributs pour déterminer l'âge de l'humain.

```

<xs:attributeGroup name="age">
    <xs:annotation>
        <xs:documentation>
            This is a set of attributes for documenting or
            Determining the age of a human
        </xs:documentation>
        <xs:appinfo>NONE</xs:appinfo>
    </xs:annotation>
    <xs:attribute name="dateOfBirth" type="xs:date" use="required"/>
    <xs:attribute name="dateOfDeath" type="xs:date" use="required"/>
</xs:attributeGroup>

```

- **Gender :** est un ensemble d'attributs pour désigner le sexe de l'humain.

```

<xs:attributeGroup name="gender">
    <xs:annotation>
        <xs:documentation>
            Gender
            This is a set of attributes for document the gender
            of a human.
        </xs:documentation>
    </xs:annotation>

```

```

    <xs:appinfo>NONE</xs:appinfo><
  </xs:annotation>
  <xs:attribute name="genderAtBirth" type="xs:string"/>
  <xs:attribute name="currentGender" type="xs:string"/>
<xs:attribute name="impersonator" type="xs:boolean"/>
</xs:attributeGroup>

```

- **physicalDescriptors** : le terme *physicalDescriptors* est un ensemble d'attributs destiné pour décrire le physique de l'humain.

NOTE: La taille et le poids devraient être des nombres mais doivent tenir compte de différents systèmes de mesure.

```

<xs:attributeGroup name="physicalDescriptors">
  <xs:annotation>
    <xs:documentation>
      Physical Descriptors
      This is a set of attributes for a physical description of
      A human
      NOTE: The height and weight types should be numbers but
      have to allow for different measurement systems.

      NOTE: The hair color, eyeColor, build attributes need code
      lists for values.

      NOTE: Scars, Marks and Tattoos should be a complex type
      for graphic, location, body part
    </xs:documentation>
  <xs:appinfo>NONE</xs:appinfo>
</xs:annotation>
  <xs:attribute name="height" type="xs:string"/>
  <xs:attribute name="weight" type="xs:string"/>
  <xs:attribute name="hairColor" type="xs:string"/>
  <xs:attribute name="eyeColor" type="xs:string"/>
  <xs:attribute name="build" type="xs:string"/>
  <xs:attribute name="scarsMarksTattoos" type="xs:string"/>
</xs:attributeGroup>

```

5.12 Le langage HPCDML:

Le sous-comité du HPCDML du comité technique d'HumanMI est chargé de développer et mettre en oeuvre les spécifications du langage (HPCDML). Ce sous-comité est présidé par Rex Brooks. HPCDML est désigné pour fournir une description standard pour les caractéristiques physiques des humains, en prenant la responsabilité de les harmoniser et de les partager avec les organismes de la santé publique, le corps médical, le domaine biométrique, archéologique... [50]

HPCDML inclus la description de l'anatomie de l'humain en se basant sur le vocabulaire existant en anatomie et en médecine, et l'animation des humanoïdes sur le web [56].

Le Sous comité de HPCDML cherche actuellement de nouveaux membres pour participer comme représentants des communautés d'intérêt liées à l'anatomie et à la physiologie humaines, Biométrie, biomécanique, informatique médicale, archéologie, anthropologie biologique, culturelle, et linguistique, santé publique, sûreté publique, Education/Training, et sciences comportementales et cognitives humaines [56].

HPCDML est utilisé pour représenter les modèles 3D et les animations des avatars.

Les rôles et les cas d'utilisation du HumanML et du HPCDML ont augmenté. Ils jouent un rôle important dans l'étude des reconstructions virtuelles, pour la représentation biomécanique [56].

Chapitre 6: L'utilisation du langage HumanMI dans les SMA

Parmi les objectifs du langage HumanMI, l'intention d'humaniser le web en permettant aux humains de communiquer avec les machines d'une manière plus facile et moins ambiguë. Pour arriver à réaliser ce langage, les membres d'OASIS Consortium invitent tous les intéressés à apporter leurs contributions dans le but de mettre en œuvre ses bases et en proposant les domaines où le langage sera utilisé.

Dans ce chapitre, nous essayons de répondre à cette demande et d'apporter une contribution concernant l'utilisation du langage HumanMI dans la communication entre agents. Notre travail consiste à présenter les principes du langage HumanMI et son utilisation comme support de communication en SMA et particulièrement dans le cas des agents conversationnels, un domaine qui est actuellement en plein essor.

Ces agents peuplent actuellement les interfaces. Sur le net, ils ont déjà été introduits dans de multiples applications. Nous les trouvons sous forme d'avatar, de chatterbots et d'ACA. Dans les applications du e-commerce et du e-learning par exemple, ces agents sont utilisés comme assistants, tuteurs ou accompagnants. Dans les IHM, l'utilisation de ces agents virtuels permet d'améliorer les échanges avec les utilisateurs. Dans de tels systèmes, les agents et les humains doivent interagir ensemble en vue de communiquer et de coopérer afin d'atteindre leurs objectifs.

Cependant, lorsque l'humain et la machine communiquent ensemble, le risque d'avoir des malentendus ou des conflits est majeur. Un agent logiciel et un utilisateur humain ne partagent ni le même vocabulaire, ni les mêmes perceptions. Dans ce cas, il est très important de trouver un point commun entre les deux participants.

HumanML est un langage de représentation capable d'inclure d'importantes informations liées aux caractéristiques humaines, tels que le langage non-verbal, les expressions faciales, les émotions...de telles informations permettent une transmission plus précise et plus individuelle des messages. Ainsi, l'utilisation de ce langage dans la communication entre humain et agents conversationnels peut apporter une nette amélioration dans les échanges entre les deux. Le pouvoir du langage HumanMI à introduire les caractéristiques humaines dans le message digital permet d'ajouter une touche d'humanité dans les échanges entre humain et agent artificiel.

Or, les ambiguïtés dans la transmission des messages ne se produisent pas uniquement dans le cas de l'interaction humain-agent. Ces conflits peuvent même arriver entre les humains qui se communiquent à travers le net (Chat, e-mail, forum...). Ceci est dû en fait à l'absence des facteurs humains nécessaires dans la communication du face à face.

L'objectif des recherches actuelles dans la communication Humain-Machine ou Humain-Machine-Humain est d'arriver à humaniser le net et créer une voie standard pour représenter les émotions, le langage du corps et les caractéristiques humaines. Ce qui conduit à une amélioration de la communication en la rendant plus compréhensible et moins ambiguë.

Utiliser le langage HumanMI avec les agents conversationnels facilite la tâche de la représentation du comportement humain par la machine et perfectionne les échanges entre l'utilisateur et l'ordinateur de manière à ce que ces échanges soient des plus naturels possible. HumanMI peut perfectionner les échanges entre les internautes par l'intermédiaire de la machine jusqu'à arriver aux interactions proches de la nature humaine.

Nous proposons d'utiliser ce langage comme un support permettant d'améliorer les échanges dans les cas suivants :

- HumanMI dans un système de Chat
- HumanMI dans les IHM

Ainsi, nous présentons dans ce chapitre non pas un modèle complet d'agent, mais uniquement les éléments du langage HumanMI à ajouter ou à inclure à un modèle quelconque d'agent.

6.1 HumanMI dans un système de Chat :

Le réseau Internet est devenu ces dernières années un moyen de communication à grande échelle grâce à la diversité des services qu'il offre. C'est le cas particulier du courrier électronique, des forums et du Chat. Les espaces de discussion éprouvent actuellement un besoin énorme pour communiquer d'une manière naturelle et intuitive vu la diversité et la nature des usagers. Le nombre d'internautes en 2007 dépasse 1 milliard 350 millions [Source : *Computer Industry Almanac*], 75% parmi eux sont issus des pays non-occidentaux [04]. Cette diversité dans l'origine, la culture et les langages des usagers va certainement créer des difficultés lors de l'interprétation des informations transmises par l'échange du message via le réseau mondial.

Parfois, l'interaction se trouve fragilisée par rapport à une situation où les deux partenaires sont face à face. De nombreux obstacles peuvent nuire à l'intercompréhension et une situation de malentendus peut s'installer ce qui peut mener à l'achèvement rapide d'une relation. Quelques études sur l'interaction à distance exposent que l'absence des aspects humains élémentaires rencontrés normalement dans une situation du face à face constitue un problème majeur lors des échanges de messages entre les intervenants.

Ces aspects humains sont représentés par le non verbal, le para-verbal, les émotions, les postures corporelles, les expressions faciales, les intentions,...

En effet, le rôle de la communication non verbale est très important dans la communication humaine. Ceci explique pourquoi le langage du corps, la tonalité de la voix, des mouvements des yeux et les expressions du visage transmettent un message plus riche et plus fort qu'un message comprenant uniquement du verbal. Ces ajouts permettent également au receveur du message de connaître quelques traits de la personnalité de l'envoyeur [55]. Le professeur Albert Mehrabian, chercheur en communication, estime que chaque message transmis comporte 55% d'expressions faciales, 38% du vocal et seulement 7% du verbal [55].

Or, la communication naturelle est celle qui ressemble à la communication du face à face, là où les interactants utilisent les différents aspects humains d'une manière intuitive. Elle permet aux interactants d'appréhender une situation sans risque d'ambiguïté et autorise une transmission efficace de l'information.

Dans une situation de Chat, une autre source de conflit provient à partir des différentes interprétations du même événement. L'interprétation fautive d'une plaisanterie par exemple peut nuire aux relations interpersonnelles. Ceci se produit parce que des processus humains nécessaires pour la communication ne sont pas explicitement exprimés (émotions, non verbal,...) [51] en plus de la rigidité du médium [55].

L'acte naturel pour parer à ce problème est de raffiner la communication. L'ajout d'informations supplémentaires au message initial lorsque qu'il est possible, permet de lever certains malentendus.

Pour remédier à cela, les logiciels mettent généralement à la disposition des interactants une panoplie de signes graphiques, les smilies (“:-)”) pour “joyeux”) destiné à compenser le non

présentiel, et inclure des aspects humains élémentaires dans la communication afin de traduire les sentiments et les gestes. Ceux-ci ont certainement aidé à donner l'expression humaine dans les messages.

Cependant, leur utilisation manque de la standardisation [51]. Ce manque de standardisation exige l'adoption d'un standard permettant d'offrir une meilleure communication parmi la population d'Internet et aider à améliorer les échanges entre les utilisateurs de différentes langues et cultures.

HumanML, peut être adopté comme un standard de communication sur le web. Il espère combler ce manque, et traiter de façon automatisée les informations qui concernent l'utilisateur.

Dans cette partie, nous proposons d'utiliser le langage HumanML comme solution aux problèmes liés à la communication notamment l'ambiguïté et les mauvaises interprétations dans la communication en ligne.

Avec l'utilisation du langage HumanML dans un système de Chat, nous pouvons arriver à transmettre des émotions à l'aide d'un avatar capables de traduire les émotions et les intentions de l'interactant. L'utilisation du langage HumanML deviendra de plus en plus efficace et utile dans ce cas car il peut apporter une solution aux problèmes de conflits situés dans la communication à travers le net en introduisant les aspects humains dans les messages transmis.

- **Le rôle des émoticônes dans les systèmes de Communication en ligne actuels :**

En communiquant avec une autre personne par e-mail, SMS ou par Chat, il est très difficile d'exprimer les émotions et les gestes du corps. Cette incapacité mène aux malentendus dans la communication.

Les émoticônes ont enrichi la communication humaine sur le net. Le plus souvent, les smileys sont décrits comme des conventions utilisées pour compenser l'absence d'indices paralinguistiques, comme la mimogestualité ou l'intonation. Il s'agit de combinaisons de caractères permettant de représenter de manière schématique (si on les incline à 90 degrés) des mimiques faciales comme des sourires, des clins d'oeil, des moues de colère ou de tristesse. Les smileys sont très souvent utilisés dans les forums de discussion. Les smileys sont des phénomènes qui s'inscrivent dans un processus de cadrage plus large : faire du face à face avec de l'écrit [26].

Un exemple d'émoticônes est montré dans le tableau 7. Lorsque ces émoticônes sont employées dans des messages de textes, le récepteur peut utiliser cette information ajoutée pour comprendre mieux le sens du message. Les smileys permettent de rendre plus accessibles les sentiments et les émotions de l'auteur du message, comme la joie, la tristesse ou la colère. Par exemple, Le smiley souriant (-: ou ☺ peut jouer un rôle expressif lorsqu'il semble apporter une information sur l'état émotionnel de l'émetteur d'un message, qui n'est pas accessible par son contenu verbal.

Le smiley peut permettre d'explicitier la dimension émotionnelle d'un message lorsque son contenu verbal rend possibles plusieurs interprétations.

Exemple, tiré du forum AudioFanzine [26].

Le trou est fait proprement, il fait 15cm de diamètre environ, plein axe avec un des HP. Le but était de placer un micro en oppo de phase pour des séances d'enregistrement, il y a de ca 5 ou 6 ans. Depuis il est rebouché avec du contreplaqué mais c'est du provisoire qui dure :-)

« Du provisoire qui dure » peut être compris comme une appréciation positive ou négative. Le smiley triste oriente vers la seconde interprétation.

:-) ou ☺	Smile
;-) ou	Wink
:-(ou ☹	Sad

Tableau 7. : quelques émoticônes les plus utilisées

Ou exagéré

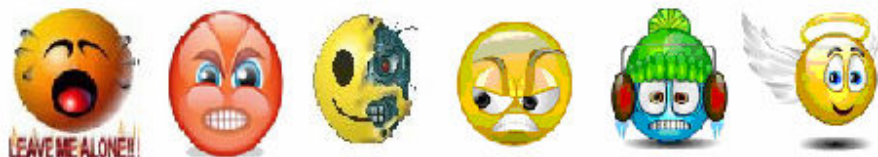


Figure 28. Smileys

On peut faire relever diverses fonctions des smileys:

- Un smiley peut être expressif : il sert à décrire l'état d'esprit du locuteur (la joie, la colère, la tristesse).
- Le smiley peut être une aide apportée au destinataire pour qu'il puisse aisément interpréter les énoncés. On peut parler de smiley interprétatif. Le smiley « clin d'oeil » a généralement cette fonction car il permet de lever les ambiguïtés des énoncés ironiques ou humoristiques.
- Le smiley permet au locuteur d'indiquer la relation qu'il désire instaurer avec son lecteur. Utiliser un smiley « souriant » peut être ainsi à la fois une manière d'exprimer son émotion et de donner une tonalité particulière à l'échange.
- Le smiley est aussi un procédé de politesse, un moyen de désamorcer le caractère offensant d'un message [25].

Les usagers de la communication par internet croient que sans l'utilisation des émoticônes dans les messages, l'incompréhension dans la communication est inévitable [56].

Par exemple, pour indiquer de la tonalité dans les messages instantanés, la plupart les utilisateurs ne prennent pas le temps d'ajouter les indicateurs paralinguistiques aux messages textuels. Cependant, avec l'utilisation des émoticônes, on peut éviter l'incompréhension entre les interactants sur le net. Jusqu'ici, la perception des émotions est exprimée par une ponctuation exagérée ou à l'aide de capital.

Cependant, puisqu'il n'y a aucun standard qui gère l'usage des émoticônes, leur signification reste différente d'un utilisateur à l'autre [55] car leur avantage n'est pas officiel et reste finalement limité.

Évolué à partir de la pratique simple de l'utilisation des émoticônes pour améliorer la communication dans les e-mail, des newsgroups et le Chat, HumanML vise vers une représentation plus officielle des caractéristiques humaines (par exemple, culturel, physique, psychologique, etc.) [31]. Il possède un but semblable aux émoticônes existantes actuellement comme :) pour un sourire, ou :(pour montrer la tristesse.

Introduire le langage HumanMI dans un SMA permet d'ajouter des caractéristiques humaines aux applications sur le web. Dans ce sens, nous proposons un système dans lequel les

utilisateurs peuvent communiquer non seulement avec le texte, mais aussi au moyen des avatars, tant verbalement que non-verbalement en exprimant les émotions capable de pallier les incompréhensions et les malentendus. Le support pour arriver à ce résultat est l'usage du langage HumanMI dans ce système de Chat.

En effet, certains systèmes récents utilisent les agents conversationnels ou des Avatars comme de nouveaux moyens de communication afin d'améliorer les interactions dans les systèmes de Chat. Les plus connus, le système de BodyChat utilise des agents conversationnels animés pour enrichir la communication Humain-Humain sur le web [27]. Récemment, une thèse de doctorat présentée à l'Université du Pays basque avance que l'utilisation d'avatars pourrait être une forme efficace de communication non-verbale, qui se concentre principalement sur les aspects émotionnels [58].

Dans le paragraphe suivant, nous décrivons brièvement notre système de Chat dans lequel des agents animés incarnent le rôle d'avatars d'utilisateur et de messagers conversationnels.

Le travail présenté dans cette partie est inspiré des travaux de Rex Brooks [61] un fondateur du langage HumanMI et le chargé de développement du langage HPCDML.

6.1.1 Description du système :

Le système de Chat que nous proposons permet à des internautes de communiquer à travers le net en utilisant des avatars qui possèdent des expressions faciales et des attitudes chez les personnages virtuels de manière intuitive, en fonction des conversations tenues et des contextes. Le schéma primaire du langage HumanMI décrit les six émotions de base qui sont : la joie, la colère, le dégoût, la peur, la surprise et la tristesse. Ces émotions traduisent même les différents niveaux d'intensité émotionnelle.

Les discussions entre les interactants se passent comme dans un système de Chat classique. Mais, puisque les émotions et les expressions faciales sont généralement perdues lors des échanges en ligne, nous proposons d'introduire des avatars dans le système de Chat. Leur rôle principal est d'enrichir les échanges entre les usagers par l'expression des aspects humains nécessaires dans la communication du face à face ainsi que pour remplacer les émoticônes souvent utilisés de manière non standard.

Le système se sert du langage HumanMI comme support pour effectuer cette tâche. Le résultat obtenu est une meilleure communication et une élimination des conflits entre les interactants.

Le but de ce système est de permettre aux internautes d'exprimer leurs émotions non seulement d'une manière verbale mais aussi d'une manière non verbale à travers un avatar qui ajoute des expressions faciales selon le texte envoyé. C'est le travail effectué par le langage HumanMI.

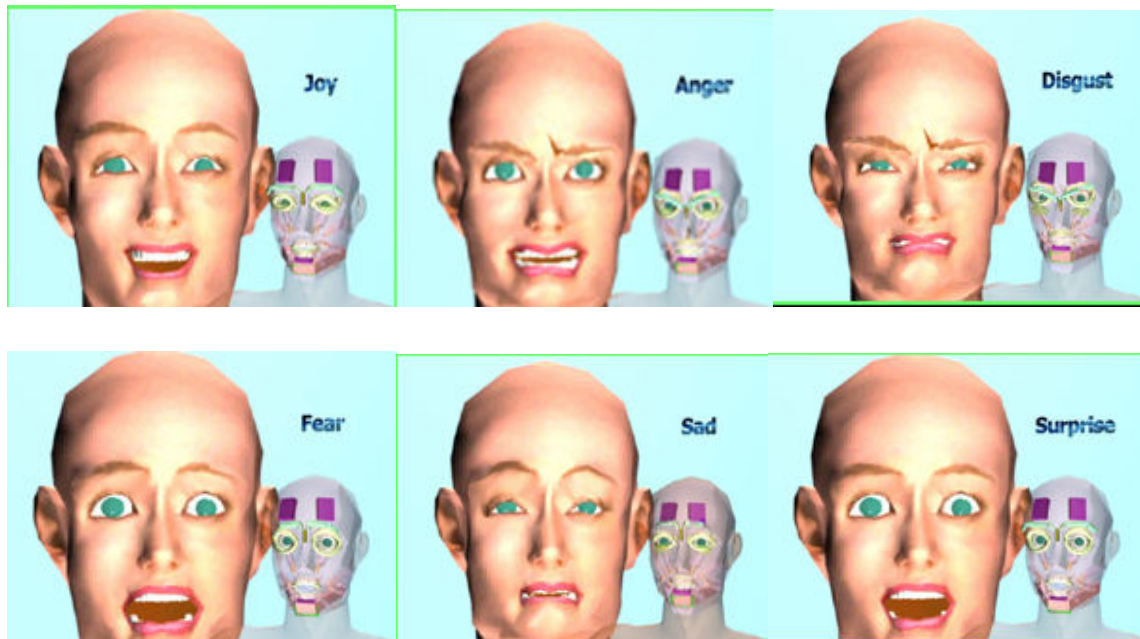


Figure 29. Les six émotions de base

6.1.2 L'architecture du système de Chat :

Dans ce paragraphe, nous présentons notre vision de l'architecture fonctionnelle du système de Chat et nous détaillons son fonctionnement. Cette architecture est illustrée par la figure 31. Le système se compose de trois parties:

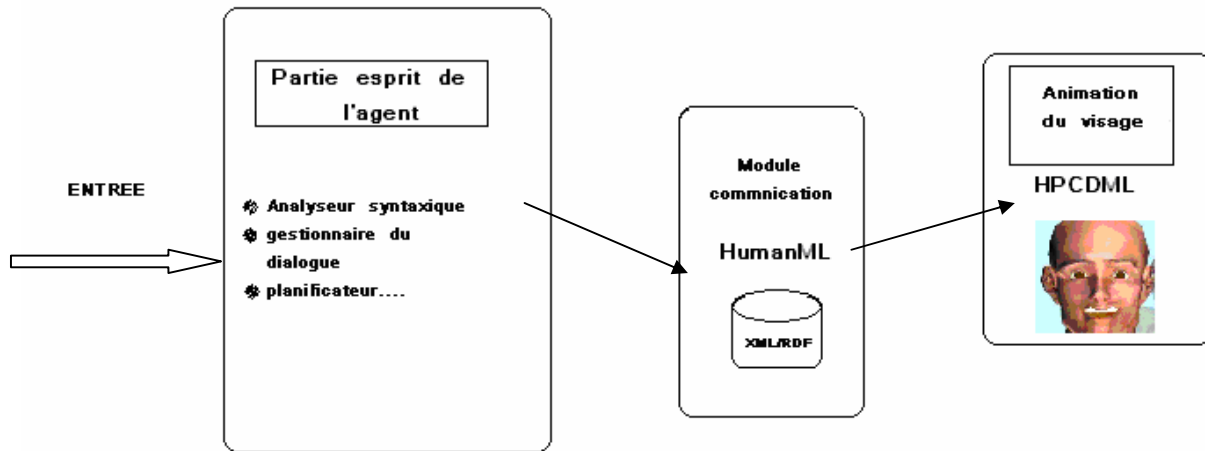


Figure 30. Architecture du système de Chat

- La première partie représente l'esprit de l'agent. L'agent peut avoir une architecture simple. Les détails de l'agent ne nous intéressent pas. Beaucoup de travaux ont été réalisés dans ce sens [30]. Cette partie permet de recevoir le texte en entrée et l'analyser afin d'extraire les émotions de bases qui sont présentes dedans.
- La deuxième partie est le corps : qui peut avoir une représentation en 2D, ou en 3D ou simplement une représentation comique. Son rôle est de représenter l'humain dans l'environnement du Chat.

Rex Brooks essaye par ses travaux sur un système d'un squelette du visage d'animer les principaux muscles responsables de produire les six expressions faciales humaines de base (la joie, la tristesse, la crainte, la colère, le dégoût et la surprise) [53]. Le système de squelette du visage est illustré sur la figure 30.

L'animation appropriée pour chaque émotion est envoyée sur le réseau dans des fichiers format .bvh [53]. Cependant, les tailles de fichiers engendrés par les logiciels de conception et le contrôle des animations posent des problèmes majeurs d'architecture web et c'est l'obstacle majeur rencontré lors du développement du langage HumanML.

Actuellement, Rex Brooks invite les intéressés à participer au développement du schéma secondaire dont l'objectif principal est le développement du langage Human

Physical Characteristics Descriptions Markup Language (HPCDML) afin d'améliorer la représentation de l'anatomie du corps humain pour ensuite l'utiliser dans des applications d'anthropologie, de médecine, ...[53].

- Le corps et l'esprit sont interfacé par un module de communication intégrant dans sa composition le langage HumanMI. Son rôle principal est d'ajouter du non verbal à la conversation. Cette partie qui est représentée par le langage HumanMI permet ainsi de baliser le texte en fonction de l'acte déterminé. Il convient ensuite d'ajouter de l'intensité à l'émotion choisie suivant le contexte de la conversation.

6.1.3 Le fonctionnement du système:

Dans cette partie, nous présentons une vision générale de notre système.

1. L'utilisateur introduit un message à travers une zone texte.
2. Le système qui est l'esprit de l'agent effectue une analyse sémantique de l'énoncé afin d'extraire des indices lui permettant de connaître l'état émotionnel de l'interlocuteur. Et détermine ensuite l'émotion que doit exprimer l'avatar.
3. Après avoir déterminer l'émotion à envoyer avec le texte, le rôle de HumanML dans cette phase est de baliser le texte en temps réel de telle sorte à produire, par exemple, une réponse non verbale (joie, peur, surprise) rapide.

Le premier schéma de base de HumanML [54] défini les six émotions de base en terme XML qui permettent de mettre en relation des états émotionnels spécifiques avec certaines formes d'action.

4. HumanMI utilise des fichiers XML pour envoyer le texte ou les images.
5. Le logiciel client interprète les balises HumanMI et produit les expressions qui correspondent à l'état émotionnel de l'envoyeur. L'avatar réalise les animations en temps réel.

Exemple:

Pour comprendre comment l'application de Chat fonctionne, nous donnons un exemple de deux internautes entrain de parler à distance.

L'utilisateur en disant : « je suis content de te voir » veut exprimer sa joie avec un sourire. Habituellement, les internautes dans un système de Chat classique, ajoutent des

émoticônes. Mais avec l'utilisation de HumanMI, le message envoyé par l'internaute est accompagné automatiquement avec l'animation correspondante.

L'agent après analyse de la phrase détecte le mot content et désigne que l'état dans lequel se trouve l'utilisateur nécessite l'expression d'un sentiment de joie. HumanMI décide de quelle émotion il s'agit, et balise le texte avec la balise suivante :

```
<huml :emotion>  
<joy intensity=1> je suis content de te voir </joy>
```

Cette balise appartient au schéma XML désigné par le huml *emotion* (voir chapitre précédent). HumanMI envoie ensuite l'animation appropriée à partir des fichiers. bvh à l'autre internaute.

Dans cette application de Chat, l'usager n'utilise pas seulement du verbal (je suis content de te voir) mais aussi du non verbal grâce aux expressions faciales associées au message. Cet ajout donne au message plus d'expressivité, et permet de lever certaines ambiguïtés. Au moins si le récepteur ne comprend pas bien la langue, il va certainement savoir que l'émetteur a envoyé un message positif, ce qui engendre un confort dans le rapport entre les intervenants.

6.2 L'utilisation du langage HumanML dans les IHM :

L'amélioration des IHM est devenue la préoccupation majeure de plusieurs chercheurs. Le but de ces recherches est de perfectionner les échanges entre l'utilisateur et l'ordinateur de manière à ce que ces échanges soient des plus naturels possible.

Actuellement, sur le web, plusieurs applications complètent leurs interfaces Homme-Machine graphiques usuelles avec des personnages virtuels utilisant le langage non-verbal qui a un rôle capital dans la communication humaine. Pour que ces agents soient plus naturels et plus convaincants dans leur discours, ils sont présentés à l'utilisateur sous forme d'humanoïde. L'agent doit comprendre ce que l'utilisateur dit, par sa parole, par ses comportements non verbaux et réciproquement doit répondre à l'utilisateur par la parole et des comportements non verbaux.

En effet, une grande partie des malentendus dans la communication est due au manque d'informations transmises avec le message. L'ajout des facteurs de communication non verbale

est essentiel dans les échanges entre les individus. Plusieurs recherches de psychologie ont montré que dans une interaction entre deux personnes, 70% des messages passent par le non-verbal. Notons simplement que les développements de ces recherches sont aujourd'hui nombreux, jusqu'aux recherches récentes sur les agents conversationnels en IHM.

Dans les dernières années, un nombre important de recherches a été mené pour spécifier et développer un mécanisme pour doter les agents conversationnels d'émotions et de comportement humain [30]. Le langage HumanML en est un exemple [47]. Son but est de donner de la crédibilité aux agents conversationnels pendant l'interaction Homme-Agent en le dotant de qualités humaines telle que la personnalité, la sympathie, les émotions, l'animation du corps etc.... afin d'éviter tout malentendus et ambiguïtés dans le message transmis. Ce qui donne comme résultat l'accroissement de la crédibilité dans le processus de la communication Human-Human ou Human-Agent.

Dans ce cadre, nous proposons d'utiliser le langage HumanMI pour introduire des aspects humains sur les interfaces Homme-Machine. Nous proposons deux applications pouvant introduire HumanMI dans leurs fonctionnements :

- Un agent présentateur du contenu d'une page web.
- Un agent conversationnel pour l'e-commerce utilisant la sémiotique propre aux clients.

6.2.1 Agents présentateurs d'informations sur internet :

La première application que nous proposons est d'utiliser le langage humanMI pour réaliser un agent conversationnel pour la présentation d'information sur le Web. Ce genre d'agents capable de lire et d'interpréter le contenu des pages Web en temps réel peut être appliqué à des contextes d'assistance technique, ou encore au domaine touristique ou encore lire une histoire illustrée d'image pour une personne qui est d'une langue ou culture différente.

L'introduction du langage HumanMI, permet au système de générer automatiquement des gestes expressives et les coordonner avec la parole. Le langage dispose d'un ensemble de balises pour définir la kinésique. Dans un souci de standardiser les gestes et les réutiliser dans d'autres applications sur le web, HumanMI propose le type complexe « Kinesic » dans son schéma primaire. Le schéma secondaire travaille sur la définition de chaque type de geste dans le but de créer une base de gestes universelle.

La kinésique communicationnelle désigne une partie du vocabulaire qui constitue le langage du corps utilisé pour décrire les comportement et les émotions et pour ajouter des poses ou des arrêts à la communication verbale.

- **Kinesic: Human Movements**

Ce type complexe définit tous les gestes qui appartiennent à l'humain (déictiques, paraverbaux, iconiques, métaphoriques, de battement) dans le but de les échanger entre les différentes applications sur le web. Ceci pourrait alors être représenté dans un browser HumanML comme images, son ou mots [52] ou animations.

Le schéma RDF correspondant au huml *Kinesic* est le suivant :

```
RDF <=> XML Schema

<?xml version="1.0" encoding="UTF-8"?>

<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
xmlns:xs="http://www.w3.org/2001/XMLSchema-datatypes">

  <rdfs:Property rdf:id="id">
    <rdfs:domain rdf:resource="#Semiosis"/>
    <rdfs:domain rdf:resource="#Kinesic"/>
    <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema-
datatypes#ID"/>
  </rdfs:Property>
  <rdfs:Property rdf:id="id">
    <rdfs:domain rdf:resource="#Semiosis"/>
    <rdfs:domain rdf:resource="#Kinesic"/>
    <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema-
datatypes#ID"/>
  </rdfs:Property>
  <rdfs:Property rdf:id="humlName">
    <rdfs:domain rdf:resource="#Semiosis"/>
    <rdfs:domain rdf:resource="#Kinesic"/>
    <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema-
datatypes#string"/>
  </rdfs:Property>
  <rdfs:Property rdf:id="fromDate">
    <rdfs:domain rdf:resource="#Semiosis"/>
    <rdfs:domain rdf:resource="#Kinesic"/>
    <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema-
datatypes#dateTime"/>
  </rdfs:Property>
  <rdfs:Property rdf:id="toDate">
    <rdfs:domain rdf:resource="#Semiosis"/>
    <rdfs:domain rdf:resource="#Kinesic"/>
    <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema-
datatypes#dateTime"/>
  </rdfs:Property>
  <rdfs:Property rdf:id="toDate">
    <rdfs:domain rdf:resource="#Semiosis"/>
```

```

    <rdfs:domain rdf:resource="#Kinesic"/>
    <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema-
datatypes#dateTime"/>
  </rdfs:Property>
  <rdfs:Property rdf:id="intensity">
    <rdfs:domain rdf:resource="#Kinesic"/>
  </rdfs:Property>
  <rdfs:Class rdf:ID="Semiosis">
    <rdfs:label xml:lang="en">Semiosis</rdfs:label>
    <rdfs:comment>
      Semiotic Communication Mode.
      Semiosis is a meaningful exchange of signs, signals and symbols
among cognitive agents.
      NOTE: This process is the model of the human communication
process upon which HumanML is based. It can be, and we expect that it will
be further enumerated by semiotic types and extended in the Secondary Base
Schema and subsequent huml schemata.
    </rdfs:comment>
  </rdfs:Class>

  <rdfs:Class rdf:ID="Kinesic">
    <rdfs:label xml:lang="en">Semiosis</rdfs:label>
    <rdfs:comment>
      Kinesic: Human Movements
      Communicational Kinesics constitute some vocabulary of body
language used to portray moods and emotions and to add emphasis to verbal
communication. As a study concerned with how bodily and facial gestures
function as a factor in communication, kinesics is fairly well understood.
For our purposes we expect enumeration of body language gestures to be
included in culture-specific subsets.
    </rdfs:comment>
  </rdfs:Class>
</rdf:RDF>

```

- **Le principe de fonctionnement :**

Notre système travaille de la manière suivante :

Le personnage est affiché à côté d'une page web ou intégré dans un environnement virtuel en 2D ou en 3D. Il a la capacité de parler et d'incorporer le contenu complet de la page.

Le personnage vise à fournir des signes de communication verbale et non verbale (regard, expressions faciales, gestes de la main, postures). Chaque mouvement, chaque posture corporelle, chaque expression faciale, chaque geste de l'agent a ainsi une fonction communicative.

Ce système n'est pas destiné pour converser avec les usagers mais uniquement pour présenter un contenu. L'utilité de ce système est de minimiser les ambiguïtés souvent rencontrées lorsque l'utilisateur ne comprend pas la langue ou il n'est pas de la même culture. Les gestes qui accompagnent les paroles peuvent guider et donner un peu d'éclaircissement à l'usager pour comprendre le contenu de la page.

Lorsqu'une page web est affichée, un agent humanoïde apparaît sur l'écran. Il commence par lire le contenu de la page en accompagnant son discours par des gestes et des expressions faciales, dans le but d'aider l'utilisateur à comprendre le contenu de la page.

L'agent peut se déplacer sur l'écran et faire des références par la parole et par des gestes déictiques aux objets tout en effectuant la démonstration du contenu de la page.

Grâce à la kinésique contenu dans le vocabulaire de HumanMI, l'Agent est également capable d'inclure dans son discours les gestes déictiques appropriés en fonction de la proximité et du nombre d'objets désignés (là, là-bas, cela, ...). Il peut également faire des gestes iconiques des objets dont il parle. Les gestes sont générés à partir du contenu sémantique du discours.

HumanMI cherche dans la base des ontologies RDF le geste approprié à la fonction communicative énoncée. En effet, chaque geste peut être généré suivant la sémantique contextuelle. L'animation correspondante sera automatiquement générée. Comme dans le cas de l'application du Chat, le langage HPCDML sera appliqué pour réaliser des séquences animées pour chaque geste.

Exemple :

Par exemple, quand l'agent rencontre un texte qui a le sens de montrer un objet, il désigne l'objet en disant 'cet objet' et il regarde en direction de l'objet et emploie un geste déictique pour désigner l'objet.

6.2.2 Agents conversationnels dans l'e-commerce:

Depuis quelques temps, certains sites web utilisent des agents conversationnels ou chatterbots pour conseiller les visiteurs. Ils jouent presque le même rôle qu'un vendeur en magasin. L'objectif principal de cette initiative est d'humaniser la relation avec le client sur le site web. Les entreprises cherchent à faire de leur site Internet, un point de vente à part entière. Pour cela, elles transposent les techniques du marketing au web.

On définit généralement l'atmosphère d'un point de ventes par des facteurs de design : l'ergonomie et l'attractivité visuelle du site et des facteurs d'ambiance : qui se traduisent sur le web par de la musique et même parfois des diffuseurs d'odeurs connectés à l'ordinateur. La dernière composante de l'atmosphère d'un point de vente sont les facteurs dits "Humains": les

vendeurs. Les agents cherchent à remplacer dans la mesure du possible le contact Humain immanquablement absent d'un site web.

Exemple dans la figure 32, l'agent présente des articles aux clients en les conseillant et parfois même en leur proposant des articles selon leurs préférences.



Figure 31. Agent e-commerce

Sur le site, les agents virtuels ont le but d'accueillir le visiteur, l'assister dans sa navigation au sein de sites complexes, et le conseiller sur les produits. L'agent permet à l'internaute paresseux ou peu initié à l'informatique de contourner les moyens de recherche classiques et parfois fastidieux (moteur de recherche, FAQ, navigation par le plan du site) pour poser simplement sa question en langage naturel.

Certains sites ont réalisé ce défi en plaçant un agent conversationnel sur la page d'accueil. Le visiteur pose une question à l'agent. Si la question n'est pas claire, l'agent se charge de proposer ou demander une reformulation, jusqu'à comprendre la demande de l'internaute et lui ouvrir la page qu'il recherche.

L'agent analyse sémantiquement et syntaxiquement les phrases de l'internaute et lui fournit une réponse grâce à son moteur de raisonnement, à ses connaissances métier et à l'historique de la conversation pour contextualiser chaque nouvel échange.

Selon le contexte où on l'utilise, l'agent pourra ainsi assurer de nombreux rôles d'assistance sur un site marchand: agent de renseignement, vendeur virtuel, speaker ou présentateur virtuel, guide ou hôtesse virtuel, agent de renseignement, conseiller virtuel, vendeur virtuel etc.

Les fonctions essentielles d'un chatterbot sur un site marchand sont généralement les suivants :

- Rendre la relation avec les visiteurs ou utilisateurs plus humaine, interactive et attractive grâce au dialogue ;
- Animer le site web ;
- Personnaliser au mieux l'accueil avec les visiteurs en collectant des données précieuses sur ces derniers (profiling) ;
- Accompagner les visiteurs tout au long de leur navigation ou de leur recherche ;
- Apporter une réponse rapide et pertinente aux questions des visiteurs en langage naturel ;
- Peut se référer à une base de connaissance précise sur un sujet particulier ;
- Permet, en lien avec des moteurs de recherche, une recherche vague ou approximative, lorsque l'utilisateur ne sait pas avec précision ce qu'il recherche ;

Supporter la communication avec un langage comme humanMI permet de dissimuler les malentendus et de concrétiser les échanges entre humain et agent grâce aux éléments du schéma primaire et du schéma secondaire du langage HumanMI visant à introduire le langage non verbal dans la communication la rendant ainsi plus conviviale et plus naturelle.

La personnalisation de l'IHM dans les échanges entre humain et agent conversationnel est une autre possibilité pour utiliser HumanMI avec un agent conversationnel dans le e-commerce. En effet, la personnalisation des interfaces est un domaine peu exploré. Adapter le site web aux visiteurs en fonction de leur profil et de leurs origines est un gain pour les entreprises marchandes. Selon une étude réalisée par intelliquest [20], 45% des visiteurs reviennent sur un site parce que l'information proposée est adaptée à leurs besoins.

Le but de la personnalisation donc est d'intégrer des informations personnelles à la présentation et au contenu de tout contact avec le client. La personnalisation permet aux entreprises d'interagir de façon plus personnelle avec leurs clients en taillant le contenu, la publicité et les services sur mesure pour des individus spécifiques [20]. Beaucoup de sites ont recours à cette technique pour se rapprocher des internautes.

Cependant, la mondialisation des marchés économiques a engendré l'accroissement des sites marchands qui attirent de plus en plus des populations issues de contextes linguistiques et

socioculturels extrêmement variés. Les interactants se trouvent parfois dans des situations d'incompréhension totale à cause des handicaps d'ordre linguistique ou culturel. Cet handicap de langue et de culture peut engendrer un problème de dépaysement dans le site. Or, si le client a l'impression que l'entreprise en ligne le comprend, alors le client aura un avantage réel à poursuivre sa relation avec ce site.

C'est pour des raisons d'ergonomie, que nous proposons de personnaliser le site web en introduisant les caractéristiques humaines dans la communication. Dans ce sens, humanMI permet d'humaniser les échanges suivant les caractéristiques des visiteurs en utilisant un agent conversationnel sur les interfaces de ce site ce qui permet d'accroître sa valeur et le rendre plus attractif. Avec HumanMI, personnaliser la communication entre individus issus de cultures différentes est un atout dont le but est de dissimuler les malentendus créés par les différences culturelles et linguistiques.

Nous proposons d'introduire dans le site marchand un agent conversationnel qui a le pouvoir de dialoguer avec les visiteurs du site grâce à la sémiotique et la gestualité propre à leur culture.

Dans le schéma primaire, HumanMI propose la balise <huml:culture> qui permet de définir toutes les cultures existantes :

L'XML schéma associé à la culture est :

```
<xs:complexType name="Culture" abstract="true">
  <xs:annotation>
    <xs:documentation>
      Cultural characteristics emerge as the defining set of elements
      which a community shares and which distinguish the community as
      a culture per se. This partakes of the semiotic basis for
      communication processes as a contextual element.
    </xs:documentation>
    <xs:appinfo>NONE</xs:appinfo>
  </xs:annotation>
  <xs:attributeGroup ref="humlTemporalAtts"/>
  <xs:attributeGroup ref="humlIdentifierAtts"/>
</xs:complexType>
```

En effet, certains sites ont récemment choisi de proposer un service personnalisé à leurs visiteurs. Connaître l'origine du client permet à l'agent conversationnel d'orienter le discours vers la culture de celui-ci et peut même lui proposer des produits sur mesure. Les gestes utilisés par l'agent conversationnel sont liés à la culture du client.

Cependant, le décodage du non-verbal est complexe. Il faut toujours tenir compte du contexte pour interpréter les éléments du non-verbal. L'utilisation du formalisme du RDF schema permet de décrire les langages et les cultures et désigner pour chaque culture les gestes appropriés. Les ontologies RDF ont été utilisées dans HumanMI pour bien interpréter les énoncés mais aussi comme aide au raisonnement. Cela permet de dire que ce geste appartient à tel culture, ou tel geste est utilisé dans telle situation [56]. En effet, le but principal d'une ontologie est de permettre le partage et la réutilisation d'une connaissance qui a été explicitée de façon formelle. Les composantes clés d'une ontologie sont un vocabulaire des termes de base et des spécifications précises de ce que signifie chaque terme. Les ontologies dans notre système servent à interpréter sémantiquement le contexte du message envoyé et l'associer avec le geste approprié.

Ainsi, lorsqu'un visiteur d'un site converse avec un agent qui s'exprime en utilisant les gestes de sa culture, il va se sentir probablement à l'aise et continue à visiter le site en toute confiance.

En effet, tout malentendu commence lorsqu'un message est interprété de plusieurs manières. Les incompréhensions et les ambiguïtés dans la communication se produisent souvent à cause des différences linguistiques et culturelles. La signification de gestes identiques peut différer d'un pays à l'autre. Par exemple, en comptant sur ses doigts, un Américain commence souvent par l'index, tandis qu'un français commence par le pouce pour indiquer le même nombre "1".

Les signes et les symboles sont des éléments importants des langues humaines. Évidemment ils changent avec l'individu, avec la culture, et avec le temps. Ces variations peuvent engendrer comme conséquence le malentendu. Le malentendu et la perception erronée sont la racine malheureuse du conflit, des dommages et douleur. Un exemple : une utilisation iconique dans une culture peut être facilement considérée comme idolâtrie dans d'autres cultures.

HumanMI dans le schéma primaire propose l'élément *semiosis* qui est utilisé pour échanger les signes, les signaux et les symboles à travers des agents cognitifs. Avec la sémiotique proposée dans HumanMI, on peut gérer les conflits d'ordre culturels et apporter une meilleure interprétation du contenu [43].

L'XML schéma associé à l'élément *Semiosis* est :

```

<xs:complexType name="Semiosis" abstract="true">
  <xs:annotation>
    <xs:documentation>
      Semiosis is a meaningful exchange of signs, signals
      and symbols among cognitive agents...
    </xs:documentation>
    <xs:appinfo>NONE</xs:appinfo>
  </xs:annotation>
  <xs:attributeGroup ref="humlTemporalAtts"/>
  <xs:attributeGroup ref="humlCommAtts"/>
</xs:complexType>

```

La sémiotique est l'étude de la manière dont les différents systèmes de signes permettent aux individus et aux collectivités de communiquer. C'est la théorie des signes, de la manière dont ils fonctionnent, de leur sens [50].

La non-identité des codes non verbaux est à l'origine des malentendus "interculturels". Ainsi les gestes varient selon le contexte et la culture de l'intervenant [46].

Les recherches dans le domaine de la sémiotique essaient d'identifier et de comprendre les signes spécifiques d'une culture. Les chercheurs ont identifié de grands ensembles de signes qui pourraient être codés et communiqués par l'intermédiaire des moyens standard. Les fondateurs de HumanMI espèrent réaliser ce défi en proposant à des participants de collecter toutes les cultures du monde et d'associer à chaque culture et gestes, les signes et les symboles utilisés dans la communication.

Ainsi, d'après les fondateurs du langage HumanMI un sourire large est différent d'un fou rire ; un bonjour indien est différent d'un bonjour occidental. Les gestes utilisés pour exprimer sa colère chez un italien ne sont pas les mêmes chez un chinois [46].

L'agent sémiote actuellement en cours de développement par OASIS consortium est l'exemple que HumanMI vise à introduire comme support pour l'utilisation des gestes spécifiques [54].

L'XML schéma associé à l'agent *Semiote* est :

```

<xs:complexType name="Semiote" abstract="true">
  <xs:annotation>
    <xs:documentation>
      A Semiote is a cognitive agent who participates in meaningful
      signal exchange among cognitive agents.
    </xs:documentation>
  </xs:annotation>
</xs:complexType>

```

```
</xs:documentation>
  <xs:appinfo>NONE</xs:appinfo>
</xs:annotation>
<xs:attributeGroup ref="humlCommAtts"/>
</xs:complexType>
```

- **Le principe de fonctionnement du système :**

Le système travaille de la manière suivante :

Nous proposons d'introduire dans l'interface d'un site marchand un agent conversationnel capable d'exprimer des gestes non-verbaux, des émotions et de personnaliser l'interface en communiquant avec les visiteurs grâce à la sémiotique de leurs cultures. Cette proposition est une tentative de personnaliser l'interface suivant les visiteurs et leurs origines dans le but de mettre le client à l'aise surtout lorsqu'il est un peu initié ou ne comprend pas bien le langage proposé sur le site.

L'agent est représenté par un humanoïde en 2D ou 3D. Son apparence peut être critiquée de plusieurs manières, puisqu'il va être en contact direct avec le visiteur car la relation qu'il va créer avec eux dépend en quelques sortes de son apparence physique. La discussion dépasse le cadre de notre mémoire.

La principale fonction d'un tel agent est de dialoguer avec l'utilisateur suivant sa culture et ses origines, ce qui accroît grandement l'interactivité. L'utilisateur entre un type de question ou d'affirmation quel qu'il soit, et le chatterbot y répond, comme le ferait une personne, instaurant un dialogue intelligent et en exprimant des gestes issus de la culture du client afin de lever les incompréhensions due aux différences culturelles. Ainsi, pour chaque visiteur, l'interface change de forme et propose un agent qui sera proche dans son comportement et ses gestes de ceux du client.

Un exemple d'un site qui utilise un agent pour assister l'utilisateur dans sa navigation est illustré sur la figure 33.

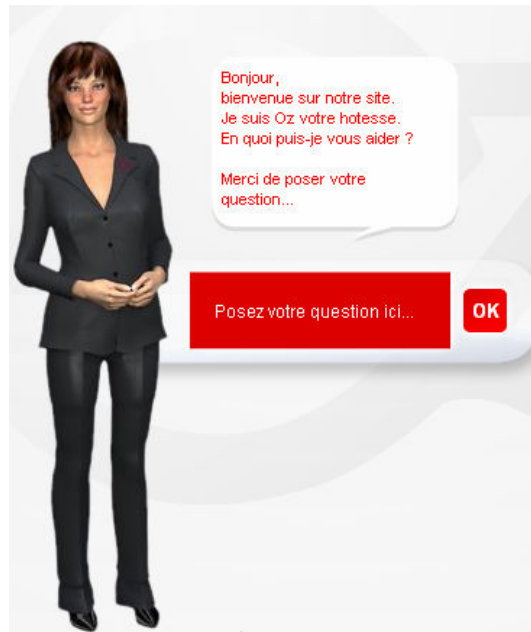


Figure 32. L'agent Oz du site Virtuos.com

- **Le processus de fonctionnement :**

L'agent attire tout d'abord l'attention de l'internaute par un petit message de bienvenue. En suivant le dialogue avec l'internaute, l'agent collecte les informations nécessaires pour le connaître mieux. Il essaie de connaître son identité culturelle, sa langue et ses origines. Les données collectées sur le visiteur permettront d'affiner la personnalisation de la relation.

Après avoir fait une image sur le visiteur, l'agent oriente le discours en utilisant les gestes, les signes et les symboles propres aux origines de l'internaute. Ainsi, certains malentendus dus aux différences culturelles seront dissimulés. L'agent peut ensuite continuer à dialoguer avec le client tout en utilisant la sémiotique de sa culture.

- **Architecture du système :**

L'agent proposé dans ce cas possède une architecture plus complexe que les deux agents cités dans l'application du Chat et dans l'application du présentateur du contenu sur le web. Puisque cet agent aura la faculté de dialoguer avec le client, son système cognitif doit être plus complexe. Il sera composé de plusieurs modules nécessaires pour planifier et générer le dialogue.

L'agent Sémiote fournit avec le langage HumanMI peut être utilisé dans ce cas puisqu'il est conçu pour présenter des gestes et du non verbal.

L'architecture du système sera comme suit:

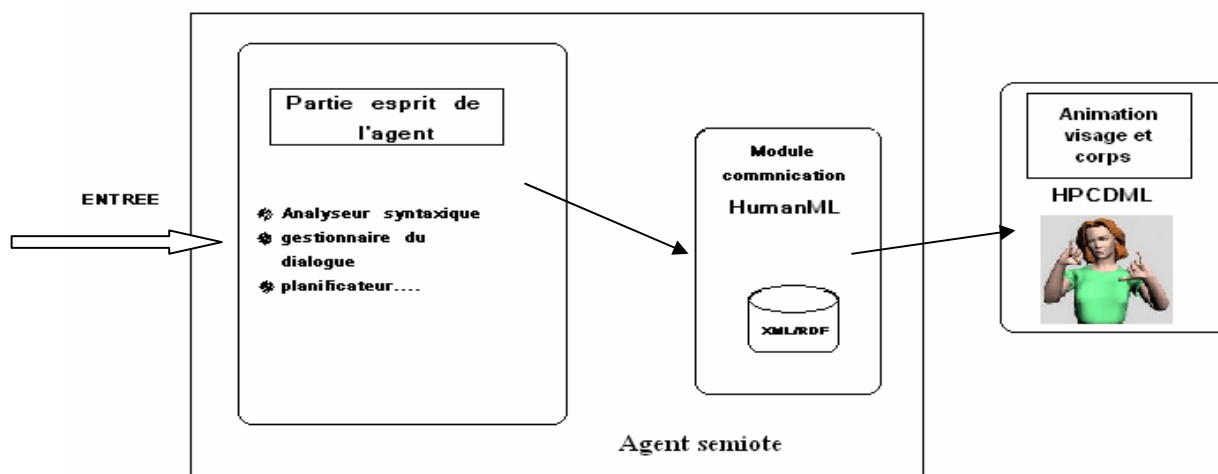


Figure 33. Architecture du l'agent pour le e-commerce

L'agent doit traiter le flux d'information en entrée en temps réel. Les différents modules et agents constituant le système travaillent ensemble dans le but de produire une réponse non verbale (écarquille des yeux, geste de retrait ou d'acquiescement) ou verbale (une exclamation par exemple) rapide. Le texte en entrée subi une analyse sémantique afin d'interpréter les actes.

Dans le but de montrer un comportement d'un assistant commercial, ou d'un vendeur, l'agent doit incarner ce rôle en adoptant la personnalité, le langage et les buts d'une personne travaillant dans le marketing. Ceci est très important, surtout lorsque le site investissant dans ce domaine veut garder sa clientèle. Donc l'architecture du système sera plus délicate, l'ajout d'un module ou d'un agent possédant cette personnalité est très important.

Durant la conversation, l'esprit de l'agent décide ce qu'il doit communiquer en prenant en considération les buts et le contexte du dialogue. Lorsque l'acte est déterminé, celui-ci est envoyé au générateur qui aura en charge de lui donner sa forme finale. Celle-ci se fonde sur une ontologie pour fournir l'énoncé de réponse.

Le deuxième agent introduit sur la plateforme, l'agent sémiotique. Il va collaborer avec le premier agent pour essayer de trouver les gestes et les signaux propres à la culture de ce client. Il

se sert dans ce cas des ontologies RDF définissant les gestes propres à chaque culture, le contexte de leurs utilisations. En effet, chaque geste peut être généré suivant la sémantique contextuelle.

L'animation correspondante sera automatiquement générée. Comme dans le cas de l'application du Chat, le langage HPCDML peut être appliqué pour réaliser des séquences animées pour chaque geste.

Conclusion :

Les applications présentées dans ce chapitre sont une proposition des cas d'utilisation du langage HumanMI et de son sous-langage le HPCDML. Le but était de montrer l'utilité de la communication non verbal et de la gestualité dans le cas des interactions humains et agents logiciels afin de lever certaines ambiguïtés dans le rapport entre eux. Nous avons limité notre travail dans la seule optique de présentation du langage. Quant à la réalisation des applications, il faut attendre que le langage soit présenté au grand public pour pouvoir bénéficier de tous les avantages que nous avons présentés le long de ce travail.

Chapitre 7 : Conclusion et perspectives

7.1 Conclusion :

Dans le cadre de ce mémoire, nous avons présenté les fondements du langage HumanML et son utilisation comme support de communication entre agents. Notre objectif initial est d'apporter une contribution concernant l'utilisation du langage HumanML dans les situations où des 'Humains' et des 'agents artificiels' communiquent sur le web en soulignant l'intérêt, mais aussi les difficultés que cela apporte. C'est la problématique autour de laquelle s'est déroulée notre recherche. L'objectif est de faire communiquer des agents de nature très différente alors qu'ils n'ont pas a priori de structure, ni de langage commun.

Nous avons commencé notre travail par la présentation des notions d'agents et des systèmes multi-agents en vue de fixer les concepts utilisés dans ce domaine. La notion des SMA met l'accent sur l'interaction permettant le passage des comportements individuels à un comportement collectif. Nous avons montré que l'humain est devenu un partenaire incontournable dans ces systèmes notamment ceux utilisés dans les interfaces Homme-Machine ou dans les environnements virtuels. Cependant, les protocoles de communication classiques entre agents tel que KQML et FIPA-ACL sont inadaptés pour la communication entre humains et agents car il ne prennent pas en considération tout les aspects humains nécessaire (tel que les émotions et les gestes, les expressions faciales...) pour atteindre une communication sans ambiguïtés. Nous avons montré que ces modes de communication ne sont pas suffisants pour les systèmes composés d'un collectif d'agents artificiels et d'agents naturels. Nous avons également traité l'importance du langage non-verbal et son impact sur la communication entre humains dans le sens qu'il permet d'enrichir le dialogue. Nous avons montré que ce type de langage est peu utilisé dans le cas des agents artificiels. Ceci nous a permis de mettre en évidence la capacité de la gestualité et des émotions à compléter le sens des énoncés. L'interaction d'agent conversationnel avec les humains pose des problèmes de représentation des émotions et de la gestualité. Cela nécessite la définition de modèles du comportements, issus des humains, pour exprimer différentes fonctions comme les actes de dialogue ou les émotions à travers différentes modalités (geste, parole, expressions faciales, postures).

C'est à la suite de ce constat que nous avons cherché de trouver un moyen pour représenter les caractéristiques humaines nécessaires pour la communication du face à face. Nous

avons choisi HumanML vu que l'objectif initial de ce langage est d'enrichir la communication par des aspects non verbaux et par la suite améliorer les contacts et lever toutes ambiguïtés dans le langage. Le chapitre 5 a été consacré pour la présentation des fondements du langage. Sa relation avec le web sémantique et les schémas XML/RDF permettent de montrer la capacité de ce langage à promouvoir la communication des humains sur le web ainsi que celle des humains avec les agents artificiels notamment l'amélioration des nouvelles Interfaces Homme-Machine. Ceci donnera aux usagers la possibilité de communiquer leurs émotions et gestualité à d'autres utilisateurs à travers Internet. Dans le chapitre 6, nous avons donné une architecture générale de notre proposition. Nous avons montré comment introduire le langage dans un système multi-agent et comment il doit réagir. Les applications qui pourront bénéficier des avantages de ce langage ont été présentées d'une manière générale. Ces applications visent en premier lieu à montrer l'importance du langage non verbal dans la communication sur le web et c'est l'objectif que HumanML est capable de concrétiser. Ces applications sont :

- Application de Chat pour remplacer automatiquement les émoticônes et trouver une voie standard pour les introduire dans un message.
- Agent présentateur d'information qui peut utiliser la gestualité lors du discours.
- Agent pour le e-commerce qui a la possibilité de personnaliser l'interface en introduisant la sémiotique et la gestualité propre à la culture du client.

Les deux dernières applications concernent l'utilisation du langage HumanML dans les Interfaces Homme-Machine.

Au cours de notre travail, nous avons cependant été limités par le manque de documentation sur le langage. Puisque le comité technique d'OASIS est fermé, actuellement il n'y a aucun organisme officiel qui se charge de développement du langage. Le peu de travaux disponibles et les publications effectuées jusqu'à ce jour ont été réalisés par les créateurs mêmes du langage au sein d'OASIS Consortium. Ce qui explique que le langage n'a pas encore atteint un niveau de maturité pour être présenté au grand public. Actuellement, les travaux en cours concernent le développement du schéma secondaire. L'avancement très long des travaux, risque lui aussi d'être arrêté par manque de contributions dans le domaine.

Nous avons également trouvé une difficulté majeure concernant la traduction des balises HumanML. Les logiciels de navigation actuelle ne sont pas ménagés pour prendre en charge ce langage. Le langage a été écrit uniquement dans un format abstrait c'est à dire la description de ces composants a été indiquée d'une manière générale. Il n'y a cependant aucun outil publiquement disponible pour la génération et le traitement des documents HumanML. Le langage tel qu'il est décrit dans ses spécifications initiales n'est pas prêt pour être utilisé pour le moment. Il nécessite plus de contributions afin d'effectuer un plug-in capable d'interpréter les balises de HumanML. En effet, puisque l'aspect humain est très difficile à modéliser et travailler sur un langage de représentation universel est une tâche extrêmement délicate, chaque équipe recourt à développer son propre langage de représentation suivant ces besoins. C'est pour cette raison que les contributions pour enrichir le langage HumanML sont faibles.

Etant conscients que la réalisation d'un agent conversationnel animé est une tâche très difficile à entreprendre, nous avons limité notre travail dans la seule optique de la présentation du langage. Quant à la réalisation des applications, il faut attendre que le langage soit présenté au grand public pour pouvoir bénéficier de tous les avantages que nous avons présentés le long de ce travail.

De la même façon, il est difficile de proposer un système sans disposer de traces d'interactions. Nous avons donc été quelque peu déçus de n'avoir pu mener à bien l'expérience sur des données concrètes pour obtenir un système avec un minimum de réalisme.

7.2 Perspectives :

Puisque le langage est encore au cours de développement, plusieurs perspectives nous paraissent intéressantes à explorer.

- ✦ Une des premières perspectives est d'apporter plus de contribution concernant les pistes qui pourraient faire l'objet de futures recherches. Les voies ainsi ouvertes sont essentiellement :
- ✓ L'e-learning par la génération automatique d'explication, avec une application possible dans les environnements virtuels de formations, introduire les émotions et le langage du corps en standardisant les gestes pédagogiques employées au cours de la présentation du cours ;

- ✓ Les applications de la réalité virtuelle ;
- ✓ Le CSCW : en améliorant la communication humain-agent-humain dans le but de travailler ensemble et de coopérer afin d'atteindre les objectifs visés par les interactants humain et agents.
- ✦ Les comportements des avatars peuvent également être améliorés. Dans ce cas, une extension de l'architecture est nécessaire pour ajouter des possibilités de décision proche de celle chez les humains. Nous proposons de situer donc au sein de l'agent un système (l'architecture BDI).

Le langage HumanMI est à ses débuts, il a besoin de plus de participation pour arriver à atteindre son but pour enrichir la communication humaine. Les responsables d'Oasis déclarent qu'il vont continuer à travailler jusqu'à la suppression totale des malentendus dans la communication humaine. Le monde sans conflits de communication sera sans doute meilleur.

Si l'idée est séduisante et qu'elle correspond effectivement à une réelle attente, bénéficier des avantages que présente le langage ne sera sûrement pas pour demain.

Bibliographie

- [01] Yasmine ARAFA et al, «*Two approaches to scripting character animation*», International Conference on Autonomous Agents, Proceedings of the second international joint conference on Autonomous agents and multiagent systems , Melbourne, Australia 14 – 18 Juillet, 2003
- [02] François BERTEL, «*Animation d'humanoïde dans un contexte conversationnel impliquant un dialogue verbal et non verbal*», Université de Rennes 1, 2003
- [03] François BOUCHET, «*Conception d'un langage de requêtes pour un agent conversationnel assistant* », Université Paris XI, 2006
- [04] Rex Brooks and Russell Ruggiero, «*HumanML & Government Related IT Directives (Part One)* », 2003
- [05] Stéphanie BUISINE, «*Conception et Evaluation d'Agents Conversationnels Multimodaux Bidirectionnels*», Université Paris 5 – René Descartes, 2005
- [06] Brahim CHAIB-DRAA et al, «*Système multi-agents : principe généraux et applications* », Université Laval, 2001
- [07] Romanic CHARTON, «*Des agents intelligents dans un environnement de communication multimédia : vers la conception de services adaptatifs*», Ecole Doctorale IAE+M Lorraine, décembre 2003
- [08] Guillaume CHICOISNE, «*Dialogue entre agents naturels et agents artificiel une application aux communautés virtuelles*», Institut National Polytechnique de Grenoble, Décembre 2002
- [09] «*Human Markup Language Brings Cultural Understanding To The Web*», <http://www.dclab.com/dclnews0506.asp>, visité en mai 2007
- [10] <http://jove.eng.yale.edu/pipermail/eas-info/2001/000173.html>
- [11] «*HumanML code les émotions dans les documents*», <http://www.01net.com/article/157369.html>
- [12] «*Cover Pages: Human Markup Language (HumanML)*», <http://xml.coverpages.org/humanML.html>
- [13] Pierre-Alexandre FAVIER, «*La notion d'intentionnalité dans la spécification de comportements d'agents autonomes situés en environnement virtuel*», Université de Bretagne Occidentale, Décembre 2004
- [14] Marie PHILIPPE, «*Xml schéma* », INSA lyon, 2002
- [15] Hondjack DEHAINSALA, «*Explicitation de la sémantique dans les bases de données : Base de*

données à base ontologique et le modèle OntoDB », Université de Poitiers, 2002

- [16] Alexis DROGOUL, « *Intelligence collective et systèmes multi-agents* », Université Paris 6, 2003
- [17] Denio DUARTE, « *Une méthode pour l'évolution de schémas XML préservant la validité des documents* », Université François-Rabelais – Tours, 2004
- [18] Mehdi El Jed, « *Interactions sociales en univers virtuel Modèles pour une interaction située* », Université Paul Sabatier, Septembre 2006
- [19] Jacques FERBER, « *Les systèmes multi-agents, vers une intelligence collective* », Interédition, 1995
- [20] Gwenaël GAVRAY, « *Personnalisation des sites Web: Elaboration d'une méthodologie de mise en oeuvre et application au cas DGTRE* », Université catholique de Louvain, 2001
- [21] Zahia GUESSOUM, « *Modèles et architectures d'agents et les systèmes multi-agents adaptatifs* », Université Pierre et Marie Curie, 2003
- [22] Paul GUYOT, « *Simulations multi-agents participatives : Faire interagir agents et humains pour modéliser, explorer et reproduire les comportements collectifs* », Université Paris 6, juin 2006
- [23] Isabelle JARS, « *Contribution des Sciences Sociales dans le domaine de l'Intelligence Artificielle Distribuée : ALONE, un modèle hybride d'agent apprenant* », Université Claude Bernard, novembre 2005
- [24] Marc-André LABRIE, « *Description de comportements d'agents autonomes évoluant dans des mondes virtuels* », Université Laval, janvier 2004
- [25] Michel MARCOCCIA, « *La communication écrite médiatisée par ordinateur : Faire du face à face avec de l'écrit* », Journée d'étude de l'ATALA, ENST Paris, 5 juin 2004
- [26] Michel Marcoccia, Nadia Gauducheau, « *L'analyse du rôle des smileys en production et en réception : un retour sur la question de l'oralité des écrits numériques* », GLOTTOPOL Revue de sociolinguistique en ligne, juillet 2007
- [27] Mitsuru Ishizuka et al, « *A Chat System Based on Emotion Estimation from Text and Embodied Conversational Messengers* », Active Media Technology, (AMT 2005). Proceedings of the 2005 International Conference, 2005
- [28] Philippe PASQUIER, « *Aspects Cognitifs des Dialogues Entre Agents Artificiels L'approche par la cohérence cognitive* », Université Laval Québec, Août 2005
- [29] Alexandre PAUCHET, « *Modélisation cognitive d'interactions humaines dans un cadre de planification coopérative* », Université Paris 13, septembre 2004

- [30] Catherine Pelachaud et Co., « *AGENTS CONVERSATIONNELS : Systèmes d'animation Modélisation des comportements multimodaux Applications : agents pédagogiques et agents signeurs* »
- [31] Jay Peltz and Ranjeeth Kumar Thunga, « *HumanML: the Vision* », 2005
- [32] Sylvie Pesty, Luca Bisognin, « *Agents, Langage et Emotions : un prototype d'agent émotionnel* », AGENTAL Agents et Langue, Journée ATALA, Paris, 13 mars 2004
- [33] Sylvie Pesty, Luca Bisognin, « *Vers une cognition incorporée chez les Agents Conversationnels Animés* », 2005
- [34] Nadine RICHARD, « *Langage de communication agent basé sur les engagements par l'entremise des jeux de dialogue* », École Nationale Supérieure des Télécommunications, octobre 2001
- [35] S.Labidi,W.Lajouad, « *De l'intelligence artificielle distribuée aux systèmes multi-agents* », Institut National Polytechnique de Grenoble, août 1993
- [36] Jean-Paul SANSONNET, « *Argumentaire pour une équipe pluridisciplinaire multi laboratoires (EPML) sur les Agents Conversationnels Animés (ACA)* », janvier 2004
- [37] Amandine SCHUURMAN, « *Recherche de services bioinformatiques dans une ontologie Investigation de la relation part-whole* », Facultés Universitaires Notre-Dame de la Paix, Institut d'Informatique, 2004
- [38] Gayo DIALLO, « *Une Architecture à Base d'Ontologies pour la Gestion Unifiée des Données Structurées et non Structurées* », Université Joseph Fourier – Grenoble 1, Ecole Doctorale MSTII, 2006
- [39] Vincent THOMAS, « *Proposition d'un formalisme pour la construction automatique d'interactions dans les systèmes multi-agents réactifs* », Université Henri Poincaré – Nancy 1, Novembre 2005
- [40] « *Les chatterbots : où, pourquoi, comment ?* », http://www.agentland.fr/cgi-bin/relocationFR.cgi?http://www.agentland.fr/pages/learn/chatterbox_challenge/chatterbot1.html,
- [41] « *Introduction à Xml* », <http://www.commentcamarche.net/xml/xmlintro.php3>
- [42] « *huml Use Case Description - Paper and Slide Show* », <http://lists.oasis-open.org/archives/huml/200303/msg00002.html>
- [43] « *Human Markup Language (HumanML): Humanness Content and Sharing across Perspective Shift* », http://cidoc.ics.forth.gr/symposiunabstracts/doc_files/landr/humlacrossPerspectives.html,
- [44] « *new OASIS discussion list: HumanML* », <http://lists.xml.org/archives/xml-dev/200103/msg00322.html>

- [45] «*Programmers to encode human behaviour*», <http://www.w3c.org/TR/1999/REC-html401-19991224>, New Scientist - 23 August 2001
- [46] «*Working on a Unified Code for 'LOL' or :)* », <http://www.w3c.org/TR/1999/>
- [47] «*Human Markup Language Requirements*», <http://www.oasis-open.org/committees/humanmarkup/documents/HM.Requirements.html>
- [48] «*oasis_news_08_21_01*», <http://www.w3c.org/TR/1999/REC-html401-19991224>
- [49] «*HumanMarkup Language*», <http://mailman.lbo-talk.org/2001/2001-August/016721.html>
- [50] «*Sémiotique*», <http://www.mediadico.com/dictionnaire/lecture.asp/definition>
- [51] «*HumanML -- Short Introduction / call for support*», <http://lists.oasis-open.org/archives/members/200102/msg00001.html>
- [52] «*HumanML: Emotional Metadata*», <http://www.oreillynet.com/onlamp/blog/2001/09/>
- [53] «*H-anim Facial Musculoskeletal System*», <http://lists.oasis-open.org/archives/huml/200307/msg00000.html>
- [54] «*Human Markup Language Primary Base Specification 1.0*», <http://www.oasis-open.org/committees/download.php/60/HM.Primary-Base-Spec-1.0.html>
- [55] «*prospectus*», <http://catchsafety.com/prospectus.htm>
- [56] «*Incubating New Kinds of Collaborations with Emerging XML and RDF Technologies*», <http://www.oasis-open.org/committees/download.php/4621/HumanMLinCollaborations.doc>.
- [57] Juan C. Dürsteler «*Human Markup Language*», The digital magazine of InfoVis.net, Message n° 64, October, 2001, http://www.infovis.net/infovis.ant/E-zine/num_64.htm
- [58] Amalia Ortiz Nicolás, «*Avatars as communicators of emotions*», www.physorg.com/news134820702.html
- [59] «*Initiation aux Schema XML*», <http://www.gchagnon.fr/cours/xml/schema.html>
- [60] «*HumanML code les émotions dans les documents*», <http://www.01net.com/article/157369.html>
- [61] «*HumanML Human Physical Characteristics Description SC*», http://www.oasis-open.org/committees/tc_home.php
- [62] José Rouillard, «*Hyperdialogue avec un agent animé sur le Web* », Université Joseph Fourier, 2000