HAL
open science

Actes de conversation multimodaux et emotions

Jérémy Rivière, Sylvie Pesty

To cite this version:

Jérémy Rivière, Sylvie Pesty. Actes de conversation multimodaux et emotions. WACA 2010 - 4e Workshop sur les Agents Conversationnels Animés, Nov 2010, Lille, France. pp.39-46. hal-00950904

HAL Id: hal-00950904

https://inria.hal.science/hal-00950904v1

Submitted on 24 Feb 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Actes de conversation multimodaux et émotions

J. Rivière
Jeremy.Riviere@imag.fr

S. Pesty
Sylvie.Pesty@imag.fr

Laboratoire d'Informatique de Grenoble (LIG)
Maison Jean Kuntzmann - 110 av. de la Chimie - Domaine Universitaire de Saint-Martin-d'Hères
38041 Grenoble cedex 9 – FRANCE

Résumé :
Cet article se place dans le cadre de dialogues entre un humain et un Agent Conversationnel Animé (ACA). Nous discutons ici de la nécessité de mettre en place un langage spécifique aux interactions entre l'homme et l'ACA, et du rôle central que jouent les émotions au sein de ces interactions. Dans ce but, nous proposons une bibliothèque d'*actes de conversations multimodaux* basée sur les actes de discours et l'expression des émotions liées aux actes. Nous nous intéressons plus particulièrement aux émotions dites *complexes*, résultant d'un raisonnement de l'agent. Ce langage de conversation multimodal sera le moyen pour l'ACA d'exprimer son intention communicative de façon multimodale : verbale et non verbale.

Mots-clés : Interaction, Dialogue, Actes de discours, Agents Conversationnels Animés, Emotions

Abstract:
This article falls within the realm of dialog between human and Embodied Conversational Agent(ECA). We discuss here about the need to have a language which is peculiar to human-ECA interaction and based on the main role of emotion in human communication. In order to define this language, a multimodal conversation acts library is proposed, based notably on speech acts and the expression of emotions linked to these acts. This study focuses in particular on complex emotions, arising from agents' reasoning. This multimodal conversation language will be used in future to express the communicative intention of an Embodied Conversational Agent able to communicate through all modalities, verbal and non-verbal.

Keywords: Interaction, Dialog, Speech act, Embodied Conversational Agent, Emotion

1 Introduction

De nos jours, l'augmentation des interactions entre l'humain et la machine et au sein même de la machine entre agents autonomes, fait du dialogue un verrou scientifique majeur. De ce fait, certaines recherches dans le domaine des systèmes de dialogue humain-agent (et plus généralement de dialogue homme-machine) se sont orientées vers l'annotation de dialogue ([8], [19]) pour la compréhension du discours, la reconnaissance de structures et la génération de dialogue, tandis que d'autres recherches s'orientent vers la communication en langage naturel. Avec l'arrivée des agents virtuels et dans le contexte du dialogue humain-agent virtuel, où les aspects expressifs et multimodaux sont premiers, il s'agit maintenant d'augmenter les capacités conversationnelles des agents en cherchant à étroitement relier les aspects verbaux aux aspects non-verbaux (expressions faciales, gestes, regard, ...). Notre problématique se concentre sur le rôle central de l'émotion dans la communication et plus généralement dans les dialogues, ce rôle ayant été montré dans de nombreux travaux tant en psychologie qu'en informatique ([3], [20], [21]). Pour améliorer ces interactions tant en termes de crédibilité que d'efficacité, c'est-à-dire de succès de la communication, et rendre les agents virtuels plus crédibles et convaincants pour l'humain, de récents travaux ont doté les agents de capacités d'exprimer les émotions, pour en faire des Agents Conversationnels Animés (ACA ou ECA- Embodied Conversational Agent en anglais) ([22], [9]).

L'objectif de nos travaux est la mise en place d'un langage qui dépasse le simple échange d'information (Inform - Request) dans les dialogues entre agents virtuels et humains, en incluant la dimension émotionnelle. La proposition ici présentée consiste à concevoir pour les agents virtuels une bibliothèque d'*actes de conversation multimodaux*, basés d'une part sur les actes de conversation d'Alexandra Berger ([5]), dans la lignée de la théorie des actes de discours ([26]), et basés d'autre part sur les émotions dans l'interaction et la multimodalité de leur expression. Cette bibliothèque doit fournir à l'agent virtuel les émotions liées à un acte, et plus particulièrement les émotions que nous appellerons *complexes*, définie notamment sur le raisonnement contrefactuel détaillé en 4.1. Ainsi, lorsque l'agent choisit un acte de conversation, il formulera non seulement un énoncé mais exprimera également, de manière multimodale, l'ensemble des émotions intimement liées à cet acte. Ce *langage de conversation multimodal* s'intègrera par la suite dans l'ACA Greta, développé au LTCI[1] à TélécomParisTech par Catherine Pelachaud ([22]), et dont nous détaillons l'architecture en 5.1.

[1] Laboratoire Traitement et Communication de l'Information

La première partie de cet article présente les principaux moyens d'expression des émotions utilisés par les ACA. Nous verrons que, dans la majorité des travaux, la modalité langagière ne fait pas le lien entre émotions et dialogue : les émotions de l'ACA sont principalement exprimées par la prosodie, mais sont indépendantes de l'énoncé. Dans le but de faire le lien entre dialogue et émotions, nous nous intéressons dans la deuxième partie à deux des principales méthodes d'annotation de dialogue et d'émotions entre humain et agent virtuel. En effet, ces travaux nous renseignent sur les émotions d'une part et sur les actes d'autre part qui interviennent dans les interactions ACA-humains. Dans la troisième partie, nous rappelons succinctement les actes de conversation entre agents développés par A. Berger sur lesquels nous basons notre proposition, puis nous détaillons nos *actes de conversation multimodaux* en montrant le lien entre les actes de langage et les *émotions complexes* dont nous donnons la définition. Enfin, nous montrons une première implémentation de plusieurs actes de conversations multimodaux avec l'ACA Greta et les principales perspectives de travail.

2 Expression d'émotions chez les ACA

La plupart des travaux sur l'expression des émotions par un ACA se focalisent sur les modalités faciale et gestuelle ([9], [4]). L'ACA Greta ([22]) que nous utilisons en est un bon exemple. Greta se base sur le système des FACs[2] de Ekman et Friesen ([10]) pour définir les expressions faciales, et permettre de contrôler l'animation indépendamment de la géométrie du visage. Les émotions ainsi exprimées sont les émotions "de base" de Ekman([11]) ou des émotions dérivées, facilement reconnaissables et universelles. Le regard et les mouvements de la tête sont également pris en compte. Les gestes de Greta, et notamment ceux des bras et des mains, sont synchronisés à la parole et définis en quatre phases : la préparation, l'apogée (le *stroke*), la suspension puis la rétraction du mouvement. Le stroke doit spécifiquement correspondre à l'intensité maximale de l'intonation. Une des difficultés à ce niveau consiste à lier différents mouvements entre eux d'une manière fluide et naturelle.

On peut distinguer deux niveaux dans la modalité langagière qui nous intéresse : l'énoncé en lui-même et la prosodie, qui est l'intonation de la voix lors de l'énoncé suivant une ou plusieurs émotions. La majorité des travaux sur les ACA se concentrent sur la prosodie, en ajustant l'intensité de la voix suivant l'énoncé d'une part et les émotions d'autre part. Les émotions et l'énoncé sont ainsi indépendants l'un de l'autre, ce qui ne garantie pas la cohérence d'un énoncé par rapport à une émotion : par exemple, on pourrait avoir un agent qui affirme être très heureux tout en exprimant une émotion de grande tristesse.

3 Annotation d'émotion et annotation de dialogue

Cette partie présente deux travaux majeurs dans les domaines d'annotation d'émotions et d'annotation de dialogue à partir d'actes de discours. Ces deux approches nous éclairent respectivement sur le rôle des émotions et celui des actes dans les interactions, ce qui met en évidence l'absence de lien entre actes et émotions.

Le groupe de travail "Emotion Markup Language" (EmotionML [2]) du W3C[3] propose un langage adapté à l'annotation d'émotions, la reconnaissance automatique d'émotions au travers du comportement de l'utilisateur, et la génération d'émotions. Ce langage est basé sur des balises xml permettant de classer les émotions suivant les théories de l'évaluation ([23]), les théories dimensionnelles ([17]) ou encore la théorie de Frijda, qui définit une émotion suivant une tendance à l'action ([12]). Ainsi, l'émotion de peur peut se traduire par exemple :

- En théorie de l'évaluation :
  "*<appraisal name="novelty" value="0.9"/> <appraisal name="intrinsic-pleasantness" value="0.1"/>*"
- Dans la théorie de Frijda :
  "*<action-tendency name="avoidance" value="0.8"/>*

De plus, à ces émotions peuvent être liées des modalités d'expression ("*<...modality="face voice"> <category name="satisfaction"/>*"), et une intensité ("*<intensity value="0.1" confidence="0.8"/>*), à laquelle est attribuée ici une certaine confiance. Ces différentes balises offrent ainsi un langage générique, permettant d'annoter les émotions suivant les principales

[2] Facial Action Coding System

[3] World Wide Web Consortium - http ://www.w3.org/

théories et de spécifier les modalités d'expression de ces émotions.

Cependant, dans le cas de la génération d'émotions, l'énoncé et l'émotion sont indépendants : l'extrait de code figure 1 illustre cet absence de lien avec un exemple de génération d'émotions. "Do you need help ?" est le texte à énoncer avec une émotion de doute dont la catégorie appartient à l'ensemble défini par "everydayEmotions". Cette émotion est à exprimer avec une intensité de 0.4, par le ton de la voix. Dans cet exemple, on peut imaginer que le texte pourrait être remplacé par une phrase qui serait incohérent de prononcer avec une émotion de doute : par exemple, "there is no doubt in my mind that I'm right".

```
<speak>
<s>
<emo:emotion>
<emo:category set="everydayEmotions"
                name="doubt"/>

<emo:intensity value="0.4"/>
</emo:emotion>

Do you need help?
</s>
</speak>
```

FIG. 1 – Exemple de génération d'émotion en EmotionML

L'identification des actes de discours reste une des méthodes les plus utilisées dans l'annotation et l'analyse de dialogue. Plusieurs travaux ([7], AMI Project[4]) s'inspirent de la théorie des actes de discours de Searle et de la taxonomie qu'il propose pour l'analyse du dialogue. Cette taxonomie est formée de cinq classes d'actes : les actes assertifs, directifs, engageants, déclaratoires et expressifs. Elle varie selon les travaux, mais une correspondance peut généralement s'établir entre les taxonomies. Dans les travaux basés sur cette taxonomie, l'identification et la classification des actes dans un énoncé s'effectuent de manière unidimensionnelle : à un énoncé correspond une seule annotation, un seul acte. Ce type d'annotation a été remis en cause notamment par Allen et Core ([8]), qui ont proposé une taxonomie multidimensionnelle, DAMSL. Un énoncé peut ainsi être lié à plusieurs actes, qui appartiennent à des fonctions communicatives "forward" ("Assert", "Offer", "Commit",...) ou "backward" ("Accept", "Signal-Understanding", "Agreement",...). Harry Bunt montre quant à lui ([19]) que le fait d'avoir plusieurs dimensions facilite l'annotation et permet une analyse plus fine du discours. Ces différentes annotations indiquent le rôle des émotions dans la structure du dialogue et la gestion du sujet de la conversation, mais ne permettent pas d'identifier les émotions associées aux actes de discours.

Dans l'optique de mettre en place une base pour un langage émotionnel, et en lien avec la théorie des actes de discours de Searle ([24]), il est nécessaire de faire un lien entre les actes de discours, qui sont les unités premières de signification du langage, et les émotions.

4 Proposition : *Actes de Conversation Multimodaux*

Notre proposition s'appuie sur les actes de discours, et se situe à la croisée de l'expression d'émotion et de la capacité langagière de l'agent. En effet, dans les travaux ci-dessus il n'y a pas de lien entre ce que dit l'agent et les émotions qui sont exprimées : le lien entre énoncé et émotion n'est pas clairement établi ou tout du moins il n'est pas formalisé. Nous cherchons à éclairer ce lien et ainsi donner la possibilité à l'agent, lors de l'usage d'un acte de conversation d'exprimer également les émotions intrinsèques à cet acte. Nous entendons par "émotions intrinsèques" les émotions inhérentes à l'acte, indépendamment de tout facteur extérieur. Nous verrons par exemple que l'acte *S'excuser* "porte" l'émotion de regret ou encore que l'acte *Se plaindre* "porte" l'émotion de déception. Ainsi, nous tentons de répondre aux questions suivantes : comment va s'exprimer multimodalement un agent qui souhaite promettre quelque chose ? Comment peut-il être convaincant ? Quelles sont les conditions requises sur son état émotionnel pour qu'il soit sincère ?

La première partie de cette section introduit les actes de conversation définis par A. Berger suivant la théorie des actes de discours de Vanderveken. Notre proposition, présentée en deuxième partie, s'appuie sur ces actes pour identifier les composantes qui portent les émotions. Des premiers *actes de conversation multimodaux* sont alors définis à partir de la formalisation des *émotions complexes* que nous avons développée. Nous nous intéressons aux *émotions complexes* car, contrairement aux émotions primaires, qui sont réactives et qui se

[4] http ://corpus.amiproject.org/

traduisent le plus souvent par des expressions faciales prototypiques, les *émotions complexes* sont le résultat d'un raisonnement réfléchi et s'expriment le plus souvent selon plusieurs modalités et principalement par la modalité langagière qui nous intéresse ici.

4.1 Définition et formalisation des *émotions complexes*

En ce qui nous concerne, nous nous intéressons à une catégorie bien particulière des émotions : celles basées d'une part sur le raisonnement contrefactuel de l'agent, et d'autre part sur son raisonnement à propos de sa responsabilité, de ses capacités et celles des autres agents, et des normes sociales. Le raisonnement contrefactuel consiste à comparer la réalité et une vue imaginaire de ce qui aurait pu se produire ([14]), et induit des émotions telles que le regret ou la culpabilité. Ce choix s'explique par l'importance de ces émotions dans le raisonnement de l'agent, comme le montrent les théories de l'évaluation ([23], [25]).

La formalisation proposée a été réalisée en coopération avec l'équipe LiLaC de l'IRIT[5], dans le cadre du projet ANR CECIL[6]. Elle se base sur trois opérateurs : l'opérateur $Goal_i\varphi$ (l'agent *i* a pour but que φ soit vrai), l'opérateur $Ideal_i\varphi$ (idéalement pour *i*, φ devrait être vrai) et l'opérateur $Resp_i\varphi$ (l'agent *i* est responsable de φ) ([6], [15]). Un quatrième opérateur, $Bel_i\varphi$, représente la croyance qu'a l'agent *i* en φ. Nous définissons explicitement huit *émotions complexes* : le regret, la déception, la culpabilité, le reproche, la satisfaction morale, l'admiration, la réjouissance et la reconnaissance ou gratitude. Par exemple, l'émotion de culpabilité peut correspondre à un regard porté par l'agent sur ce qu'il a fait et ce qu'il aurait pu faire à un instant donné. De même, le reproche peut être dû à la comparaison des capacités d'un agent et de ce qu'il a fait par un autre agent. Le tableau 1 montre la formalisation de ces *émotions complexes* pour un agent *i*, concernant une situation φ.

Comme nous le verrons par la suite, nous définissons les *actes de conversation multimodaux* à partir de cette formalisation des *émotions complexes*. Cette formalisation apporte la preuve théorique de notre proposition et permettra par la suite de servir de base au raisonnement de l'agent : à partir de ses connaissances, son environnement et des événements passés , un agent BDI[7] ou EBDI[8] ([18], [1]) peut déduire l'état émotionnel dans lequel il se trouve et communiquer de façon émotionnelle.

4.2 Actes de Conversation

Les travaux d'A. Berger ([5]) ont été menés dans le but de dépasser le simple échange d'information entre agents. Ils ont permis de mettre en place un *Langage de Conversation*, basé sur la théorie des actes de discours de Searle ([24]) et composé de trente-deux actes de conversation définis formellement. Ces actes ont été sélectionnés à partir des définitions des verbes performatifs données par Vanderveken ([26]) pour leur pertinence dans l'utilisation qu'un agent peut en faire, dans des situations comme l'assistance à l'utilisateur, la négociation, la coopération etc... Ces actes appartiennent aux cinq classes d'actes de discours définies par la théorie, qui correspondent aux cinq différentes façons fondamentales d'utiliser le langage.

Chacun de ces actes est de la forme *F(P)*, où P est la proposition énoncée et F la *force illocutoire* qui détermine dans quel but l'énoncé est formulé ([26]). La force illocutoire F est constituée de six composantes qui expriment chacune un aspect différent de l'acte. Ces composantes sont les conditions à respecter pour que l'acte soit performé avec succès et satisfaction, suivant *le but illocutoire* de l'acte qui représente l'usage que l'on fait de l'acte, l'état mental du locuteur s'il est sincère (*conditions de sincérité*), le contexte d'énonciation (*conditions préparatoires, conditions sur le contenu*), la manière dont le but doit être atteint (*mode d'atteinte*) et le *degré de puissance* de l'acte. De cette façon, les actes de conversation basiques tels que *Informer* et *Demander* ont été définis, mais aussi *Conseiller*, *Approuver* ou encore l'acte *S'excuser* dont le but illocutoire est expressif.

4.3 Lien avec les émotions et définition des *Actes de Conversation Multimodaux*

Notre travail a permis d'identifier le lien qui existe entre les actes de conversation détaillés ci-dessus et les émotions. Ce lien est obtenu

[5] Institut de Recherche en Informatique de Toulouse - http ://www.irit.fr/-Equipe-LILaC-

[6] projet CECIL - Complex Emotions in Communication, Interaction, and Language - ANR-08-CORD-005

[7] Belief, Desire, Intention

[8] Emotional BDI

$\wedge$	$Goal_i\varphi$	$Goal_i\neg\varphi$	$Ideal_i\varphi$	$Ideal_i\neg\varphi$
$Bel_iResp_i\varphi$	$Rejouissance_i\varphi$	$Regret_i\varphi$	$SatisfactionMorale_i\varphi$	$Culpabilite_i\varphi$
$Bel_iResp_j\varphi$	$Reconnaissance_{i,j}\varphi$	$Deception_{i,j}\varphi$	$Adhesion_{i,j}\varphi$	$Reproche_{i,j}\varphi$

TAB. 1 – Formalisation des émotions *complexes* pour un agent *i* vis-à-vis de sa responsabilité (ou la responsabilité d'un agent *j*) par rapport à une situation φ.

à partir de l'étude de la théorie des actes de discours et consiste en trois composantes de la force illocutoire F de ces actes. Ainsi, si l'on considère par exemple l'acte de conversation *S'excuser*, l'émotion du regret est exprimée par **la condition de sincérité** de l'acte, qui est la condition minimale pour que le locuteur soit sincère. D'après Vanderveken ([26]) : *S'excuser, c'est exprimer ses regrets à un allocutaire (condition de sincérité) à propos de l'état de chose représenté par p en présupposant (condition préparatoire) que l'on est responsable de cet état de chose et qu'il est mauvais pour l'allocutaire.* Cet acte est un cas particulier de l'expression de l'*émotion complexe* du regret : vis-à-vis d'autrui, il s'agit d'exprimer du regret *de ne pas avoir satisfait le but d'autrui p*. On peut également exprimer du regret vis-à-vis de soi-même par rapport à la situation présente p.

Nous nous basons sur la formalisation logique des *émotions complexes* présentée précédemment pour exprimer l'émotion portée par la condition de sincérité. Ainsi, la condition de sincérité de l'acte *S'excuser* se traduit formellement par :

$$Regret_i(Goal_j\varphi \wedge \neg\varphi)$$

i se croit responsable du fait que le but de *j* ne soit pas atteint.

De manière plus générale, les conditions de sincérité portent les émotions à exprimer lors de l'acte. Ces conditions forment un lien permanent entre émotions et actes dans le sens où l'expression de l'émotion lors de l'acte est nécessaire et suffisante pour que le locuteur soit sincère. Nous définissons les *actes de conversation multimodaux* que nous appelons "neutres", car exprimant une seule *émotion complexe* avec un degré neutre, contrairement à des actes qui pourraient exprimer une composition de ces émotions ou des actes avec un *degré de puissance* plus élevé, comme détaillé ci-dessous. L'équipe LiLaC a défini l'opérateur $Assert_{i,j}\varphi$, représentant l'usage assertif de la théorie des actes de discours, car exprimer un état mental peut se traduire comme informer quelqu'un de son état mental. Cet opérateur permet donc de traduire formellement : "i informe j que φ" ou, dans la logique des annonces ([13], symbolisée par **!**), "le but de i est que j croit φ".

$$Assert_{i,j}\varphi \stackrel{déf}{=} Goal_iBel_j\varphi\,!$$

L'*acte de conversation multimodal S'excuser* s'écrit donc formellement :

$$S'excuser_{i,j}\varphi \stackrel{déf}{=} Assert_{i,j}(Regret_i(Goal_j\varphi \wedge \neg\varphi))$$

Nous proposons une taxonomie des *actes de conversation multimodaux* "neutres" suivant les différentes *émotions complexes* qu'ils expriment. Chaque acte est défini formellement suivant l'opérateur $Assert_{i,j}$ et l'*émotion complexe* que l'agent *i* exprime à un autre agent *j* : comme ces actes sont des actes de communication, il y a toujours un locuteur et un allocutaire. Le tableau 2 présente les huit actes ainsi définis, dont la seconde forme de l'expression de l'émotion du regret (regretter vis-à-vis de soi-même).

L'émotion à exprimer lors de l'acte est donc incluse dans la condition de sincérité, mais celle-ci n'est pas la seule composante à porter les émotions dans les actes de conversation. En effet, une seconde composante permet de déterminer la force avec laquelle s'exprime une émotion, par rapport à un acte plus modéré : c'est la composante de **degré de puissance**. En général, le degré de puissance d'un acte s'exprime dans un énoncé en utilisant des adverbes tels que sincèrement, franchement... Il forme des liens entre différents actes qui expriment une même émotion : par exemple, récriminer est plus fort que protester qui lui-même est plus fort que désapprouver ([26]). Le degré de puissance peut donc nous permettre d'enrichir notre bibliothèque *d'actes de conversation multimodaux* : on peut imaginer qu'une émotion de reconnaissance avec une forte intensité amènera

$Assert_{i,j}(...)$	Acte de conversation multimodal
$Goal_i\varphi \wedge Bel_i Resp_i\varphi$	$\stackrel{déf}{=} Assert_{i,j}(Rejouissance_i\varphi) \stackrel{déf}{=} \mathbf{SeFeliciter}_{i,j}\varphi$
$Goal_i\varphi \wedge Bel_i Resp_j\varphi$	$\stackrel{déf}{=} Assert_{i,j}(Reconnaissance_{i,j}\varphi) \stackrel{déf}{=} \mathbf{Remercier}_{i,j}\varphi$
$Goal_i\neg\varphi \wedge Bel_i Resp_i\varphi$	$\stackrel{déf}{=} Assert_{i,j}(Regret_i\varphi) \stackrel{déf}{=} \mathbf{Regretter}_{i,j}\varphi$
$Goal_i\neg\varphi \wedge Bel_i Resp_j\varphi$	$\stackrel{déf}{=} Assert_{i,j}(Deception_{i,j}\varphi) \stackrel{déf}{=} \mathbf{SePlaindre}_{i,j}\varphi$
$Ideal_i\varphi \wedge Bel_i Resp_i\varphi$	$\stackrel{déf}{=} Assert_{i,j}(SatisfactionMorale_i\varphi) \stackrel{déf}{=} \mathbf{Reconnaitre}_{i,j}\varphi$
$Ideal_i\varphi \wedge Bel_i Resp_j\varphi$	$\stackrel{déf}{=} Assert_{i,j}(Adhesion_{i,j}\varphi) \stackrel{déf}{=} \mathbf{Adherer}_{i,j}\varphi$
$Ideal_i\neg\varphi \wedge Bel_i Resp_i\varphi$	$\stackrel{déf}{=} Assert_{i,j}(Culpabilite_i\varphi) \stackrel{déf}{=} \mathbf{Culpabiliser}_{i,j}\varphi$
$Ideal_i\neg\varphi \wedge Bel_i Resp_j\varphi$	$\stackrel{déf}{=} Assert_{i,j}(Reproche_{i,j}\varphi) \stackrel{déf}{=} \mathbf{Reprocher}_{i,j}\varphi$

TAB. 2 – Actes de conversation multimodaux définis avec l'opérateur Assert et exprimant une émotion complexe

l'agent à utiliser un acte tel que complimenter, vanter ou louer plutôt que remercier.

Une dernière composante, **le mode d'accomplissement**, détermine comment le but illocutoire doit être accompli et les modalités d'expression des émotions contenues dans l'acte. En effet, si l'on considère l'exemple du verbe performatif Protester définis par Vanderveken ([26]) : *Protester, c'est exprimer d'une façon formelle (mode d'accomplissement) sa désapprobation.... D'une façon formelle* peut se traduire par le langage (tournures élaborées, rigueur,...) mais également par la gestuelle, la posture ou l'expression du visage. Ainsi, les modes d'accomplissement permettent de déterminer les modalités d'expression des différentes émotions, que ce soit la modalité langagière ou d'autres modalités. On peut imaginer que cette composante pourra être par la suite influencée par exemple par la personnalité de l'ACA, le type de relation qu'il entretient avec l'utilisateur ou simplement par les modalités qu'il peut exprimer. Un agent peut donc, à la sortie de son raisonnement, choisir l'acte qui correspond à sa volonté d'action : à cet acte sont liées les émotions et les modalités d'expression de celles-ci, que l'agent va alors exprimer de façon multimodale.

Nous proposons ici neuf premiers *actes de conversation multimodaux* "neutres" exprimant les *émotions complexes*, dont la forme particulière de l'expression de l'émotion du regret qui induit deux actes, *S'excuser* et *Regretter*. Ces actes constituent les premiers actes de notre bibliothèque. De prochains actes, construits à partir des degrés de puissance et des compositions d'émotions, restent à définir. L'étape suivante sera la formalisation des autres classes d'actes de conversation d'A. Berger, c'est-à-dire les actes engageants tels que promettre et accepter, les actes directifs tels que exiger et conseiller, etc... La différence entre ces actes et ceux que nous avons définis est que ces derniers ont pour but d'exprimer l'*émotion complexe* liée, alors que pour les autres classes l'émotion pourra être liée mais ne sera pas le but de l'acte : on peut imaginer par exemple que des actes comme exiger ou promettre ne sont pas détachés de toute émotion pour être convaincants. Les quatre *actes de conversation multimodaux S'excuser*, *Reprocher*, *Se plaindre* et *Culpabiliser* ont été expérimentés avec l'agent Greta, dans le but d'identifier les limites et les prochaines perspectives de travail concernant la mise en place de notre bibliothèque dans l'architecture de Greta, laquelle est détaillée dans la section suivante.

5 Première implémentation avec Greta et perspectives de travail

Cette partie présente l'architecture de Greta et l'implémentation de l'*acte de conversation multimodal S'excuser*. Cette première implémentation ne prend pas en compte la formalisation proposée qui pourra être utilisée par la suite comme base de raisonnement de l'ACA, mais a pour but de déterminer les modifications à effectuer sur l'architecture de Greta de manière à y ajouter notre bibliothèque d'*actes de conversation multimodaux*.

5.1 Architecture de Greta et implémentation

L'ACA Greta est conforme au standard SAIBA ([27]) qui est une initiative de recherche internationale dont le but est de définir une architecture logicielle commune pour la génération du comportement d'agents virtuels. L'architecture SAIBA est composée de trois modules. Le premier module *Intent Planning* décide de l'intention communicative de l'agent à partir de ses buts, de ses croyances et de son état émotionnel. Cette intention communicative est encodée en langage FML[9] ([16]) qui est un langage de balisage pour être interprété par le module *Behavior Planning*. Ce module gère un certain nombre de modalités de communication dans le but d'exprimer l'intention communicative de l'agent : la parole, l'expression du visage et les gestes sont ainsi calculés. Ces différentes modalités traduisent le comportement verbal et non-verbal de l'agent, qui, une fois ces modalités encodées en BML[10] ([27]), est réalisé par le dernier module *Behavior Realization*.

Notre travail consiste à intervenir au niveau des deux premiers modules. Lors de la planification de l'intention communicative, l'agent choisit, dans la bibliothèque d'*actes de conversation multimodaux*, l'acte qui correspond le mieux à son intention de communiquer. A cet acte est associé un ensemble d'émotions qui seront exprimées de manière concomitante. Nous avons implémenté les quatre *actes de conversation multimodaux* que nous avons définis. L'extrait de code en langage FML figure 2 illustre ici comment a été implémenté l'acte *S'excuser* et l'*émotion complexe* du regret associée.

```
      <tm id="tm1"/>
       I apologise
      <tm id="tm2"/>
      for my absence
      <tm id="tm3"/>
<pitchaccent id="xpa1" type="Lstar" start="s1:tm1"
end="s1:tm2" level="high" importance="1" />
<performative id="p1" type="announce"
start="s1:tm1" end="s1:tm3"/>
<emotion id="e1" type="remorse"
start="s1:tm1" end="s1:tm3"
regulation="felt" importance="1" intensity="1" />
```

FIG. 2 – Texte à énoncer divisé par des marqueurs de temps

Ces informations sont ensuite traduites par le module *Behavior Planning* en BML, pour la réalisation du comportement effectuée par le dernier module. Ainsi, l'émotion *remorse* correspond à une certaine expression du visage, à une certaine posture du corps, et est exprimée durant l'énonciation de l'acte.

5.2 Conclusions et perspectives de travail

Cette implémentation sur l'ACA soulève un certain nombre de problèmes qui illustrent une partie des modifications à réaliser dans l'architecture de Greta. Dans l'exemple de l'acte *S'excuser*, la séparation de l'énoncé en plusieurs parties par les marqueurs de temps est à faire de façon manuelle, ainsi que la spécification des parties à accentuer. L'*émotion complexe* du regret n'est pas intrinsèque à l'acte et doit également être explicitée manuellement. Par ailleurs, un certain nombres d'approximations sont faites pour les besoins de l'implémentation (émotion de *remords*, performatif *announce*). En l'état actuel donc, le langage FML ne permet pas d'exprimer toutes les informations définies dans les *actes de conversation multimodaux*, comme celles portées par le mode d'accomplissement et le degré de puissance. La création de nouvelles balises dans le langage FML est donc nécessaire pour transmettre ces nouvelles informations au module *Behavior Planning*, et pouvoir choisir directement pour l'intention communicative un *acte de conversation multimodal* de la bibliothèque, ce qui permettra d'obtenir implicitement les émotions liées à cet acte. De ce fait, des modifications sur le module *Behavior Planning* sont à réaliser, dans le but de le rendre capable d'interpréter les nouvelles balises qu'il reçoit en FML, et traduire les émotions dans plusieurs modalités.

Le travail décrit dans cet article n'est que le premier pas vers l'implémentation d'une bibliothèque la plus complète possible d'*actes de conversation multimodaux*, générique et ouverte. Nous avons actuellement défini neuf *actes de conversation multimodaux* présentés en 4.3 et implémenté quatre d'entre eux avec Greta, dans le but d'identifier les modifications à faire dans son architecture. Cette bibliothèque devra rendre possible par la suite l'expression de l'ensemble des trente-deux actes du *langage de conversation*. Elle sera ensuite implémentée dans l'architecture SAIBA, ce qui permettra l'évaluation de cette bibliothèque auprès d'utilisateurs grâce à des scenarii de tests. Cette évaluation devra vérifier d'un côté la crédibilité de l'agent en s'appuyant sur des critères de re-

[9] Function Markup Language
[10] Behavior Markup Language

connaissance d'émotions, et de l'autre l'amélioration de l'interaction entre agents humains et virtuels, c'est-à-dire la performance de l'ACA, et pourra permettre de valider notre hypothèse théorique.

Remerciements

Ce travail a bénéficié d'une aide de l'Agence Nationale de la Recherche portant la référence ANR-08-CORD-005.

Références

[1] Carole Adam. *The emotions : from psychological theories to logical formalization and implementation in a BDI agent*. Thèse de doctorat, Institut National Polytechnique de Toulouse, Toulouse, France, juillet 2007.

[2] Paolo Baggia and al. Emotion markup language (emotionml) 1.0. 2009.

[3] Joseph Bates. The role of emotion in believable agents. *Commun. ACM*, 37(7) :122–125, 1994.

[4] Christian Becker and Ipke Wachsmuth. Simulating the emotion dynamics of a multimodal conversational agent. In *In Proceedings Tutorial and Research Workshop on Affective Dialogue Systems (ADS-04), LNAI 3068*, pages 154–165. Springer, 2004.

[5] Alexandra Berger. *La communication entre agents de communautés mixtes : Un Langage de Conversation Expressif pour agents artificiels*. Thèse de doctorat, Institut National de Polytechnique de Grenoble et Université du Québec à Trois-Rivières, décembre 2006.

[6] Jean-François Bonnefon, Dominique Longin, and Manh Hung Nguyen. A Logical Framework for Trust-Related Emotions. *Electronic Communications of the EASST, Formal Methods for Interactive Systems 2009*, 22, 2009.

[7] Jean Caelen. Stratégies de dialogue. In Chaib-Draa Brahim Philippe Mathieu Herzig, Andreas, editor, *Actes des Secondes Journées Francophones des Modèles Formels de l'Interaction*, page 29/39. Cépaduès, 2003.

[8] Mark G. Core and James F. Allen. Coding dialogs with the DAMSL annotation scheme. In *Proc. of the Working Notes of the AAAI Fall Symposium on Communicative Action in Humans and Machines*, Cambridge, MA, 1997.

[9] Matthieu Courgeon, Jean-Claude Martin, and Christian Jacquemin. Marc : a multimodal affective and reactive character. In *Proceedings of The 1st workshop on Affective Interaction in Natural Environments (AFFINE)*, 2008.

[10] P. Ekman and W.V. Friesen. The facial action coding system : A technique for the measurement of facial movement. In *Consulting Psychologists*, 1978.

[11] Paul Ekman. An argument for basic emotions. *Cognition & Emotion*, 6(3) :169–200, 1992.

[12] Nico H. Frijda. *The Emotions*. Cambridge University Press, 1986.

[13] Nadine Guiraud, Andreas Herzig, and Emiliano Lorini. Speech acts as announcements (regular paper). In *Dagstuhl Seminar on Information processing, rational belief change and social interaction , Dagstuhl, Germany*, number 1862-4405. Science Publications, novembre 2009.

[14] Daniel Kahneman and Dale T. Miller. Norm theory : Comparing reality to its alternatives. *Psychological Review*, 93(2) :136–152, 1986.

[15] Emiliano Lorini and François Schwarzentruber. A logic for reasoning about counterfactual emotions. In C. Boutilier, editor, *International Joint Conference on Artificial Intelligence (IJCAI), Pasadena (USA)*, pages 867–872. AAAI Press, 2009.

[16] M. Mancini and C. Pelachaud. The FML - APML language. In *The First FML workshop, AAMAS'08, Estoril, Portugal*, 2008.

[17] Albert Mehrabian. Pleasure-arousal-dominance : A general framework for describing and measuring individual differences in temperament. *Current Psychology*, 14(4) :261–292, December 1996.

[18] David Pereira, Eugénio Oliveira, and Nelma Moreira. Formal modelling of emotions in bdi agents. pages 62–81, 2008.

[19] Volha Petukhova and Harry Bunt. A multidimensional approach to multimodal dialogue act annotation. In *Proceedings Seventh International Workshop on Computational Semantics (IWCS-7)*, pages 142–153, 2007.

[20] J. Piaget. *Les emotions*, pages 75–95. Delachaux et Niestle, Neuchâtel-Paris, 1989.

[21] Rosalind W. Picard. *Affective Computing*. The MIT Press, Cambridge, Massachusetts, 1997.

[22] I. Poggi and al. Greta : A believable embodied conversational agent, 2005.

[23] Klaus R. Scherer. *Emotion*, pages 151–195. Wiley-Blackwell, 2001.

[24] John R. Searle. *Speech acts : an essay in the philosophy of language / John R. Searle*. Cambridge University Press, London :, 1969.

[25] Aaron Sloman. Beyond shallow models of emotion. In *Cognitive Processing : International Quarterly of Cognitive Science*, pages 177–198, 2001.

[26] Daniel Vanderveken. *Les actes de discours*. Editions Pierre Mardaga, Liège, Bruxelles, 1988.

[27] Hannes Vilhjalmsson and al. The behaviour markup language : recent developments and challenges. *Lecture notes in artificial intelligence*, 4722 :99–11, 2007.