



HAL
open science

Définition d'un agent pédagogique par analyse et annotation de corpus vidéo

Mohamedade Farouk, Jean-Hugues Réty, Nelly Bensimon, Jean-Claude Martin

► **To cite this version:**

Mohamedade Farouk, Jean-Hugues Réty, Nelly Bensimon, Jean-Claude Martin. Définition d'un agent pédagogique par analyse et annotation de corpus vidéo. Jun 2007. hal-00161411

HAL Id: hal-00161411

<https://hal.science/hal-00161411>

Submitted on 10 Jul 2007

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.

Définition d'un agent pédagogique par analyse et annotation de corpus vidéo

Mohamedade Farouk*, **Jean-Hugues Réty***, **Nelly Bensimon***,
Jean-Claude Martin*,**

* *LINC – Université Paris 8*

IUT de Montreuil

140, rue de la Nouvelle France

93100 Montreuil

{m.farouk, jh.rety, n.bensimon, jc.martin}@iut.univ-paris8.fr

***LIMSI-CNRS*

B.P. 133

91403 ORSAY CEDEX

martin@limsi.fr

RÉSUMÉ. Les agents pédagogiques se situent à la croisée de deux domaines : les EIAH et les IHM. Nous décrivons dans cet article l'utilisation d'une approche par annotation et analyse statistique de corpus vidéo pour étudier le comportement d'enseignants en situation réelle d'enseignement. L'objectif à terme de ce travail est d'élaborer un modèle de comportement permettant de piloter un agent pédagogique. Nous décrivons la constitution du corpus vidéo, le schéma d'annotation que nous avons défini, ainsi que les premières mesures statistiques réalisées.

MOTS-CLÉS : IHM, Agents Pédagogiques, Annotation des corpus vidéo

1. Introduction et objectifs de recherche

Un agent conversationnel animé [CASSELL 2000] est une interface homme-machine utilisant un personnage animé affiché à l'écran doté de capacités communicatives inspirées de la communication humaine. On distingue généralement deux niveaux de spécification : celui des fonctions communicatives (par exemple gestion des tours de parole) et celui des signes multimodaux (par exemple regard) [POGGI 2003].

Les agents conversationnels appliqués à l'enseignement (ou agents pédagogiques) résultent de la jonction entre le domaine des Interactions Homme-Machine (IHM) et celui des Environnements Informatiques d'Apprentissage Humain (EIAH). Les recherches en IHM apportent des principes généraux et des outils informatiques pour la communication verbale, non verbale et émotive via des personnages virtuels. Les recherches en EIAH doivent analyser la spécificité du contexte lié à l'apprentissage, et en particulier étudier le comportement de l'apprenant et de l'enseignant en matière de communication verbale et non verbale.

Les agents pédagogiques, visent donc à incorporer dans les EIAH des actes communicatifs et des actes d'enseignement qui ne sont pas exploités dans les EIAH classiques.

Le travail présenté ici s'inscrit dans un objectif à plus long terme. Cet objectif concerne la spécification et le développement d'un agent pédagogique dont le rôle est d'effectuer un retour personnel vers l'élève de ses résultats obtenus au test d'évaluation effectuée sur la plateforme Pépite. Le projet Pépite [pepite] a pour objet l'évaluation des connaissances en algèbre d'élèves à la sortie du collège (niveau fin de troisième). La plateforme Pépite est composée principalement d'un module de test qui propose à l'élève un ensemble structuré d'exercices et d'un module d'analyse des réponses de l'élève produisant un diagnostic des compétences de l'élève en algèbre.

Dans le retour qu'un enseignant donne à un élève qui a effectué le test Pépite, nous trouvons essentiellement deux parties. Dans la partie que nous nommons *partie bilan*, l'enseignant montre à l'élève ses réussites et ses échecs et lui explique ce que, d'après l'analyse faite par Pépite de ses résultats au test, il maîtrise bien, partiellement, ou bien pas du tout. Dans la partie que nous nommons *partie correction*, l'enseignant prend appui sur les réponses de l'élève aux exercices ainsi que sur les faiblesses et leviers déterminés par le module de d'analyse pour tenter d'amener l'élève à remettre en cause certains acquis erronés et bâtir de nouveaux acquis corrects.

Bien que plus simple d'un point de vue des interactions enseignant-élève, la partie bilan est très riche en terme de communication multimodale. Nous avons décidé d'y concentrer dans un premier temps l'essentiel de notre travail

Notre objectif de recherche est d'apporter des connaissances sur le comportement multimodal d'un enseignant lors d'une interaction dyadique avec un élève dans le contexte décrit précédemment.

L'approche que nous avons utilisée repose sur une analyse par annotation de corpus vidéo. Nous renvoyons le lecteur intéressé à [BAKEMAN & GOTTMAN 97] et [MONTADA 2005] pour une présentation détaillée de cette méthode. Les étapes

principales de ce processus sont : la collection des données, l'annotation des données collectées, la validation des annotations et l'exploitation des données annotées.

La section 2 de cet article décrit la constitution de notre corpus vidéo. La section 3 présente le schéma d'annotation que nous avons défini et décrit le travail d'annotation. La section 4 présente les premiers résultats obtenus par analyse statistique du corpus annoté.

2. Constitution du corpus vidéo

Huit enregistrements ont été réalisés en décembre 2005 et janvier 2006. Ces enregistrements ont impliqué deux enseignantes et huit élèves de classe de troisième, et ont été effectués comme suit.

Chaque enseignante a travaillé avec quatre élèves. Les élèves ont réalisé le test Pépite. Les enseignantes ont ensuite disposé de quelques jours afin de pouvoir étudier les diagnostics produits par Pépite ainsi que le détail des réponses des élèves aux exercices, et se préparer aux entretiens individuels. Enfin, une plage d'une heure par élève (une demi-heure dans certains cas) a été réservée pour que l'enseignante qui lui a été affectée effectue un retour vers l'élève au cours d'un entretien individuel. Le dispositif d'enregistrement est décrit ci-dessous.

L'enseignante dispose d'un document de travail. Celui-ci comprend l'ensemble des réponses de l'élève aux exercices. L'enseignante peut se référer à ce document, mais aussi y écrire et y faire écrire l'élève. L'enseignante dispose d'autre part d'un autre document constitué par le diagnostic de l'élève produit par le module d'analyse de Pépite.

L'enseignante et l'élève sont assis à côté l'un de l'autre. Les documents sont disposés sur une table basse. Cette situation d'entretien fait que l'enseignante et l'élève se positionnent naturellement partiellement tournés l'un vers l'autre.

Deux caméras ont été utilisées, l'une positionnée dans l'axe de l'enseignante et l'autre dans l'axe de l'élève, formant un angle à environ 60 degrés comme représenté par la figure ci-dessus. Chaque caméra enregistre l'ensemble de la scène et permet de capter efficacement les expressions faciales ainsi que les gestes et postures de la personne sur laquelle elle est axée.

Nous avons sélectionné trois enregistrements pour effectuer le travail détaillé d'annotation décrit dans les deux sections suivantes. Les parties bilan de ces enregistrements que nous avons annotées forment un corpus de plus de 54 minutes.

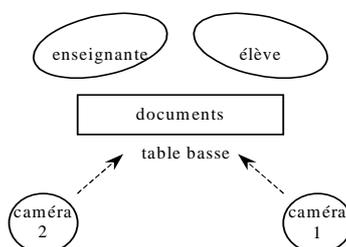


Figure 1. Description du dispositif d'enregistrement des séances

3. Annotation des vidéos

3.1 Schéma d'annotation

Nos objectifs de recherche, l'étude de la littérature dans les domaines de la communication non verbale et de l'enseignement, ainsi qu'une analyse informelle des vidéos que nous avons collectées, nous ont amené à définir un schéma de codage à plusieurs niveaux. Nous nous focalisons ici sur trois niveaux que nous détaillons ci-dessous.

Nous avons dans un premier temps effectué une annotation informelle des vidéos selon des composantes que nous nommons actes généraux et actes pédagogiques pour ensuite regrouper et classifier les éléments observés. Le schéma d'annotation obtenu recoupe pour partie certaines classifications ou taxonomies existantes comme par exemple [SEARLE 1975] concernant les actes de parole, [CORE & ALLEN. 1997] concernant les actes de dialogue ou bien [PARIES 2004] qui définit un ensemble de fonctions communicatives liées aux différentes phases de discours des enseignants.

Le niveau des *actes généraux*, regroupe des actes de communication pouvant être utilisés dans le contexte spécifiquement pédagogique qui est celui de notre étude, mais ne relevant pas directement de celui-ci. On trouve par exemple dans ce groupe des actes de communication comme *interroger, informer, demander de faire, valider, invalider, valoriser, motiver*, ou bien *signaler un malentendu*.

Le niveau des *actes de communication à caractère pédagogique* regroupe des actes de communication ayant une fonction spécifiquement pédagogique au sein du discours de l'enseignant. On trouve par exemple dans ce groupe : *donner un résultat quantitatif, donner un résultat qualitatif, présenter un point fort de l'élève, présenter un point faible de l'élève, signaler le traitement correct d'une question, signaler le traitement incorrect d'une question, signaler un point à travailler, signaler un levier*.

Le troisième niveau concerne la communication non verbale. Nous avons pour l'instant focalisé notre travail sur la direction du regard de l'enseignant. Le contexte expérimental induit quatre valeurs possibles. Dans un contexte d'entretien interpersonnel mettant en jeu des documents, les deux points principaux de focalisation du regard de l'enseignant sont bien sûr d'une part l'élève, et d'autre part les documents. Dans le contexte spécifique de Pépite, nous avons distingué le document de travail et document contenant le diagnostic de l'élève. Il peut de plus arriver que l'enseignant ne dirige son regard ni sur l'élève ni sur les documents, mais regarde « ailleurs ».

3.2 Travail d'annotation et validation

Le corpus a été annoté par Mohamedade Farouk avec le logiciel ANVIL [anvil], selon le schéma de codage décrit dans la section précédente.

Le temps de travail nécessaire pour annoter une minute de vidéo est variable selon la densité des annotations. La durée moyenne peut être estimée à une heure de travail d'annotation par minute de vidéo.

Un travail de validation a été effectué sur un extrait de trois minutes. Cet extrait a été annoté par une autre personne. Nous avons ensuite appliqué le test statistique de l'accord interjuges Kappa proposé par [COHEN 1960] qui permet de chiffrer

l'accord entre deux (ou plusieurs) observateurs qui font référence à un même objet. Le résultat de l'application de cette mesure correspond, selon la classification de [LANDIS & KOCH 1977], à un accord bon à excellent.

4. Analyse statistique des annotations

Nous avons étudié la corrélation entre d'une part la direction du regard, et d'autre part les actes généraux et les actes pédagogiques de l'enseignant. L'objectif de cette étude est de recueillir des connaissances sur le comportement de l'enseignant en ce qui concerne la direction du regard, avec pour perspective la définition d'un modèle de comportement permettant de piloter la direction du regard d'un agent pédagogique.

Pour chaque piste et chaque valeur du schéma d'annotation, nous avons calculé pour chacune des quatre directions du regard le pourcentage de cooccurrence. Nous discutons ci-dessous les premiers résultats de cette étude.

Acte général poser une question

Selon notre étude, lorsque l'enseignant pose une question, l'élève est le point de focalisation du regard avec une probabilité de seulement 0,4, contre 0,6 pour les documents (document de travail et diagnostic Pépite de l'élève). Dans une situation de communication enseignant-élève avec support d'enseignement, il n'est pas surprenant que ce dernier représente un point d'attention. Nous constatons que le temps d'attention porté sur le support d'enseignement peut prendre le pas sur celui porté sur l'élève.

Acte général demander de faire

Nous annotons ainsi les demandes de l'enseignant à l'élève. Une demande peut avoir pour objectif de reformuler, de donner une explication, d'accomplir une tâche, de traiter un exercice, de corriger une erreur, etc.

Notre étude montre une probabilité de regarder les documents supérieure à 0,8. Cette probabilité élevée peut s'expliquer par la nécessité de bien cibler la tâche demandée et d'éviter toute incompréhension ou confusion. Dans le domaine des Mathématiques, le support d'enseignement joue en cela un rôle essentiel.

Actes pédagogiques signaler un traitement incorrect et signaler un traitement correct

Selon notre étude, la probabilité que l'enseignant fixe son attention sur les documents est très forte dans les deux cas. Cela peut s'expliquer par la nécessité de pointer sur les réponses de l'élève aux exercices source de l'erreur, ou bien l'exercice correctement traité. Nous observons cependant une nette variation : les chiffres obtenus sont en effet de 0,91 pour signaler un traitement incorrect et de 0,79 pour signaler un traitement correct. Nous pouvons supposer deux raisons à cela. D'une part en cas de réussite l'enseignant encourage l'élève, et en l'encourageant il le regarde. D'autre part dans le cas d'un traitement incorrect la zone pointée sur le document contenant les réponses de l'élève aux exercices est probablement plus précise. L'enseignant, cherchant à cerner au mieux l'origine de l'erreur, insisterait alors plus sur ce document.

5. Conclusion

Nous présentons dans cet article les premiers résultats d'une étude menée par analyse de corpus vidéo. Partant de situations réelles de dyades enseignant-élève dans lesquelles l'enseignant fournit à un élève un retour de ses résultats obtenus avec le logiciel Pépite, nous étudions la cooccurrence de la direction du regard de l'enseignant avec des actes généraux de communication et des actes pédagogiques. Ces premières mesures montrent que la gestion des tours de parole n'est pas le seul critère pour orienter le regard de l'enseignant dans la situation étudiée. Celui-ci sert à l'enseignant à diriger l'attention de l'élève vers des points importants des supports d'enseignement.

Ce travail va être poursuivi avec pour objectif d'élaborer un modèle de comportement permettant de piloter la composante communication multimodale d'un agent pédagogique associé à la plateforme Pépite.

Remerciements

Nous remercions Elisabeth Delozanne et Brigitte Grugeon qui ont activement participé au travail de recherche présenté dans cet article. Nous tenons à remercier Catherine Pelachaud pour sa contribution dans le suivi général du projet et l'élaboration des schémas de codage, ainsi que Marie-France Delord pour son aimable participation au protocole expérimental.

Ce travail a été financé partiellement par l'Agence Universitaire de la Francophonie (AUF).

Références

- [BAKEMAN & GOTTMAN 97] Bakeman, R. and Gottman, J. M. *Observing Interaction. An introduction to sequential analysis*. Second edition., Cambridge University Press 1997
- [CASSELL 2000] Cassell J., Sullivan J., Prevost S. and Churchill E. *Embodied Conversational Agents*, MIT Press.
- [CORE & ALLEN. 1997] Core M. G. and Allen J. F. *Coding Dialogues with the DAMSL Annotation Scheme*. AAAI Fall Symposium on Communicative Action in Humans and Machines, Menlo Park, California, American Association for Artificial Intelligence. citeseer.ist.psu.edu/core97coding.html. 1997
- [COHEN 1960] Cohen J. : "A coefficient of agreement for nominal scales", *Educ. Psychol. Meas.*: 20, 27-46. 1960
- [LANDIS & KOCH 1977] Landis, J. R. and Koch, G. G. "The measurement of observer agreement for categorical data". *Biometrics* 33: 159-174. 1977
- [MONTADA 2005] Montada, L., Video analysis and the interactional relevance of bodily details. Panel presentation. "Interacting Bodies", 2nd ISGS Conference, Lyon, France, 15-18 June, 2005
- [PARIÉS 2004] Chappet Pariès M., « Comparaison de pratiques d'enseignants de mathématiques relations entre discours des professeurs et activités potentielles des élèves ». In *Recherches en didactique des mathématiques* : vol. 24, n°2-3, pages 251-284 . 2004
- [POGGI 2003] Poggi, I. "Mind Markers". *Gestures. Meaning and Use*. M. Rector, I. Poggi and N. Trigo. Oporto, Portugal University Fernando Pessoa Press: 119-132. 2003
- [SEARLE 1975] Searle J.R. "A taxonomy of illocutionary acts". In K. Gunderson editor, *Language, mind and knowledge*, Minneapolis, University of Minnesota Press, pages 334 - 369. 1975
- [pepite] <http://pepite.univ-lemans.fr/> (visité le 06/02/07 à 11h10)
- [anvil] <http://www.dfki.de/~kipp/anvil/> (visité le 06/02/07 à 11h20)