Modélisation de profils d'élèves à partir de traces de résolution de problèmes

Vivien Robinet, Gilles Bisson, Benoît Lemaire, Mirta B. Gordon.
Email : {prénom.nom@imag.fr}
Equipe Apprentisage - Laboratoire Leibniz-IMAG, UMR 5522
46 Avenue Félix Viallet - 38031 GRENOBLE Cedex 1

Résumé : La modélisation automatique des connaissances de l'apprenant est nécessaire dès lors que l'on veut analyser un apprentissage humain dans le cadre des environnements d'enseignement intelligents. Dans cet article, nous présentons un travail en cours dans lequel des techniques d'apprentissage automatique permettent d'engendrer des modèles d'élèves interprétables par un enseignant ou un didacticien et utilisables éventuellement par un « tuteur » automatique. L'approche proposée a été appliquée au domaine de la résolution d'équations ou d'inéquations par des classes de 4ème et 3ème. Dans les expérimentations qui servent de support à l'apprentissage automatique, l'élève résout de manière libre un grand nombre de problèmes (comme 5x-3<-2x+11) et nous cherchons à repérer et modéliser ses connaissances à partir de ses traces de résolutions.

Mot-clés : EIAH, Modélisation didactique, Construction de profils d'élèves, Classification automatique, Apprentissage de l'algèbre, APLUSIX.

1 Introduction

Les Environnement Informatiques pour l'Apprentissage Humain (EIAH) ont pour objectif de faciliter l'acquisition de nouvelles connaissances. Dans un tel contexte, il est important d'avoir la capacité de diagnostiquer automatiquement les acquis et déficits de l'apprenant afin de faciliter et de guider cet apprentissage. Cet article présente les premiers résultats[1] d'une méthode relativement générale de caractérisation du profil cognitif d'un apprenant à partir des traces de ses résolutions de problèmes. Les traces sont ici vues comme des suites d'exemples de comportements élémentaires, représentés sous la forme d'un couple formé d'un *contexte* et d'une *action*, tous deux caractérisés par un ensemble de descripteurs. La méthode cherche à repérer des régularités par des techniques de classification ascendante. Cela conduit à l'élaboration d'« attitudes » qui correspondent à des comportements généraux mis en oeuvre par l'apprenant face à un contexte.

Prenons un exemple. Dans le cas de la conduite automobile, le fait de mettre son clignotant gauche, de jeter un coup d'oeil à gauche et de prendre la file intérieure pour faire le tour d'un rond-point constitue une attitude, dès lors que suffisamment d'exemples de ce comportement auront pu être repérés. Différentes variables peuvent caractériser une attitude :

- tout d'abord, une attitude peut être *correcte* ou *incorrecte*. L'intérêt de repérer automatiquement les attitudes incorrectes est important dans le contexte éducatif ;
- ensuite une attitude peut également être caractérisée par un *degré de confiance*, dépendant du nombre d'exemples qui la fondent ;
- enfin, une attitude (qu'elle soit correcte ou non), peut être soit *cohérente* si l'apprenti l'a systématiquement mise en oeuvre soit *incohérente*. C'est le cas de l'apprenti qui prend parfois la file intérieure du rond-point et parfois la file extérieure dans des situations qui nous paraissent similaires. La difficulté dans ce cas sera de faire la part entre une vraie incohérence chez l'apprenti et un défaut dans notre description du problème.

[1] Ce travail a été partiellement financé via l'ACI « Cognitique » du Ministère de la recherche.

Dans le cas de la conduite automobile, le contexte pourrait être caractérisé par des descripteurs comme le densité du flux automobile, la présence de panneaux, la nature du carrefour, la direction souhaitée, etc. L'action serait, elle, décrite par des descripteurs comme l'usage des clignotants, la réduction de la vitesse, le choix de la file, le freinage ou non, etc.

L'article est structuré de la manière suivante. Dans une première partie (§ 2) nous allons présenter le contexte du travail, qui porte sur l'enseignement de l'algèbre et plus particulièrement la résolution d'équations ou d'inéquations. Nous ferons ensuite (§ 3) un bref état de l'art sur les systèmes d'EIAH capables de modéliser le comportement d'un apprenant. Les parties 4 et 5 présenteront successivement le contenu et la représentation de l'ensemble des données pour effectuer l'apprentissage automatique puis la méthode de classification utilisée pour engendrer les attitudes. Enfin après avoir montré comment ces attitudes peuvent être mises sous une forme intelligible pour un enseignant ou un didacticien (§ 6) nous conclurons en discutant de la généralité de notre approche et de son utilisation possible pour mettre en œuvre des tuteurs automatiques.

2 Contexte du travail

2.1 L'éditeur Aplusix

Le logiciel APLUSIX (Nicaud et al. 2002) est un logiciel d'EIAH ayant pour objectif de faire acquérir et de renforcer les connaissances de base d'un élève en le confrontant à un environnement d'apprentissage de l'algèbre de type « micro-monde », c'est-à-dire, dans lequel l'élève se forme en explorant librement le domaine dans lequel il travaille. L'acquisition et le renforcement des connaissances s'effectuent tout naturellement en résolvant des exercices d'algèbre. L'élève résout chaque problème *étape par étape* comme s'il le faisait sur une feuille de papier (cf. figure 1). Selon les *consignes* adoptées par l'enseignant pour chaque session de travail, l'élève peut en outre demander au logiciel de vérifier son travail en continu, à la demande ou rétroactivement à la fin de son exercice. Ce logiciel impose néanmoins de ne pouvoir passer à l'étape suivante que lorsque la ligne en cours est correctement formatée (parenthèses fermées, expression complète, …).

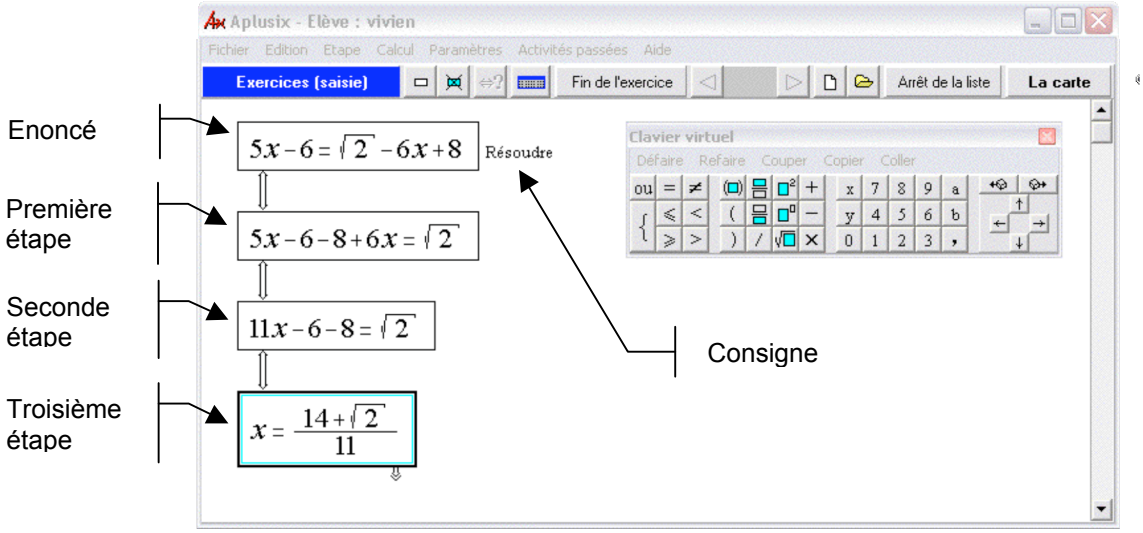

Figure 1 : Environnement APLUSIX.

Outre l'enseignement de l'algèbre, l'objectif d'APLUSIX est d'offrir une plateforme d'expérimentation didactique. A cette fin, la séquence des actions élémentaires effectuées par l'élève pour résoudre un exercice est enregistrée dans un fichier texte qui est appelé le *protocole brut*. Dans le cadre d'expérimentations effectuées dans les années 2003-2005 en France et au Brésil, l'équipe d' APLUSIX a collecté un grand nombre de traces d'élèves résolvant des problèmes. C'est ce matériau qui constitue le point d'entrée de notre approche.

2.2 Analyse des protocoles

Chaque ligne d'un fichier de protocole brut représente une action « élémentaire » effectuée par l'élève, qui peut être une frappe de touche, une utilisation de la souris, ou une action prévue dans les menus d'APLUSIX. Il conserve également d'autres informations, telles l'expression en cours et la durée en secondes écoulée depuis la précédente action. Si les protocoles bruts permettent d'analyser l'interaction d'un élève avec le système (temps de résolution, commande utilisée, …), ils sont toutefois de trop bas niveau pour étudier facilement les connaissances algébriques d'un élève.

Aussi, un module d'interprétation symbolique des *protocoles bruts,* ANAIS, a été développé par l'équipe d'APLUSIX. Ce module est composé de deux parties. D'une part, d'une base de règles élémentaire de réécriture permettant de caractériser le passage d'une expression algébrique à une autre ; ces règles peuvent être « correctes» (par exemple : $(a+b)^2 \rightarrow a^2+2ab+b^2$), ou « erronées ». Ces dernières permettent de représenter les erreurs des élèves (par exemple : $x^n \rightarrow nx$). D'autre part, ANAIS contient un moteur de recherche travaillant en stratégie *meilleur d'abord*. Il permet de retrouver de manière heuristique, à partir du protocole brut, une séquence de règles qui, appliquées successivement, constituent un chemin logique entre une étape de l'élève et l'étape suivante. Cette séquence correspond à une interprétation[2] du raisonnement de l'élève.

Ainsi dans la figure 2 ci-contre, on voit que le logiciel ANAIS interprète la transformation de l'élève comme l'application successive de deux règles, R1 et R2.

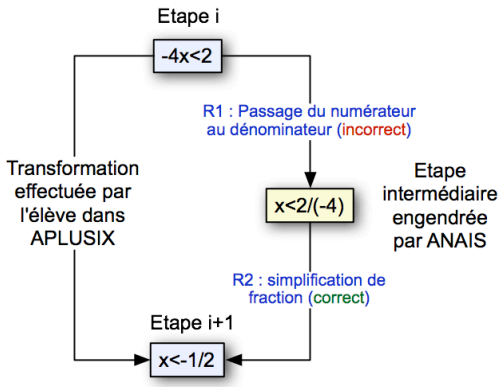

Figure 2 : analyse des protocoles par ANAIS

Ce système fournit plusieurs informations utiles. D'une part, il permet d'identifier les règles, et donc les types d'opérations, implicitement utilisées par l'apprenti en indiquant si elles sont algébriquement correctes ou non. D'autre part, *et c'est l'aspect que nous allons exploiter dans ce travail*, le logiciel ANAIS subdivise les étapes produites par l'élève en introduisant des étapes intermédiaires pouvant contenir simultanément des actions justes et/ou erronées. Il permet ainsi de déceler plus précisément les (mé)connaissances de l'élève. Toutefois, les informations fournies pas ANAIS sont de nature « locale » (il s'agit d'une sous-étape d'un exercice) et pour faire une modélisation de l'élève il est nécessaire d'agréger ces informations.

3 Modélisation de l'élève

3.1 Etat de l'art

Généralement, les traces de résolution de problèmes enregistrées par un EIAH sont, comme dans le cas d'APLUSIX, de longues séquences de micro-comportements qu'il s'agit d'interpréter en changeant de niveau de granularité (Mostow, 2004). Plusieurs travaux ont déjà utilisé des techniques d'apprentissage automatique pour en extraire des régularités afin de découvrir, derrière une succession de clics de souris, les connaissances du domaine bien ou mal maîtrisées par l'élève. Cependant, dans tous les cas, l'extraction de ces régularités nécessite une redescription préalable des productions des élèves en termes de plus haut niveau, définis par des experts humains, et qui sont spécifiques du domaine d'application visé par l'EIAH. C'est également ce que nous faisons dans ce travail, en nous appuyant sur les interprétations produites par ANAIS.

[2] Dans ce travail, nous n'analysons que les opérations de déplacement de termes dans les expressions (passage d'un terme du membre gauche vers le membre droite, ou inversement). Dans un tel contexte, les interprétations effectuées automatiquement par ANAIS sont validées par les didacticiens dans plus de 80% des cas.

Toutefois, beaucoup de travaux construisent les modèles d'élèves avec un apprentissage supervisé qui s'appuient sur des profils pré-établis par des experts du domaine. C'est par exemple le cas de Profile Extractor (Esposito et al., 2004) qui utilise des données comportementales d'élèves, issues de leurs interactions avec un système de e-learning, pour construire un arbre de décisions capable d'assigner chaque élève à une classe prédéfinie. Nous nous distinguons de ces approches en ce sens que notre objectif est plutôt de découvrir ces profils.

AnimalWatch (Beck & Woolf, 2000) qui concerne le domaine de l'arithmétique est plus proche de nos considérations. Le système analyse les données comportementales d'un élève pour prédire s'il sera capable de résoudre un problème donné et le temps qu'il y consacrera. Un tel modèle est très utile puisqu'il permet de fournir à l'élève des problèmes adaptés à son niveau de compétences et qui lui permettront de progresser. Les descripteurs de haut niveau utilisés par *AnimalWatch* appartiennent à quatre catégories :

- *Student* : le sexe, le score à différents tests ;
- *Topic* : le type d'opérateur (addition, soustraction, etc.) et le type d'opérande ;
- *Problem* : le degré de difficulté du problème, le nombre de prérequis nécessaires pour le résoudre (par exemple : addition de fractions, simplification de fractions, calcul du PPCM, etc.), et enfin le degré de difficulté des prérequis ;
- *Context* : le nombre d'erreurs antérieures, le niveau des conseils donnés par le système, etc.

Dancs *AnimalWatch,* différents algorithmes ont été utilisés pour prédire la probabilité de succès de l'élève et le temps consacré, tels qu'un classifieur bayésien et un arbre de décisions, mais les auteurs se sont finalement rabattus sur une simple régression linéaire. La encore nous nous distinguons de cette approche par le domaine d'application et surtout la finalité : nous cherchons à construire un profil cognitif et non à prédire la valeur de deux variables. Cependant, comme nous le verrons nous utilisons certains des descripteurs de *AnimalWatch* notamment ceux de la catégorie *Topic*.

Dans le domaine de la modélisation d'élèves en mathématique, on peut aussi citer *Web-EasyMath* (Tsiriga et Virvou, 2004). Plutôt que d'assigner une classe prédéfinie à chaque nouvel élève, ce système lui fait passer un test préliminaire permettant d'extraire des descripteurs pertinents (auto-estimation de ses compétences dans le domaine et taux de réponses correctes aux quatre opérations de base). En appliquant un algorithme des k plus proches voisins, il le positionne par rapport aux autres élèves en calculant leurs similarités. Ceci permet d'initialiser le modèle de l'élève, représenté *in fine* par un ensemble de critères accompagnés par la probabilité que l'élève les connaissent.

Sison & Shimura (1998) font un état de l'art des travaux qui appliquent des techniques d'apprentissage automatique dans les EIAH et ils énoncent plusieurs axes pour classer les différentes manières d'apprendre un modèle d'élève à partir de ses « comportements », qui peuvent être des traces de résolution ou plus simplement des réponses à des questions

- Complexité du comportement étudié : depuis une simple valeur jusqu'à des expressions élaborées. Dans notre cas, les comportements de l'élève sont relativement complexes puisqu'il s'agit de la manipulation d'expressions algébriques ;
- Multiplicité des comportements : certaines approches peuvent se satisfaire d'analyser un comportement unique pour inférer des connaissances sur l'élève, alors que d'autres, comme la nôtre, nécessitent un grand nombre d'exemples de comportements ;
- Connaissances du domaine : elles peuvent être codées dans le système ou elles-mêmes apprises automatiquement ; dans notre cas, elles sont en partie définies automatiquement mais aussi représentées avec des descripteurs sélectionnés de manière « ad hoc » ;
- Utilisation de « misconceptions » (bibliothèque d'erreurs) : là encore, cet ensemble peut être établi à l'avance ou appris. Dans notre approche les règles erronées sont apprises par le système. La notion de correction d'une étape de raisonnement de l'élève est toutefois fournie par ANAIS en comparant les « domaines solution » avant et après chaque étape.

- Construction du modèle de l'élève : elle peut être analytique (en essayant de retrouver les étapes qu'a pu suivre l'élève) ou synthétique (par généralisation des comportements). La détection des étapes intermédiaires effectuée par ANAIS se situe dans la première modalité, et notre approche se place dans la seconde.

4 Constitution de l'ensemble d'apprentissage

4.1 Représentation "contexte-action"

Les données sur lesquelles nous avons travaillé proviennent d'expérimentations effectuées au Brésil au printemps 2004. Le niveau des classes étudiées était équivalent à la 4ème et 3ème de nos collèges. Les expérimentations ont concerné 2700 élèves environ, chaque élève ayant traité en moyenne entre 3 et 10 problèmes (résolution d'équations et/ou inéquations)

Le modèle de l'élève que nous construisons s'appuie sur les observables fournis par ANAIS, c'est-à-dire sur la décomposition des étapes de l'élève en étapes plus élémentaires correspondant à l'application d'une règle de réécriture. Notre modèle de l'élève se situe donc au niveau des manipulations algébriques élémentaires et nous n'aborderons pas ici la question de la modélisation des connaissances de plus haut niveau, telle la stratégie de résolution utilisée par l'élève. Par ailleurs, dans la présente étude, nous ne nous intéressons qu'aux seules règles qui traitent du déplacement des termes d'un côté à l'autre d'une expression algébrique (équation ou inéquation), ce que nous appelerons par la suite un *mouvement*. En voici deux exemples :

$$\begin{cases} 2x - x < 2 \rightarrow 2x < 2 + x \\ -3x(x-5) < 4 \rightarrow (x-5) > -4/3x \end{cases}$$

Ce choix est motivé, de manière pragmatique, par le fait que ANAIS contient l'ensemble des règles définissant ce type d'opération (ce qui est un prérequis). Mais il convient de souligner également que la compréhension des mécanismes de déplacement de termes constitue les fondements de l'apprentissage de l'algèbre dans les classes de 4ème et 3ème. Chaque exemple, que nous appelerons désormais un « cas[3] », dans notre ensemble d'apprentissage est donc constitué ainsi :

$$Cas_i : Expression\ \ de\ \ départ \xrightarrow{R_j} Expression\ \ d'arrivée \qquad (R_j \text{ étant une règle de mouvement})$$

Nous n'allons toutefois pas travailler avec cette représentation pour deux raisons. D'une part, il est délicat de travailler directement sur des expressions algébriques dont la sémantique est complexe, en particulier il est clair que deux équations syntaxiquement voisines peuvent être sémantiquement très différentes. D'autre part, il est nécessaire d'expliciter la signification des informations véhiculées par la règle. Aussi comme dans (Renaudie 2005) nous allons modéliser chaque cas par deux ensembles d'attributs A_i désignant respectivement le *contexte* et *l'action* effectués par l'élève. Le partie contexte décrit l'expression avant l'application du mouvement et la partie action décrit l'expression une fois ce mouvement effectué[4]. L'information véhiculée par la règle se trouve alors directement exprimée dans ces attributs :

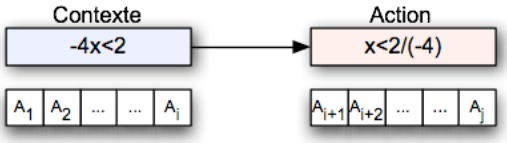

Figure 3 : Représentation d'un cas sous la forme d'un couple <*contexte*, *action*>

[3] Il s'agit en fait d'un acronyme « ContexteActionS » dont la signification est explicite dans la suite de l'article.

[4] Dans cette représentation on considère que le contexte est imposé à l'élève et qu'il doit réagir par une action. En réalité, le contexte n'est pas totalement imposé, puisqu'il est dépendant des actions précédentes. Mais ce type de considération fait partie de la stratégie mise en œuvre par l'élève ce qui n'est pas l'objectif de la présente étude.

Ainsi que nous l'avons évoqué dans l'introduction, un autre avantage de cette représentation très générale est qu'elle nous permet de mettre en œuvre un mécanisme d'apprentissage qui est dans une certaine mesure indépendant du domaine étudié et donc transposable à d'autres problème. Cet aspect sera discuté plus en détail dans la conclusion de l'article. Muni de cette représentation, les modèles que nous allons apprendre (cf. §3) visent à déterminer des comportements stables que l'on appelle des « attitudes » et qui permettent de caractériser les connaissances des élèves. D'un point de vue didactique, deux catégories d'attitudes opposées sont intéressantes à rechercher :

- les « théorèmes en actes » (Vergnaud 2001), qui correspondent à des attitudes dans lesquels les élèves font toujours le même type d'actions dans un ou plusieurs contextes. Ces théorèmes en actes peuvent soit correspondre à des règles algébriques correctes ou à la mise en œuvre de connaissances partiellement ou totalement erronées.
- les « irrégularités » (Renaudie 2005) qui correspondent à des attitudes dans lesquelles les élèves effectuent des actions différentes dans des contextes mathématiquement identiques.

4.2 Descripteurs utilisés

Les descripteurs retenus pour décrire les expressions de départ et d'arrivée doivent caractériser à la fois la réalité mathématique de la transformation effectuée (en termes d'opérandes et d'opérateurs mis en jeu), mais également rendre explicites l'ensemble des « perceptions » de l'élève. En effet, un débutant en algèbre peut être perturbé par la présence (ou l'absence) de certains symboles même s'ils sont non signifiants, d'un point de vue mathématique, vis-à-vis de l'opération effectuée (nous en verrons une illustration au §6.2). Ainsi, l'expression « -1x » peut être percue comme étant de nature différente de « -x » ou encore la présence d'une racine carré dans une expression, même si elle n'est pas impliquée dans le mouvement, peut avoir un effet perturbant pour l'élève.

Nous avons donc défini un ensemble assez large de descripteurs, avec parfois des redondances[5], afin de prendre en compte cette double dimension mathématique et perceptive. Soulignons bien ici qu'il ne s'agit pas de réaliser une description complète des expressions algébriques mais simplement de ne garder que les informations qui sont soit impliquées dans la transformation, soit susceptibles de changer le regard que porte l'élève sur l'expression.

Parallèlement, dans la description des contextes, certains descripteurs ont été dédoublés afin de prendre en compte la « localisation » de l'information qu'ils véhiculent. Nous distinguons ainsi dans le contexte et l'action quatre localisations différentes dans une formule :

- Argument : partie de l'expression qui est directement affectée par le mouvement
- Terme : l'ensemble des éléments liés à l'argument par des opérations de type ×, ÷
- Membre : partie droite ou gauche de l'équation (ou inéquation)
- Expression : ensemble de l'expression

$$\overbrace{\underbrace{\overbrace{\underbrace{\mathbf{5}}_{\text{Argument}}x}^{\text{Terme local}} + 4 - \frac{7}{3}}_{\text{Membre Gauche}} = \underbrace{9x^2 - 10}_{\text{Membre Droit}}}^{\text{Expression globale}} \quad \rightarrow \quad x + 4 - \frac{7}{3} = 9x^2 - 10 - \mathbf{5}$$

Figure 4 : Exemple de mouvement (incorrect) avec la description des différentes parties du contexte.

[5] On utilise notamment cette redondance afin de découper des domaines de valeurs des descripteurs de différentes manières. Cela permet d'exprimer des taxonomies de valeurs et d'avoir plusieurs niveaux de généralisation possibles.

Finalement, les descripteurs de contexte retenus sont au nombre de 25 (cf. annexe A) et ceux liés aux actionx sont au nombre de 12. Les descripteurs sont majoritairement de type booléen. La table 1 montre un exemple de vecteur <contexte, action> qui est obtenu avec le cas décrit dans la figure 4. Pour simplifier le tableau les descripteurs non significatifs ont été supprimés.

Contexte		Action	
argument.coté	*gauche*	argument.opérateur	-
argument.position	*début*	argument.catégorie	*add*
argument.polynôme	*vrai*	argument.mêmeOpérateur	*faux*
argument.coefficient	*vrai*	argument.mêmeCatégorie	*faux*
argument.signeImplicite	*vrai*	argument.négatif	*vrai*
argument.opérateur	×	argument.mêmeSigne	*faux*
argument.catégorie	*mult*	expression.type	*équation*
argument.négatif	*faux*	expression.mêmeType	*vrai*
terme.polynôme	*vrai*	transformation.juste	*faux*
expression.type	*equation*		
expression.polynôme	*vrai*		

Table 1 : Exemple de description d'un cas.

5 Construction des modèles

5.1 Introduction

La détermination des attitudes ne peut se faire qu'en utilisant une approche de type classificatoire (apprentissage non-supervisé) car nous n'avons de concept cible à apprendre. Certes, on pourrait construire un classifieur permettant de déterminer les contextes dans lesquels les élèves obtiennent un résultat correct ou non mais le but de notre étude n'est pas de prévoir le succès ou l'échec.

La classification est effectuée élève par élève sur l'ensemble des cas qui lui sont associés, c'est-à-dire l'ensemble des mouvements qu'il a pu faire dans les exercices qu'il a traités. En moyenne, comme nous l'avons dit, les élèves ont traité entre 3 et 10 exercices ce qui permet à ANAIS d'extraire entre 50 et 100 « cas »[6] (nos exemples) pour chaque élève. Expérimentalement, cela s'est avéré suffisant pour faire émerger des attitudes dans lesquelles on retrouve un motif stable de paire <contexte, action> caractérisant les « réflexes mathématiques » de l'apprenant. Notons que l'on pourrait également envisager d'effectuer ce traitement au niveau d'un ensemble d'élèves de même niveau pour faire ressortir un modèle collectif comme dans (Bisson et al. 2003).

Les attitudes ne correspondent pas nécessairement à quelque chose qui a été vu en cours. Elles pourront être soit justes, soit fausses, soit partiellement exactes selon les situations. De surcroit, il est important de tenir compte du fait que les êtres humains peuvent faire des erreurs d'inattention, voire même avoir un comportement « pseudo-aléatoires », dans le sens où, pour un même contexte, plusieurs actions différentes peuvent être effectuées par l'élève selon le moment où il travaille avec APLUSIX. Ces situations sont potentiellement révélatrices d'une difficulté qu'il faut caractériser. Du point de vue de l'apprentissage cela signifie que les cas étudiés vont être « bruités » et qu'il faut introduire une dimension statistique dans le mécanisme de classification afin de pouvoir décider à partir de quel niveau un comportement devient réellement significatif.

[6] Il y a évidemment plusieurs mouvements algébrique effectués dans chaque exercice.

5.2 Classification des cas

Pour construire les attitudes, nous allons utiliser une méthode de type CAH (Classification Ascendante Hiérarchique, voir entre autres : Berkhin 2002, Jain et al. 1999) qui consiste à aggréger itérativement les exemples (ici il s'agit des *cas*) de manière à construire un arbre binaire de *classes*. Ces classes sont ici les *attitudes*, qui seront plus ou moins générales suivant à quel niveau de l'arbre on les examine. L'utilisation de la CAH nécessite la définition préalable de deux distances :

- Une distance entre les exemples (ex: distance de Minkowski, …)
- Une distance entre les classes (ex : saut minimal/maximal, distance de Ward, …)

Toutefois, nous n'avons besoin ici que d'une seule distance qui porte à la fois sur les *cas* et les *attitudes* car nous utilisons une méthode d'aggrégation des classes spécifique.

5.2.1 Méthode d'aggrégation

Dans l'algorithme de CAH « standard » les classes construites n'ont pas de description explicite[7] : à chaque itération le système regroupe les 2 instances les plus proches voisines (exemples et/ou classes) puis remet à jour l'ensemble des distances en s'appuyant sur les valeurs existantes. Dans notre travail, cas et attitude ont une représentation commune et lors de l'agrégation on va calculer explicitement la distance entre la nouvelle attitude et les autres entités de la matrice[8].

Dans la suite, chaque élève est représenté par un vecteur dont chaque composant corrrespondent à une modalité possible du couple <contexte-action> et qui a pour valeur le nombre d'occurrences du descripteur correspondant. Pour un *cas* ce nombre est forcément soit 0 soit 1, mais dans une *attitude* (qui agrège des informations de plusieurs *cas*) la valeur proportionnelle aux *cas* qui la constituent. Grâce à cette représentation nous pouvons donc prendre en compte la variabilité des comportements que décrite dans l'introduction (§ 5.1). Lorsque deux cas/attitudes sont agrégées la nouvelle attitude est calculée en faisant la somme des occurrences de chacune des modalités (figure 5) :

Descripteurs
arg.coté
arg.position
arg.complexe
arg.polynôme
…

Cas 12			
gauche	droite		
0	1		
début	milieu	fin	isolé
0	1	0	0
vrai	faux		
1	0		
vrai	faux		
1	0		
…	…		

+

Attitude 6			
gauche	droite		
1	3		
début	milieu	fin	isolé
0	1	3	0
vrai	faux		
0	1		
vrai	faux		
3	0		
…	…		

=

Attitude (6&12)			
gauche	droite		
1	4		
début	milieu	fin	isolé
0	2	3	0
vrai	faux		
1	1		
vrai	faux		
4	0		
…	…		

Figure 5: Exemple d'agrégation de deux cas/attitudes avec la remise à jour du nombre d'occurrences.

5.2.2 Distance entre deux attitudes

Chaque attitude A_i est composée de deux termes complémentaires, le contexte (*ctx* (A_i)) et l'action (*act* (A_i)). La distance globale $Dist_T$ est une somme pondérée de la distance entre ces deux termes :

$$Dist_T\left(A_i, A_j\right) = \alpha \times dist\left((ctx(A_i), ctx(A_j)\right) + (1-\alpha) \times dist\left(act(A_i), act(A_j)\right)$$

En faisant varier le paramètre α on donne une importance plus grande à l'un ou l'autre de ces termes. Si le coefficient α est petit ($\alpha \approx 0{,}20$) le terme dominant est le poids relatif de la distance entre actions, ce qui favorise le regroupement d'actions voisines même si les contextes sont

[7] On peut bien sur toujours construire une description en intention de la classe mais ce n'est qu'à des fins explicatives.

[8] Notons qu'avec cette approche il est possible que les distances trouvées successivement lors de la classification ne soient plus monotones croissantes. Cela se traduit par des inversions de nœuds dans la hiérarchie indicée, cependant elle présente l'avantage de fournir des distances plus en adéquation avec la représentation dans le haut de l'arbre.

différents. Ceci permet de mettre en évidence des *théorèmes en actes* (cf. §4.1). Inversement si le coefficient α est grand on favorise le regroupement de contextes voisins, et par conséquent, la mise en évidence d'*attitudes* ayant des contextes similaires, c'est-à-dire d'*irrégularités*. Pour calculer la distance dist(V_i, V_j) entre deux vecteurs quelconque de données (*ctx* (A_i)) ou (*act* (A_i)), nous comparons, descripteur par descripteur (*d*), le profil de répartition fréquentiel des modalités (*m*).

$$dist(V_i, V_j) = \sum_d \left(p_d \times \sum_m \left| \frac{\# V_{i,d,m}}{\Sigma V_{i,d}} - \frac{\# V_{j,d,m}}{\Sigma V_{j,d}} \right| \right) \quad avec \begin{cases} \# V_{i,d,m} & \text{: nombre d' occurrences de la modalité } m \text{ du descripteur } d \\ \Sigma V_{i,d} & \text{: somme des occurrences des modalités du descripteur } d \end{cases}$$

Dans cette formule, on observe que chaque descripteur peut-être pondéré par un coefficient p_d. Dans nos expérimentations les poids ont été ajustés de la manière suivante :

- l'attribut « *Transformation.juste* » n'est pas pris en compte (p=0) car, ainsi que nous l'avons déjà dit, nous ne cherchons pas à construire des attitudes basées sur le fait que l'élève a réussi ou échoué. Toutefois comme on le verra dans la la classification obtenue cette distinction apparaît parfois sous l'influence des autres descripteurs.
- on met un fort poids sur « *Expression.type* » car on ne souhaite pas mélanger dans les attitudes les équations et les inéquations qui représentent des problèmes assez différents.
- hormis les attributs ci-dessus, tous les attributs de la partie action ont le même poids. Dans la partie contexte c'est également le cas sauf pour les attributs « Argument.opérateur » et « Argument.négatif » dont le poids a été augmenté après les premières expérimentations de manière à maintenir séparée, lors de la classification, les opérations (et signes) entre elles.

6 Résultats obtenus

6.1 Exemples d'attitudes

Nous avons utilisé cette approche classificatoire sur l'ensemble des élèves de l'expérimentation effectuée au Brésil. La figure 6 présente un exemple de hiérarchie d'attitudes obtenue pour un élève donné (dont l'identifiant est 1497) avec une valeur $\alpha = 0{,}80$ (cf § 2.5.2), ce qui correspond donc à privilégier la détection d'irrégularités (cela n'empêche toutefois pas d'identifier des théorèmes en acte comme nous le verrons dans la suite). Cet élève avait produit cinquante couples <contexte-action>, nous n'avons représenté dans cette figure que les attitudes qui avaient une distance non nulle entre elles. Afin de visualiser plus rapidement la signification des classes, lors des expérimentations nous avons associé les valeurs de certains attributs importants à des codes visuels :

Attribut graphique	Descripteur associé			
Couleur de fond	Action.Transformation.juste	vrai : vert/clair	faux: rouge/foncé	mixte: orange/moyen
Couleur du contour	Action.Argument.négatif	vrai : bleu/foncé	faux: jaune/clair	mixte: vert/moyen
Couleur du nombre	Contexte.Argument.négatif	vrai : bleu/foncé	faux: jaune/clair	mixte: vert/moyen
Forme du noeud	Contexte.Argument.opérateur	+ : triangle | - : carré | × : pentagone | ÷ : cercle		
Nombre de contours	Action.Argument.opérateur	+ : 1 | - : 2 | × : 3 | ÷ : 4		
Orientation de la forme	Contexte.Expression.type	Equation : 0°	Inéquation : 45°	
Anomalie de diagnostic ANAIS		10 traits de contour autour du nœuds		

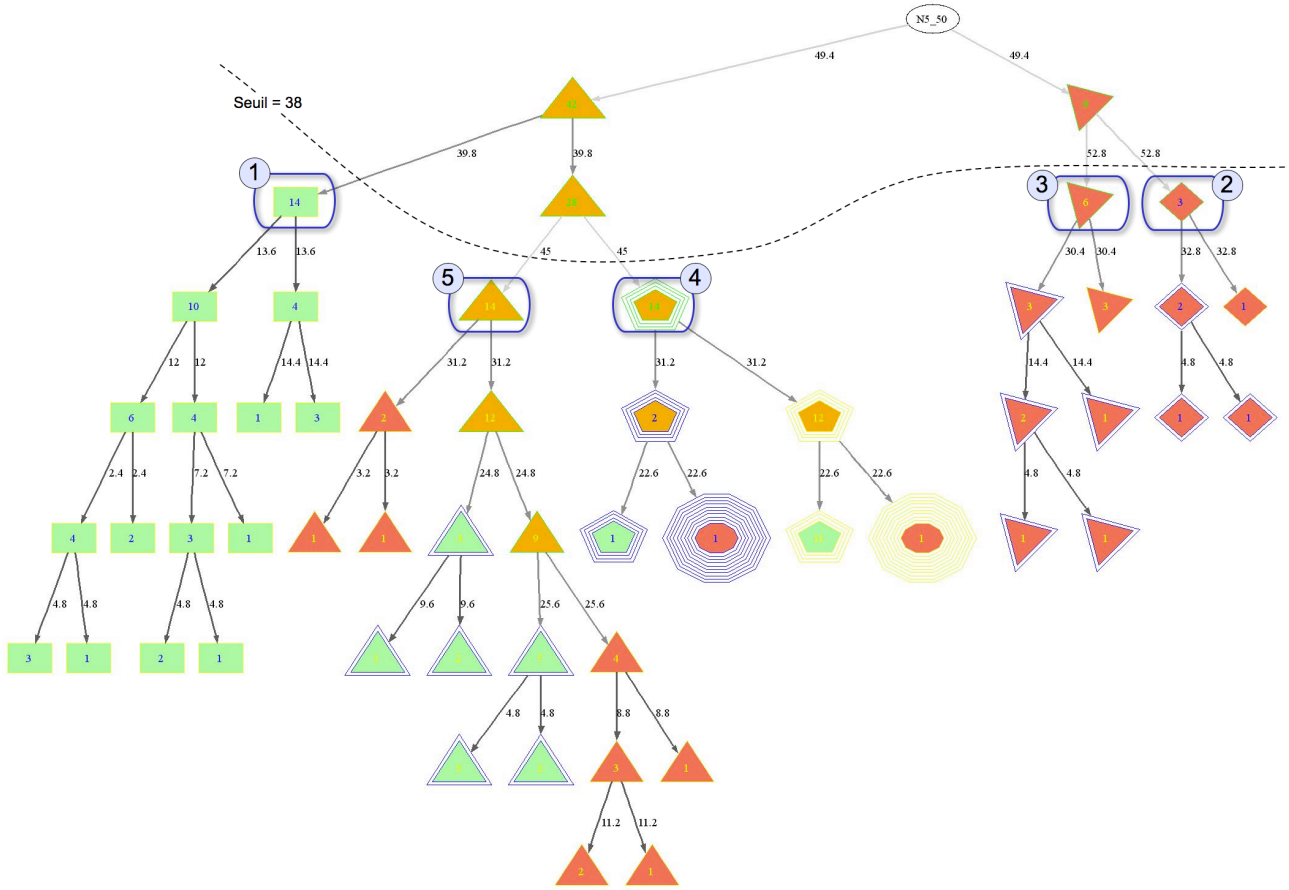

Figure 6: Classification des cas associés à l'élève id=1497 et numéro des cinq attitudes retenues. La représentation graphique est produite avec Graphviz[9]. Les distances entre noeuds sont indiquées sur les arcs.

[9] Ce logiciel, maintenant en license « open-source », a été développé par ATT : http://www.graphviz.org/

Comme toujours dans le cas de la classification hiérarchique, se pose le problème de l'identification des nœuds à conserver, ou en d'autres termes la définition du *seuil de coupure* dans l'arbre. Il s'est avéré expérimentalement, qu'avec la représentation et la distance utilisée, la valeur pertinente du seuil était relativement stable pour l'ensemble des élèves de l'expérimentation et que la valeur 0,38 constituait un bon compromis. Dans la figure 6, cette valeur de seuil conduit à l'identification de 5 attitudes diférentes qui sont décrites dans la table 2. Les cases colorées correspondent à des descripteurs pour lesquels plusieurs modalités simultanées ont été repérées dans les attitudes.

	Descripteur	**Attitude 1**	**Attitude 2**	**Attitude 3**	**Attitude 4**	**Attitude 5**
Contexte	Arg.coté	gauche, droit	gauche	gauche	gauche	gauche, droit
	Arg.position	début,fin,milieu,isolé	début,fin	fin	début	début,fin,milieu
	Arg.complexe	faux	faux	faux	faux	faux
	Arg.polynôme	vrai, faux	vrai, faux	faux	faux	vrai, faux
	Arg.racine	faux	faux	faux	faux	faux
	Arg.puissance	faux	faux	faux	faux	faux
	Arg.fraction	faux	faux	faux	faux	faux
	Arg.coefficient	vrai, faux	vrai	vrai	vrai	vrai, faux
	Arg.négatif	vrai	vrai	faux	vrai, faux	faux
	Arg.signeImplicite	faux	faux	faux	vrai, faux	vrai, faux
	Arg.opérateur	-	-	+	×	+
	Arg.catégorie	add	add	add	mult	add
	Terme.polynôme	vrai	vrai	vrai	vrai	vrai
	Terme.fraction	faux	faux	faux	faux	faux
	Terme.puissance	faux	faux	faux	faux	faux
	Terme.racine	faux	faux	faux	faux	faux
	Exp.type	équation	inéquation	inéquation	équation	équation
	Exp.sens	NA	gauche, droit	gauche, droit	NA	NA
	Exp.large	NA	faux	faux	NA	NA
	Exp.Polynôme	vrai	vrai	vrai	vrai	vrai
	Exp.Fraction	faux	faux	faux	faux	faux
	Exp.Puissance	faux	faux	faux	faux	faux
	Exp.Racine	faux	faux	faux	faux	faux
	Exp.nulGauche	faux	faux	faux	faux	vrai, faux
	Exp.nulDroite	faux	faux	faux	faux	faux
Action	Arg.opérateur	+	+, -	+, -	÷	+,-
	Arg.catégorie	add	add	add	mult	add
	Arg.mêmeOpérateur	faux	vrai, faux	vrai, faux	vrai, faux	vrai, faux
	Arg.mêmeCatégorie	vrai	vrai	vrai	vrai	vrai
	Arg.négatif	faux	vrai, faux	vrai, faux	vrai, faux	vrai, faux
	Arg.mêmeSigne	faux	vrai, faux	vrai, faux	vrai	vrai, faux
	Exp.type	équation	équation	équation	équation	équation
	Exp.sens	NA	NA	NA	NA	NA
	Exp.large	NA	NA	NA	NA	NA
	Exp.mêmeType	vrai	faux	faux	vrai	vrai
	Exp.mêmeSens	NA	NA	NA	NA	NA
	Transformation.juste	vrai	faux	faux	vrai, faux	vrai, faux
	Effectifs	*14*	*3*	*6*	*14*	*14*

Table 2 : Attitudes produites pour l'élève id=1497

Détaillons quelques unes de ces attitudes. Notons tout de suite que la seconde attitude a des effectifs trop réduits (dernière ligne du tableau) pour qu'elle soit significative.

L'attitude numéro 1 correspond à un comportement *correct*. Elle indique que cet élève maîtrise la résolution d'équations (Exp.type) simples (Arg.racine, Arg.puissance et Arg.fraction sont faux) dans lesquelles un argument négatif doit être déplacé (arg.négatif=vrai). Il s'agit par exemple d'équations comme : 6x**-3**=2x+4 → 6x=2x+4+**3**. L'élève déplace correctement cet argument de l'autre côté du signe égalité, quel que soit sa position dans l'expression (Arg.côté et Arg.position sont généralisés) et la valeur du coefficient (Arg.coefficient), en conservant la catégorie additive (Arg.mêmeCatégorie=vrai) et en changeant le signe (Arg.mêmeSigne=faux). C'est donc un *théorème en acte* : dans des contextes différents, l'élève applique toujours la même règle.

L'attitude numéro 5 est en revanche *incohérente* : dans des contextes similaires, mais avec un argument positif (Arg.négatif=faux), l'élève a un comportement moins systématique et il échoue

parfois à changer le signe de l'argument. Il y a deux sous-attitudes qui ont été regroupées , l'une correspondant à un comportement *correct*, l'autre à un comportement *incorrect*, et c'est justement l'intérêt de la méthode que d'avoir explicité le fait que cet élève n'a pas un comportement identique face à des situations similaires. Il est toutefois possible comme nous le verrons dans la section suivante § 6.2 de redescendre dans l'arbre jusqu'à la bifurcation entre une attitude généralisée correcte et une attitude généralisée incorrecte, et d'isoler les quelques descripteurs qui peuvent être la cause de ces comportements différents.

6.2 Exploitation des attitudes

Les attitudes ainsi générées peuvent avoir une double utilisation dans un EIAH. Tout d'abord, on peut les utiliser simplement pour rappeler à l'élève qu'elle est la règle *correcte* qu'il aurait du appliquer dans ce contexte, mais il est plus intéressant encore d'exploiter les valeurs des descripteurs afin d'engendrer automatiquement des exercices permettant à l'élève de corriger une attitude *incorrecte* et/ou *incohérente*. Par exemple, dans l'attitude 5 ci-dessus, on utiliserait les valeurs des descripteurs pour construire un exercice contenant un argument positif à déplacer de l'autre côté de l'équation, par exemple : 7x+4=11.

La seconde application de cette approche est à destination des enseignants ou des didacticiens qui sont intéressés à connaître les comportements récurrents des élèves. Il est clair que la représentation des attitudes sous la forme d'un vecteur n'est pas directement lisible, aussi nous avons conçu un module qui convertit automatiquement ces informations en langue naturelle.

La méthode repose sur une simple concaténation de morceaux de textes instanciés par les valeurs des descripteurs. Lorsque les descripteurs ont été généralisés, c'est-à-dire qu'ils possèdent plusieurs modalités, le morceau de texte correspondant est omis. De surcroît, deux exemples générique sont également engendrés pour illustrer l'attitude. Enfin, dans le cas d'une attitude représentant à la fois des transformations correctes et incorrectes, nous redescendons dans l'arbre comme indiqué précédemment afin d'indiquer les descripteurs qui discriminent les comportements *corrects* de ceux *incorrects*. Voici un exemple de texte correspondant à l'attitude 5 présentée plus haut (les textes correspondant aux trois autres attitudes se trouvent dans l'annexe B) :

```
Attitude numero 40, effectif 14.
Correcte : 8 Incorrecte : 6.

Diagnostic :
Attitude incoherente consistant en un mouvement dans une equation d'un terme,
initialement positif en position additive, ce mouvement s'effectuant avec ou
sans changement de signe, la position finale etant additive.

Exemples :
r - v = nx ----------> r - v - nx = 0
v - r = n ----------> v - r + n = 0

Explication :
L'eleve ne semble pas avoir une attitude coherente face a ce type de contexte,
ce qui peut traduire une incomprehension plus profonde.
On peut neanmoins chercher certains descripteurs de contexte qui pourraient
permettre d'expliquer le comportement de l'eleve.

Les causes possibles de l'erreur sont :

->    l'argument se trouve du cote droit
->    l'argument comporte une partie polynomiale
->    l'argument comporte une partie numerique implicite
->    le signe de l'argument est explicite
->    l'argument initial est en derniere position
->    le membre de gauche est nul
```

7 Conclusion

Dans cet article, nous avons présenté une méthode permettant d'analyser automatiquement des comportements d'élèves, représentés sous la forme de descripteurs de contexte et d'action, afin de construire des profils exprimés sous la forme *d'attitudes*. Dans son développement actuel, cette approche offre plusieurs propriétés intéressantes :

- Le résultat est directement interprétable par un expert et apporte aux didacticiens une vision globale et synthétique de l'état des connaissances d'un élève.
- Le système sait prendre en compte les connaissances contradictoires, représentées sous la forme d'irrégularités, et permet de faire la distinction entre le *régulier* et *l'occasionnel*.
- Les diagnostics effectués dans le domaine de l'algèbre peuvent être facilement utilisés pour paramétrer un générateur d'exercices destiné à entrainer les élèves.

L'un des avantages de ce modèle de l'élève réside dans sa dimension cognitive car il représente la logique propre à l'élève et non la justesse mathématique de ses actions. Ceci rejoint les avis de (Dillenbourg et al. 1992) et de (Balacheff 1992), qui préfèrent la vision d'une utilisation rationnelle, mais non appropriée, de processus cognitifs à la notion d'« erreur de connaissance ». Ainsi, les théorèmes-en-acte de statut mixte sont le signe de connaissances qui peuvent être erronées, mais qui ont par ailleurs un domaine de validité : utiles dans certaines situations particulières, ils peuvent se révéler inadaptés dans un autre contexte. « Dans ce sens, une conception même fausse est une véritable connaissance » (Balacheff, 1992). D'un point de vue qualitatif, les modèles produits mettent en évidence des compétences et des erreurs plutôt classiques, que les enseignants connaissent bien. Ceci dit, il n'en reste pas moins intéressant qu'un processus automatique conduise à des conclusions congruentes avec celles de l'expert.

Le domaine d'application étudié est limité aux mouvements algébriques dans les équations et inéquations, mais nous envisageons, à court terme, l'extension de ces possibilités. En effet, l'équipe Did@tic du laboratoire Leibniz-IMAG a récemment développé un nouvel ensemble de règles pour ANAIS permettant l'identification et l'interprétation automatique des étapes de *développements* et de *réductions* d'expressions. Comme pour les mouvements, ces règles semblent caractériser de manière satisfaisantes environ 80% des productions d'élèves. Pour analyser ces nouveaux cas il sera bien sur nécessaire d'introduire de nouveaux descripteurs de nos couples <contexte, actions>. Une étude préliminaire semble montrer que cette tâche est réalisable, les descripteurs supplémentaires étant qualitativement et quantitativement voisins de ceux qui sont actuellement utilisés.

Plus fondamentalement, l'une des limites de notre approche réside dans le fait que les comportements d'élève sont considérés indépendemment du temps. Il serait donc intéressant d'essayer de modéliser la *trajectoire d'apprentissage* entière de l'élève depuis ses premiers exercices, en examinant comment les différentes attitudes apparaissent, disparaissent et se stabilisent. Pour effectuer une telle étude nous pensons utiliser une classification incrémentale du type de COBWEB (Fisher, 1987) , qui permet de traiter les exercices séquentiellement.

Enfin, on peut s'interroger sur la possibilité d'appliquer notre méthode de génération de profils à d'autres domaines, ce qui n'est pas toujours une chose facile dans les EIAH. Tout d'abord, il faut souligner que nous avons introduit relativement peu de mécanismes « ad-hoc ». Certes les descripteurs utilisés sont spécifiques au domaine (comment ne pas l'être), mais les concepts généraux tels que : la notion d'étape élémentaire de traitement[10], la représentation par des couples <contexte-action>, le mécanisme de classification, et la notion et l'explication des attitudes restent indépendants du domaine traité. On peut donc considérer que la portabilité de notre approche est conditionnée à la définition de deux critères. D'une part il est nécessaire que des comportements élémentaires de l'élève puissent être repérés, codés et mémorisés, sous la forme de couple

[10] Notons que nous n'utilisons pas directement les règles présentes dans ANAIS. Ce logiciel est utilisé pour faire apparaître les étapes élémentaires à partir desquelles nous travaillons.

<contexte-action>, ce qui n'est pas le cas de tous les EIAH. D'autre part, ces comportements doivent posséder une granularité qui permette l'analyse. En d'autres termes, il faut une certaine variabilité dans les exemples pour disposer de suffisamment de descripteurs : un EIAH dans lequel l'élève rentrerait directement des entiers en réponse à des questions de calcul ne conviendrait pas. Mais inversement, il faut que ces comportements n'aient pas un grain trop fin qui rendrait difficilement intelligibles les résultats : un EIAH qui ne stockerait comme trace de l'élève que des positions de clics de souris ne permettrait pas un diagnostic pertinent.

Remerciements
Nous remercions les membres de l'équipe Did@tic du laboratoire Leibniz-IMAG travaillant sur le projet APLUSIX qui ont permis la réalisation de ce travail en nous fournissant les données et surtout en nous apportant l'expertise didactique nécessaire à l'avancement de nos travaux.

8 Références

BISSON G., BRONNER A., GORDON M.B., NICAUD J.F., RENAUDIE D. (2003) : *Analyse statistique de comportements d'élèves en algèbre*. Actes de la conférence EIAH'2003 (Strasbourg), pp. 67-78.

BALACHEFF N. (2002) : Contribution à la réflexion sur la recherche sur les EIAH. Les Technologies en éducation : perspectives de recherche et questions vives, Baron G.-L. & Bruillard E. Eds, Paris.

BECK, J.E., WOOLF, B.P (2000). High-level Student Modeling with Machine Learning, Proceedings of the Intelligent Tutoring Systems Conference, 584-593.

BERKHIN, P. (2002). Survey of clustering data mining techniques. Tech. rep., Accrue Software, San Jose, CA. http://citeseer.nj.nec.com/berkhin02survey.html

DILLENBOURG P., SELF J.A. (1992) : A framework for learner modelling. Interactive Learning Environments Vol 2, pp 111-137.

ESPOSITO, F., LICCHELLI O., SEMERARO, G. (2004). Discovering Student Models in e-learning Systems. Journal of Universal Computer Science 10(1), 47-57.

FISHER, D.H. (1987), Knowledge Acquisition via Incremental Conceptual Clustering, Machine Learning 2:139-172

JAIN A.K., MURTY M.N., FLYNN P.J. (1999). Data Clustering: A Review. ACM Computing Surveys, vol. 31, pages 265-322, September 1999

NICAUD J-F, BOUHINEAU D., HUGUET T. (2002). « The Aplusix-Editor: A New Kind of Software for the Learning of Algebra » . Proceedings of the 6th International Conference on Intelligent Tutoring Systems. Lecture Notes In Computer Science, Springer Vol. 2363. Biarritz, France and San Sebastian, Spain, June 2-7, 2002 Pages: 178 – 187

MOSTOW, J. (2004). Some useful design tactics for mining ITS data, Proceedings of the ITS2004 Workshop on Analyzing Student-Tutor Interaction Logs to Improve Educational Outcomes, August, 2004.

RENAUDIE D (2005). Méthodes d'apprentissage automatique pour la modélisation de l'élève en algèbre. Thèse de l'INPG soutenue le 14 janvier 2005.

SISON R., SHIMURA, M. (1998). Student Modeling and Machine Learning. International Journal of AI and Education 9, 128-158.

TSIRIGA V., VIRVOU, M. (2002). Initializing the Student Model using Stereotypes and Machine Learning, In Proceedings of the 2002 IEEE International Conference on Systems, Man and Cybernetics.

VERGNAUD G. (2001) : *Forme opératoire et forme prédicative de la connaissance*. Actes du colloque GDM 2001, Portugais J. Eds, pp 6-27.

Annexe A : Descripteurs utilisés pour caractériser les « mouvements »

Descripteurs de contexte		
Description de l'argument sur lequel porte le mouvement		
Argument.coté	{gauche, droite}	Place de l'argument dans l'expression
Argument.position	{début, milieu, fin, isolé}	Sa position dans le membre, la valeur « isolé » signifiant que le membre ne contient que lui.
Argument.complexe	{vrai, faux}	Présence d'un opérateur + ou –, ex : • 2x (Argument.complexe = faux) • (3x+8) (Argument.complexe = vrai)
Argument.polynôme	{vrai, faux}	Présence d'une variable, ex : 2**x**
Argument.racine	{vrai, faux}	Présence d'une racine, ex : √3x
Argument.puissance	{vrai, faux}	Présence d'un exposant, ex : x^2
Argument.fraction	{vrai, faux}	Présence d'une fraction, ex : -3/4
Argument.coefficient	{vrai, faux}	Coefficient numérique explicite, ex : **3**x
Argument.négatif	{vrai, faux}	Argument de signe négatif, ex : -4
Argument.signeImplicite	{vrai, faux}	Signe de l'argument est implicite, ex : **(+)**4x
Argument.opérateur	{+, -, ×, ÷, ^, √}	Opérateur de plus haute priorité dans l'argument
Argument.catégorie	{add, mult, puiss}	Catégorie de cet opérateur : • add={+, -} ; mult={×, ÷} ; puiss={^, √}
Description du terme impliqué par le mouvement.		
Terme.polynôme	{vrai, faux}	Même signification que pour l'argument
Terme.fraction	{vrai, faux}	idem
Terme.puissance	{vrai, faux}	idem
Terme.racine	{vrai, faux}	idem
Description de l'expression		
Expression.type	{équation, inéquation}	Type de l'expression
Expression.sens	{gauche, droite, NA}	Orientation de l'inéquation (NA si équation)
Expression.large	{vrai, faux, NA}	Inégalité large ou stricte (NA si équation)
Expression.Polynôme	{vrai, faux}	Même signification que pour l'argument
Expression.Fraction	{vrai, faux}	idem
Expression.Puissance	{vrai, faux}	idem
Expression.Racine	{vrai, faux}	idem
Expression.nulGauche	{vrai, faux}	Indique que le membre gauche = 0
Expression.nulDroite	{vrai, faux}	Indique que le membre droit = 0

Descripteurs d'action		
Description de l'argument sur lequel porte le mouvement		
Argument.opérateur	{+, -, ×, ÷, ^, √}	Même signification que pour le contexte
Argument.catégorie	{add, mult, puiss}	idem
Argument.négatif	{vrai, faux}	idem
Argument.mêmeOpérateur	{vrai, faux}	Indique si l'opérateur a été modifié
Argument.mêmeCatégorie	{vrai, faux}	Indique si la catégorie a été modifié
Argument.mêmeSigne	{vrai, faux}	Indique si le signe de l'argument a été modifié
Description de l'expression		
Expression.type	{équation, inéquation}	Même signification que pour le contexte
Expression.sens	{gauche, droite, NA}	idem
Expression.large	{vrai, faux, NA}	idem
Expression.mêmeType	{vrai, faux}	Indique si le type de l'expression a été modifié
Expression.mêmeSens	{vrai, faux, NA}	Indique si le sens de l'expression a été modifié
Information fournie par ANAIS		
Transformation.juste	{vrai, faux}	indique la validité du mouvement

Annexe B : Exemples de diagnostics automatiques

```
Attitude numero 44, effectif 6.
Correcte : 0 Incorrecte : 6.

Diagnostic :
Attitude incorrecte consistant en un mouvement dans une inequation au sens
stricte d'un terme , initialement positif en position additive ce mouvement
s'effectuant avec changement du type de l'equation avec ou sans changement de
signe , la position finale etant additive.

Exemples :
n <= t - w ----------> 0 = t - w - n
n > w - t ----------> 0 = w - t + n

Explication :
L'eleve a adopte une attitude incorrecte, mais neanmoins coherente, ce qui
correspond a l'application rigoureuse d'une regle fausse.
```

```
Attitude numero 48, effectif 14.
Correcte : 14 Incorrecte : 0.

Diagnostic :
Attitude correcte consistant en un mouvement dans une equation d'un terme,
initialement negatif en position additive ce mouvement s'effectuant avec
changement de signe , la position finale etant additive.

Exemples :
r - v =  - nx ----------> r - v + nx = 0
v - r =  - n ----------> v - r + n = 0

Explication :
L'eleve a adopte une attitude a la fois correcte et coherente pour ce type de
contexte, ce qui correspond a la connaissance et l'application d'une regle
juste.
```

```
Attitude numero 50, effectif 14.
Correcte : 12 Incorrecte : 2.

Diagnostic :
Attitude incoherente consistant en un mouvement dans une equation d'un terme ,
initialement positif ou negatif en position multiplicative ce mouvement
s'effectuant sans changement de signe , la position finale etant divisive.

Exemples :
r - v = n*x ----------> (r - v) / (n) = x
v - r =  - n*x ----------> (v - r) / (n) = x

Explication :
Le rapport des effectifs corrects-incorrects etant tres desequilibre, on peut
considerer cette attitude comme etant coherente.
```