



HAL
open science

Un système d'assistance au diagnostic de compétences en algèbre élémentaire

Stéphanie Jean-Daubias

► **To cite this version:**

Stéphanie Jean-Daubias. Un système d'assistance au diagnostic de compétences en algèbre élémentaire. Sciences et Techniques Educatives, 2002, STE - Sciences et techniques éducatives, 9 (1-2), pp.171-200. hal-01574293

HAL Id: hal-01574293

<https://hal.science/hal-01574293>

Submitted on 12 Aug 2017

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.

Un système d'assistance au diagnostic de compétences en algèbre élémentaire

Stéphanie Jean-Daubias

*Laboratoire d'Ingénierie des Systèmes d'Information¹
Université Claude Bernard - Lyon 1
Bâtiment Nautibus, 8 bd Niels Bohr, Campus de la Doua
69622 Villeurbanne Cedex
Stephanie.Jean-Daubias@bat710.univ-lyon1.fr*

RESUMÉ. La recherche que nous présentons dans cet article concerne la conception d'un environnement capable d'aider les enseignants dans l'évaluation des compétences des élèves en algèbre élémentaire. Afin d'obtenir une modélisation cognitive des connaissances des élèves, nous avons fait appel à une expertise didactique existante. Cette approche nous a conduit à automatiser un outil papier-crayon, mais également à adopter une démarche de conception interdisciplinaire, faisant intervenir didacticiens, enseignants et élèves. L'outil que nous proposons ne prétend pas être un système de diagnostic automatique, remplaçant l'enseignant dans sa tâche, mais au contraire un système d'assistance au diagnostic, aidant l'enseignant dans son travail d'évaluation des compétences des élèves. Nous présentons ici les résultats obtenus concernant le recueil des observables, le diagnostic de compétences, ainsi que la transmission des profils à l'enseignant.

ABSTRACT. In this paper, we present a research aiming to build an environment able to help teachers in assessing students' competencies in elementary algebra. To obtain a cognitive model of students' knowledge, we use an existing didactical expertise. This approach leads us not only to automate a pencil and paper tool, but also to adopt a multidisciplinary design method, involving educational scientists, teachers and students. The tool we propose does not pretend to be an automatic diagnosis, but a system that assists the teacher in the assessment of students' competencies. We present here the results we obtained concerning data gathering, diagnosis of competencies and transmission of the profiles to the teacher.

MOTS CLÉS : EIAH (Environnement Informatique pour l'Apprentissage Humain), diagnostic de compétences, IHM, pluridisciplinarité, algèbre.

KEY WORDS: ILE (Interactive Learning Environment), diagnosis of competences, HCI, multidisciplinary approach, algebra.

¹ Les travaux présentés ici ont été réalisés dans le cadre de notre thèse en informatique effectuée au Laboratoire d'Informatique de l'Université du Maine.

1. Introduction

Le projet PÉPITE² vise à étudier comment concevoir un système d'évaluation des connaissances des élèves, réellement utilisable par les enseignants dans le cadre de leur pratique professionnelle courante, en automatisant un outil papier-crayon existant [JEA 99]. Cet outil permet d'évaluer les compétences des élèves en algèbre élémentaire par rapport aux objectifs d'apprentissage lors de la transition entre deux cycles d'enseignement.

Ainsi, nous avons conçu un système opérationnalisant les résultats de recherches en didactique des mathématiques avec pour préoccupation, l'intégration du système conçu à l'enseignement. Ce souci d'usage se retrouve aux différentes étapes de notre recherche. Il se retrouve tout d'abord à travers la recherche d'une fiabilité maximale des observables (qui ne peuvent pas se limiter à des QCM) sans laquelle le diagnostic ne peut être fiable, ce qui passe par un travail important autour de la conception de l'interface de test et de son adaptation aux habitudes des élèves. Ensuite, le diagnostic doit être multidimensionnel pour prendre en compte à la fois l'analyse de différents aspects de l'algèbre et l'analyse de différents types de réponses proposées par les élèves, en particulier des réponses à des questions ouvertes. Notre préoccupation d'usage se retrouve enfin au niveau de la conception de l'interface enseignant, qui doit permettre à ces derniers de s'approprier les profils construits automatiquement et ainsi d'intégrer l'utilisation du système à leurs pratiques.

Dans cet article, nous présentons tout d'abord l'outil didactique sur lequel nous nous appuyons, ainsi que l'architecture du système conçu. Nous exposons ensuite les résultats de notre recherche en suivant les différentes phases de construction du profil des élèves.

2. Le travail didactique à l'origine du projet

Dans le projet PÉPITE, pour assurer la validité cognitive et la pertinence du diagnostic, nous nous sommes appuyée sur une étude didactique et cognitive validée proposant un modèle de la compétence algébrique [GRU 95]. Cette étude a donné lieu à la conception d'un outil papier-crayon. Dans cette section, nous décrivons cet outil papier-crayon, les utilisations qui en sont faites, ainsi que les raisons qui ont conduit à son automatisation.

2.1. L'outil de diagnostic papier-crayon

Grugeon propose dans [GRU 95] un modèle de la compétence algébrique pour la transition entre les classes de troisième et de seconde de l'enseignement français

² PÉPITE est un projet pluridisciplinaire impliquant le Laboratoire d'Informatique de l'Université du Maine en informatique avec Martial Vivet, Élisabeth Delozanne, Pierre Jacoboni et le laboratoire DIDIREM de Paris VII en didactique des mathématiques avec Michèle Artigue et Brigitte Grugeon.

(élèves de 15 ans). Pour tenir compte de la complexité de la compétence algébrique et afin de mettre en évidence des cohérences dans le fonctionnement des élèves, elle propose une structure d'analyse multidimensionnelle. Ce modèle est opérationnalisé dans un outil de diagnostic des compétences algébriques qui comporte trois éléments :

– Un ensemble de tâches papier-crayon : les exercices (exercices techniques, exercices de mathématisation et exercices de reconnaissance) proposés aux élèves lors d'un test (la Figure 1 présente un exemple d'exercice de mathématisation qui permet d'étudier comment l'élève passe du cadre algébrique au langage naturel).

Complète le tableau en écrivant une phrase traduisant chaque étape du programme de calcul en face de l'expression algébrique correspondante.		
Etape 1	Soit un nombre de départ désigné par x	x
Etape 2	<i>Ajouter 3 à l'inverse du nombre de départ</i>	$-x + 3$
Etape 3	<i>Mettre le tout au carré</i>	$(-x + 3)^2$
Etape 4	<i>Ajouter 4 et diviser 1 par le nouveau nombre</i>	$\frac{1}{(-x + 3)^2 + 4}$
<i>Légende :</i>		
Énoncé		<i>Réponse de l'élève</i>

Figure 1 : Un exemple d'exercice de mathématisation.

– Une grille d'analyse multidimensionnelle analysant les tâches selon différents axes (les composantes) et permettant à l'enseignant d'interpréter les productions des élèves en fonction du modèle de la compétence algébrique. Cette grille, sorte de guide indiquant comment interpréter les réponses d'élèves, comporte six composantes (type de traitement, utilisation des lettres, calcul algébrique, type de conversion, type de justification et connaissances numériques). Par exemple trois composantes sont concernées pour l'étape 2 de l'exercice présenté ci-dessus (cf. Figure 1) : avec la réponse donnée par l'élève pour cette étape (***Ajouter 3 à l'inverse du nombre de départ***), la grille propose l'analyse suivante : traitement incorrect pour la composante type de traitement, conversion incorrecte pour la composante type de conversion et règles de transformation incorrectes pour la composante type de justification).

– Les profils cognitifs des élèves sont des descriptions de plus haut niveau, rendant compte des connaissances des élèves. Ils sont établis en recoupant les résultats du diagnostic obtenus pour l'élève, par analyse transversale, composante par composante. L'analyse transversale permet d'identifier des cohérences dans le fonctionnement de l'élève et fait ressortir les principaux traits de son fonctionnement cognitif. Les profils comportent trois niveaux de description : une description quantitative des compétences algébriques en terme de taux de réussite ; une

4 Sciences et techniques éducatives.

description qualitative des cohérences de fonctionnement, composante par composante, en terme de modes de fonctionnement ; une description de la flexibilité entre les cadres (entre le cadre algébrique et les autres cadres étudiés dans le test : cadre graphique, cadre numérique, cadre géométrique et autres contextes), représentée par un diagramme. La Figure 2 donne un exemple de profil cognitif présentant ces trois niveaux de description. Nous avons choisi de présenter un exemple produit par PÉPITE afin de faciliter la lecture du profil.

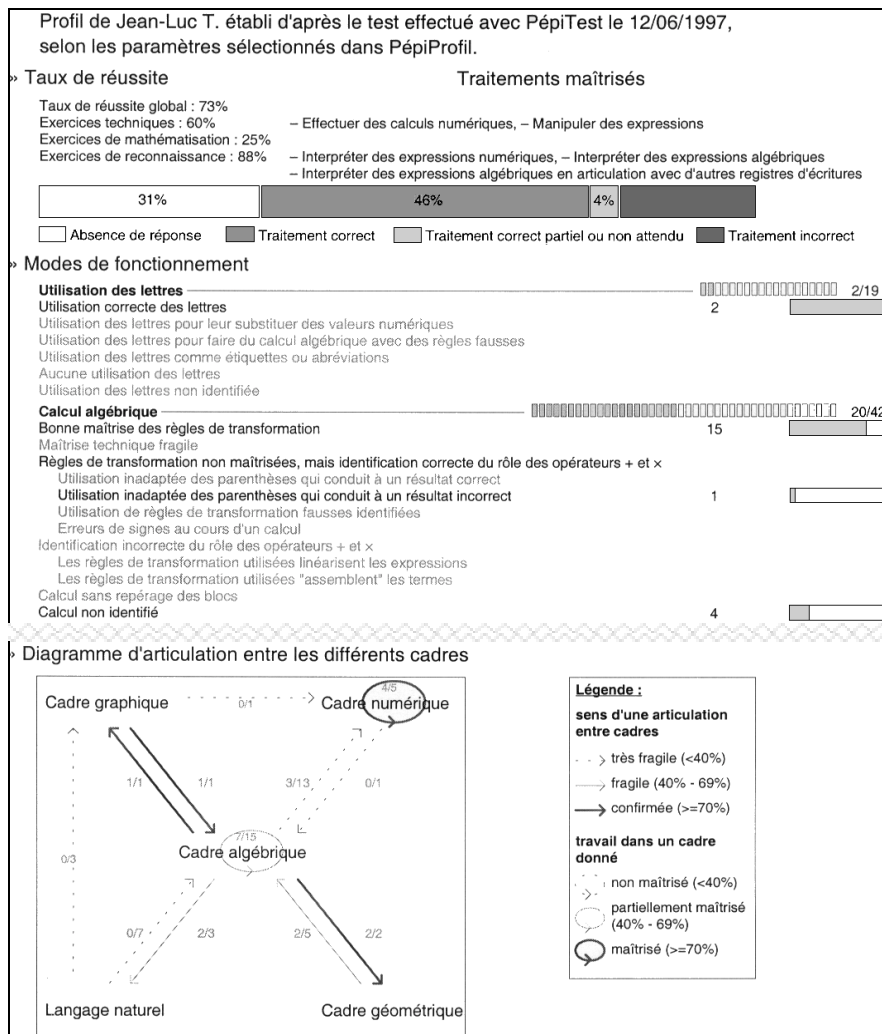


Figure 2 : Un extrait de l'impression du profil de Jean-Luc établi automatiquement.

Cet outil papier - crayon, conçu dans le cadre d'une recherche en didactique des mathématiques est utilisé en classe par quelques enseignants. La construction des profils par l'enseignant se fait en trois temps. Les élèves passent tout d'abord le test. L'enseignant analyse ensuite manuellement leurs productions en utilisant la grille d'analyse de la tâche et la grille d'analyse associée à chaque exercice. L'enseignant procède enfin à une analyse transversale pour construire les profils.

2.2. Utilisation de l'outil papier-crayon et nécessité de son opérationnalisation

L'objectif de l'outil est d'identifier les compétences des élèves, pour permettre à l'*enseignant* de s'appuyer sur ces compétences pour aider l'élève à construire de nouvelles connaissances en adaptant son enseignement et les activités qu'il propose. Cette outil peut également être utilisé par l'*équipe enseignante* qui peut exploiter les profils pour former des groupes d'élèves différenciés, fondés non plus seulement sur les performances des élèves en termes de notes, mais également sur leurs compétences identifiées grâce au test.

Même si quelques enseignants l'utilisent déjà, la diffusion de l'outil de diagnostic à plus grande échelle est toutefois limitée par les difficultés de sa mise en œuvre. Le travail d'analyse des tests est en effet beaucoup plus lourd que celui demandé par la correction de devoirs "classiques", car l'enseignant doit non seulement corriger la réponse de l'élève, mais également identifier tous les modes de fonctionnement mis en œuvre par l'élève pour chaque question, en se référant à la grille d'analyse. Outre l'investissement en temps, l'utilisation de l'outil de diagnostic demande un apprentissage du modèle de la compétence algébrique qui le sous-tend. Ces contraintes rendent difficile l'utilisation courante de l'outil de diagnostic papier-crayon par des enseignants en dehors d'un cadre de recherche et d'expérimentation. Par ailleurs, les enseignants qui ont connaissance de cet outil souhaitent vivement pouvoir en utiliser les résultats sans pour autant accepter le lourd travail que sa mise en œuvre nécessite ; d'autres ressentent le besoin d'un tel outil d'évaluation sans connaître de travaux pouvant y répondre. Le projet PÉPITE répond ainsi à un besoin d'opérationnalisation d'un travail de recherche exprimé par les didacticiens, mais aussi aux besoins de nombreux enseignants.

De notre point de vue d'informaticienne, automatiser un outil de diagnostic papier-crayon existant nous permet, d'une part, de disposer d'un corpus important dès la phase de conception (600 tests papier-crayon sont disponibles). Ce corpus sert de point de départ à la conception, à la fois du logiciel élève (en donnant une idée des réponses qui pourront être saisies à l'interface) et du module de diagnostic (en permettant d'identifier les types de réponses qui devront être analysées). D'autre part, l'automatisation d'un outil existant, en nous permettant de nous appuyer sur une expertise didactique, nous permet de partir de tâches spécialement prévues pour construire le profil des élèves. Enfin, la réutilisation d'un outil papier-crayon validé nous permet de définir précisément des critères de validation pour le système à concevoir s'appuyant sur la comparaison de l'outil papier - crayon à sa version

informatisée. Pour ces raisons, nous avons adopté pour le système réalisé une démarche adaptée de la structure de l'outil papier-crayon.

3. Architecture du système conçu

PÉPITE comporte trois modules (cf. Figure 3) qui correspondent aux trois étapes de la construction manuelle des profils : passation du test par les élèves, analyse automatique des réponses, puis analyse transversale et présentation des profils à l'enseignant.

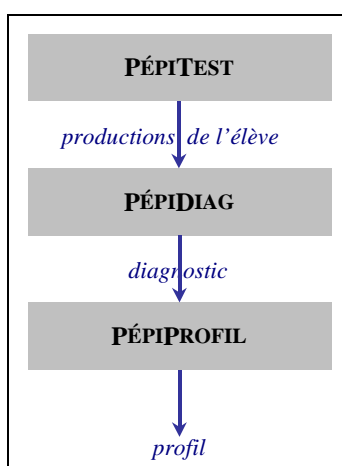


Figure 3 : L'architecture de PÉPITE.

Détaillons le rôle de chaque module :

– **PÉPITEST**, le logiciel élève, propose aux élèves les exercices qui sont une adaptation des tâches papier-crayon au support informatique, et recueille leurs réponses. Ce logiciel est amené à être utilisé une seule fois par les élèves, en début de classe de seconde.

– **PÉPIDIAG** analyse les productions des élèves, à partir des observables fournis par PÉPITEST, selon la grille d'analyse multidimensionnelle, en remplissant une matrice correspondant aux différents modes de fonctionnement, pour les composantes mises en jeu.

– **PÉPIPROFIL** établit les profils des élèves par analyse transversale de la matrice de diagnostic complétée par PÉPIDIAG et les présente aux utilisateurs, c'est-à-dire aux enseignants.

Dans la suite de l'article nous présentons les principaux résultats de notre recherche pour chaque étape de la modélisation des compétences : recueil des observables, diagnostic de compétences, analyse transversale et présentation des profils à l'enseignant.

4. Le recueil des observables par le logiciel - élève

Du point de vue du diagnostic, le rôle du logiciel élève, PÉPI TEST, outre la présentation des exercices aux élèves, est de recueillir leurs réponses et de les stocker de sorte qu'elles soient utilisables par le module de diagnostic. Les exercices de PÉPI TEST comportent des questions fermées apportant des informations sur les connaissances des élèves grâce à l'analyse didactique poussée réalisée en amont. Ces informations doivent toutefois naturellement être complétées par une étude des réponses produites spontanément par les élèves (étude capitale, notamment concernant les justifications). C'est pourquoi PÉPI TEST comporte également des questions entièrement ouvertes, sans contrôle sur la réponse de l'élève.

La fiabilité des observables est primordiale. Les productions doivent être représentatives des compétences des élèves, il nous faut donc évaluer les élèves dans un environnement leur permettant d'exprimer leurs compétences comme ils le font en environnement papier - crayon. Ce besoin de compatibilité entre l'environnement habituel papier-crayon et l'environnement d'évaluation informatisé a des conséquences sur la conception de PÉPI TEST. La « médiatisation » d'un outil papier-crayon est un domaine qui pose des questions à la fois en informatique et en didactique. Plusieurs travaux, parmi lesquels [ART 95] [BAL 94] [CHE 96] [NAN 90] [RAB 95], qu'ils correspondent à des problématiques d'ordre didactique, informatique ou résolument orientées EIAH, mènent à des conclusions communes : le passage d'un environnement papier-crayon à un environnement informatisé n'est pas anodin, il pose des problèmes plus complexes que ceux attachés généralement à la notion de médiatisation [JEA 00a]. Cette articulation entre les deux types d'environnements fait en effet intervenir non seulement un processus de médiatisation, mais également un processus de transposition informatique [BAL 94] et de genèse instrumentale [RAB 95]. C'est pourquoi nous préférons utiliser par la suite le terme de *transfert* pour désigner cette articulation.

4.1. Le transfert de l'environnement et des exercices

Afin de faciliter la tâche de l'élève, nous proposons des outils informatisés qui tendent à remplacer les outils « physiques » disponibles dans l'environnement papier-crayon. L'indispensable mise en place de ces outils est l'occasion de transferts. En automatisant le test papier-crayon, nous ne transférons pas seulement un ensemble d'exercices, mais également les différents outils familiers [MAC 97a] utilisés pendant le test : feuilles, crayons, gomme, ainsi que calculatrice et cahier de brouillon. Ceci est d'autant plus important que dans le domaine sur lequel porte le test, les mathématiques, ces outils ont un rôle majeur.

Le transfert de l'environnement pose deux problèmes. Tout d'abord, la mise à disposition des outils peut influencer sur le comportement de l'élève (certains élèves peuvent se dire : « puisqu'un tel outil existe, c'est qu'il doit servir quelque part »). Mais surtout, l'ensemble des outils virtuels proposant dans PÉPI TEST de remplacer les outils familiers sont différents de ceux utilisés habituellement par les élèves (par

exemple, la gomme virtuelle ne peut pas fonctionner exactement comme la gomme « physique »).

Ce délicat travail de transfert a demandé une analyse fine des besoins des élèves (facilitée par l'existence d'un important corpus de réponses d'élèves au test papier-crayon), ainsi qu'un travail important au sein de l'équipe pluridisciplinaire, il a également donné lieu à plusieurs tests du logiciel. Le travail réalisé nous a permis de déterminer un certain nombre de règles à suivre lors d'un transfert d'exercices (cf. Figure 4).

①	S'approcher au mieux des tâches papier-crayon en tenant compte de l'analyse didactique afin d'extraire de la session les mêmes informations que celles obtenues avec le test papier-crayon.
②	Éviter les manipulations de pages en proposant toutes les informations nécessaires sur un même écran.
③	Décomposer les exercices en questions indépendantes lorsque cela est possible afin que les exercices ne soient pas trop longs et qu'ils puissent être présentés sur un écran unique.
④	Décomposer les questions lorsqu'elles demandent à l'élève de faire plusieurs tâches en même temps.
⑤	Modifier la tâche quand l'informatique peut apporter un <i>plus</i> ou proposer une autre façon de présenter les choses, sans appauvrir le diagnostic (par exemple en proposant les schémas que l'élève peut manipuler).
⑥	Rendre les écrans plus attrayants que les exercices papier-crayon, en particulier en utilisant la couleur et des schémas.
⑦	Éviter au maximum la saisie de texte et d'expressions algébriques par l'élève pour limiter la perte de temps, la gêne de l'élève vis-à-vis du dispositif informatique.
⑧	Favoriser l'utilisation du copier - coller et du glisser - déplacer quand c'est possible.
⑨	Ne pas proposer de zones de saisie de grande taille alors que l'on attend une réponse brève (certains élèves peuvent avoir tendance à « remplir » la zone de saisie).
⑩	Ne pas proposer de zone de saisie lorsque la réponse ne sera pas analysée.

Figure 4 : Règles appliquées pour le transfert des exercices.

Ces règles sont indépendantes du domaine d'application du logiciel. Nous avons traité séparément les problèmes plus spécifiques à l'algèbre, en particulier la manipulation des expressions algébriques, que nous ne traitons pas ici.

4.2. Effets producteurs et réducteurs du transfert

Même si notre but était de limiter le plus possible les perturbations, le transfert du test papier-crayon a eu, nous nous y attendions, des conséquences sur les productions des élèves. Les expérimentations conduites ont permis d'identifier un certain nombre d'effets réducteurs et d'effets producteurs du transfert sur les observables. Que PÉPITEST facilite le travail de l'élève en lui proposant par exemple des schémas cliquables ou des termes à utiliser pour construire sa réponse (sans pour autant donner d'indication quant à la réponse à donner), qu'il permette des comportements nouveaux comme l'utilisation du copier - coller ou qu'il perturbe

l'élève en contraignant la saisie de ses réponses, en particulier pour les expressions algébriques, PÉPITEST agit comme un microscope cognitif [TWI 93], c'est-à-dire qu'il met parfois en évidence des comportements qui n'apparaissent pas avec le test papier-crayon, débloquent l'élève ou mettant en évidence ses incertitudes. Ces perturbations qu'en tant qu'informaticienne nous considérons comme des biais pouvant perturber le diagnostic, sont au contraire considérées par les didacticiens de l'équipe comme des révélateurs des connaissances des élèves, allant dans le sens d'une évaluation plus précise des difficultés des élèves. Citons deux exemples. Le premier concerne une élève qui travaillait sur les expressions algébriques par équivalence en environnement papier – crayon et qui, avec PÉPITEST a produit une écriture en pas à pas enchaînés [GRU 95]. PÉPITE met ainsi en évidence une connaissance fragile chez cette élève, alors que son enseignante pensait cette connaissance maîtrisée. Le deuxième exemple concerne l'exercice présenté à la Figure 8 dont le fonctionnement est décrit dans la section 5.2.3. Dans ce cas, la représentation de l'exercice à l'interface, en débloquent un grand nombre d'élèves, a permis d'obtenir un taux de réponse beaucoup plus élevé avec le logiciel que dans l'environnement papier – crayon.

4.3. *Évaluation de PÉPITEST*

Pour la validation de PÉPITEST, nous nous sommes appuyée sur la distinction faite par Senach entre utilisabilité et utilité [SEN 93], tout en étant consciente que ces deux points sont liés. L'utilisabilité concerne la capacité du logiciel à permettre à l'utilisateur d'atteindre facilement ses objectifs : pour PÉPITEST l'utilisabilité concerne donc la qualité de l'interface du point de vue de l'élève. L'utilité s'intéresse à l'adéquation du logiciel aux objectifs de haut niveau du client : pour PÉPITEST, il s'agit d'étudier la capacité du logiciel à rendre compte du comportement de l'élève pour établir le diagnostic demandé par l'enseignant.

Du point de vue de l'**utilisabilité**, l'interface de PÉPITEST a été évaluée à toutes les étapes de la conception : application des standards des interfaces graphiques, application des recommandations ergonomiques [BAS 93], soumission des prototypes à des jugements d'experts, à des tests informels d'acceptabilité auprès d'utilisateurs hors contexte d'utilisation, et enfin à des expérimentations en situation réelle dans des classes. La référence à des travaux d'IHM sert de base à notre réflexion sur l'ergonomie. Nous avons toutefois dû adapter les résultats de ces travaux à nos besoins spécifiques. PÉPITEST est en effet un EIAH dédié à l'évaluation des connaissances, ce qui fait de lui un cas particulier à différents niveaux. Les résultats de cette étude sur la nécessité d'une adaptation des résultats des recherches en IHM aux spécificités des EIAH est détaillée dans [JEA 00b].

L'évaluation de PÉPITEST du point de vue de son **utilité** comporte deux parties. La première partie consiste à vérifier que les observables obtenus à partir de PÉPITEST sont représentatifs du comportement de l'élève. À un niveau micro, il s'agit d'établir par comparaison si les productions d'un élève avec PÉPITEST correspondent à ses productions habituelles (c'est-à-dire aux copies qu'il rend

habituellement ou à ses cahiers). À un niveau macro, il s'agit de comparer l'ensemble des productions des élèves avec PÉPITEST aux productions correspondant au test papier-crayon, pour voir si le changement d'environnement provoque ou non une réduction du spectre des réponses, c'est-à-dire si on retrouve dans les productions avec PÉPITEST, l'ensemble des productions trouvées dans le test papier-crayon. La seconde partie de l'évaluation de PÉPITEST consiste à vérifier que ces observables permettent à un didacticien (et plus tard à PÉPIDIAG) de construire des profils et que ces profils sont équivalents aux profils papier-crayon. Plus précisément, nous cherchons à montrer que les profils construits manuellement à partir des traces de PÉPITEST sont équivalents aux profils construits manuellement à partir des observables du test papier-crayon.

PÉPITEST est entièrement réalisé, il a été expérimenté à plusieurs reprises en situation réelle en classe, avec au total une centaine d'élèves. Les résultats obtenus sont très satisfaisants, aussi bien au niveau du respect des critères de validation concernant l'utilité, qu'au niveau de l'utilisabilité de l'interface. Les différentes expérimentations ont en effet d'une part, permis de valider l'interface, d'autre part, elles ont montré qu'il n'y a pas de réduction du spectre des réponses (au niveau macro, on retrouve dans les corpus issus de l'utilisation de PÉPITE, tous les types de réponses prévus dans l'analyse a priori et trouvés dans les corpus papier - crayon) et que les observables fournis par PÉPITEST sont suffisamment riches et fiables pour permettre le diagnostic.

4.4. Bilan concernant le recueil des observables

Nous considérons le recueil des observables comme une étape cruciale de la modélisation des compétences. Pour cette raison, nous avons porté un soin particulier à la conception de l'interface du logiciel élève. Nous avons tout d'abord travaillé avec des didacticiens, des enseignants, mais également des élèves. Nous avons ensuite caractérisé le transfert qui intervient lorsqu'on informatise un test papier-crayon, en particulier en définissant des règles à suivre lors d'un transfert et lors de la conception d'un éditeur d'expressions algébriques. Nous avons enfin conçu, nous le verrons dans la section suivante, des exercices qui facilitent le recueil des observables, et même parfois le diagnostic. Ce dernier point confirme le rôle important que peut jouer l'interface dans un dispositif de diagnostic.

5. Le diagnostic de compétences

Le diagnostic de compétences est l'objectif principal de notre système. La difficulté du travail d'analyse réside à la fois dans la diversité des questions (portant sur plusieurs cadres : non seulement le cadre algébrique, mais également les cadres numérique, géométrique, graphique et le langage naturel), dans la diversité des réponses que peuvent proposer les élèves (comportant par exemple à la fois du langage naturel et des expressions mathématiques) et dans la complexité du modèle

de la compétence algébrique sur lequel nous nous appuyons. Rappelons en effet que pour des raisons didactiques, le diagnostic ne peut se limiter à l'analyse de questionnaires à choix multiples, ou de réponses préformées, il doit permettre l'analyse de réponses ouvertes. À la diversité des réponses proposées par les élèves, nous répondons en proposant différents types d'analyse adaptés aux différents types de réponses d'élèves : réponses en langage naturel, réponses sous forme d'expressions algébriques, réponses combinant langage naturel et expressions algébrique.

Suivant la suggestion de Baron qui indique que « des approches possibles reposent probablement sur la détermination d'objectifs de modélisation raisonnables, selon les situations d'apprentissage visées et l'utilisation effective des informations du modèle de l'apprenant » [BAR 98], dans PÉPITE, nous ne cherchons pas à modéliser le raisonnement de l'expert dans sa tâche de diagnostic des connaissances des élèves, mais plus raisonnablement à opérationnaliser le diagnostic en recueillant et en procéduralisant l'expertise didactique.

Le diagnostic est construit localement en fonction des questions et des éléments à identifier, il est ensuite testé puis, en fonction des résultats des tests, le diagnostic est affiné et éventuellement généralisé. Cette démarche incrémentale nous a permis d'obtenir des résultats testables avec des enseignants sans attendre que le diagnostic soit complet.

Dans cette section, nous commençons par décrire la matrice de diagnostic qui rassemble les résultats de l'analyse, nous présentons ensuite brièvement les différentes techniques d'analyse mises en œuvre, détaillées dans [JEA 00a] et [JEA 02]. Avant de conclure, nous mettons en évidence les liens qui existent entre le module de diagnostic et l'interface élève.

5.1. La matrice de diagnostic

Le résultat du diagnostic est porté dans une *matrice de diagnostic* (cf. Figure 5). Les 40 colonnes correspondent aux modes de fonctionnement des six composantes étudiées dans le test (par exemple, pour la composante *type de traitement*, T0 correspond à *absence de traitement*, T1 à *traitement correct*, T2 à *traitement correct partiel ou correct non attendu* et T3 à *réponse incorrecte*). Les codes X? correspondent à des modes de fonctionnement non identifiés par PÉPIDIAG. Les 55 lignes correspondent aux questions (1a pour exercice 1 question a...). Dans les cases, 0 indique un mode de fonctionnement possible dans la question, 1 indique que le mode de fonctionnement y est mis en œuvre par l'élève. Les cases vides correspondent aux modes de fonctionnement non étudiés dans la question.

Numéros des questions

Composantes

Traitement Utilisation des lettres Calcul algébrique Conversion Type de justification Connaissances numériques

Q	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10
Q1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
Q2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
Q3	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
Q4	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
Q5	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
Q6	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
Q7	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
Q8	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
Q9	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
Q10	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	

Figure 5 : Extrait de la matrice de diagnostic.

En reprenant l'étape 2 dans l'exemple de la Figure 1, pour la ligne correspondant à la question, PÉPIDIAG porterait un 1 dans les cases correspondant aux colonnes T3 (traitement incorrect), C3 (conversion incorrecte) et M3 (règles de transformation incorrectes).

5.2. L'analyse des réponses

À chaque question correspond un ensemble de composantes étudiées, donné par l'analyse a priori. Pour chaque composante, PÉPIDIAG doit identifier le ou les modes de fonctionnement mis en œuvre par l'élève dans la question. L'analyse de la tâche donne, pour chaque question, la liste des réponses généralement proposées par les élèves, ainsi que les modes de fonctionnement qui doivent y être associés. PÉPIDIAG ne cherche pas à comprendre la réponse de l'élève, mais à identifier des éléments particuliers, déterminés a priori. Selon le type d'éléments à identifier, le logiciel utilise des méthodes de diagnostic différentes. PÉPIDIAG compare successivement la réponse de l'élève à plusieurs réponses préenregistrées, réponses qui correspondent aux réponses fréquemment rencontrées et auxquelles l'analyse a priori associe un ou plusieurs modes de fonctionnement. Si aucune réponse préenregistrée ne correspond, le fonctionnement de l'élève est considéré comme non identifié par le système. Les modes de fonctionnements non identifiés pourront éventuellement être ensuite complétés par l'enseignant à l'aide de PÉPIPROFIL.

5.2.1. Les données à analyser

Les productions d'élèves issues de PÉPITEST que PÉPIDIAG analyse, comportent différents type de réponses : les réponses aux questions fermées, et parmi les réponses aux questions ouvertes, les réponses en langage naturel, les réponses sous forme d'expressions algébriques et les réponses mixtes faisant intervenir du langage naturel et des expressions algébriques.

Pour les questions ouvertes, les élèves sont totalement libres dans leur réponse. Ils peuvent entrer aussi bien des réponses en langage naturel que des expressions algébriques. Dans les justifications, avec PÉPITEST comme sur papier-crayon, ces deux types de réponses sont souvent combinés. La réponse de l'élève est alors sous une forme que nous appelons *langage « mathurel »*, mélange de langage naturel et de langage mathématique. Par ailleurs, produire des réponses sur ordinateur est, nous l'avons vu, un facteur perturbant pour un élève. D'une part la production d'expressions algébriques est rendue difficile par le dispositif disponible, d'autre part, comme le montrent des corpus de dialogue homme - machine [LUZ 95], certains utilisateurs ont tendance à modifier leur syntaxe en saisissant des phrases au clavier (que ce soit par appauvrissement ou par enrichissement), afin que l'ordinateur puisse « comprendre » leurs entrées. En fait, l'étude du corpus du test papier-crayon montre que dans leurs copies, les élèves ont plutôt tendance à utiliser une combinaison de langage naturel parlé (à la syntaxe et l'orthographe approximatives, auxquelles peuvent s'ajouter des fautes de frappe) et de langage pseudo-mathématique. Le *langage « mathurel »* utilisé habituellement dans un environnement papier-crayon par les élèves, devient donc *langage « informathurel »* une fois médiatisé, c'est-à-dire lorsque les élèves doivent produire leurs réponses sur ordinateur.

5.2.2. Les techniques d'analyse

À chacun de ces types de réponse correspond une technique d'analyse.

Pour les **réponses aux questions fermées**, à chaque réponse proposée à l'élève dans l'interface, correspond un ensemble de modes de fonctionnement, fourni par l'analyse didactique a priori. L'analyse de PÉPIDIAG consiste à associer à chaque réponse ou ensemble de réponses, les modes de fonctionnement correspondant à la mise en œuvre de l'élève.

Pour les **réponses aux questions ouvertes** (réponses en langage naturel, réponses sous forme d'expressions algébriques et réponses mixtes), l'analyse commence par une normalisation (suppression des espaces multiples, suppression des parenthèses inutiles, ajout des multipliés implicites...). L'analyse diffère ensuite selon qu'il s'agit de réponses en langage naturel, sous forme d'expressions algébriques ou en langage informathurel.

Les **réponses en langage naturel** sont celles pour lesquelles l'élève utilise uniquement des phrases (sans expression algébrique), qui peuvent être correctes ou non, que ce soit du point de vue de la syntaxe et de l'orthographe ou du point de vue mathématique. L'expression des élèves y est entièrement libre et le corpus montre une très grande diversité dans les réponses. L'analyse est effectuée par recherche de mots-clés, elle dépend des questions, des points particuliers que l'on cherche à identifier (précisés par l'analyse didactique a priori).

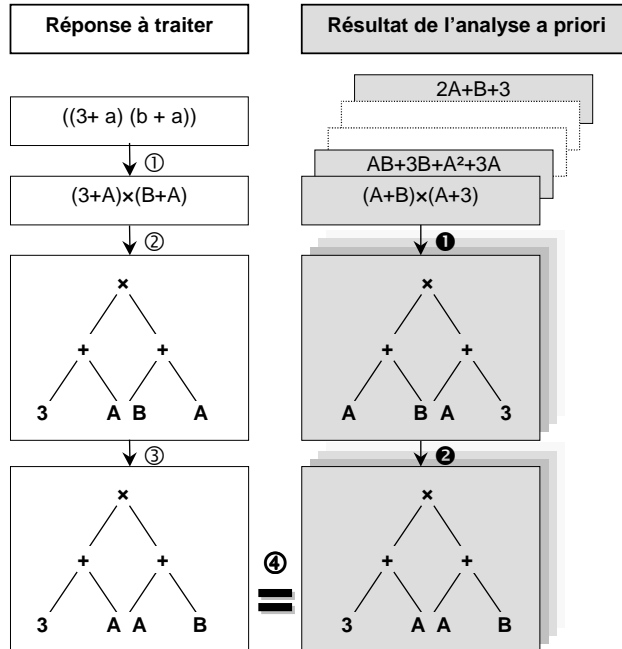


Figure 6 : Étapes de l'analyse d'une réponse sous forme d'expression algébrique.

De nombreuses productions sont sous forme **d'expressions algébriques**, elles correspondent aux exercices pour lesquels la réponse de l'élève est le résultat de calculs. L'analyse de ces réponses comporte plusieurs étapes : pré-traitements sur les expressions algébriques (cf. ① Figure 6), représentation des expressions sous forme d'arbres d'expression (②), normalisation des arbres (③), puis comparaison (④) aux arbres normalisés correspondant aux réponses attendues données par l'analyse a priori (cf. ① et ② Figure 6).

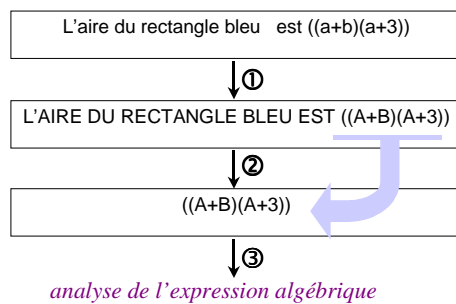


Figure 7 : Étapes de l'analyse d'une réponse en langage "informathurel".

Lorsque PÉPIDIAG cherche à identifier des expressions algébriques dans une **réponse en langage « informathurel »**, l'analyse comporte plusieurs étapes : après une normalisation de la réponse (cf. ① sur la Figure 7), l'expression algébrique est extraite de la chaîne (②). La démarche est ensuite la même que celle adoptée pour l'analyse des expressions algébriques (③) décrite ci-dessus.

5.2.3. Lien entre interface élève et diagnostic

Dans de nombreux cas, des modifications de l'interface élève ont facilité le diagnostic. Nous présentons ici deux exemples mettant en évidence le lien entre interface élève et diagnostic.

The screenshot shows the PÉPIDIAG software interface. At the top, there is a menu bar with 'Fichier', 'Édition', 'Outils', and 'Aide'. Below the menu is a toolbar with mathematical symbols like '+', '-', 'x', '÷', '√', '1', '2', '3', 'a', '=', '≠', '≈', '[', ']', and a calculator icon. A row of numbered tabs from 1 to 22 is visible, with tab 11 selected.

The main content area is divided into several sections:

- Exemple d'un programme de calcul :**
 - Étape 1 : Soit un nombre de départ désigné par x x
 - Étape 2 : Ajouter 4 à ce nombre $x + 4$
 - Étape 3 : Multiplier le résultat par -7 $-7(x + 4)$
- 1ère partie** / **2ème partie** (selected)
- Instruction :** Pour chaque étape, traduisez l'expression algébrique par une phrase. Vous pouvez cliquer sur les termes proposés ci-contre.
- Étape 1 :** Soit un nombre de départ désigné par x x
- Étape 2 :** ajouter 3 à l'inverse du nombre de départ $-x + 3$
- Étape 3 :** mettre le résultat au carré $(-x + 3)^2$
- Étape 4 :** ajouter 4 au résultat et prendre l'opposé de l'expression $\frac{1}{(-x + 3)^2 + 4}$

On the right side, there are three panels of suggested terms:

- Verbes:** ajouter, diviser, retrancher, prendre, multiplier, mettre.
- Noms:** nombre de départ, expression, carré, nombre, cube, résultat, exposant, l'inverse, double, l'opposé, triple.
- Mots de liaison:** et, par, le, de, à, puis, sur, l', du, au.
- Chiffres:** A grid of numbers from 0 to 9.

Figure 8 : Exercice donnant lieu à des réponses en langage naturel contraint, la réponse d'Amandine.

L'exercice présenté à la Figure 8 demande à l'élève de produire une phrase correspondant à chaque étape d'un programme de calcul donné. Dans cet exercice, le langage naturel est contraint, puisqu'il passe principalement par l'utilisation de termes préformés proposés à l'interface (l'élève peut cependant taper du texte librement en complément), ce qui incite les élèves à utiliser un lexique restreint et des tournures de phrases prédéfinies, rendant l'analyse plus simple. Les termes proposés permettent de construire les différentes réponses correctes (comme *mettre le résultat au carré*, *élever l'expression au carré* et *prendre le carré de ce nombre*), ainsi

que les réponses erronées attendues ou non par l'analyse didactique. L'analyse par PÉPIDIAG de ces différentes réponses ne pose pas de problème, car dans ces réponses en langage naturel contraint, comme pour les réponses en langage naturel, on cherche à identifier des éléments particuliers.

Cet exercice donne un exemple de situation où le travail de conception de l'interface, en améliorant les observables, permet de faciliter le diagnostic.

Dans certains cas, une modification de l'interface, en explicitant ce sur quoi porte la réponse demandée, permet de lever le problème de l'extraction de l'expression algébrique dans la réponse de l'élève (cf. Figure 9). Ce remplacement n'est fait que s'il ne gêne en rien l'étude du comportement de l'élève ?

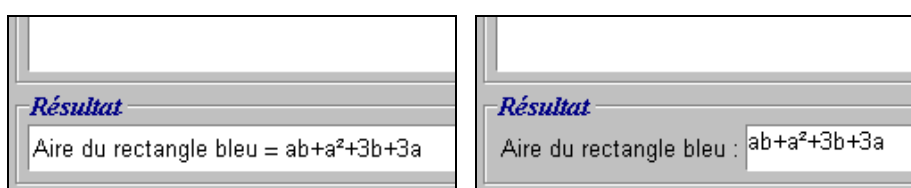


Figure 9 : Modification de l'interface simplifiant l'analyse des réponses.

5.2.4. Bilan sur le fonctionnement de PÉPIDIAG

Pour tous les types de réponse et pour tous les modes de fonctionnement étudiés dans le test, PÉPIDIAG propose une analyse. Toutefois, tous les exercices ne sont pas encore entièrement analysés. Le test comporte 55 questions, qui se décomposent en 91 items. Nous entendons par item, un élément de réponse (un ensemble de cases à cocher, une zone de saisie pour une justification ou un résultat, etc.), une question peut comporter plusieurs items (par exemple des boutons radio pour répondre par vrai ou faux et une zone de saisie pour les justifications).

Le tableau donné en Figure 10 récapitule pour chaque type de réponse les techniques d'analyse mises en œuvre dans PÉPIDIAG et présente la répartition des 91 items du test entre ces différents types, ainsi que le taux d'analyse obtenu par PÉPIDIAG. Le taux de réussite de 45% pour les réponses en langage informathurel correspond aux questions dont l'analyse n'est pas encore entièrement effectuée par PÉPIDIAG (seules certaines composantes y sont traitées). Il s'agit principalement de questions donnant lieu à des calculs comportant plusieurs étapes.

Globalement, ce sont 83% des items qui sont entièrement analysés, ce qui correspond à 71% des questions. Les items partiellement analysés donnent toutefois également des informations intéressantes. Du point de vue des enseignants, l'information la plus pertinente est le taux d'analyse en termes de modes de fonctionnement : 87% des modes de fonctionnement sont attribués.

Type de réponse	Technique d'analyse mise en œuvre	Répartition des items	Taux d'analyse
Questions fermées	Association à chaque réponse ou à chaque ensemble de réponses, des modes de fonctionnement mis en œuvre par l'élève.	42%	100%
Questions ouvertes	Langage naturel Recherche de mots clés pour identifier des points particuliers. Mise en place d'un langage naturel contraint pour faciliter le diagnostic.	3%	100%
	Expressions algébriques Pré-traitements sur les expressions, représentation sous forme d'arbres d'expression, normalisation des arbres et comparaison aux arbres normalisés correspondant aux réponses attendues données par l'analyse a priori.	23%	100%
	Langage informathurel Extraction de l'expression algébrique contenue dans la réponse de l'élève, puis traitement de cette expression selon la méthode décrite précédemment.	32%	45%

Figure 10 : Les techniques d'analyse mises en œuvre dans PÉPIDIAG, par types de réponse, leur importance relative dans PÉPITEST et le taux d'analyse obtenu.

5.3. Évaluation de PÉPIDIAG

La validation du module de diagnostic consiste à vérifier l'exactitude et la pertinence du diagnostic établi par PÉPIDIAG. Nous cherchons donc à montrer que les profils construits automatiquement par PÉPIDIAG à partir des productions de PÉPITEST sont équivalents aux profils construits manuellement à partir des mêmes productions de PÉPITEST.

Techniquement, l'évaluation que nous avons effectuée a consisté à comparer la matrice de diagnostic proposée par PÉPIDIAG à celle, construite à la main, proposée par un expert. Étant donnée la charge de travail demandée par l'analyse manuelle des productions d'élèves, une comparaison à grande échelle était impossible, nous avons donc procédé à quelques études de cas. Nous avons choisi cinq élèves de niveaux différents (d'après les indications données par leur enseignant) afin de pouvoir identifier les écarts d'analyse pour différents types de productions.

Globalement, les différences entre matrices complétées automatiquement et matrices complétées manuellement sont faibles. Les différences constatées sont de trois types :

- L'expert propose un diagnostic, mais pas PÉPIDIAG. Ceci correspond aux cas où PÉPIDIAG n'analyse pas encore ou ne réussit pas à analyser la production de l'élève.

– Des différences non significatives apparaissent entre l'analyse de l'expert et celle de PÉPIDIAG, elles correspondent à des modes de fonctionnement proches (par exemple réponse correcte / réponse correcte non attendue). Une recherche didactique, menée par Lenfant, a montré qu'à partir des productions d'un même élève, le diagnostic établi manuellement peut varier sensiblement d'un enseignant à l'autre, en raison de différences d'appréciation [LEN 97]. Les profils construits par analyse transversale des différentes matrices de diagnostic sont pourtant quasiment identiques. Ceci montre la tolérance de l'outil de diagnostic vis-à-vis de certaines variations dans l'analyse.

– Des erreurs sont identifiées dans le diagnostic de PÉPIDIAG, elles correspondent à des incohérences entre le programme et l'analyse a priori. Les différences mises en évidence par cette évaluation ont alors permis de corriger le programme.

Globalement, la validité de PÉPIDIAG n'est donc que partiellement assurée, principalement en raison du faible taux d'analyse obtenu pour le langage informathurel.

5.4. Bilan concernant l'analyse des réponses

Pour la conception du module de diagnostic de notre système, nous avons adopté une démarche incrémentale, qui replace le diagnostic dans le contexte de l'ensemble du dispositif d'évaluation des connaissances, dispositif qui inclut l'interface élève. Cette démarche incrémentale nous permet de tester, de compléter et d'améliorer le diagnostic ou l'interface du logiciel élève, étape après étape dans notre projet.

L'existence du prototype actuel et les premiers résultats que nous obtenons montrent déjà la faisabilité du diagnostic.

6. L'analyse transversale

La matrice de diagnostic complétée par PÉPIDIAG donne des informations sur les procédures appliquées ponctuellement par les élèves dans une question mais ne rend pas compte de leurs conceptions de façon plus globale. Ces informations ne sont donc pas directement exploitables par un enseignant. Afin de rendre ces informations opérationnelles en mettant en évidence les cohérences dans le fonctionnement de l'élève, il est indispensable d'effectuer une analyse transversale du contenu de la matrice de diagnostic sur l'ensemble des exercices, calcul indépendant des modules de test et de diagnostic.

6.1. Méthode pour l'analyse transversale

La Figure 11 montre le fonctionnement général de ce calcul, qui établit des recoupements entre la matrice de diagnostic et les caractéristiques des différentes questions. Ces recoupements indiquent pour chaque question comment lire la matrice

de diagnostic afin d'obtenir des informations selon les types d'exercices et les types de traitement, ainsi que les renseignements concernant les articulations entre cadres.

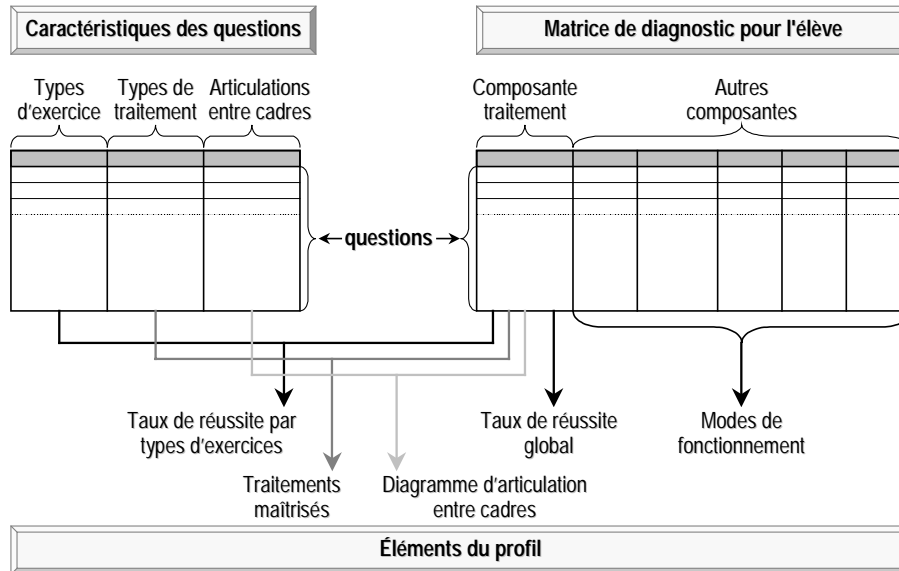


Figure 11 : Calcul des différents éléments du profil à partir des caractéristiques des exercices et de la matrice de diagnostic.

Pour chaque item du profil (taux de réussites, répartition des modes de fonctionnement pour chaque composante, articulations entre cadres...), le taux de réussite correspondant est calculé pour l'élève. À chacun des éléments calculés, sont ensuite appliqués des seuils indiquant le niveau de maîtrise de l'élève pour cet item (maîtrisé, partiellement maîtrisé ou non maîtrisé) ou, dans le cas des modes de fonctionnement, le mode de fonctionnement privilégié de l'élève. Ces seuils sont paramétrables, permettant à chaque enseignant d'adapter les seuils à ses exigences.

6.2. Évaluation de l'analyse transversale

L'analyse transversale correspond à un algorithme de calcul, sa validation consiste donc uniquement à vérifier la validité de cet algorithme. Nous cherchons à montrer que les profils construits automatiquement par analyse transversale de la matrice complétée par PÉPIDIAG sont équivalents aux profils construits manuellement à partir de la même matrice de PÉPIDIAG, c'est-à-dire que nous cherchons à montrer l'équivalence de la construction manuelle des profils et de la construction automatique des profils par PÉPIPROFIL.

Techniquement, l'évaluation s'est faite en comparant, pour tous les items du profil (taux de réussite, modes de fonctionnement et articulations entre cadres), la valeur calculée par l'expert à la valeur donnée par le logiciel. Les tests effectués ont montré que la construction automatique des profils est possible. Il y a correspondance entre les profils construits manuellement et les profils construits automatiquement.

7. Le logiciel enseignant

Une des originalités de PÉPITE est de proposer, avec PÉPIPROFIL un outil destiné directement aux enseignants. Les EIAH s'adressent en effet quasi exclusivement aux élèves, même si l'enseignant, en le choisissant, en le paramétrant, en assistant les élèves pendant son utilisation, est lui aussi souvent utilisateur du système (le meilleur exemple d'un EIAH laissant une place importante à l'enseignant est ROBOTEACH [LER 95]).

La problématique de notre travail sur le logiciel enseignant concerne l'intégration de PÉPIPROFIL aux usages des enseignants. Contrairement à PÉPITEST qui n'est utilisé qu'une seule fois par les élèves, PÉPIPROFIL est destiné à être utilisé à plusieurs reprises par les enseignants (au moins à chaque rentrée scolaire pour les enseignants de classes de seconde ou à chaque fin d'année pour les enseignants de classes de troisième). Une genèse instrumentale [RAB 95] va s'instaurer : les enseignants pourront, au fil des utilisations, s'approprier l'instrument qu'est pour eux PÉPIPROFIL, en s'adaptant à son fonctionnement et en lui conférant des usages qui iront certainement au-delà de ceux que nous envisageons actuellement.

Cette appropriation de PÉPIPROFIL et cette intégration aux pratiques passent par l'acceptation et l'appropriation du profil d'élève par l'enseignant. Non seulement la présentation du profil doit être claire et efficace, mais le système ne peut pas se contenter de présenter le profil qu'il a construit. PÉPIPROFIL doit pouvoir justifier le diagnostic qu'il a établi, permettre à l'enseignant de s'approprier le profil. PÉPITE est bien un *assistant* au service de l'enseignant dans sa tâche de diagnostic de compétence, ce n'est en aucun cas un système imposant un modèle figé des connaissances de l'élève. Cette nuance entre système de diagnostic automatique et assistant au diagnostic est cruciale pour l'acceptation du système par les enseignants, étant donnée leur réserve légitime à l'égard de l'évaluation automatique.

7.1. Présentation globale du profil

Le premier niveau de description de PÉPIPROFIL propose une vision globale du profil à travers trois parties : les taux de réussite et les traitements maîtrisés, les modes de fonctionnement et le diagramme d'articulation entre les cadres mis en jeu dans le test. Ces éléments sont proposés sous forme à la fois graphique et numérique. Cette présentation globale visuelle vise à permettre à l'enseignant d'appréhender rapidement le profil de l'élève (la Figure 12 présente une copie

d'écran de la première partie du profil, la Figure 2 donne un exemple d'impression de l'ensemble de cette présentation globale).

En plus de ces trois parties qui correspondent au profil papier-crayon, PÉPIPROFIL propose un résumé du profil sous forme textuelle. Le résumé est généré par le système à partir de phrases pré-enregistrées comportant des variables à instancier. Ce résumé est une aide apportée à l'enseignant dans l'interprétation du profil, puisqu'il reprend les points significatifs du profil. Mais notre but en proposant ce résumé est également de présenter le profil en utilisant une représentation complémentaire aux représentations graphiques et numériques : à l'enseignant de choisir celle qu'il préfère, ou d'en utiliser plusieurs.

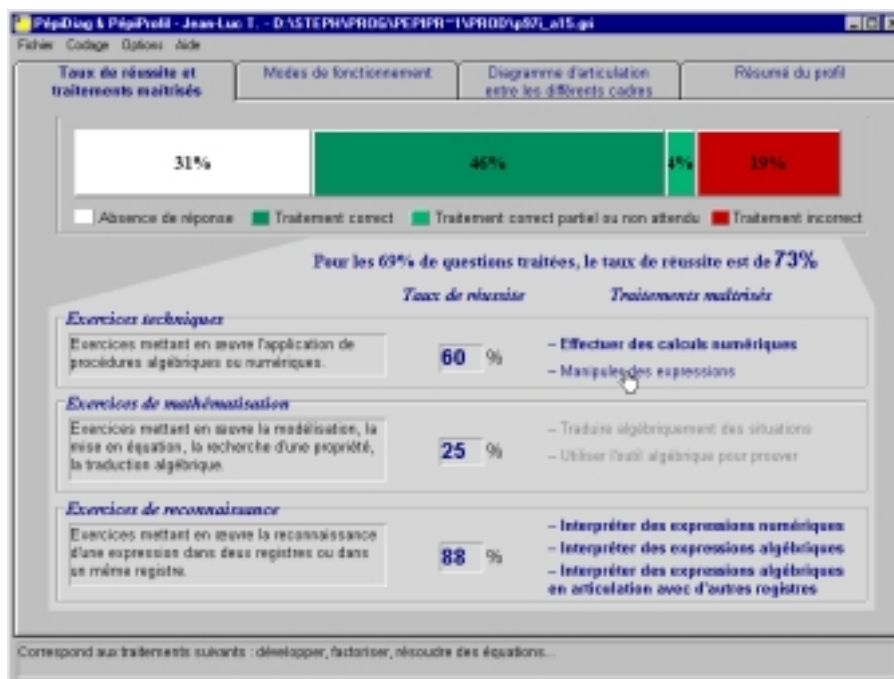


Figure 12 : PÉPIPROFIL : taux de réussite et traitements maîtrisés.

7.2. Présentation détaillée du profil

La présentation globale du profil peut être complétée, si l'enseignant le souhaite, par une présentation des détails du diagnostic proposé par PÉPIDIAG, détails justifiant les valeurs du profil. Pour chaque item du profil, il est ainsi possible de visualiser les détails du diagnostic dans un écran présentant, question par question, l'énoncé, la réponse de l'élève et son éventuelle justification, ainsi que le diagnostic

proposé par PÉPIDIAG (sous forme plus accessible que la matrice de diagnostic),
composante par composante (cf. Figure 13).

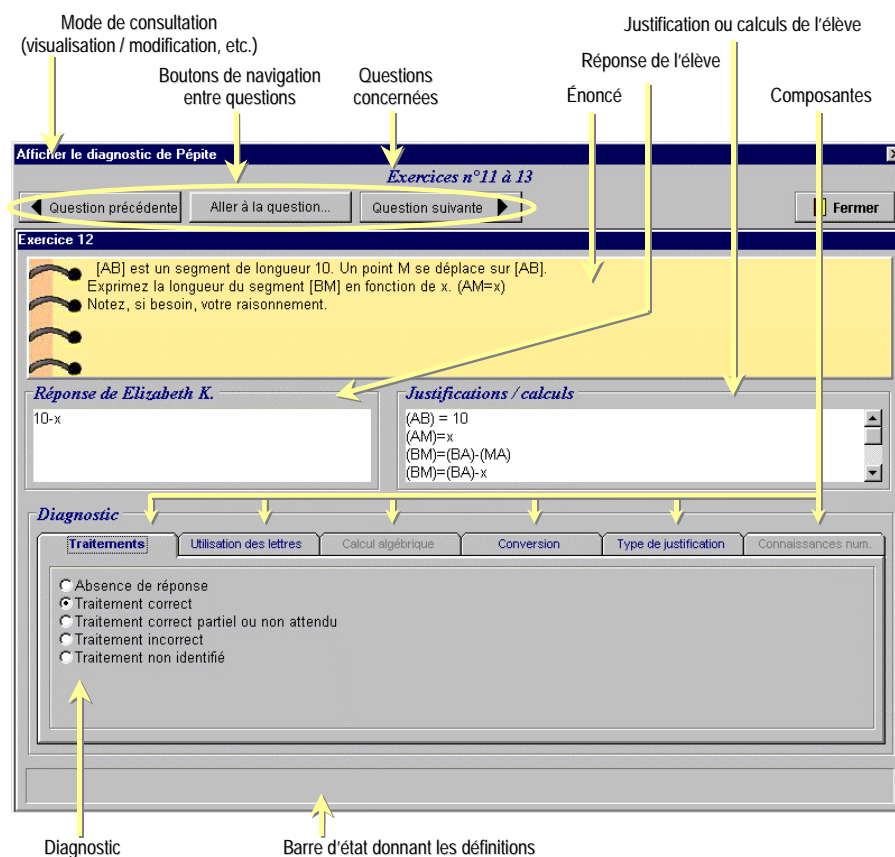


Figure 13 : PÉPIPROFIL, le détail du diagnostic.

7.3. Des profils paramétrables

Chaque enseignant a ses habitudes de travail et souvent des exigences particulières vis-à-vis des élèves. C'est pourquoi la plupart des éléments du profil (les seuils, mais aussi de façon plus anecdotique, les couleurs) sont paramétrables grâce à un écran spécifique. L'enseignant peut donc personnaliser PÉPIPROFIL pour l'adapter à ses usages et à ses exigences. Par exemple, il peut baisser le seuil à partir duquel les traitements sont considérés comme maîtrisés, ce qui revient à être moins exigeant vis-à-vis des élèves concernant ces traitements. L'existence de cette

possibilité de personnalisation peut être l'un des facteurs facilitant l'intégration de PÉPITE aux pratiques des enseignants.

7.4. Différents niveaux d'implication de l'enseignant

En proposant à l'enseignant différents niveaux d'implication dans l'appréhension des profils, PÉPIPROFIL aide l'enseignant à se construire son propre modèle de l'élève, en s'adaptant aux besoins de chaque enseignant, mais aussi, l'usage évoluant dans le temps, à chaque période. En permettant aux enseignants de choisir leur rôle dans le système, l'existence de ces différents niveaux d'implication facilite l'intégration du système aux pratiques des enseignants.

L'utilisation la plus élémentaire de PÉPIPROFIL consiste à **visualiser les différentes parties du profil**. Cette utilisation est celle, parmi les différentes utilisations proposées, qui demande le moins d'implication de l'enseignant. Toutefois, l'appropriation des différentes parties du profil demande déjà un travail important : l'enseignant doit comprendre les différentes représentations utilisées (expliquées dans l'aide).

La **consultation du détail du diagnostic** permet à l'enseignant de voir, question par question, les réponses de l'élève et le diagnostic que leur a associé PÉPIDIAG. Ces informations expliquent, pour chaque élément du profil, les raisons du résultat proposé (traitement considéré comme correct ou non, mais surtout détails sur les modes de fonctionnement concernant utilisation des lettres, calcul algébrique, conversion, type de justification et connaissances numériques). Cette utilisation correspond à la situation d'un l'enseignant qui cherche à comprendre le processus de construction du profil ou à vérifier le diagnostic.

PÉPIPROFIL comporte un mode de **complétion du diagnostic** proposé par PÉPIDIAG. En effet, pour certaines réponses d'élèves, PÉPIDIAG ne réussit pas à proposer de diagnostic. PÉPIPROFIL fait alors appel à l'enseignant en lui proposant de compléter lui-même le diagnostic. Ce mode demande une plus grande implication de l'enseignant : celui-ci doit non seulement « lire » le profil, mais aussi comprendre suffisamment le fonctionnement du diagnostic pour pouvoir le compléter.

L'enseignant peut enfin demander à visualiser le détail du **diagnostic** en mode **modification**. Dans ce cas, il consulte et vérifie le diagnostic proposé et éventuellement le modifie s'il est en désaccord avec PÉPIDIAG. L'implication de l'enseignant est alors à son maximum, pour utiliser ce mode il doit maîtriser suffisamment le fonctionnement du diagnostic pour remettre les propositions de PÉPIDIAG en question. Si l'enseignant modifie le diagnostic, la nouvelle matrice est enregistrée dans un fichier distinct. Au moment de la consultation du profil, l'enseignant peut donc choisir de visualiser le profil d'après le diagnostic automatique ou d'après son propre diagnostic.

La Figure 14 résume ces différents niveaux d'implication de l'enseignant dans l'utilisation de PÉPIPROFIL en les classant sur une échelle d'implication.

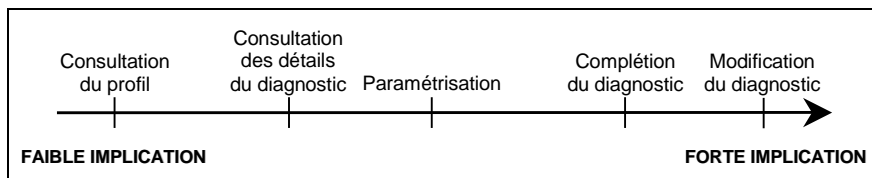


Figure 14 : Différentes utilisations de PÉPIPROFIL selon le degré d'implication de l'enseignant.

Selon nous, l'implication des enseignants dans l'utilisation de PÉPITE évoluera dans le temps. Il semble en effet naturel que, lors des premières utilisations, l'enseignant apprendra à connaître et à utiliser le système. Après cette phase de familiarisation, certains enseignants souhaiteront mieux comprendre le profil et la façon dont il a été construit et désireront donc voir le détail du diagnostic, puis progressivement le compléter et même le modifier. Une fois que les enseignants auront pris confiance en PÉPITE en consultant à plusieurs reprises le détail du diagnostic, ils accepteront sans doute les profils tels quels, sans systématiquement chercher à voir le profil détaillé, ce qui demande plus de temps. Cette mise en œuvre est toutefois capitale : de cette possibilité pour l'enseignant de remettre en cause le diagnostic dépend en partie son acceptation du système.

7.5. Évaluation de la présentation des profils

Pour l'évaluation de PÉPIPROFIL, nous avons adopté des méthodes d'évaluation d'IHM. L'évaluation de l'interface du logiciel enseignant a été faite d'abord pendant la conception, en particulier en respectant les critères ergonomiques, mais également par des jugements d'experts. Nous avons également procédé à plusieurs tests avec des didacticiens, puis avec des enseignants concernés par le projet PÉPITE. Enfin, nous avons fait des premiers tests avec des enseignants non impliqués dans le projet. De telles expérimentations sont particulièrement importantes pour tester, non seulement l'utilisabilité de l'interface, mais également l'utilité du logiciel [SEN 93].

Ces tests ont tout d'abord permis de mettre en évidence un grand intérêt des enseignants, à la fois pour le test et pour les profils construits. Ces tests ont également montré que les représentations proposées sont efficaces, même si une adaptation des termes employés reste indispensable. Nous pouvons d'ores et déjà affirmer que PÉPIPROFIL est utilisable par des enseignants, même non impliqués dans le projet. Cependant des expérimentations à long terme en situation réelle avec des enseignants correspondant à notre public cible doivent être conduites afin d'étudier de façon plus approfondie comment PÉPIPROFIL est réellement utilisé. Nous nous attendons en effet à des différences entre les utilisations que nous avons prévues et les utilisations effectives par les enseignants. Dans les premiers tests effectués, certains enseignants étaient plus intéressés par les profils de leurs élèves que par la façon dont les profils avaient été construits. Ceci nous a amenée à nous

interroger sur l'usage que les enseignants feront de PÉPIPROFIL, notamment concernant les possibilités de consultation et de modification du diagnostic.

7.6. Les utilisations possibles de PÉPITE

Afin de faciliter l'intégration de PÉPITE aux pratiques des enseignants, nous avons souhaité prévoir des utilisations qui peuvent en être faites, en essayant de déterminer le rôle de l'enseignant dans ces utilisations [VIV 90]. Nous avons conscience que les usages que nous prévoyons seront adaptés ou même ignorés par les enseignants. Ceux-ci se forgeront leurs propres usages de PÉPITE, à partir ou non des usages que nous proposons. Ils devront s'approprier le logiciel : en s'adaptant à ses possibilités et à ses contraintes, mais également en l'adaptant à leurs besoins et leurs habitudes de travail [RAB 95] [MAC 97b] [PER 89].

Nous présentons ici quelques-uns des usages envisagés.

L'utilisation « classique » des profils de PÉPITE, c'est-à-dire l'usage prévu à l'origine du projet, correspond à l'utilisation des profils par l'enseignant, en classe de seconde, pour identifier les connaissances et compétences des élèves en début d'année, et pour former des groupes d'élèves fondés non plus seulement sur les notes issues d'un devoir de rentrée, mais sur un bilan établissant plus explicitement les compétences des élèves. Un autre logiciel, PÉPICLASSE, partiellement développé, a pour objectif de faciliter cette utilisation en classe en construisant le profil de la classe, en aidant l'enseignant à constituer des groupes d'élèves selon certains critères choisis parmi les éléments du profil et en proposant des thèmes de travail adaptés à ces critères. Une autre façon d'exploiter la richesse des profils de PÉPITE, est d'utiliser les informations qu'ils donnent comme initialisation d'un modèle de l'apprenant dans un EIAH proposant des activités, des aides et des explications adaptées aux connaissances et compétences de l'élève, en s'appuyant sur le profil donné par PÉPITE.

PÉPITE proposant une possibilité d'évaluation qualitative des connaissances des élèves peu coûteuse en temps pour les enseignants, certains d'entre eux ont des idées d'utilisations qui vont au-delà des utilisations prévues à l'origine : utiliser les profils de PÉPITE comme bilan en fin de troisième ou comme test d'orientation en fin de seconde. Il semble par ailleurs intéressant de proposer aux élèves une version du profil s'adressant à eux, en reformulant les éléments du profil et en modifiant leur présentation.

Notons enfin que PÉPITE a rapidement été perçu comme un outil intéressant pour la formation des enseignants en mathématiques. Il est déjà utilisé à cette fin dans les Instituts Universitaires de Formation des Maîtres de Créteil, Beauvais et Montpellier.

8. Conclusion et perspectives

Les trois modules du système que nous avons réalisé sont développés : le logiciel élève, le module de diagnostic et le logiciel enseignant. Ils ont été testés à plusieurs reprises et donnent des résultats satisfaisants. Nous avons pu recueillir des observables fiables et représentatifs du fonctionnement des élèves. Nous avons également montré, en réalisant un prototype du module de diagnostic, que l'analyse multidimensionnelle automatique est possible pour des réponses d'élèves à différents types de questions, incluant des questions ouvertes. Nous avons enfin proposé un logiciel enseignant facilitant l'appropriation des profils par l'enseignant. Pour lever en partie les difficultés posées par la modélisation de l'apprenant, nous avons proposé une façon nouvelle d'envisager le diagnostic. D'une part, le point de départ de notre travail est une analyse didactique multidimensionnelle validée dont le résultat est l'outil papier-crayon que nous avons opérationnalisé. D'autre part, nous considérons le logiciel élève, permettant le recueil des observables, comme faisant partie intégrante du diagnostic. De la qualité des observables dépend en effet en partie la qualité du modèle de l'apprenant construit. Enfin, avec PÉPITE, nous nous adressons aux enseignants comme utilisateurs finals d'un logiciel et nous leur proposons non pas un système de diagnostic automatique, mais un système d'assistance au diagnostic.

Toutefois, nous avons fait le choix de concevoir un prototype du module de diagnostic permettant une analyse, même grossière, de tous les types de productions, en tenant compte de la spécificité de chaque exercice. Les techniques employées nous permettent certes d'obtenir un diagnostic pour la plupart des questions, mais elles sont très liées à chaque type d'exercice. La méthode adoptée conduit donc à un diagnostic dépendant du test.

Les principales perspectives de recherche qui apparaissent à l'issue de cette recherche concernent donc la réutilisabilité de notre travail, à travers la conception d'un système auteur, facilitant la conception de nouveaux tests, ainsi que l'amélioration du module de diagnostic, en faisant appel à des systèmes de diagnostic existants performants. À plus court terme, nous pouvons envisager de compléter le diagnostic en intégrant certaines informations fournies par la trace. Enfin, nous pourrions adapter les profils pour qu'ils puissent être utilisés par d'autres systèmes, en utilisant des ontologies pédagogiques.

9. Bibliographie

- [ART 95] ARTIGUE M., « Une approche didactique de l'intégration des EIAO à l'enseignement », in D. Guin, J.-F. Nicaud et D. Py, *Actes des 4^{èmes} journées EIAO*, pp. 17-28, Eyrolles, Paris, 1995.
- [BAL 94] BALACHEFF N. et VIVET M., « Didactique et intelligence artificielle », La pensée sauvage éditions, 1994.
- [BAL 98] BALACHEFF N., « Éclairage didactique sur les EIAH en mathématique », In : *Actes du Colloque de la Société de Didactique des Mathématiques du Québec*, Montréal, 1998.
- [BAR 98] BARON M., « Introduction », *Sciences et Techniques Éducatives*, numéro spécial modélisation de l'apprenant, Vol. 5, n°2, pp. 117-122, Hermès, 1998.
- [BAS 93] BASTIEN C. et SCAPIN D., « Critères ergonomiques pour l'évaluation des interfaces utilisateurs », RT n°156, INRIA, juin 1993.
- [BRU 94] BRULLARD É. et VIVET M., « Concevoir des EIAO pour des situations scolaires : approche méthodologique », in N. Balacheff et M. Vivet, *Didactique et intelligence artificielle*, pp. 273-302, La pensée sauvage éditions, 1994.
- [CHE 96] CHEVALLARD Y., « Les outils sémiotiques du travail mathématique », *Petit x*, n°42, pp33-57, 1995-1996.
- [GRU 95] GRUGEON B., « Étude des rapports institutionnels et des rapports personnels des élèves à l'algèbre élémentaire dans la transition entre deux cycles d'enseignement : BEP et Première G », thèse de doctorat, Université Paris VII, décembre 1995.
- [HAS 98] HASQUENOPH-BERNOU B., « Analyse des effets de la transposition informatique de tâches en algèbre élémentaire », Mémoire de DEA de didactique des mathématiques, Université Paris VII, 1998.
- [JEA 99] JEAN S., DELOZANNE É., JACOBONI P. and GRUGEON B., « A Diagnosis based on a Qualitative Model of Competence in Elementary Algebra », in proceedings of AI-ED 99, Le Mans, pp 491-498, 1999.
- [JEA 00a] JEAN S., « PÉPITE : un système d'assistance au diagnostic de compétences », Thèse de doctorat de l'Université du Maine, janvier 2000.
- [JEA 00b] JEAN S., « Application de recommandations ergonomiques : spécificités des EIAO dédiés à l'évaluation », *Actes des Rencontres Jeunes Chercheurs en IHM*, pp 39-42, 2000.
- [JEA 02] JEAN-DAUBIAS S., « Un système d'assistance au diagnostic de compétences », *Actes du 13^{ème} Congrès Francophone RFIA*, Vol. 3, pp.1053-1061, 2002.
- [LEN 97] LENFANT A., « Étude sur la transposition d'un outil de recherche destinée aux enseignants », Mémoire de DEA de didactique des mathématiques, Université Paris VII, 1997.
- [LER 95] LEROUX P., « Conception et réalisation d'un système coopératif d'apprentissage, étude d'une double coopération : maître/ordinateur et ordinateur/groupe d'apprenants », Thèse de doctorat, Université Paris VI, juin 1995.

- [LUZ 95] LUZZATI D., « Le dialogue verbal homme - machine – études de cas », Masson (eds.), 1995.
- [MAC 97a] MACKAY W., « Réalité augmentée : le meilleur des deux mondes », *La Recherche*, n°285, pp. 32-37, mars 1997.
- [MAC 97b] MACKAY W. et FAYARD A.-L., « Radicalement nouveau et néanmoins familier : les strips papiers revus par la réalité augmentée », *Actes d'IHM '97*, Poitiers, 1997.
- [NAN 90] NANARD J., « La manipulation directe en interface homme – machine », thèse d'état, Université de Montpellier II, 1990.
- [PER 89] PERRIAULT J., « La logique de l'usage, essai sur les machines à communiquer », Flammarion, 1989.
- [RAB 95] RABARDEL P., « Les hommes et les technologies - approche cognitive des instruments contemporains », Série psychologie, Armand Colin, 1995.
- [SEN 93] SENACH B., « L'évaluation ergonomique des interfaces homme – machine », in J.-C. Sperandio éditeur, *L'ergonomie dans la conception des projets informatiques*, pp. 69-122, Octares éditions, 1993.
- [SHN 97] SHNEIDERMAN B., « Designing the user Interface: Strategies for Effective Human-Computer Interaction », Addison-Wesley Publishers, 1997 (troisième édition).
- [TWI 93] TWIDALE M., "Redressing the Balance: the Advantages of Informal Evaluation Techniques for Intelligent Learning Environments", *Journal of Artificial Intelligence in Education*, Vol. 4, n°2, pp. 155-178, 1993.
- [VAN 99] VAN LABEKE N., « Prise en compte de l'utilisateur enseignant dans la conception des EIAO, Illustration dans Calques 3D », Thèse de doctorat de l'Université Henri Poincaré, Nancy I, 1999.
- [VAN 88] VANLEHN K., "Student Modeling", in *Foundations of Intelligent Tutoring Systems*, Chapter 3, pp. 55-78, 1988.
- [VIV 90] VIVET M., "Uses of ITS: Which role for the teacher?", in *New Directions for Intelligent Tutoring Systems*, NATO ASI series, Vol. F91, Springer-verlag, Sintra, 1990.
- [WEN 87] WENGER E., "Artificial Intelligence and Tutoring Systems, Computational and Cognitive Approaches to the Communication of Knowledge", Morgan Kaufmann Publishers Inc., Los Altos, 1987.