



HAL
open science

Un système d'assistance au diagnostic de compétences

Stéphanie Jean-Daubias

► **To cite this version:**

Stéphanie Jean-Daubias. Un système d'assistance au diagnostic de compétences. RFIA 2001, 2001, Angers, France. pp.1053-1061. hal-01575482

HAL Id: hal-01575482

<https://hal.science/hal-01575482v1>

Submitted on 20 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

A system aiding diagnosis

Stéphanie Jean-Daubias

Laboratoire d'Ingénierie des Systèmes d'Information¹

Bâtiment Nautibus - UFR d'Informatique - Université Claude Bernard - Lyon 1
8, Bd Niels Bohr - 69622 Villeurbanne cedex - France
Stephanie.Jean-Daubias@lisi.univ-lyon1.fr

Résumé

Ce travail, dans le cadre des recherches en EIAH (Environnement Informatiques d'Appren-tissage Humain) a pour objectif de modéliser les connaissances des élèves en construisant leur profil cognitif. Nous considérons que le diagnostic commence dès la création des productions à analyser et continue jusqu'à la phase de transmission du profil à l'enseignant.

Dans cet article, nous ne présentons toutefois que la phase de diagnostic elle-même, en exposant les différentes techniques de diagnostic mises en œuvres pour analyser les différents types de réponses, que ces réponses soient sous forme de QCM, sous forme d'expressions algébriques, en langage naturel ou sous forme mixte, mélangeant algèbre et langage naturel.

Mots Clef

Diagnostic, EIAH.

Abstract

This research concerning ILE (Interactive Learning Environments) aims to model students' knowledge by building their cognitive profile. We consider that diagnosis begins with the building of productions to analyse and continues until the communication of the profile to the teacher.

In this paper, we just present the diagnosis stage, by showing the different techniques used to analyse different types of students' answers; whether they are answers to MCQ, algebraic expressions, in natural language or in an intermediate form, mixing natural language and algebraic expressions.

Keywords

Diagnosis, ILE.

1 Introduction

Ce travail de recherche, qui s'inscrit dans le cadre du projet pluridisciplinaire PÉPITE, a pour objectif de modéliser les connaissances des élèves en algèbre élémentaire en construisant leur profil cognitif (cf. [4] et [3] pour une présentation plus complète de ces travaux).

L'élaboration du profil de l'élève se fait en trois étapes auxquelles correspondent trois modules dans notre système : le recueil des observables avec le logiciel élève, PÉPITEST, le diagnostic établi par PÉPIDIAG, puis la présentation des profils à l'enseignant, avec PÉPIPROFIL. Chaque étape pose des problèmes qui lui sont propres.

Pour lever en partie les difficultés posées par la modélisation de l'apprenant, nous proposons une façon nouvelle d'envisager le diagnostic, prenant en compte chacune des trois étapes de l'élaboration du profil. D'une part, pour assurer la qualité et la validité du diagnostic, nous nous appuyons sur une analyse didactique multidimensionnelle validée dont le résultat est un outil papier - crayon qui aboutit à la création de profils cognitifs d'élèves, outil que nous avons informatisé. D'autre part, nous considérons le logiciel 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, PÉPIPROFIL s'adresse aux enseignants pour leur transmettre les profils construits automatiquement, en leur laissant la possibilité de remettre en cause ce diagnostic : nous proposons non pas un système de diagnostic automatique, mais un système d'assistance au diagnostic.

Le système conçu a été expérimenté à plusieurs reprises en classe auprès d'une centaine d'élèves. Un prototype du module de diagnostic permet d'analyser environ 75% des réponses et de construire les profils

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

qui correspondent aux connaissances des élèves. Pour le logiciel enseignant, les résultats des premiers tests auprès d'enseignants montrent que les représentations utilisées sont efficaces, mais les tests doivent être poursuivis.

Dans cet article, nous présentons de façon plus détaillée le module de diagnostic, même si, nous l'avons vu, la qualité du diagnostic dépend également de la qualité des données (et donc du module de test) et de la qualité de la transmission du profil à l'enseignant.

2 Le diagnostic dans PÉPITE

La difficulté du travail d'analyse réside à la fois dans la diversité des questions (portant sur l'algèbre, mais aussi sur les domaines qui y sont liés : cadre numérique, géométrique, graphique et 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. C'est en quoi notre travail se distingue des travaux proposant une analyse de réponses en EAO (Enseignement Assisté par Ordinateur).

2.1 Les données à analyser

Afin de construire des profils cognitifs riches, comportant à la fois des informations quantitatives et des informations qualitatives sur le fonctionnement cognitif des élèves (cf. Figure 3 pour un exemple de profil), les productions des apprenants qui servent de base au diagnostic ne peuvent se limiter à des questions fermées.

Le logiciel de test de PÉPITE propose donc à la fois des questions fermées qui assurent une analyse simple et parfaitement fiable, et des questions ouvertes, dont l'analyse est plus complexe, mais qui, en laissant l'apprenant entièrement libre de ses réponses, permettent de connaître les compétences qu'il peut mettre en œuvre spontanément, sans être guidé, aidé ou perturbé par les réponses des questions fermées. Ces deux types de questions se complètent pour obtenir un diagnostic le plus riche possible. D'autres questions sont intermédiaires et proposent des questions ouvertes auxquelles l'élève répond en utilisant des éléments de réponse préformés (cf. Figure 2 pour un exemple d'exercice de ce type et Figure 1 pour des exemples de réponses d'élèves).

ajouter à l'opposé du nombre de départ le nombre 3
l'inverse du nombre de départ puis ajouter 3
prendre l'opposé du nombre de départ et l'ajouter à 3
ajouter 3 à l'opposé du nombre de départ.
ajouter 3 à l'inverse nombre de départ

Figure 1 : Productions d'élèves pour l'exercice 11.

2.2 Le fonctionnement du diagnostic

Dans PÉPITE, nous automatisons un outil papier-crayon existant, issu des recherches en didactique des mathématiques de Brigitte GRUGEON [1]. Cet outil comporte, comme le système que nous avons réalisé, trois parties : le test proposé aux apprenants, une grille d'analyse multidimensionnelle permettant d'analyser les productions des élèves, et les profils eux-mêmes, construits ensuite par analyse transversale.

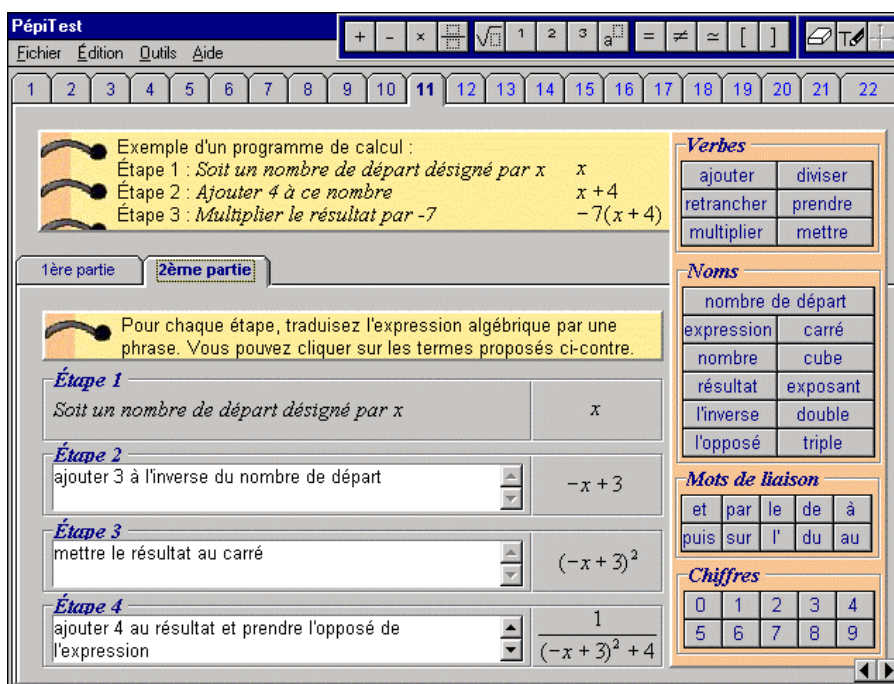


Figure 2 : Exercice donnant lieu à des réponses en langage naturel contraint, la réponse d'Amandine.

Le diagnostic automatique s'appuie donc sur cette grille d'analyse papier – crayon, explicitation de l'expertise didactique de Brigitte GRUGEON pour l'analyse des productions d'élèves (cf. Figure 4). Notre travail, pour le module de diagnostic a donc consisté en une opérationnalisation de cette expertise.

Suivant la suggestion de Monique Baron indiquant 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 » [1], 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 procéduralisant l'expertise didactique. L'idée étant de commencer par montrer la faisabilité du diagnostic avant de proposer des techniques de diagnostic performantes.

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.

Pour le module de diagnostic, notre questionnement initial comporte deux points principaux : tout d'abord, est-il possible d'expliciter l'expertise didactique de diagnostic ? Ensuite, étant donné la diversité des réponses des élèves et la diversité des éléments à identifier dans le diagnostic, est-il possible d'analyser automatiquement les productions d'élèves ? 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. C'est cette distinction que nous utiliserons pour la présentation des méthodes d'analyse employées dans PÉPIDIAG. Avant cela, nous décrivons la matrice de diagnostic qui rassemble les résultats de l'analyse.

2.3 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 pour *absence de traitement*, T1 correspond 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.

Figure 5 : Extrait de la matrice de diagnostic.

2.4 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 les modes de fonctionnement mis en œuvre par l'élève. 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 teste successivement plusieurs hypothèses sur le fonctionnement de l'élève, hypothèses qui correspondent aux réponses identifiées par les didacticiens dans l'analyse a priori. Si aucune hypothèse n'est vérifiée, le fonctionnement de l'élève est considéré comme non identifié. Les modes de fonctionnements non identifiés pourront ensuite être complétés par l'enseignant à l'aide de PÉPIPROFIL.

L'analyse des réponses aux questions fermées

Les réponses à des questions fermées regroupent : les réponses binaires de type vrai / faux ou oui / non (qui sont complétées dans PÉPITEST par des réponses ouvertes) ; les réponses aux questions à choix multiples ainsi que les réponses à sélectionner dans une liste (la réponse peut comporter un seul élément ou plusieurs) ; les réponses préformées à ordonner, ce cas est apparenté aux questions fermées car l'élève n'a pas de réponse à taper, il utilise les éléments qui lui sont proposés pour former sa réponse en les ordonnant.

Pour ce type de questions, 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. Cette technique d'analyse permet d'obtenir un diagnostic fiable pour toutes les réponses aux questions fermées.

Les pré-traitements sur les réponses aux questions ouvertes

Pour les trois types de réponses aux questions ouvertes (réponses en langage naturel, réponses sous forme d'expressions algébriques et réponses mixtes), l'analyse commence par des pré-traitements identiques. Ces pré-traitements consistent en une normalisation des chaînes de caractères constituant la réponse de l'élève, utile quelle que soit l'analyse qui en sera faite ensuite. En effet, qu'il s'agisse de langage naturel ou d'expressions algébriques, une réponse, doit donner lieu à la même interprétation, qu'elle soit donnée par exemple en caractères majuscules ou minuscules. Par conséquent, les caractères sont convertis en lettres majuscules, les espaces multiples sont supprimés et les caractères accentués sont remplacés par les caractères non accentués correspondants. L'analyse diffère ensuite selon le type de réponse.

L'analyse des productions en langage naturel

Les réponses en langage naturel sont celles pour lesquelles l'élève utilise uniquement des phrases (sans expressions algébriques), 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. Ce type de réponses concerne peu de questions dans PÉPITEST : essentiellement certaines justifications, ainsi que l'exercice 11, cas particulier qui fait l'objet d'une explication détaillée dans la section suivante.

①	on a deux même valeurs a, on a deux puissances, ils faut donc les multipliers
②	$a^{[3]} \times a^{[2]} = a^{[6]}$

Figure 6 : L'expression $a^3 a^2 = a^6$ est-elle vraie ? : exemples de justifications d'élèves.

Dans les justifications, l'expression des élèves est entièrement libre et les corpus montrent une très grande diversité dans les réponses. L'analyse des justifications consiste en grande partie à identifier le type de justification de l'élève : *justification en langage naturel* ou *par l'algèbre*, puis plus précisément *appel au légal* (mis en évidence par les termes « faut » et « donc » dans la justification ① de la Figure 6) ou *preuve par l'exemple* (cf. justification ② dans la Figure 6) ou le *contre-exemple*.

L'analyse est effectuée par recherche de mots clés, elle dépend des questions. Il ne s'agit en effet pas de comprendre le sens de ce qui a été écrit, mais

d'identifier les points particuliers étudiés dans l'exercice, précisés par l'analyse didactique a priori.

L'utilisation de cette technique permet d'obtenir un résultat fiable à condition d'exprimer toutes les possibilités et de tenir compte des variations possibles dans l'écriture des termes par les élèves. Cette méthode est adaptée à PÉPIDIAG dans la mesure où les réponses possibles sont en partie données dans l'analyse didactique a priori. Nous sommes consciente des limites de ces méthodes. Nous les avons utilisées afin de montrer la faisabilité du diagnostic.

L'analyse des productions en langage naturel, le cas de l'exercice 11

L'exercice 11 demande à l'élève d'écrire une phrase correspondant à chaque étape d'un programme de calcul donné (cf. Figure 2).

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.

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

L'analyse des productions sous forme d'expressions algébriques

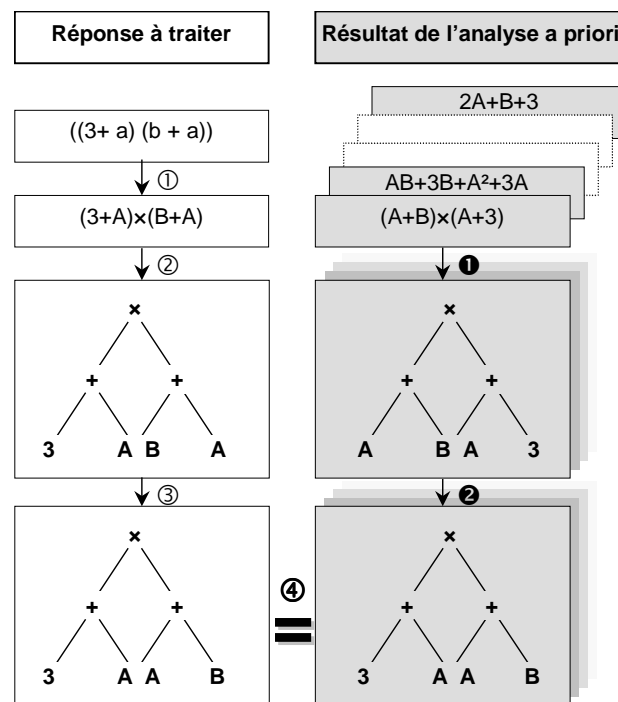


Figure 7 : Les étapes de l'analyse d'une réponse sous forme d'expression algébrique.

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 commence par des pré-traitements sur les expressions algébriques (cf. ① Figure 7), les expressions sont ensuite représentées sous forme d'arbres d'expression (②). Les arbres sont normalisés (③) avant d'être comparés (④) aux arbres normalisés correspondant aux réponses attendues données par l'analyse a priori (cf. ① et ② Figure 7).

Les *pré-traitements* sur les expressions algébriques, qui s'ajoutent aux pré-traitements appliqués à toutes les réponses, permettent de travailler sur des expressions sans caractères inutiles et sans implicites. Ils consistent en une suppression de tous les espaces, ainsi que des parenthèses et crochets inutiles, c'est-à-dire redondants sur une expression correctement parenthésée. Le système procède ensuite à une « normalisation » de la représentation de l'expression en homogénéisant la représentation des opérateurs, en ajoutant des signes multipliés dans les cas où ce signe est implicite, en parenthésant les expressions contenant des signes moins et des racines carrées. Remarquons que ces pré-traitements ne sont pas utilisés systématiquement. En effet, dans les exercices où PÉPIDIAG cherche à étudier l'utilisation des parenthèses par l'élève, la modification du parenthésage serait particulièrement malvenue.

Les *expressions algébriques* sont représentées sous forme d'arbres binaires étiquetés dont les nœuds internes sont les opérateurs et les sous-arbres gauches et droits sont les opérandes. L'étiquette des feuilles contient les symboles de variables ou de constantes. Ces constantes peuvent être aussi bien positives que négatives. Le signe moins n'est donc pas considéré comme opérateur. Les opérateurs sont donc les opérateurs binaires plus, multiplié, divisé, exposant, l'opérateur unaire racine carrée ainsi que les signes différent et égal qui sont également considérés comme des opérateurs puisqu'ils peuvent être nœuds internes. La construction d'un arbre à partir d'une expression algébrique se fait linéairement en respectant les priorités opératoires.

La *normalisation* des arbres d'expression a pour but de rendre comparables les arbres représentant différentes expressions. Elle consiste en l'application de plusieurs règles parmi lesquelles la commutativité de l'addition, de la multiplication et de l'égalité, la neutralité du zéro pour l'addition, de 1 pour la multiplication. L'application de ces règles se fait par simplification ou modification de l'arbre d'expression. Comme pour les pré-traitements sur les réponses d'élèves, ces normalisations ne sont pas

faites systématiquement, leur application peut dépendre des éléments à étudier. Selon la question, on indique donc la liste des règles à appliquer. Les règles mises en œuvre dans PÉPIDIAG permettent une comparaison fiable pour les expressions algébriques rencontrées dans le test.

L'analyse des productions en langage « informathurel »

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 PÉPIDI-TEST 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 un facteur perturbant pour un élève. D'une part la production d'expressions algébriques est rendue difficile par le dispositif disponible, d'autre part, des corpus de dialogue homme - machine [6] montrent que 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 de 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 » lorsqu'il est médiatisé, c'est-à-dire lorsque les élèves doivent produire leurs réponses sur ordinateur (la Figure 8 donne des exemples de telles productions).

a x a est une multiplication donc: 3+2= 5
c' est "a+a"qui est egal à 2a
L'expression a ² équivaut à axa et non pas à 2xa.
2a ² =2xa ² et non pas (2a) ² car cela serait égal à 4a ² .

Figure 8 : L'expression $a^2 = 2a$ est-elle vraie ? : exemples de justifications en langage « informathurel ».

Lorsque PÉPIDIAG cherche à identifier des expressions algébriques dans une réponse en langage « informathurel », l'analyse comporte plusieurs étapes : la réponse, sous forme de chaîne de caractères, subit tout d'abord les pré-traitements de normalisation appliqués à toutes les réponses aux questions ouvertes. L'expression algébrique est ensuite extraite de la chaîne. La démarche est ensuite la même que celle adoptée pour l'analyse des expressions algébriques décrite dans la section précédente.

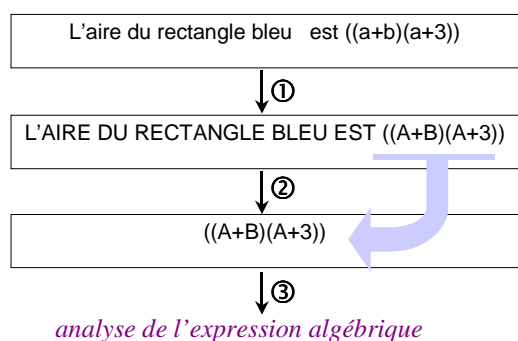


Figure 9 : Les étapes de l'analyse d'une réponse en langage "informathurel".

L'extraction d'une expression algébrique à partir d'une réponse en langage « informathurel » est un problème non trivial. Dans le cas où seule l'expression algébrique doit être analysée, il s'agit d'identifier dans la réponse de l'élève, l'élément de la réponse qui doit être analysé, sans tenir compte des termes ajoutés, qui peuvent être considérés comme du bruit. Ainsi, dans l'exemple présenté ci-dessus (cf. Figure 9), les termes ajoutés par l'élève pour annoncer sa réponse ("L'aire du rectangle bleu est") constituent du bruit, dans la mesure où seule l'expression algébrique $((a+b)(a+3))$ doit être analysée. L'extraction consiste à ignorer les mots qui précèdent l'expression.

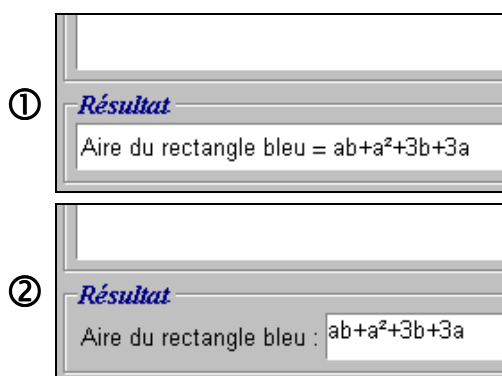


Figure 10 : Exemple de modification de l'interface simplifiant l'analyse des réponses.

Dans certains cas, une modification de l'interface, en explicitant sur quoi porte la réponse demandée, permet de lever le problème de l'extraction de l'expression algébrique de la réponse de l'élève (cf. Figure 10).

2.5 Évaluation 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, élément de réponse (un ensemble de cases à cocher, une zone de saisie pour une justification ou un résultat, etc.), une question pouvant 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 11 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. 55% items donnant lieu à des productions de langage « informathurel » ne sont pas encore analysés, le reste des items est entièrement analysé.

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.

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.

Type de réponse	Technique d'analyse mise en œuvre	Répartition des items	Taux d'analyse
Questions fermées	Association à chaque réponse ou chaque ensemble de réponse, les 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%
Questions "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 11 : Les techniques d'analyse de PÉPIDIAG.

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 donné la charge de travail demandée par l'analyse manuelle des productions d'élèves, une comparaison à grande échelle était impossible, nous n'avons donc pour l'instant procédé qu'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 Agnès 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 [5]. Les profils construits par analyse transversale des différentes matrices de diagnostic sont pourtant quasiment identiques. Ceci montre la tolérance de l'outil de diagnostic – traitant l'analyse des productions et l'analyse transversale – vis-à-vis de certaines variations dans le diagnostic.

- 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. Une évaluation à plus grande échelle reste souhaitable.

3 L'analyse transversale

La matrice de diagnostic complétée par PÉPIDIAG donne des informations sur les procédures appliquées par les élèves mais ne rend pas compte de leurs conceptions. 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, faisant ainsi ressortir les principaux traits de son fonctionnement cognitif, 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. Cette analyse transversale est effectuée par le module PÉPIPROFIL. L'outil papier - crayon ne comportait pas d'explicitation de cette analyse transversale. Nous avons donc dû travailler avec notre expert en didactique des mathématiques afin d'explicitier les techniques qui permettent le calcul des profils.

3.1 Méthode pour l'analyse transversale

Pour chaque item du profil (taux de réussites, répartition des modes de fonctionnement pour chaque composante, articulations entre cadres...) – cf. Figure 3 pour un exemple de profil –, 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.

3.2 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 donc 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.

4 Conclusion et perspectives

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

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.

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 manquent toutefois de généralité. Ces techniques nous permettent certes d'obtenir un diagnostic pour la plupart des questions, mais elles sont liées à chaque type d'exercice et même, dans certains cas, à un exercice. La méthode adoptée conduit donc à un diagnostic dépendant du test. Les principales perspectives qui apparaissent à l'issue de cette recherche concernent ainsi la réutilisabilité de notre travail, à travers la conception d'un système auteur permettant de créer de nouveaux tests (interface élève et module d'analyse) à partir d'exercices types à paramétrer. Nous pourrions

également améliorer du module de diagnostic en faisant appel à des systèmes de diagnostic performants existants, spécialistes des différents domaines traités dans le test.

Références

- [1] M. Baron, *Introduction*, Sciences et techniques éducatives, numéro spécial modélisation de l'apprenant, Vol. 5, n°2, pp. 117-122, Hermès, 1998.
- [2] B. Grugeon, *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.
- [3] S. Jean, *PÉPITE : un système d'assistance au diagnostic de compétences*, Thèse de doctorat, Université du Maine, 298 p, 2000.
- [4] S. Jean, É. Delozanne, P. Jacoboni and B. Grugeon, *A diagnosis based on a qualitative model of competence in elementary algebra*, AIED' 99, S.P. Lajoie & M. Vivet (eds.), Le Mans, pp. 491-498, 1999.
- [5] A. Lenfant, *Étude sur la transposition d'un outil de recherche destinée aux enseignants*, Mémoire de DEA de didactique des mathématiques, Université Paris VII, 1997.
- [6] D. Luzzati, *Le dialogue verbal homme - machine - études de cas*, Masson (eds.), 1995.