

Extraction de pépites de connaissances dans des réponses d'élèves en langage naturel

Sylvie Normand-Assadi*, Lalina Coulange*
Elisabeth Delozanne**, Brigitte Grugeon***

* IUFM de Créteil, Rue Jean Macé, 94861 BONNEUIL Cedex, France
(sylvie.normand, lalina.coulange)@creteil.iufm.fr

** CRIP5 - Paris V, 45-46 rue des Saints-Pères, 75 006 PARIS, France
elisabeth.delozanne@math-info.univ-paris5.fr
<http://www.math-info.univ-paris5.fr/~delozanne>

*** IUFM d'Amiens 49, boulevard de Châteaudun 80044 AMIENS CEDEX,
brigitte.grugeon@creteil.amiens.fr

Résumé. Le projet Pépite a pour objectif la construction d'un diagnostic des compétences d'élèves en algèbre élémentaire permettant aux enseignants de gérer la diversité cognitive de leurs élèves. Dans cet article, nous présentons une étude pluridisciplinaire (linguistique, didactique et informatique), s'appuyant sur un corpus de productions d'élèves utilisant le logiciel Pépite. Le corpus est analysé selon les points de vue croisés de la linguistique et de la didactique. L'objectif de cette démarche est d'améliorer l'évaluation des réponses d'élèves aux questions ouvertes quand ces derniers répondent avec leurs propres mots. Après avoir situé notre étude, nous présentons la méthodologie retenue et les premiers résultats. Nous montrons ensuite la pertinence de ces résultats avec le point de vue de la recherche sur l'enseignement des mathématiques. Nous terminons par les perspectives ouvertes par ce travail en nous interrogeant sur les apports d'approches EGC à la problématique du diagnostic de compétences.

1 Introduction

L'analyse des réponses d'élèves exprimées dans leurs propres mots est un verrou sur lequel butent de nombreux projets en EIAH (Rosé et al 2003a). Beaucoup de travaux s'intéressent à l'analyse automatique des explications produites en langage naturel et des interactions verbales d'étudiants (avec des pairs, avec l'enseignant), ou se centrent sur l'utilisation de modèles linguistiques dans la conception de logiciels d'apprentissage (Aleven et al 2002), (Arroyo et al 2001), (Rosé et al 2003). Comment prendre en compte la diversité des productions spontanées d'élèves « dans leurs propres mots » dans une analyse informatisée ? En quoi la nature de ces productions peut-elle nous informer sur les apprentissages ou leurs dysfonctionnements ? Notre étude vise à apporter des éléments de réponses à ces questions, dans le contexte très spécifique d'un logiciel d'évaluation diagnostique en algèbre élémentaire. Cette étude se situe dans le cadre du projet Pépite qui vise à concevoir un logiciel pour évaluer des compétences en algèbre élémentaire. Il s'agit d'analyser les erreurs et les cohérences de raisonnement d'élèves en algèbre en tenant compte de leurs réponses à un test couvrant les principales dimensions de la compétence algébrique de base. L'objectif est, à partir de ce diagnostic de proposer des situations d'apprentissage adaptées aux profils d'élèves ainsi repérés. Le logiciel Pépite ne se contente pas de répertorier des erreurs, mais