



STAGE MASTER DE RECHERCHE



RAPPORT BIBLIOGRAPHIQUE

Interprétation temps réel de la production de schémas géométriques pour la génération de e-feedback

Domaine: Technology for Human Learning - Document Processing

Auteur:
Omar Krichen

Superviseurs:
Nathalie Girard
Eric Anquetil
Damien Simonnet
Equipe Intuidoc

Résumé

L'adaptation des principes pédagogiques aux spécificités du numérique est un enjeu majeur. Ce stage s'inscrit dans le projet ACTIF (projet e-Fran) qui vise à concevoir et expérimenter des outils et méthodes pédagogiques qui facilitent l'apprentissage actif et collaboratif au collège à partir de tablettes numériques orientées stylet, en donnant un rôle important au feedback délivré à l'élève. Dans ce travail, nous nous intéressons à l'apprentissage par le dessin (generative drawing) pour les cours de géométrie au collège. L'objectif de ce stage est de concevoir un système capable d'interpréter à la volée les tracés de l'élève et de générer des feedbacks correctifs et de guidage. Nous présentons dans cette étude bibliographique les principes de l'apprentissage par le dessin, l'importance du feedback, et un état de l'art des méthodes d'interprétation de documents manuscrits.

Table des matières

1	Introduction	1
2	Sciences de l'éducation	2
3	Systèmes interactifs d'apprentissage avec feedback	4
4	Interprétation de documents structurés : définitions	6
5	État de l'art des méthodes d'interprétation	8
5.1	Langage Ladder	8
5.2	Interprétation à base de grammaires bidimensionnelles hors contexte	9
5.3	Interprétation à base de grammaires multi-ensembles à contraintes pilotées par le contexte	10
5.3.1	La grammaire GM-PC	10
5.3.2	Le processus d'analyse à la volée	11
5.4	Interprétation a posteriori, Imisketch	12
5.5	Bilan	13
6	Interprétation de documents et feedback	13
7	Conclusion	14

1 Introduction

Contexte

L'éducation étant un enjeu majeur dans le monde actuel, il est devenu primordial d'investir dans des méthodes pédagogiques permettant d'améliorer la capacité d'apprentissage des élèves. Les outils numériques sont de plus en plus utilisés en raison des nouvelles modalités d'interactions (stylet, touch). Dans ce contexte, plusieurs travaux ont été réalisés pour faire le lien entre les théories des sciences de l'éducation et les spécificités du numérique.

Ce stage s'inscrit dans le projet ACTIF (projet e-Fran) qui vise à concevoir et expérimenter des outils et méthodes pédagogiques qui facilitent l'apprentissage «actif» et «collaboratif» au collège à partir de tablettes numériques orientées «stylet». Une attention particulière est portée sur les feedbacks délivrés à chaque élève, à l'ensemble de la classe ou à des élèves réunis en équipes. Ce projet s'appuie sur des travaux récents qui examinent les effets de la production de schémas à partir d'un texte («Generative Drawing») sur l'apprentissage [Leutner & Schmeck, 2014]. Les hypothèses sous-jacentes à ces travaux sont que les étudiants amenés à schématiser le contenu d'un texte en cours de lecture seraient :

- encouragés à traiter plus activement ce matériel pédagogique que lorsque le schéma est fourni d'emblée ;
- incités à mettre en place plus de régulations métacognitives visant à mieux comprendre le texte ;

Objectifs

Ce stage sera focalisé sur l'interprétation temps réel de la production de schémas sur tablette stylet. Plus précisément, l'objectif est de permettre à un élève, à partir d'une consigne textuelle de l'enseignant, de composer à main levée un schéma ou une figure structurée. Cette action sera focalisée dans un premier temps sur des figures de géométrie telle que celle présentée dans la Figure 1. Le système devra donc interpréter automatiquement et au fur et à mesure, les tracés manuscrits afin de fournir à l'élève, en temps réel, un retour visuel de l'interprétation de ses tracés, ainsi qu'un feedback correctif si sa réalisation n'est pas conforme au résultat attendu et un guidage sur les prochaines étapes de sa réalisation.

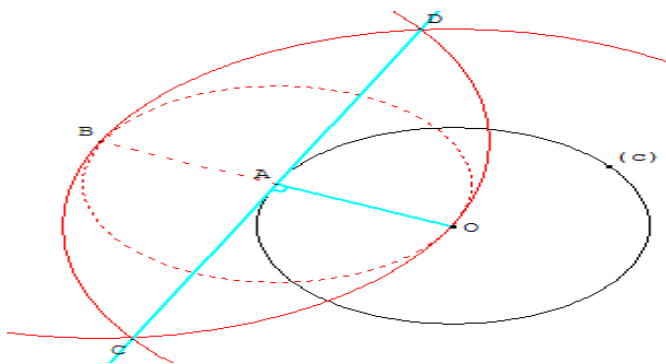


FIGURE 1: figure géométrique à analyser

Organisation du rapport

Les travaux qui nous intéressent dans ce projet sont relatifs à trois domaines que nous allons adresser dans cet état de l'art. En premier lieu, nous présentons les motivations de notre sujet par rapport aux sciences de l'éducation, ensuite nous parlerons de systèmes d'apprentissage interactifs et de feedback. Enfin, nous présenterons un état de l'art sur l'interprétation de documents structurés et la reconnaissance de schémas.

2 Sciences de l'éducation

Introduction

Les sciences de l'éducation se réfèrent à l'étude scientifique du processus d'apprentissage des humains [Mayer & Fiorella, 2015]. Dans cette section, nous définissons quelques concepts utiles, et nous présentons quelques travaux qui ont motivé l'approche du projet ACTIF dans laquelle s'inscrit cette étude.

Processus cognitif de l'apprentissage actif

L'apprentissage actif, tel que défini dans [Bonwell & James, 1991], caractérise le fait que les élèves participent dans le processus d'acquisition de la connaissance et ne se contentent pas uniquement d'entendre passivement le cours de l'enseignant. Les activités telles que la lecture, le débat, le dessin, ou la prise de note stimulent la motivation et permettent le développement des aptitudes de l'élève. Dans [Mayer & Fiorella, 2015], trois types d'apprentissage sont présentés : l'apprentissage génératif, l'apprentissage par coeur et l'apprentissage par association. L'apprentissage est génératif quand l'élève essaye activement de comprendre le matériel qui lui est proposé en s'impliquant dans un traitement cognitif de l'information. Ce traitement se compose de 3 étapes :

- La sélection : prêter attention aux informations pertinentes (mémoire sensorielle) ;
- L'organisation : organiser les informations sélectionnées dans des structures cognitives cohérentes (mémoire opérationnelle) ;
- L'intégration : intégrer ces structures avec des connaissances acquises (mémoire long terme).

La mémoire sensorielle permet de retenir des copies sensorielles des images visualisées et des sons perçus pendant une fraction de seconde. La mémoire opérationnelle permet de manipuler ces informations mais a une capacité limitée. Quant à la mémoire long terme, c'est une sorte d'entrepôt permanent de connaissances. Ce processus cognitif permet à l'élève de saisir le sens de l'information et de résoudre de nouveaux problèmes, contrairement à l'apprentissage par coeur et l'apprentissage par association. L'apprentissage génératif est équivalent à l'apprentissage actif. Nous nous concentrons dans notre projet sur l'aspect génératif, car avec la disponibilité de masses volumineuses d'informations, notamment sur

support numérique, l'intérêt de l'apprentissage dans notre époque ne consiste pas à générer chez l'élève la capacité à apprendre par coeur, mais à générer la capacité à traiter progressivement l'information pour faciliter le processus de l'acquisition des connaissances.

Dans [Mayer & Fiorella, 2015], 8 stratégies d'apprentissage génératif sont proposées. Nous nous intéressons ici à la stratégie de l'apprentissage par le dessin (generative drawing) qui est liée à notre sujet de stage.

Apprentissage par le dessin

L'apprentissage par le dessin consiste à demander à l'élève d'illustrer une leçon à contenu textuel par un dessin, une figure ou un schéma. L'intérêt de cette stratégie est le fait qu'elle stimule les processus cognitifs d'apprentissage génératif :

- processus de sélection : l'apprenant choisit les éléments à inclure dans son dessin/schéma.
- processus d'organisation : l'apprenant réalise une organisation spatiale des éléments dans le dessin.
- processus d'intégration : l'apprenant traduit la représentation verbale de l'information en une représentation visuelle et un modèle mental.

La figure 2 caractérise l'activité cognitive qui s'établit lorsque la stratégie de l'apprentissage par le dessin est appliquée.

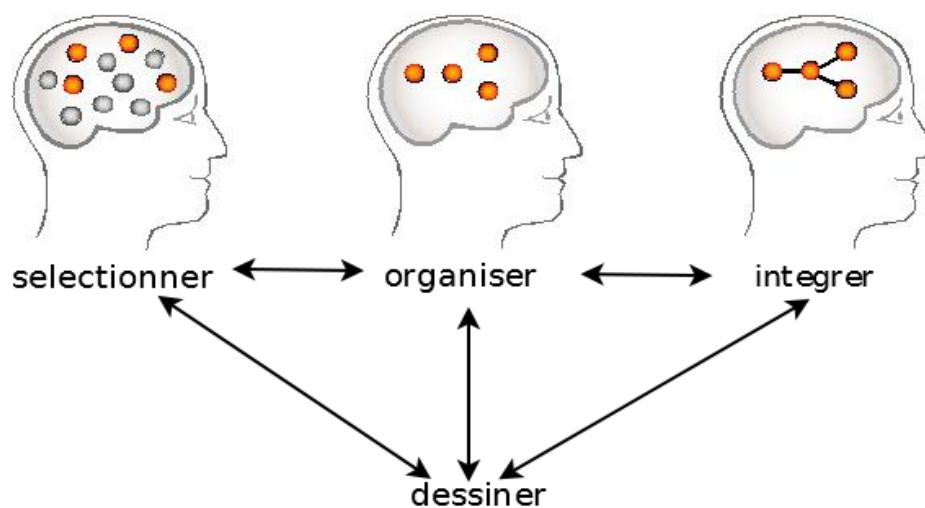


FIGURE 2: processus cognitif de l'apprentissage par le dessin [Quillin & Thomas, 2015]

Toujours dans [Mayer & Fiorella, 2015], une liste de 28 expériences variées (selon le domaine d'étude, la population, le support) qui visent à étudier l'impact de la stratégie d'apprentissage par le dessin sur les performances des étudiants est présentée. Le principe de base consiste à distribuer un même texte à tous les étudiants. Alors qu'un groupe doit schématiser les idées clés du texte, l'autre groupe (appelé groupe de contrôle), traite l'information d'une autre manière (exemple : par simple relecture).

Pour 26 expériences sur les 28, les étudiants qui dessinent ont de meilleures performances que le groupe de contrôle. C'est sur ces travaux que reposent les fondements pédagogiques du projet ACTIF.

3 Systèmes interactifs d'apprentissage avec feedback

L'importance du feedback

En classe, l'élève a besoin d'opportunités fréquentes pour participer et recevoir des feedbacks sur sa performance [Chickering & Ehrmann, 1996]. Ceci lui permet de tester sa compréhension du cours, et de savoir ce qu'il lui reste à comprendre. Dans cette section, nous présentons quelques systèmes d'apprentissage numériques et nous parlons du rôle du feedback dans ces systèmes.

L'environnement d'apprentissage "Diagram" [Py & AI, 2013]

Diagram est un environnement d'apprentissage pour la modélisation de diagrammes UML. Une base de 30 exercices classiques est constituée par des énoncés textuels et des diagrammes de références fournis par un enseignant. L'environnement est ouvert, c'est à dire qu'il est possible d'ajouter des énoncés sans contraintes sur le vocabulaire et la taille du diagramme. L'étudiant modélise son diagramme à l'aide d'une palette d'outils graphique, présentée dans la figure 3, puis valide son modèle pour que le système puisse vérifier son exactitude.

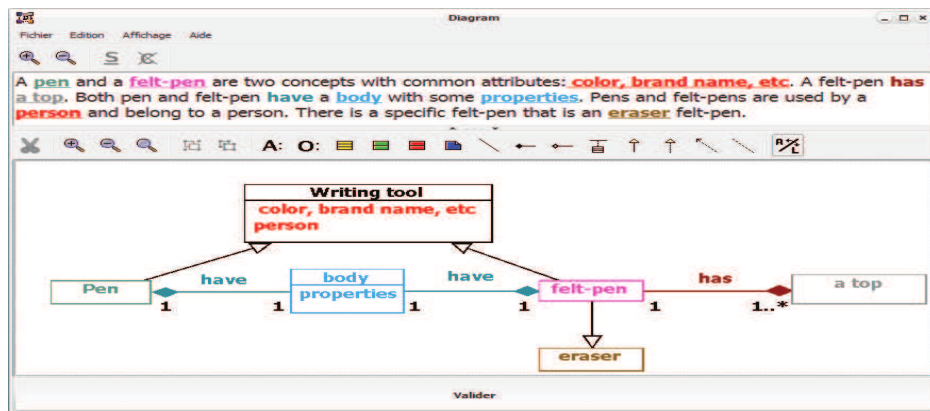


FIGURE 3: interface diagram [Py & AI, 2013]

L'analyse et la vérification du diagramme de l'étudiant se font par la mise en correspondance de ce dernier avec le diagramme de référence. Pour ce faire, les diagrammes sont considérés en tant que graphes caractérisés par des attributs spécifiques. La sortie de l'algorithme de diagnostic est une liste des différences entre le diagramme de l'étudiant et le diagramme de référence, par exemple l'omission d'une classe ou l'inversion d'une relation. Cette liste est le critère sur lequel se base le feedback correctif du système. Chaque feedback se compose de 3 niveaux d'intervention, qui vont du plus général au plus précis.

- *notification* : consiste à attirer l'attention de l'étudiant sur une partie de son diagramme ;
- *question* : consiste à poser des questions sur les propriétés du diagramme ;
- *proposition* : modalité plus directe, consiste à suggérer une façon de corriger le diagramme ;

Discussion

Les auteurs ont testé cet environnement sur un groupe de 18 étudiants qui ont effectué 86 exercices. Après avoir lu les feedbacks générés lors de la validation de leur modèle, 49.7% des étudiants ont corrigé leur diagramme en adéquation avec le diagramme de référence, 3.8% l'ont modifié mais les différences persistent, et 46.5 % ignorent le message et n'effectuent pas de changement. Ceci s'explique par le fait qu'avant la session, il a été signifié aux étudiants qu'ils pouvaient ignorer les feedbacks s'ils étaient sûrs de leur travail. Une des limites de cette méthode réside dans le fait que 11% des messages sont inappropriés, car dans certains cas le module de diagnostic n'a pas correctement mis en correspondance les diagrammes. Soulignons que notre projet se place dans un contexte de schémas réalisés à main levée avec un stylet directement sur l'écran de la tablette, avec un plus grand degré de liberté de l'élève, alors que dans "Diagram", l'étudiant modélise son diagramme à partir d'une palette d'outils à sélectionner (boutons, menus...). Par contre, on pourra retenir de cette approche notamment la gradation des feedbacks qui semble vraiment pertinente pour notre projet.

Le tuteur intelligent Mentoniezsh [Py, 2001]

Mentoniezsh est un tuteur intelligent d'aide à la démonstration de preuves en géométrie. Ici, les figures ne sont pas manipulées. Le domaine couvert par ce système est celui de la géométrie affine.

La représentation des connaissances géométriques est basée sur de la logique du premier ordre, avec le langage HDL, pour Hypothesis Description Language.

Les objets manipulés dans ce langage sont les droites, les cercles et les points. Un prédicat se présente par exemple sous cette forme : alignés(a,b,c) pour signifier que a, b et c appartiennent à la même droite, ou encore triangle-rectangle(a,b,c) pour un triangle rectangle en a.

La base des règles correspond aux théorèmes et définitions du programme des classes de quatrième. Par exemple, le théorème "Un triangle qui a deux côtés perpendiculaires est un triangle rectangle", se traduit par la clause :

$$\forall a,b,c : \text{triangle-rectangle}(a,b,c) \implies \text{perpendiculaire}(\text{droite}(a,b), \text{droite}(a,c)).$$

Un énoncé est constitué d'une conjonction de termes HDL représentant les données et d'un terme appelé but représentant la preuve à démontrer. Mentoniezsh est muni d'un démonstrateur qui, à partir des données et du but, extrait l'ensemble des faits démontrables et établit les pas nécessaires à la réalisation de la démonstration. Du point de vue de l'élève, l'énoncé est présenté sous forme textuelle, accompagné de la figure géométrique. Dans un premier temps, l'élève doit extraire les données de l'énoncé à l'aide d'une palette d'outils graphiques, sous le contrôle du tuteur qui lui indique les données manquantes

si elles existent. Dans une deuxième étape, il effectue les pas relatifs à la démonstration, et il a la possibilité d'interagir avec le système pour vérifier la cohérence de sa démarche. L'établissement de la démonstration est considéré comme un problème de planification, où les pas (théorèmes, hypothèses et faits) sont les actions. Grâce à un algorithme de reconnaissance de plan, le système est capable de reconnaître l'intention de l'élève et de le guider dans sa démarche. Le diagnostic et la génération de messages correctifs se font sur la base d'un catalogue contenant les erreurs de syntaxe, les erreurs de logique, et les erreurs de sémantique.

Discussion

La capacité du système à corriger les erreurs de l'élève avant qu'il ne termine sa démonstration s'approche de notre idée d'interagir avec l'apprenant en temps réel. La représentation des connaissances géométriques en langage HDL apparaît pertinente. Néanmoins, une limite est que ce langage n'est pas assez expressif pour couvrir tous les rapports métriques et, de plus, le système est centré sur la démonstration de preuves et non sur la composition de la figure.

Nous avons donc défini les concepts pédagogiques relatifs au projet Actif dans lequel s'inscrit notre stage. Nous nous sommes ensuite intéressés aux systèmes interactifs d'apprentissage avec feedback. Nous abordons maintenant l'axe central de ce stage : analyser en temps réel la composition manuscrite d'un schéma géométrique sur tablette. Nous commencerons par quelques définitions relatives au domaine de l'interprétation des documents structurés, puis nous présenterons un état de l'art sur les méthodes de reconnaissance et d'interprétation de schémas.

4 Interprétation de documents structurés : définitions

Document structurés

Un document structuré, de par son nom, est caractérisé par une structure prédéfinie. Il se constitue d'un ensemble de symboles organisés spatialement de manière à permettre l'extraction de sens à partir du dit document. Par exemple, un schéma de circuit électrique ou un plan architectural sont considérés comme des documents structurés. Un document structuré peut se présenter sous des formes diverses : il peut être imprimé, ou sous forme numérique manipulable par ordinateur (un fichier constitué d'un ensemble de vecteurs), ou bien composé à main levée. Dans ce stage, nous nous intéressons aux schémas géométriques.

Composition de documents structurés

On distingue trois façons de composer un document structuré numérique. La première consiste à utiliser des logiciels basés sur une interaction graphique orientée souris et boutons (interface WIMP). Une liste complète des symboles est disponible dans l'interface et l'utilisateur va cliquer sur le symbole et le placer dans son schéma. Tel est le cas de l'environnement Diagramm présenté dans la section précédente.

Une deuxième façon de le composer consiste à dessiner le schéma sur une feuille de papier pour ensuite la numériser et l'interpréter. Ce type de document est appelé document "hors ligne". Typiquement, un signal hors-ligne est une matrice de pixels en niveaux de gris. La troisième, qui nous intéresse dans ce projet, est la composition orientée stylet. Le signal généré est dit "en-ligne", et peut contenir des informations telles que l'ordre et le sens du tracé ou encore la vitesse de l'écriture. L'élément le plus important de ce signal est le tracé, qui représente une succession de coordonnées entre un poser de stylo et un lever de stylo. Un symbole est alors représenté par une séquence de tracés. Ces informations (sens, ordre des tracés) peuvent faciliter l'analyse et l'interprétation du schéma.

Approches d'interprétation

Au delà des manières de composer des documents structurés, il existe aussi différentes approches pour les interpréter et les analyser.

Interprétation a posteriori

L'étape d'interprétation du document ne commence qu'à la demande de l'utilisateur, lorsque la composition du document est terminée. En termes de reconnaissance de formes, l'avantage est que l'agencement spatial complet des tracés, ou le contexte structurel, est connu par le système avant l'étape d'interprétation. L'inconvénient majeur est la propagation de l'erreur, puisque la mauvaise interprétation d'un tracé entraînera forcément d'autres erreurs.

Interprétation à la volée

Les tracés sont interprétés au fil de l'eau, chaque tracé ajouté est analysé en "temps réel". De plus, un retour visuel est fourni à l'utilisateur pour chaque symbole interprété, par exemple une remise au propre ou une "beautification" (lissage du symbole mais en gardant dans une certaine mesure la spécificité de l'écriture de l'utilisateur). L'avantage de cette technique est que l'utilisateur devient acteur du processus d'interprétation, puisqu'il peut valider implicitement l'interprétation du système ou bien la rejeter explicitement en l'effaçant par exemple, ce qui règle le problème de la propagation des erreurs. L'inconvénient est que l'utilisateur peut être gêné s'il est sollicité un grand nombre de fois dans le processus d'interprétation.

Discussion

Puisque nous nous plaçons dans un contexte pédagogique, l'interprétation à la volée semble la plus pertinente. La remise au propre donne à l'élève un certain degré de liberté dans son dessin (le système doit être capable de reconnaître deux droites parallèles même si elles ne sont pas exactement dans le dessin). De plus, l'interprétation à la volée est plus intéressante pour la génération de feedbacks correctifs en temps réel (plus les corrections interviennent tôt mieux c'est), et aussi pour les feedbacks prédictifs

(reconnaître une figure géométrique telle qu'un cercle avant que l'élève ne termine son dessin). Différents travaux autour des méthodes d'interprétation sont présentés dans la section suivante.

5 État de l'art des méthodes d'interprétation

5.1 Langage Ladder

Ladder [Hammond & Davis, 2005], est une approche pour l'interprétation générique de documents manuscrits en ligne provenant de plusieurs domaines. Ladder se base sur un langage qui permet de décrire comment les symboles tracés spécifiques à un domaine peuvent être dessinés, affichés ou modifiés. Le langage se constitue de formes primitives prédéfinies (ligne, cercle, arc, bézier, polygone..), de contraintes géométriques prédéfinies (parallèle, perpendiculaire, colinéaire..), de fonctions d'affichage et de modification. Ces deux dernières fonctions sont intéressantes puisqu'elles permettent la remise au propre d'un tracé et sa manipulation par l'utilisateur. La figure 4 représente la définition du symbole "flèche" composé de 3 lignes, avec les contraintes géométriques qui régissent l'agencement spatial de celles-ci.

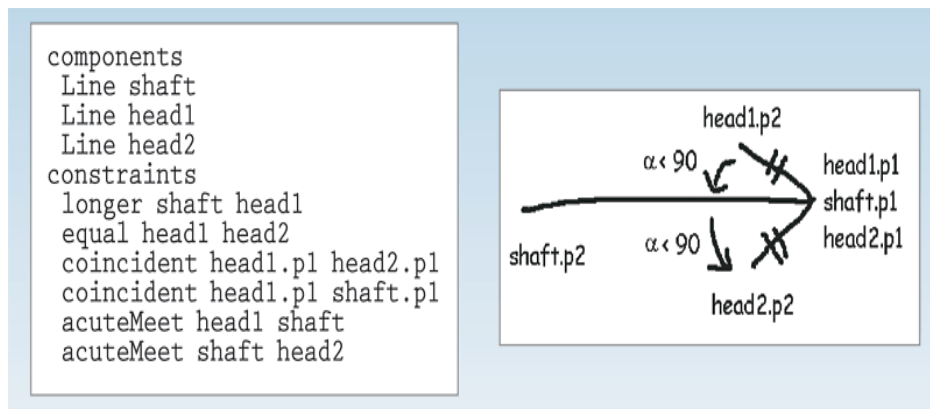


FIGURE 4: définition du symbole flèche dans ladder [Hammond & Davis, 2005]

Pour l'interprétation de la forme quand un tracé est dessiné, une reconnaissance bas-niveau est réalisée pour déterminer s'il peut être classifié en tant que ligne, arc, ellipse, point ou bien une combinaison de ces formes primitives. Ensuite, une reconnaissance haut-niveau basée sur les informations contextuelles (contraintes géométriques) est effectuée sur le groupement de tracés constituant un symbole. La reconnaissance est donc *ascendante*, des formes primitives vers les formes complexes. Les auteurs ne fournissent pas plus de détails quant au fonctionnement de leur analyseur. Une limite de cette méthode d'interprétation est que si un tracé est mal classifié, l'analyseur ne pourra pas reconnaître le symbole correctement.

Ladder a été exploité pour implémenter Paleosketch [Paulson & Hammond, 2008], un système de reconnaissance et de "beautification" automatique de schémas. Pour chaque tracé, un vecteur de caractéristiques est défini pour tester s'il correspond à l'une des 8 formes primitives prédéfinies

dans Ladder. Ensuite, les tracés composant un symbole sont regroupés et Ladder est utilisé pour la reconnaissance haut-niveau du symbole grâce aux contraintes géométriques définies par l'utilisateur. Les résultats en terme de performance sont probants. Un ensemble de données composé de 900 schémas dessinés par 10 utilisateurs a été utilisé pour l'entraînement. Un autre ensemble de même taille a été utilisé pour le test. Le taux de bonne reconnaissance atteint est de 98.56 %.

5.2 Interprétation à base de grammaires bidimensionnelles hors contexte

Définition formelle

Une grammaire hors contexte est un tuple $G=(N, \Sigma, S, P)$ avec

- N l'ensemble de symboles non terminaux ;
- Σ l'ensemble fini de symboles terminaux ;
- $S \subset N$ le symbole de départ ;
- P l'ensemble fini de règles $A \rightarrow \alpha$, $A \in N$, $\alpha \in (N \cup \Sigma)^+$;

Une grammaire stochastique hors contexte (Stochastic Context Free Grammar en anglais) est un couple (G, p) où G est la grammaire hors contexte, et $p : P \rightarrow [0, 1]$ est une fonction de probabilité de l'application d'une règle tel que $\sum_{i=1}^{n_A} p(A \rightarrow \alpha_i) = 1$, où n_A est le nombre de règles associées au symbole non terminal A .

Une généralisation des SCFG est proposée dans [Muñoz & Al, 2008], par la définition d'une 2D-SCFG pour la reconnaissance des expressions mathématiques. Les symboles terminaux sont les symboles mathématiques reconnus dans les régions bidimensionnelles qui correspondent à des hypothèses de segmentation des symboles. Les règles ont un paramètre qui décrit la relation spatiale entre les régions. Cette relation est de la forme $A \xrightarrow{spr} \alpha$, ou spr correspond à la relation spatiale qui modélise la règle. Typiquement, une relation spatiale peut être *horizontale* (BC), *verticale* ($\frac{A}{B}$), *dans* (\sqrt{A}). C'est cette relation spatiale qui marque la différence entre la grammaire 1D et la grammaire 2D.

Cette méthode combine deux approches pour reconnaître une expression mathématique complète :

- L'apprentissage d'un modèle pour chaque symbole mathématique en utilisant les modèles de Markov cachés ;
- L'apprentissage de la grammaire stochastique 2D SCFG pour la classification des relations spatiales entre symboles en utilisant les SVM ;

Les modèles de Markov ont souvent été utilisés pour la classification de symboles mathématiques. Muñoz & Al utilisent ici des ensembles de caractéristiques couramment utilisées dans le domaine de la reconnaissance de l'écriture manuscrite pour entraîner leur modèle.

Les résultats en terme de performance sont probants, et le système implémenté a été déclaré vainqueur

dans la compétition CROHME 2011 [Mouchere & Al, 2011] (Competition on Recognition of Online Handwritten Mathematical Expressions).

5.3 Interprétation à base de grammaires multi-ensembles à contraintes pilotées par le contexte

La méthode Dali [Macé & Anquetil, 2009], se base sur la théorie des grammaires et langages visuels pour l'interprétation à la volée de documents structurés en ligne. Pour ce faire, une nouvelle grammaire appelée GM-PC (Grammaires multi-ensembles à contraintes pilotées par le contexte) a été définie.

5.3.1 La grammaire GM-PC

Définition formelle

Une GM-PC est un tuple $G=(N, \Sigma, S, P)$ avec

- N l'ensemble de symboles non terminaux ;
- Σ l'ensemble fini de symboles terminaux, ou alphabet ;
- $S \subset N$ le symbole de départ ;
- P l'ensemble fini de règles, appelées productions (ou réécritures) ;

L'alphabet de cette grammaire a été limité à une seule primitive graphique, qui est le tracé $\Rightarrow \Sigma = \{\text{tracé}\}$. Une production p s'écrit sous la forme $\alpha \rightarrow \beta$ avec $\{\text{préconditions}, \text{contraintes}, \text{postconditions}\}$ et (D) , tel que D est la définition des attributs de α . Cela signifie qu'un multi-ensemble de symboles β peut être réécrit en un autre multi-ensemble α quand les contraintes C sont satisfaites. Ces contraintes représentent différents types d'information, telle que la relation spatiale par exemple. Le bloc de contraintes modélise une vision locale des éléments analysées, tandis que les blocs de préconditions et de postconditions modélisent une vision globale du document en se basant sur la position relative des éléments d'un symbole. Pour ce faire, la notion de contexte structurel du document (CSD) est introduite. Un CSD caractérise une position dans le document et l'élément qui est attendu dans cette position.

Un CSD s'écrit sous la forme : $\gamma[\text{position}] \delta[\text{condition}]$, tel que :

- γ ensemble d'éléments de références ;
- position est la position relative à ces éléments ;
- δ ensemble d'éléments attendus à la position ;
- condition une partie des points d'un élément de δ ;

Par exemple, le CSD (m) [sous] (s) [tous] modélise le fait que tous les points de l'élément s sont dans la

zone au dessous de l'élément m .

Les CSD sont utilisées dans les blocs de préconditions et de postcondition. Les préconditions sont un ensemble de CSD qui assurent que les éléments en cours d'analyse soient cohérents structurellement avec la production.

Les postconditions permettent de mettre à jour la structure globale du document en indiquant les CSD qui doivent être créés si la règle de production est réalisée.

Les auteurs ont défini deux types de contraintes. Les contraintes structurelles qui modélisent l'agencement relatif des éléments, et les contraintes statistiques qui visent à interpréter les formes complexes difficiles à reconnaître structurellement à partir de classifieurs de symboles.

5.3.2 Le processus d'analyse à la volée

La structure d'analyse, qui est un graphe de dérivation, est mise à jour chaque fois qu'un tracé est dessiné par l'utilisateur. Le processus consiste à remplacer des multi-ensembles de symboles en d'autres multi-ensembles en cherchant à réduire les productions qui satisfont les contraintes. Si la production $\alpha \rightarrow \beta$ est réduite, les éléments β sont remplacés en α , qui peuvent être remplacés en réduisant d'autres productions. Ce processus est répété jusqu'à ce que la stabilité soit atteinte, c'est à dire qu'il n'y a plus de réduction possible. Un élagage des éléments et des CSD à considérer est effectué pour éviter l'explosion combinatoire. L'originalité de cet analyseur est le couplage entre une analyse ascendante guidée par les éléments et une analyse descendante guidée par les contextes.

Chaque tracé a comme résultat la réduction d'une séquence de productions, qui est considérée comme l'interprétation du tracé dans le contexte du document. Il se peut que plusieurs interprétations soient possibles. La comparaison de ces interprétations pour la prise de décision se base sur la théorie de la logique floue. La règle de production $\alpha \rightarrow \beta$ est considérée comme $\alpha \rightarrow^p \beta$, où $p \in [0, 1]$ est le degré d'adéquation de β par rapport à α . Des mesures d'adéquation sont calculées sur les CSD pour choisir la séquence de production avec le meilleur score. Il y'a trois situations possibles :

- Une seule interprétation possible \Rightarrow pas d'ambiguïté \Rightarrow la séquence de production est appliquée et le retour visuel est affiché ;
- Plusieurs interprétations possibles \Rightarrow ambiguïté \Rightarrow calcul des degrés d'adéquation \Rightarrow possibilité de rejet si l'ambiguïté persiste (demande à l'utilisateur de redessiner le tracé) ;
- Aucune interprétation possible (degrés d'adéquation < seuil minimal)
 \Rightarrow rejet de distance \Rightarrow segmentation du tracé et relance de l'analyse ;

La méthode Dali a été appliquée dans différents domaines, des systèmes ont été implémentés pour la composition de diagrammes UML et la composition musicale. De plus, cette approche a été utilisée pour concevoir un système de composition de schéma électrique qui a commercialisé : "Script&Go Sketches", ce qui prouve le bien-fondé de la méthodologie.

5.4 Interprétation a posteriori, Imisketch

Imisketch [Ghorbel & Al, 2015], est une approche pour l'interprétation a posteriori de documents de différents types (manuscrits, imprimés ou vectoriels). Ce travail s'inspire de la méthode Dali [Macé & Anquetil, 2009] en se basant sur la grammaire GM-PC et son analyseur associé. Imisketch est une méthode interactive qui sollicite l'utilisateur dans la phase d'analyse du document (pour résoudre une ambiguïté par exemple). La première étape de l'analyse est l'extraction de primitives à partir du document structuré. Les primitives considérées ici sont les segments et les polygones. La deuxième étape est la modélisation de la connaissance a priori. La connaissance structurelle est modélisée par la grammaire GM-PC. Une production GM-PC peut aussi appeler un classifieur pour la reconnaissance de symboles. Cette grammaire a été adaptée au contexte de l'interprétation a posteriori. Dans le cas de l'interprétation à la volée, une seule primitive est considérée (le tracé), alors que dans ce cas, plusieurs primitives le sont, d'où le risque d'explosion combinatoire. Les auteurs proposent donc une stratégie hybride qui permet de choisir, à partir d'une production, la stratégie d'exploration de l'arbre d'analyse (recherche en largeur ou recherche en profondeur).

Processus d'analyse

Le processus d'analyse est proche de l'analyseur de la méthode Dali, à ceci près qu'une zone dite de "focus contextuel spatial" est définie avant l'interprétation d'une primitive pour limiter l'exploration. Ensuite les arbres d'analyses sont construits. L'originalité de l'analyseur Imisketch réside dans la gestion de l'ambiguïté et l'interaction avec l'utilisateur.

Plusieurs interprétations possibles \Rightarrow ambiguïté \Rightarrow calcul des degrés d'adéquation des séquences de productions envisageables. Si la différence entre les scores des deux meilleures séquences est inférieure à un seuil minimal, alors le système demande à l'utilisateur de choisir entre les deux interprétations.

De plus, le classifieur de symboles implémenté est évolutif, c'est à dire que son apprentissage est incrémental, et ce grâce à l'intervention de l'utilisateur. Par exemple, si le degré de confiance du classifieur pour un symbole est inférieur à un seuil donné, il sollicite l'utilisateur qui va soit :

- associer ce symbole à une nouvelle classe, le classifieur apprend un nouveau modèle ;
- associer ce symbole à une classe déjà existante, le classifieur réduit la distance entre deux classes ;
- ou l'ignorer (rejet), dans ce cas le classifieur ne fait rien ;

Cette méthode a été appliquée sur la composition de plan architecturaux. Pour l'évaluation, une base de données composée de 24 plans architecturaux a été créée. Chaque plan se compose de murs, de symboles de type "fourniture" (lit, toilette, etc), et de symboles de type "ouverture" (porte, fenêtre, etc). Un sous ensemble de 15 plans a été utilisé pour l'apprentissage supervisé du classifieur et le reste a été utilisé pour l'évaluation. Le taux de bonne reconnaissance des plans atteint est de 93.4%.

5.5 Bilan

Dans cette section nous avons présenté quelques formalismes et méthodes pour l'interprétation de documents structurés. Comme notre projet se place dans le contexte de l'interprétation à la volée de documents manuscrits en ligne, il est clair que la grammaire GM-PC est la plus adaptée à notre problème, d'autant plus qu'elle fait partie du savoir-faire de l'équipe Intuidoc.

6 Interprétation de documents et feedback

Il existe beaucoup d'approches dans la littérature qui abordent l'interprétation de documents structurés en ligne, de même pour les systèmes interactifs d'apprentissage avec feedback, mais à notre connaissance, peu de travaux considèrent ces deux aspects pour un même système. Nous citons ici Mechanix [Atilola & Al, 2014], un système d'interprétation de croquis et d'apprentissage interactif pour les cours de mécanique. Pour la partie reconnaissance de schémas, Mechanix se base sur le logiciel Paleosketch [Paulson & Hammond, 2008], cité plus haut dans la section 3. Le schéma considéré ici est le treillis, qui est une structure basique pouvant être utilisée dans la conception de ponts ou d'immeubles, enrichie par des vecteurs de force. Le treillis est apparenté à une collection de polygones convexes, tel que chaque polygone partage au moins un côté avec un autre polygone.

La reconnaissance de formes géométriques se fait en étapes et se base sur l'interprétation à la volée des tracés de l'étudiant. A chaque nouveau tracé, le système fait appel à Paleosketch [Paulson & Hammond, 2008], qui reconnaît la forme primitive (ligne, arc, etc). Ensuite, ces formes sont regroupées pour la reconnaissance de formes complexes. Les auteurs ne donnent pas plus de détails sur ce processus. La reconnaissance de treillis se fait par la construction de graphes connectés. Chaque arc AB est considéré comme côté partagé. Pour savoir si AB est un côté partagé, AB est enlevé et on cherche un autre chemin pour aller de A à B. Si le chemin existe, AB sera un côté partagé. La figure 5 présente le fonctionnement de l'algorithme de reconnaissance de treillis par la recherche de côtés partagés par deux polygones.

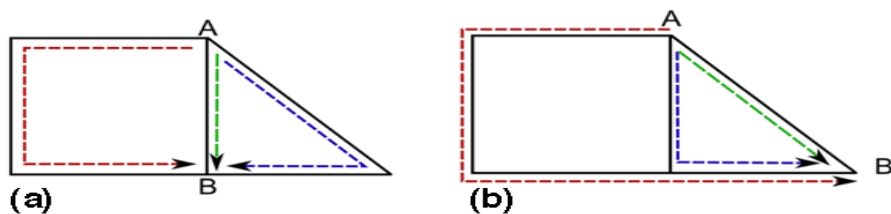


FIGURE 5: a reconnu comme côté partagé, b non reconnu [Atilola & Al, 2014]

Les consignes sont de type textuel, et le professeur fournit un treillis de référence au système. La vérification du travail de l'étudiant est réalisée lorsqu'il valide son schéma. Les feedbacks correctifs sont donc générés a posteriori. La comparaison entre les deux treillis se fait par la construction de graphes représentant les schémas. Les auteurs utilisent un isomorphisme de graphes pour déterminer leur similarité. Ensuite le système vérifie les types et les valeurs des forces. Finalement, Mechanix génère un

feedback correctif de type "il vous manque une force dans le noeud N".

Mechanix a été testé dans un cours d'ingénierie mécanique. Sur une classe de 64 étudiants, 20 ont utilisé Mechanix et les autres ont travaillé sur papier pour comparer la performance des étudiants. Mechanix a été utilisé pour "3 devoirs à la maison". Les notes pour le premier devoir étaient similaires tandis qu'elles étaient 25 % supérieures pour les utilisateurs de Mechanix pour le deuxième et troisième devoir.

Cette méthode semble prometteuse, mais les auteurs ne fournissent pas assez de détails sur leur approche, aussi bien sur le processus d'interprétation que le processus de vérification du travail de l'étudiant. Le fait que le processus d'interprétation soit en temps réel est intéressant, sauf que l'approche de génération des feedbacks a posteriori (après que l'élève valide son schéma) nous semble limitée puisqu'on veut guider l'élève dans son tracé. Pour ce faire, il faudrait que le feedback soit en temps réel.

7 Conclusion

A travers cette étude bibliographique, nous avons essayé de présenter l'aspect pluridisciplinaire de ce stage. Le projet ACTIF, dans lequel ce stage s'inscrit, vise à concevoir et expérimenter des outils et méthodes pédagogiques qui facilitent l'apprentissage « actif » et « collaboratif » au collège à partir de tablettes numériques orientées « stylet », en donnant un rôle déterminant au feedback délivré à chaque élève. Nous avons expliqué la pertinence en terme de pédagogie de l'apprentissage par le dessin (generative drawing), et l'importance des feedbacks pour l'amélioration de la compréhension des élèves. Nous avons choisi de nous orienter pour ce projet sur la production de feedbacks en temps réel, pour plus d'interaction avec l'élève. Nous nous sommes aussi penché sur différentes méthodes d'interprétation de documents manuscrits. Le bilan de cet état de l'art nous invite à utiliser la grammaire GM-PC qui semble bien adaptée à notre problème. Il faudra donc modéliser les connaissances du domaine à partir de cette grammaire et des classifieurs de formes élémentaires de tous les objets constituant une figure géométrique (ligne, cercle, triangle, angles, symboles, etc). Une des questions ouvertes est la difficulté de mise en correspondance entre les résultats de l'analyse en temps réel du tracé de l'élève et la consigne de l'enseignant pour en déduire les feedbacks de guidage ou correctifs à générer. Par ailleurs, le choix des feedbacks visuels d'interprétation temps réel des tracés est aussi un élément important sur lequel nous devons travailler : faut-il remettre parfaitement au propre le schéma tracé à main levée ou simplement le "beautifiser", c'est à dire le modifier tout en préservant la spécificité du dessin de l'élève.

Références

[Leutner & Schmeck, 2014] D. Leutner, A. Schmeck. «The generative drawing principle in multimedia learning». In R. E. Mayer (Ed.), The Cambridge handbook of multimedia learning. Cambridge : Cambridge University Press, pages 433-448, 2014.

- [Bonwell & Eison, 1991] C.C. Bonwell, J.A. Eison. «Active Learning : Creating Excitement in the Classroom». ASHE-ERIC Higher Education Report no.1, George Washington University, Washington, DC, 1991.
- [Mayer & Fiorella, 2015] L. Fiorella, R.E. Mayer. « Learning as a generative activity : Eight learning strategies that promote understanding». New York : Cambridge University Press, 2015.
- [Quillin & Thomas, 2015] K. Quillin, S. Thomas. «Drawing-to-Learn : A Framework for Using Drawings to Promote Model-Based Reasoning in Biology». CBE-Life Sciences Education, 14, pages 1-16, 2015.
- [Chickering & Ehrmann, 1996] A.W. Chickering, S.C. Ehrmann. «Implementing the Seven Principles : Technology as Lever». AAHE Bulletin, vol. 49, no. 2, pages 3-6, 1996.
- [Py & Al, 2013] D. Py, L. Auxepaules, M. Alonso. « Diagram, a Learning Environment for Initiation to Object-Oriented Modeling with UML Class Diagrams». Journal of Interactive Learning Research, Association for the Advancement of Computing in Education, 24 (4), pages 425-446, 2013.
- [Py, 2001] D. Py, « Environnements interactifs d'apprentissage et géométrie». Habilitation à diriger des recherches de l'université Rennes 1, 2001.
- [Hammond & Davis, 2005] T. Hammond, R.Davis. «LADDER, A Sketching Language for User Interface Developers». Computers & Graphics, 29, pages 518-532, 2005.
- [Paulson & Hammond, 2008] B. Paulson, T. Hammond. «Paleosketch : accurate primitive sketch recognition and beautification», Proceedings of the 13th international conference on Intelligent user, pages 1-10, 2008.
- [Muñoz & Al, 2008] Á. Muñoz, F.S. Peiró, J.B. Ruiz. « Recognition of on-line handwritten mathematical expressions using 2D stochastic context-free grammars and hidden Markov models». Pattern Recognition Letters, 35, pages 58-67, 2014.
- [Macé & Anquetil, 2009] S. Macé, E. Anquetil, «Eager interpretation of on-line handdrawn structured documents : The DALI methodology». Pattern Recognition, 42, pages 3202–3214, 2009.
- [Ghorbel & Al, 2015] A. Ghorbel, A. Lemaître, E. Anquetil, S. Fleury, E. Jamet. «Interactive interpretation of structured documents : application to the recognition of handwritten architectural plans ». Pattern recognition, 48(8), pages 2446-2458, 2015.
- [Atilola & Al, 2014] O.Atilola, S.Valentine, H.H. Kim, D. Turner, E. McTigue, T. Hammond, J. Linsey. «Mechanix : A natural sketch interface tool for teaching truss analysis and free-body diagrams». Artificial Intelligence for Engineering Design, Analysis and Manufacturing, 28, pages 169-192, 2014.
- [Mouchere & Al, 2011] H. Mouchere, C. Viard-Gaudin, U. Garain, D.H. Kim, J.H. Kim. «CROHME2011 : Competition on recognition of online handwritten mathematical expressions». 2011 International Conference on Document Analysis and Recognition, Beijing, pages 1497-1500, 2011.