

---

# SBT-IM : Un Système à Base de Traces pour le calcul des 'indicateurs d'interaction dans Moodle

Tarek Djouad\* — Lotfi Sofiane Settouti\* — Alain Mille\* —  
Christophe Reffay\*\* — Yannick Prié\*

\*Université de Lyon, CNRS

Université Lyon 1, LIRIS, UMR5205, F-69622, France

{tarek.djouad, lotfiosofiane.settouti, alain.mille, yannick.prie}@liris.cnrs.fr

\*\* Laboratoire d'informatique de l'université de Franche-Comté

16 route de Gray, 25030 BESANCON cedex

{christophe.reffay}@univ-fcomte.fr

---

*RÉSUMÉ.* Nous présentons dans ce travail un Système à Base de Trace implémenté pour le calcul des Indicateurs dans la plateforme d'apprentissage Moodle (SBT-IM). Nous présentons d'abord un cadre conceptuel permettant de penser tout système exploitant des traces modélisées en définissant les notions de trace, de modèle de trace et de transformation. Nous décrivons ensuite la mise en œuvre de ce cadre conceptuel dans le cas d'un système à base de trace dédié au calcul des indicateurs dans les situations d'apprentissage collaboratif.

*ABSTRACT.* We present in this paper an implementation of a Trace Based System used to calculate Indicators in the Moodle learning platform (TBS-IM). First we present a conceptual framework supporting any trace-based systems by defining the notions of trace, trace model and trace transformation. Afterwards we describe the implementation of this conceptual framework in the case of a trace-based system dedicated to the computation of indicators in collaborative situations.

*MOTS-CLÉS.* Systèmes à base de trace, calcul des indicateurs, modèle de trace, transformations de traces.

*KEYWORDS.* trace-based systems, indicator computation, trace model, traces transformations.

---

## 1. Introduction

Les EIAH (Environnements Informatiques pour l'Apprentissage Humains) fournissent des supports d'activités d'apprentissage et possèdent également des capacités d'évaluation et de personnalisation du processus d'apprentissage pour un apprenant individuel et/ou pour un groupe d'apprenants. Il est en effet techniquement relativement facile de « garder trace » de l'activité pour l'analyser de façon à élaborer des indicateurs d'activité d'apprentissage permettant une adaptation de l'EIAH par l'enseignant ou/et par l'EIAH lui-même selon des règles de personnalisation.

La première partie de cet article fait le point sur l'état de l'art des EIAH exploitant les traces d'activité pour des objectifs très variés, mettant en évidence la nécessité de collecter des traces d'activité et de les analyser pour produire des indicateurs permettant la personnalisation. Selon les indicateurs à produire, il s'agit de disposer d'observations de l'activité de granularités très variables en décrivant des éléments hétérogènes liés aux interactions. Cette variabilité et cette hétérogénéité des sources entraînent de grandes difficultés à réutiliser les techniques d'élaboration d'indicateurs pour des EIAH évolutifs par nature et déployés dans des contextes très variés. Pour résoudre cette difficulté, nous proposons dans la deuxième partie de considérer les traces numériques d'interaction comme des « objets informatiques » à part entière possédant des modèles explicites et associés à des modèles de transformation. Une telle modélisation permet la réalisation de « systèmes à base de traces modélisées » permettant de concevoir et de calculer les transformations nécessaires à différents types d'usage, en particulier le calcul d'indicateurs. Une ingénierie des traces permet alors une conception guidée par les modèles et facilite la réutilisation dans la conception des EIAH exploitant des traces d'activité. Le bénéfice d'une telle approche est illustré dans une troisième partie par le système SBT-IM (Système à Base de Traces- Indicateurs Moodle) dans le cadre d'un usage des traces pour la conception et l'exploitation d'indicateurs d'activité d'apprentissage collaboratif. L'ingénierie des indicateurs s'en trouve considérablement simplifiée et explicitée. Nous évoquons dans la conclusion d'autres usages de l'approche SBT pour faciliter de nombreuses activités de l'assistance à l'apprentissage.

### 1.1. Traces et EIAH : présentation générale

Il convient pour commencer de préciser quelques notions nécessaires à la considération des travaux à base de traces. Tout d'abord, nous entendons par *EIAH utilisant les traces* tout EIAH dans lequel on peut relever l'utilisation de traces numériques, de quelque manière que ce soit, avec des degrés de généralité variables de manipulation des traces.

Ensuite, il existe actuellement plusieurs points de vue sur ce que pourrait être la définition d'une trace. Au-delà des divergences entre définitions, les objets considérés comme étant des traces sont souvent différents. On s'intéresse dans cet

article à la notion de trace numérique et on considérera qu'une trace numérique est une trace – inscrite sur un support numérique – de l'activité d'un utilisateur qui utilise un outil informatique pour mener à bien cette activité. Dans le contexte des EIAH, cette notion peut signifier deux choses : (1) l'histoire interactionnelle d'un apprenant utilisant un EIAH ou (2) les productions qu'il a laissées lors de son apprentissage. On se place clairement, dans la suite de cet article, dans le premier cas, la trace étant alors une empreinte restante de l'activité (propriété résiduelle) relatant l'historique et la chronologie de l'interaction de l'apprenant (propriété descriptive). Dans le second cas, la trace (production documentaire par exemple) n'est plus trace à cause de sa capacité à « retracer » l'interaction mais uniquement parce qu'elle peut être vue et considérée comme un élément passé restant et résultant de cette activité. Ce deuxième cas nous intéresse moins la propriété essentiellement temporelle et séquentielle des traces d'activités y étant perdue.

Avec des traces « histoires interactionnelles », un enseignant-tuteur peut guider l'activité individuelle ou collaborative en essayant de comprendre les dysfonctionnements éventuels par rapport au scénario qu'il avait préconisé. Il peut alors adapter la session, introduire des aides personnalisées, fournir des supports pédagogiques adaptés aux différents publics. Un enseignant-concepteur peut lui exploiter les traces pour personnaliser un scénario pédagogique, permettant ainsi de réguler le déroulement d'une session d'apprentissage en tenant compte de certains aspects qui ne peuvent être mesurés qu'en cours de session, tel le temps de réponse à un exercice. Un apprenant peut visualiser sa trace et se faire une image de son évolution dans l'activité, ce qui lui permettra de comprendre son cheminement dans la construction de sa connaissance. Ces quelques exemples d'utilisation des traces par les différents acteurs d'un EIAH donnent une vue très parcellaire des potentialités d'utilisation qu'elles offrent pour des participants pouvant avoir des besoins très différents. Il serait donc difficile de décrire exhaustivement l'ensemble des fonctionnalités imaginables.

Dans la suite de cet article nous appellerons « trace première » une trace telle que collectée au fil de l'activité et « trace transformée » toute interprétation de cette trace première dans une intention d'analyse particulière.

### **1.2. Usage des traces dans les EIAH : un état de l'art**

Il existe dans la littérature sur les EIAH plusieurs plateformes destinées à l'apprentissage individuel dans un domaine particulier et fondées sur la didactique d'une discipline. L'outil « Aplusix » (Bouhineau *et al.*, 2001) illustre très bien cette catégorie d'EIAH. Il aide des élèves à apprendre l'arithmétique et l'algèbre. Cet outil laisse l'élève résoudre les exercices et vérifie pas à pas que les solutions sont bien correctes et complètes. Pour cela Aplusix offre un éditeur d'expressions algébriques très intuitif (en deux dimensions) avec 400 patrons d'exercices.

D'autres outils comme « REFLET » (Després *et al.*, 2004) et « FORMID » (Guéraud *et al.*, 2004) affichent l'évolution du travail de chaque apprenant. Le tuteur

peut alors voir tout le détail des activités de ses apprenants. « Cabri Géometre » (Soury-Lavergne, 2001) développé par l'Université de Grenoble et le CNRS est destiné aux collégiens hospitalisés. Il permet de faire des calculs géométriques pour construire une figure à l'aide d'un ensemble de primitives. L'outil « TACSI » (Laperrousaz, 2006), également basé sur l'exploitation des traces, s'intéresse à la façon d'instrumenter les activités des tuteurs (et non pas celles de l'apprenant).

Quelques travaux se sont intéressés à la visualisation des traces issues des plateformes de téléformation collaboratives. On peut citer le travail de (Cram, 2007) issu du projet eLycée où la trace première est transformée en trace modélisée pour une activité de co-production dans une dyade. Ce travail propose un état de l'art sur la question de la visualisation de traces réflexives.

« GISMO » (Graphical Interactive Student Monitoring System for Moodle) (Mazza *et al.*, 2005) trace les activités des apprenants sur la plateforme Moodle. C'est un outil Open source avec une architecture modulaire, développé en Java et PHP, qui affiche les traces sous forme d'un graphe en représentant le nombre d'accès aux différentes ressources Moodle par les différents apprenants. Cette trace est de type « trace brute » (sans intention explicite pour celui qui l'exploitera).

« CourseVis » (Mazza *et al.*, 2004) est utilisé pour la visualisation 3D des traces issues de la plateforme WebCT. Le but est de faire des visualisations à partir de calcul d'indicateurs et des mesures statistiques. La visualisation est précédée par quelques transformations : extraction et sélection des données significatives à partir des sources de traçage ; transformation des données pour les mettre dans le format désiré ; transformation des données sous forme de primitives graphiques qui seront les briques pour la visualisation des données aux utilisateurs ; enfin génération de la visualisation graphique à partir des primitives. Cette visualisation peut être paramétrée par les utilisateurs. CourseVis permet aux tuteurs de disposer d'une sorte de tableau de bord d'une activité d'apprentissage (*suivi d'un groupe*).

L'analyse des interactions entre acteurs d'une situation d'apprentissage est largement utilisée pour comprendre les dynamiques en place et proposer des évolutions dans la manière de concevoir, réaliser, suivre, ou encore animer une situation d'apprentissage médiée par un environnement informatique. On trouve dans la littérature EIAH des dizaines d'expérimentations liées à des situations synchrones et asynchrones tel le projet CoPEAS (Communication Pédagogique en Environnement orienté Audio Synchron) (Betbeder *et al.*, 2006) décrivant une expérimentation d'apprentissage collaboratif à distance médiatisé par l'environnement Lyceum<sup>1</sup> L'expérimentation est une formation en anglais pour 14 apprenants en master professionnel FOAD et regroupe 14 apprenants divisés en 2 groupes. Les traces audio et vidéo issues de cette expérimentation sont transcrites pour faciliter l'étude et l'analyse du comportement collaboratif des étudiants. (Greffier *et al.*, 2006) décrit une expérimentation exploitant les environnements

---

1. <http://lyceum.open.ac.uk/>

asynchrones de type forum : 84 étudiants en licence informatique 'bac+2' suivent un module de programmation orienté objet 'POO' ; le niveau des étudiants est hétérogène, et le tuteur expert du domaine est en même temps animateur du forum. Les traces résultat sont utilisés pour faire une étude statistique et une classification des apprenants (ex : un étudiant occasionnel est un étudiant qui envoie de 1 à 8 messages dans un forum). L'expérimentation synchrone de (Beldame, 2006) avec deux groupes d'apprenants qui travaillent par paires pour construire un origami a permis de mettre en évidence les effets métacognitifs de traces réflexives en situation de tâche collaborative.

Tel qu'il a été défini par (Dimitracopoulou, 2004), « un indicateur est une variable au sens mathématique à laquelle est attribuée une série de caractéristiques ». Les valeurs de l'indicateur peuvent être représentées par des formes numériques, alphanumériques ou même graphiques. La valeur possède un statut, c'est-à-dire qu'elle peut être brute (sans unité définie), calibrée ou interprétée. Le statut identifie une caractéristique bien précise : celle du type d'assistance offert par l'indicateur aux utilisateurs. Chaque indicateur peut être ou non dépendant d'autres variables comme le temps, ou même d'autres indicateurs. Pour calculer un indicateur on utilise les traces brutes, les traces premières, ou même les traces transformées à un niveau supérieur de la trace première. Chaque indicateur respecte une méthode de calcul, et il est soit visualisé, soit utilisé dans d'autres calculs.

Parmi les travaux sur des indicateurs calculés à partir des données quantitatives figurent ceux de (Reffay *et al.*, 2006) sur le calcul de la cohésion et la centralité dans les réseaux sociaux et à partir des forums de discussion ou encore la plateforme ACOLAD (Jaillet, 2005) sur laquelle le tuteur dispose d'un outil qui lui fournit des informations sur le triplet d'activité (Assiduité, Disponibilité, Implication). (Santos *et al.*, 2003) proposent un outil qui calcule à partir des interactions, le degré d'implication de chaque apprenant dans la formation. Il identifie plusieurs types d'apprenant : participatifs, perspicaces, utiles, non-collaboratifs, qui prennent des initiatives, et enfin communicatifs. D'autres indicateurs sont interprétés qualitativement comme dans (Martinez *et al.*, 2003) où l'indicateur de la densité du réseau social est interprété à l'aide des histogrammes. Dans (Tedesco, 2003), on calcule l'accord et le désaccord entre les apprenants. On trouve également dans (Dimitracopoulou, 2004), une revue de la littérature sur les différents indicateurs permettant de calculer par exemple la division des tâches, le niveau des activités, etc.

Alors que de nombreux travaux s'intéressent à l'analyse des interactions dans les situations collaboratives asynchrones (typiquement les forums) comme (Bratitsis *et al.*, 2006) ou d'autres déjà cités plus haut, on ne trouve que peu de travaux qui s'intéressent au calcul d'indicateurs dans le cas collaboratif et dans les situations synchrones. D'après la définition de (Baker, 2002), la collaboration est basée sur les deux concepts : symétrie et alignement. On dit qu'une séquence d'interactions est symétrique si les participants adoptent dans le processus d'interaction des rôles, d'égaux à égaux. Une asymétrie signifie au contraire que les participants adoptent des rôles qui les placent à des niveaux différents dans une certaine hiérarchie.

L'alignement par rapport aux actions faites par les apprenants signifie qu'il y a une certaine ressemblance dans la façon dont ceux-ci agissent : ils sont en phase. Dans un alignement négatif, on a au contraire un déphasage, par exemple dans le cas d'un premier apprenant qui essaie de comprendre un problème alors que l'autre en est déjà à générer la solution et la communiquer aux autres, ou dans celui où deux apprenants travaillent sur la même phase du problème, mais chacun n'arrive pas à comprendre et à interpréter les activités de l'autre apprenant. Une collaboration est alors une séquence d'actions symétriques et alignées, où l'accord et le désaccord ne sont pas pris en compte dans le processus de collaboration. Ce type de séquence peut être repéré, étiqueté et donner lieu à une statistique fournissant un indicateur synthétique sur la collaboration en cours.

Nous pouvons citer aussi le travail de (Barros *et al.*, 2000) où l'on trouve une relation entre la coordination, la coopération et l'argumentation, pour arriver à la collaboration. Le calcul de ces trois facteurs dépend du calcul d'autres variables et indicateurs calculés par l'outil DEGREE (Barros *et al.*, 2000).

### **1.3. Traces et interactions médiatisées et médiées**

Nous reprenons ici les éléments présentés dans (Lund *et al.*, 2008) pour situer la question des traces et des interactions à des fins de supervision d'activité d'apprentissage.

Selon (Baker, 1997), une interaction est une suite d'actions verbales ou non-verbales qui sont interdépendantes et qui s'influencent mutuellement. On distingue deux types d'interactions possibles à travers un ordinateur : soit un humain interagit avec un ordinateur, on parle alors d'une suite d'événements prenant la forme d'une action-réaction ; soit plusieurs humains interagissent au travers d'un ordinateur, on parle alors d'une interaction *médiatisée* ou d'une interaction *médiée* par ordinateur. Pour (Pera, 2003) et (Meunier *et al.*, 2004) les processus de médiatisation et de médiation portent sur des objets différents. La médiatisation désignerait le processus de mise en dispositif médiatique, incluant ainsi le choix de médias et d'une scénarisation, alors que la médiation serait le processus par lequel un artefact technique et symbolique peut modifier les processus de production et de communication des connaissances ou encore influencer sur l'apprentissage ou sur le développement des processus cognitifs des êtres humains (Charlier *et al.*, 2007). Tracer les interactions pour établir des éléments utiles à la supervision d'une activité d'apprentissage collaboratif est donc une nécessité.

On peut distinguer plusieurs modes d'utilisation des traces dans les interactions médiatisées et médiées.

*Utilisation des traces pour les études empiriques sur les caractéristiques de l'interaction humaine.* Les études empiriques se focalisent sur une description des caractéristiques, voire de la qualité de l'interaction humaine, médiatisée et médiée par ordinateur. L'objectif de cette description est souvent de différencier entre les interactions humaines produites dans deux scénarii pédagogiques différents. Les

scénarii diffèrent par exemple, dans la succession des tâches demandées aux apprenants, dans les instructions précises qui leur sont données ou encore dans les outils mis à leurs dispositions. Cette différenciation entre interactions humaines produites est donc utilisée pour évaluer lequel des deux scénarii favorise davantage l'apprentissage en question. Nous pouvons citer le travail de (Dyke *et al.*, 2007) où a été mis en évidence comment les caractéristiques d'un éditeur de texte influencent les formes de l'écriture collective effectuée avec cet outil. Treize binômes devaient concevoir un texte procédural à distance par l'utilisation de la plateforme collaborative DREW (Corbel *et al.*, 2002). Il s'agissait d'un mode d'emploi pour la fabrication d'un pliage en origami. Certains des binômes utilisaient un éditeur de texte à accès par jeton et certains, un éditeur de texte libre : les participants devaient soit négocier le temps de rédaction, soit ils pouvaient rédiger en même temps. Dans ce cas, le destinataire de la trace est le chercheur. L'objectif précis est de déceler les formes différentes d'écriture collective. En ce qui concerne une adaptation à proposer, sous l'hypothèse que la co-élaboration favorise l'apprentissage.

*Utilisation des traces pour la réingénierie d'un EIAH.* La conception d'une nouvelle version d'un EIAH à partir d'une version existante se base sur l'observation d'indicateurs soit préalablement définis et instrumentés (souvent en collaboration avec l'enseignant) soit conçus à la volée en cas de nouveaux besoins, lors du déroulement d'un scénario pédagogique. Il s'agit de tenir compte des écarts observés entre le scénario tel qu'il a été conçu pour fonctionner (scénario prédictif) et tel qu'il s'est réellement déroulé (scénario descriptif), l'idée étant d'identifier comment la réingénierie de tel ou tel aspect pourrait faire mieux atteindre les objectifs pédagogiques. Nous pouvons citer dans ce cadre le travail de (Luengo *et al.*, 2006) qui propose un travail didactique portant sur l'analyse des connaissances des apprenants en milieu hospitalier (chirurgie orthopédique) exploitant un système d'enseignement et d'apprentissage sur le geste chirurgical. L'acteur produisant la trace est un apprenant agissant seul, qui s'entraîne à poser une vis dans le bassin d'un patient grâce à un simulateur. L'acteur manipulant la trace est l'EIAH et cette manipulation se fait en temps réel. Le type de manipulation effectuée porte sur la modélisation des connaissances de l'utilisateur en fonction des actions qu'il effectue sur l'interface du simulateur pendant la pose de la vis ainsi que sur la prise de décision didactique qui suit. L'objectif précis est l'automatisation de la prise de la décision didactique en utilisant les connaissances chirurgicales représentées par des réseaux bayésiens. En termes d'adaptation proposée, les auteurs préconisent trois types : pointer vers les cours en ligne qui sont pertinents par rapport au diagnostic du travail de l'apprenant ; choisir un problème similaire et mettre le simulateur dans un état précis afin que l'apprenant puisse le prendre en main ; ou choisir un cas clinique et le présenter à l'utilisateur. Ces adaptations ne sont pas faites en temps réel mais sont proposées suite à la décision diagnostic et guident la réingénierie du dispositif.

*Utilisation des traces pour l'adaptation d'une interaction en temps réel.* Les travaux de (France *et al.*, 2007) montrent comment une visualisation de trace en temps réel permet au tuteur d'observer l'activité d'un groupe d'apprenants et

d'adapter l'activité en interagissant directement avec l'interface de visualisation. Les travaux de (May *et al.*, 2007) sont similaires, mais se focalisent sur une activité plus restreinte – l'utilisation d'un forum. L'acteur produisant la trace est l'apprenant lecteur. L'acteur manipulant la trace est l'EIAH. Le destinataire de la trace est multiple pour ces auteurs – l'apprenant, l'enseignant ou le chercheur. La manipulation de la trace est faite en temps réels et il s'agit de calculer de manière automatique le nombre de messages lus par personne ainsi que le temps qui aurait pu être passé pour la lecture d'un message. En effet, il n'est pas possible d'affirmer avec certitude qu'un apprenant a lu un message. Par contre, il est possible d'indiquer si un apprenant donné n'a pas vu un message ou une partie de message.

Les travaux cités exploitent de manière explicite ou implicite des traces d'interaction sans définir formellement la notion de trace et encore moins les opérateurs qui permettent de les utiliser pour l'analyse d'activité. Cette question d'une trace *modélisée* et des *transformations* permettant de réaliser les analyses au niveau convenable de modélisation est au cœur du travail présenté dans cet article.

## 2. Systèmes à Base de Traces Modélisées (SBT)

Nous présentons dans cette section la notion de Systèmes à Base de Traces Modélisées comme une solution générique au problème de la modélisation et de la manipulation des traces. Afin de définir de tels systèmes, nous définissons et formalisons tout d'abord les concepts de trace et de modèle de trace (section 2.1). Nous présentons ensuite la notion de SBT (section 2.2), et précisons les étapes de collecte et de transformation. Une vue d'ensemble des utilisations générales des SBT dans le contexte des EIAH permettra d'en définir les fonctionnalités (section 2.3).

### 2.1. Trace : modélisation et formalisation

Nous allons dans cette partie définir la notion de trace et de modèle de trace.

#### 2.1.1. *Observé et Trace*

Dans sa définition la plus générale<sup>2</sup>, « une trace est une chose ou une suite de choses laissées par une action quelconque et relatives à un être ou un objet ; une suite d'empreintes ou de marques que laisse le passage d'un être ou d'un objet ; ce à quoi on reconnaît que quelque chose a existé ; ce qui subsiste d'une chose passée ». Dans notre contexte, une trace numérique est issue de l'observation d'une activité, elle représente une marque rémanente d'un processus interactionnel.

Une trace numérique est composée d'objets qui sont situées les uns par rapport aux autres parce qu'on les observe et qu'on les inscrit sur un support. Cela signifie

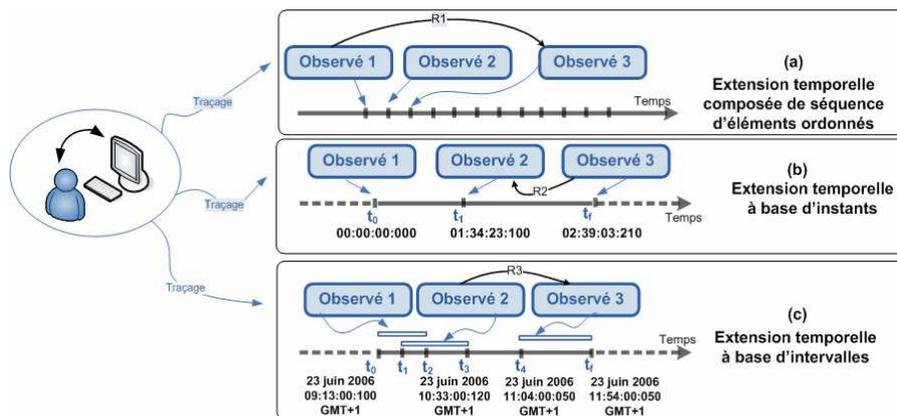
---

2. Dictionnaire Petit Robert. Version électronique du nouveau petit robert, dictionnaire alphabétique et analogique de la langue française, 2005. En ligne ; Page disponible le 22-décembre-2005.

qu'une trace est explicitement composée d'objets arrangés et inscrits par rapport à une représentation du temps de l'activité tracée (Figure 1). L'arrangement peut être séquentiel explicite (chaque observé est suivi et/ou précédé par un autre) ou découler du caractère temporalisé des objets de la trace. Pour donner notre définition de la notion de trace, nous définissons une représentation du temps de l'activité tracée que nous appelons extension temporelle. Différentes traces peuvent utiliser différentes représentations de temps, cependant, formellement, toutes les traces partagent ceci : on peut voir ce temps comme une collection ordonnée d'instants discrets, tel que chaque instant a un successeur unique. Nous nommons cette représentation du temps un domaine temporel.

**Définition 1** (Domaine Temporel). *Un domaine Temporel  $T$  est un ensemble dénombrable d'instants. Une fonction bijective  $SuccT : T \rightarrow T$  associée à chaque instant son successeur, définissant un ordre total  $\leq_T$  dans  $T$ .*

À l'intérieur d'un domaine temporel donné, une trace décrit seulement un intervalle fini de temps. Nous appelons un tel intervalle l'*extension temporelle* de la trace.



**Figure 1.** Trace, observés et extension temporelle de l'activité

**Définition 2** (Extension Temporelle). *Étant donné un domaine temporel  $T$ , une extension temporelle  $\xi_T$  est un élément quelconque de  $I(T)$  (avec  $I(T)$  l'ensemble de tous les intervalles finis sur  $T$ ).*

On appelle *trace* une collection d'observés temporellement situés. On dénote par *observé* toute information structurée issue de l'observation d'une interaction. Un observé est dit « temporellement situé » dès le moment où il est associé à une partie de l'extension temporelle de la trace à laquelle il appartient.

### 2.1.2. Trace modélisée

Une trace numérique obéit à un modèle de trace, qui décrit les objets qui en font partie. Quelque soit le niveau d'abstraction des éléments de la trace (i.e. leur proximité aux événements informatiques), le modèle de trace vient préciser comment il est possible de les comprendre et de les utiliser. Un modèle de trace peut se limiter à une description des classes d'objets, mais peut aussi prendre en compte les types de relations.

On appelle modèle de trace le vocabulaire de la trace. Le modèle de trace permet la compréhension de la trace en en décrivant abstraitement ses éléments.

**Définition 3** (Modèle de Trace). Un modèle de trace est défini comme un tuple

$$MTR = (T, C, R, A, dom_R, range_R, dom_A, Range_A)$$

Il est constitué de :

- un domaine temporel  $T$ ,
- un ensemble fini  $C$  de types d'observé (ou classes) et un ordre partiel  $\leq_C$  défini sur  $C$ ,
- un ensemble fini  $R$  de types de relations disjoint de  $C$  et un ordre partiel  $\leq_R$  défini sur  $R$ ,
- un ensemble  $A$  fini d'attributs disjoint de  $C$  et  $R$ ,
- deux fonctions  $dom_R : R \rightarrow C$  et  $range_R : R \rightarrow C$  définissant le domaine et le codomaine des types de relation,
- deux fonctions  $dom_A : A \rightarrow C$  and  $range_A : A \rightarrow D$  définissant le domaine et le codomaine des attributs,

Pour tout couple de relations  $r1$  et  $r2$  d'un modèle de trace, on doit garantir la propriété suivante:

$$r1 \leq_R r2 \Rightarrow dom_R(r1) \leq_C dom_R(r2) \wedge range_R(r1) \leq_C range_R(r2)$$

Intuitivement, un modèle de trace définit un vocabulaire pour décrire des traces en définissant : comment le temps est représenté ( $T$ ), comment les observés sont catégorisés et classés ( $C$ ), quelles relations peuvent exister entre des observés ( $R$ ), quels attributs décrivent chaque observé ( $A$ ). Les domaines de départ et d'arrivée contraignent les types de relations et les attributs qu'un observé d'un type donné peut avoir. Les ordres partiels  $\leq_C$  and  $\leq_R$  définissent une hiérarchie de types pour les observés et les relations. La dernière contrainte garantit la cohérence et consistance du domaine et du codomaine pour une relation et ses parents dans la hiérarchie.

On appelle trace modélisée (M-Trace dans la suite) l'association d'une collection d'observés temporellement situés et d'un modèle explicite de cette collection d'observés. Une M-Trace est toujours associée à un modèle de trace définissant les

éléments qui la composent : observés et relations. Un modèle de trace peut être implicite, c'est-à-dire uniquement présent implicitement dans le code de l'outil l'utilisant ; par exemple le Common Log Format est le modèle des traces du serveur Apache. Il peut au contraire être explicite, c'est-à-dire formalisé suffisamment pour permettre l'échange et la réutilisation de traces (e.g. un schéma XML décrivant et contraignant le contenu d'une trace XML).

**Définition 4** (*M-Trace*). Une *M-Trace* est un tuple

$$TR = (MTR, \xi_T, O, \lambda_C, \lambda_R, \lambda_A, \lambda_T)$$

Constitué de

- un modèle de trace  $MTR = (T, C, R, A, dom_R, range_R, dom_A, range_A)$ ,
- une extension temporelle  $\xi_T$  définit dans  $T$ ,
- un ensemble fini  $O$  d'observés, disjoint de  $C, R$  et  $A$ ,
- une fonction totale<sup>3</sup>  $\lambda_C : O \rightarrow C$  décrivant l'instanciation des observés,
- une relation  $\lambda_R \subseteq O \times O \times R$  décrivant l'instanciation des relations,
- une fonction partielle<sup>4</sup>  $\lambda_A : O \times A \rightarrow V$  décrivant l'instanciation des attributs,
- une fonction totale  $\lambda_A : O \rightarrow I(\xi_T)$  décrivant la temporalisation des observés.

De plus, une trace doit être consistante avec son modèle, i.e. vérifier les contraintes suivantes :

- $\lambda_C(o)$  (respect des types d'observés),
- $\forall (o1, o2, r) \in \lambda_R, \lambda_C(o1) \leq_c dom_R(r) \wedge \lambda_C(o2) \leq C \ range_R(r)$ ,
- $\forall (o, a, v) \in \lambda_A, \lambda_C(o) \leq_c dom_A(a) \wedge v \in range_A(a)$ .

Intuitivement, une trace représente, en respectant un modèle de trace ( $MTR$ ), le résultat d'une observation pendant une période donnée ( $\xi_T$ ). Elle contient un ensemble d'observés typés ( $O$  et  $\lambda_C$ ) situés dans le temps ( $\lambda_T$ ), en relation les uns avec les autres ( $\lambda_R$ ) et ayant des attributs valués ( $\lambda_A$ ). Chaque observé  $o$  a exactement un type direct (le  $\lambda_A$  est une fonction totale).

Notons aussi que l'ordre partiel  $\leq_c$  induit une représentation de hiérarchie de type, c'est-à-dire que chaque type  $C \leq_c$  ;  $\lambda_C(o)$  peut être considéré comme un type indirect de  $o$ . Il peut y avoir une ou plusieurs relation(s) entre deux observés donnés ( $\lambda_R$  peut être n'importe quelle relation). Les valeurs d'attribut ne sont jamais obligatoires. La *M-Trace* est compatible avec son modèle si (1) son extension temporelle appartient au domaine temporel du modèle (2) si les contraintes du domaine et codomaine des relations et des attributs sont toutes satisfaites.

---

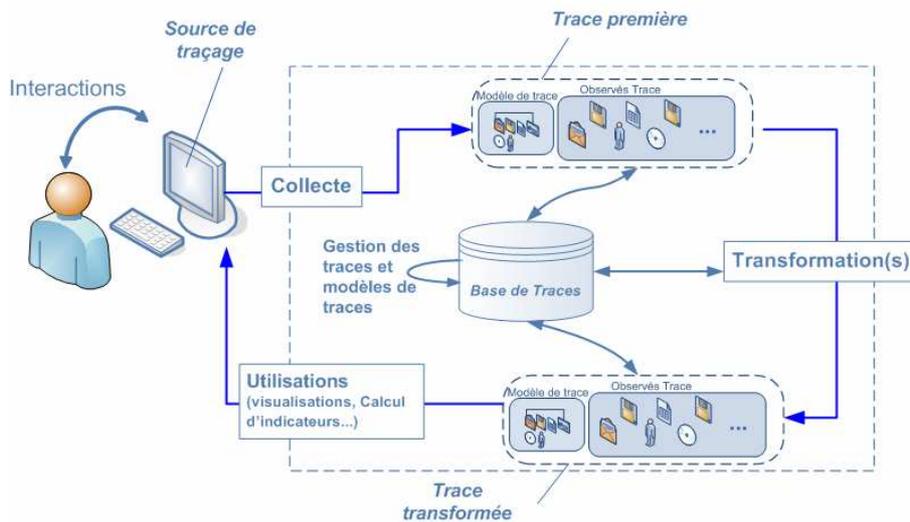
3.  $\lambda_C$  est défini sur chaque  $o \in O$

4.  $\lambda_{A(o, a)}$  peut être indéfini pour un certain  $(o, a) \in O \times A$

## 2.2. Principe général des SBT

Les concepts de trace et de modèle de traces définis dans la section précédente sont le noyau du cadre conceptuel que nous proposons (Settoui *et al.*, 2006). Un système informatique implémentant ces deux notions est un *système à base de traces modélisées* qui offrira divers services de manipulation de traces que nous décrivons dans la suite.

On appelle *système à base de traces modélisées* (voir Figure 2) tout système informatique dont le fonctionnement implique à des degrés divers la gestion, la transformation et la visualisation de *traces modélisées explicitement* en tant que telles.



**Figure 2.** Principe général d'un système à base de traces modélisées

On peut ainsi construire un ou plusieurs systèmes informatiques à base de traces qui permettent de collecter, gérer, transformer, visualiser des m-traces ayant des modèles explicites différents (Figure 2). Nous allons à présent décrire les processus de collecte et de transformation et les définitions associées.

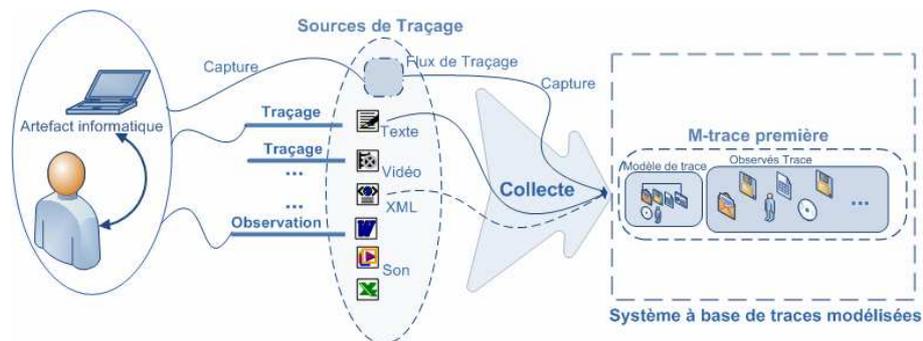
### 2.2.1. Collecte de traces modélisées

La collecte permet de mettre en place l'observation de l'utilisation d'un système à partir de *sources de traçage*. Elle consiste à transformer de façon automatique ou

semi-automatique, pendant l'activité ou *a posteriori*, des informations<sup>5</sup> générées par l'interaction utilisateur/système en une *m-trace première* du SBT (Figure 3).

On appelle *source de traçage* tout flux d'information structuré à partir duquel il est possible de mettre en place un processus de collecte de traces pour un système à base de traces.

On appelle *collecte de m-traces* le processus qui consiste à exploiter de façon automatique, semi-automatique ou manuelle un ensemble de sources de traçage pour construire une *m-trace* du SBT. La construction d'une telle *m-trace* nécessite la construction d'une collection d'observés temporellement situés associée à un modèle de *m-traces*. Une *m-trace* issue de la collecte est appelée *m-trace première* du SBT, car c'est la première à être manipulable dans ce système à l'issue du processus initial de collecte.



**Figure 3.** Collecte de *m-trace première*

Il existe deux techniques de traçage dans les EIAH. La première concerne les traces générées en interne par l'EIAH lui-même, qui embarque un module de traçage<sup>6</sup>, lequel peut être paramétrable. Ce module conçu par les concepteurs de l'EIAH donne des traces plus ou moins pertinentes au regard de ses fonctionnalités (par exemple des éléments proches du scénario). La deuxième technique consiste à instrumenter l'EIAH par des logiciels de traçage externes scrutant et espionnant tout de qui se passe (espions informatiques, *Keylogger*<sup>7</sup>, etc.). Les traces générées par cette technique sont, la plupart du temps, inexploitables dans leur état brut initial.

5. Ces informations correspondent par exemple aux logs avec le système lui-même comme source de traçage.

6. Si l'EIAH est une application Web, on pourra récupérer, en plus de ce que le module de traçage capture, les traces laissées traditionnellement par le navigateur client et le serveur web.

7. Applications installées sur des ordinateurs enregistrant toute frappe clavier dans un fichier trace.

Une des raisons vient du fait qu'elles retiennent les informations<sup>8</sup> sans intention explicitée, et sont trop/pas assez riches en détails, trop/pas assez proches des événements informatiques et ne permettent presque jamais et sans ambiguïté<sup>9</sup> de comprendre l'activité observée.

Le processus de collecte réduit en partie cette ambiguïté en prenant en compte différentes sources de traçage issues de différentes techniques de traçage afin d'obtenir une trace première conforme à un modèle explicite.

### 2.2.2. Base de m-traces et transformation de m-traces

On appelle *base de m-traces* d'un SBT l'ensemble des m-traces qui sont manipulées par le système à base de m-traces.

La trace obtenue à l'issue de la collecte est qualifiée de *m-trace première*. Celle-ci n'est cependant pas toujours exploitable directement, et il faut parfois passer par une ou plusieurs transformation(s) pour atteindre une trace d'un niveau d'abstraction cohérent avec l'activité, *i.e.* signifiant dans le contexte pour l'utilisateur.

On appelle *transformation de m-traces* tout processus qui transforme une ou plusieurs m-trace gérée(s) par un système à base de m-traces en une autre m-trace gérée par le même système. Les *m-traces premières* d'une base de m-traces d'un SBT sont les seules m-traces non transformées de ce SBT.

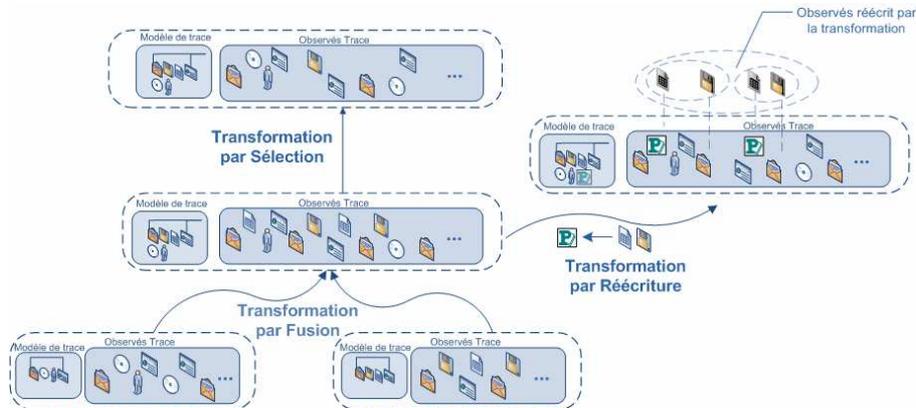
L'exploitation de la trace consiste en partie en sa transformation. Deux catégories de transformations sont prises en charge par le SBT : les transformations manuelles et les transformations automatiques utilisant un modèle de transformation.

Une transformation manuelle désigne toute modification impliquant un changement dans la composition des éléments de la trace, due à un ajout, un filtrage ou un enrichissement de ses observés. Ce type de transformation est réalisé par un utilisateur interagissant avec sa trace pour la mettre à jour ainsi que son modèle. Le SBT permet de garder la cohérence entre les traces et leurs modèles lors d'une transformation manuelle.

---

8. Conséquence du syndrome des informations manquantes, on trace et on garde toutes les informations possibles.

9. Déterminer la sémantique d'une trace brute est d'une part difficile à faire manuellement par un humain et d'autre part impossible à faire de manière directe et automatique puisque nécessitant des algorithmes spécifiques.



**Figure 4.** Exemples de transformations possibles dans le SBT

Les transformations automatiques sont appliquées dans le SBT selon des *modèles de transformation*. Un modèle de transformation est un ensemble de règles exprimant des filtres de sélection ou de réécriture de motifs<sup>10</sup>. Les transformations automatiques peuvent être combinées.

On peut distinguer trois types de transformations automatiques (Figure 4).

Les transformations de type *sélection* permettent la création d'une nouvelle trace contenant tous les observés respectant un filtre de sélection donné. Ces transformations permettent entre autres de séparer les observés pertinents de la trace du *bruit*. Le but est d'extraire un sous-ensemble d'observés de la trace approprié en exprimant le filtre adéquat qui donne les conditions de sélection sur les composants de la trace. Les filtres sont des contraintes sur les objets de la trace, le domaine temporel, et les relations structurelles.

Les transformations de type *réécriture de motifs* permettent de remplacer un ou plusieurs observés par un autre observé. Ces transformations peuvent toucher une séquence d'observés ou des observés qui ne se suivent pas directement. L'expression d'une succession d'observés représentant un motif (pattern) exprimant la séquence qu'on veut réécrire.

Les transformations par *fusion temporelle* consistent à fusionner, assembler et grouper les observés de ces traces en tenant compte de la temporalité de leurs éléments afin d'obtenir une nouvelle trace. La trace transformée par fusion suit un modèle de trace regroupant les modèles des traces fusionnées. Cette transformation est possible si les traces fusionnées possèdent des extensions temporelles homogènes<sup>11</sup>.

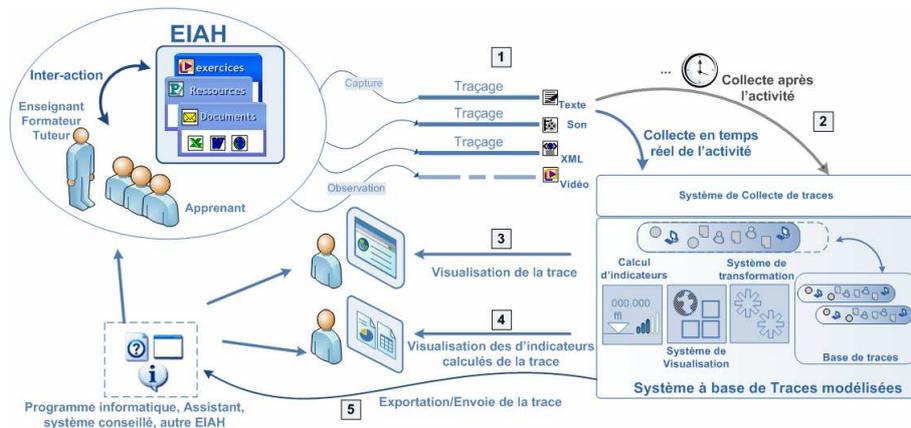
10. Une représentation d'une séquence ou d'un sous ensemble d'observés de la trace.

11. Permettant des translations et conversions de temps (intervalles vs instants).

### 2.3. EIAH et Systèmes à Base de Traces Modélisées

Les nombreuses utilisations des traces que nous avons présentées dans la première partie de cet article nous permettent d'en mettre en évidence deux dimensions importantes : le *temps de l'utilisation* et *l'acteur utilisant les traces*. L'utilisation des traces peut se faire dans le temps de leur collecte pendant une session (temps réel), ou bien on peut avoir une utilisation différée. Les traces peuvent être utilisées par l'utilisateur qui les a générées (réflexivité), ou bien par un autre utilisateur (analyste, enseignant).

La Figure 5 illustre ces différentes possibilités. Les étapes 1 et 2 montrent un EIAH traçant l'interaction de ses utilisateurs et fournissant des sources à collecter pour le SBT. Alimenté par ces sources, le SBT permet de gérer et de traiter ainsi que de visualiser les traces (étape 3). Il permet aussi d'effectuer des calculs d'indicateurs (étape 4), des exportations de traces interprétées et adaptées aux besoins d'un autre système (étape 5). Les étapes 3, 4 et 5, mettant en jeu des transformations de m-traces en m-traces, elles peuvent se succéder dans un ordre quelconque, se chevaucher ou être menées parallèlement.



**Figure 5.** Différentes utilisations d'un SBT au service d'un EIAH

Nous proposons donc trois types d'utilisation d'un SBT au sein d'un EIAH, qui correspondent aux étapes 3, 4 et 5 de la figure 5 :

- la *visualisation des traces* consiste pour un utilisateur à explorer une trace ou une base de traces, de façon plus ou moins interactive ;
- le *calcul d'indicateurs* sur les traces vise à fournir à l'utilisateur des indicateurs résumés d'une ou plusieurs traces (nombres, tableaux, etc.) ;

- la *gestion et le traitement des traces au service d'un autre système* visent, dans le cadre d'une utilisation non directe des traces par un utilisateur humain, à fournir des traces ou des indicateurs à un autre système informatique, qui à son tour les utilisera pour mener d'autres calculs. Le SBT s'intègre alors dans une chaîne d'outils, en y jouant le rôle d'outil de gestion et de traitement des traces.

### **3. SBT-IM : un SBT implémenté exploité pour le calcul des Indicateurs dans Moodle.**

Dans cet article, nous illustrons le deuxième type d'utilisation présenté précédemment : le calcul d'indicateurs. Nous présentons une architecture en trois niveaux : collecte, transformation des m-traces et calcul des indicateurs à partir de ces transformations. Nous donnons par la suite des exemples précis sur l'utilisation de SBT-IM pour le calcul d'indicateurs concrets. Pour introduire ce travail, nous présentons dans la section suivante, un aperçu sur les indicateurs à mettre en œuvre ainsi que quelques exemples issus de la littérature EIAH.

#### **3.1. Traces et indicateurs en EIAH**

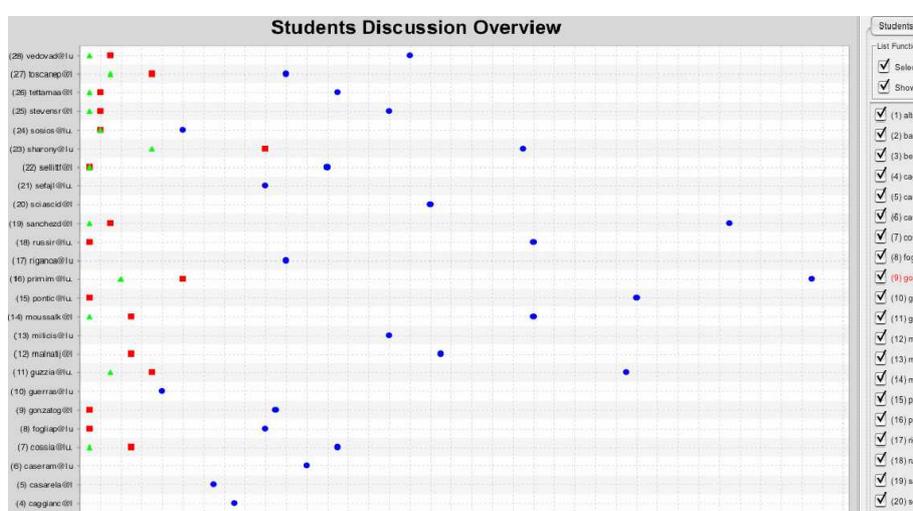
##### *3.1.1. Les indicateurs en EIAH : définition et exemples*

Une trace présentée de manière très synthétique sous forme de bilan peut permettre de fournir à l'enseignant une analyse « objective » du travail d'un apprenant (Renie, 2000). Utilisés pendant ou après l'activité d'apprentissage, des indicateurs simples peuvent décrire des informations précises et ponctuelles, comme par exemple le nombre d'accès à une ressource en ligne, le nombre d'essais pour chaque réponse, le temps passé dans chaque activité, le taux de réussite lors d'une activité.

Dans le contexte des EIAH collaboratifs, le calcul d'indicateurs est d'une autre dimension. En effet, l'indicateur pertinent doit permettre de synthétiser des statistiques sur les traces de l'espace de travail commun des apprenants (du groupe) et les traces individuelles de chacun. La définition des indicateurs dans ce contexte tente de décrire le comportement des apprenants dans la collaboration.

Le calcul d'indicateurs à partir de traces pour soutenir la collaboration et l'apprentissage collaboratif est un champ de recherche très actif. Selon (Dimitracopoulou *et al.*, 2006) un indicateur est une variable au sens mathématique à laquelle est attribué une série de caractéristiques. C'est une variable qui prend des valeurs représentées par une forme numérique, alphanumérique ou même graphique... La valeur possède un statut elle peut être brute (sans unité définie), calibrée ou interprétée. Le statut identifie une caractéristique bien précise celle du type d'assistance offert aux utilisateurs. Chaque indicateur est soit indépendant, soit dépendant d'autres variables comme le temps, ou même d'autres indicateurs.

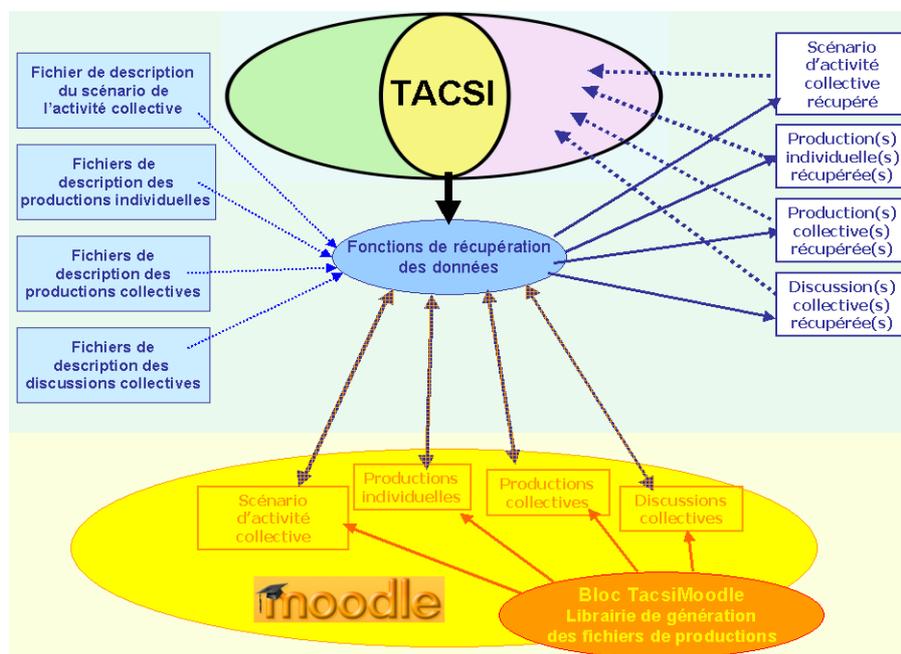
Il existe dans la littérature quelques travaux basés sur les traces Moodle<sup>12</sup>. L'outil GISMO, déjà cité page 4, (Graphical Interactive Student Monitoring System for Moodle) (Mazza *et al.*, 2005) trace les activités des apprenants sur Moodle. C'est un logiciel open source avec une architecture modulaire, développé en java et php, et affiche les traces sous forme d'un graphe en représentant le nombre d'accès aux différentes ressources Moodle par les différents acteurs. Sur la figure 6, GISMO donne pour chaque apprenant (en ordonnée), le nombre de messages lus, postés et initiant une discussion.



**Figure 6.** Présentation graphique des discussions faites dans Moodle entre les apprenants

(Laperrousaz, 2006) propose une connexion entre l'outil TACSI (Tuteur d'Activité Collective et Suivi Individualisé) et Moodle (Figure 7). Pour établir ce lien, une librairie de fonctions Moodle permet, grâce aux traces, de faire le lien entre Moodle et TACSI et d'accéder aux caractéristiques des apprenants et à leurs différentes productions. Lorsque le tuteur lance TACSI la librairie est appelée et les fichiers sont régulièrement mis à jour, ce qui permet au tuteur de suivre son groupe d'apprenants de manière synchrone.

12. Moodle : [www.moodle.org](http://www.moodle.org)



**Figure 7.** Communication entre TACSI et Moodle

Une autre librairie<sup>13</sup> permet aux utilisateurs d'indiquer manuellement qu'ils ont fini une activité et d'informer les enseignants de la progression des apprenants au cours d'une activité.

Toutefois, sans interface permettant d'y accéder, l'analyse des traces reste difficile. Comme exemple de cette difficulté, on peut citer (Rezeau, 2007) qui présente une analyse d'un corpus Moodle de productions de 106 apprenants utilisant des forums de discussion pour apprendre l'anglais : cette analyse est rendue fastidieuse par l'absence de module d'exportation de trace depuis la plateforme Moodle. Les messages postés par les étudiants dans 4 cours concernés ont été exportés au format texte, nettoyés des éventuelles balises HTML, étiquetés (nom abrégé du cours et nom de l'étudiant) et enfin concaténés en un seul fichier au format texte pur.

### 3.1.3. Calculer les indicateurs à partir des traces non modélisées

Nous présentons ci-après quelques exemples d'indicateurs implémentés sur des plateformes variées.

13. <http://tracker.moodle.org/browse/MDL-15498>

L'indicateur *proportion entre deux types d'actions*, implémenté dans la plateforme COTRAS (Jerma, 2004), compare deux types d'actions différents comme par exemple la proportion entre les actions de type chat messages privés, etc. Nous pouvons donner comme comparaison significative le calcul de l'équilibre entre les actions qui servent à la production dans la résolution d'un problème donné, et les actions classées comme actions de dialogue entre les utilisateurs de la plateforme.

L'indicateur *taux de participation* implémenté dans la plateforme FREESTYLER (Gassner, 2003) calcule le taux de participation des acteurs par rapport à un type d'action donné. Cet indicateur permet par exemple aux enseignants de repérer les apprenants actifs par rapport au nombre d'actions faites par chaque apprenant. Il permet aussi aux apprenants de régler leurs efforts par rapport à la valeur affichée par l'indicateur.

L'indicateur *nombre des actions faites par tous les acteurs* implémenté dans la plateforme DEGREE (Barros *et al.*, 2000) calcule le degré de contribution des différents acteurs dans une activité, et par rapport à une action. Il affiche pour chaque type d'action le nombre des actions faites par tous les acteurs. Cela aide l'enseignant à voir le type d'action le plus utilisé par les apprenants. Il peut ainsi choisir par exemple d'analyser les actions de type « Chat » et non pas de type « Ecrire dans un forum » au cas où le nombre des actions de type « Chat » est plus significatif.

L'indicateur *degré de centralité* proposé dans (Martinez *et al.*, 2003) est calculé par des techniques d'analyse des réseaux sociaux, et visualisé par un socio-diagramme. L'indicateur représente le nombre de liaisons qu'un participant maintient avec d'autres participants. Chaque participant est souvent représenté par un cercle ou une étoile dans le socio-diagramme, avec son nom ou le code correspondant (connu seulement par l'individu lui-même et l'administrateur du système).

### **3.2. Architecture globale du SBT-IM**

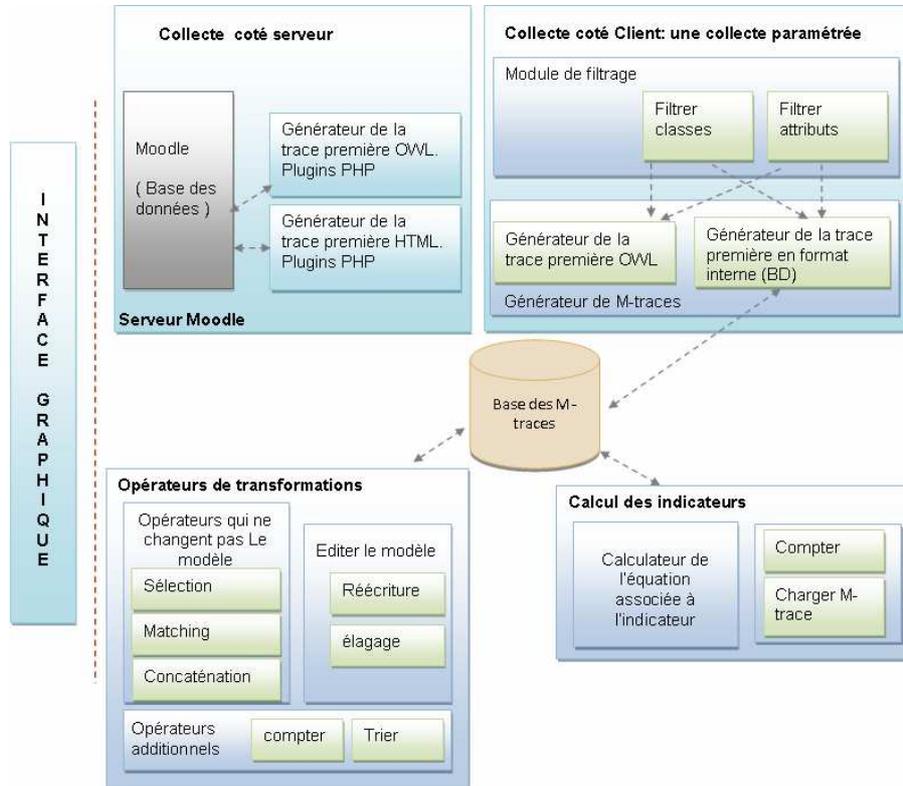
Nous avons implémenté « SBT-IM » : un Système à Base de Traces pour calculer les Indicateurs dans Moodle. Nous décomposons SBT-IM en trois sous-systèmes : le premier permettant la collecte des données à partir des plateformes d'apprentissages en une trace première ; le second gère la transformation des m-traces ; tandis que le troisième permet le calcul des indicateurs à partir des m-traces transformées.

L'objectif essentiel du *sous-système de collecte* est de récupérer à partir des sources de traçage Moodle les données nécessaires pour générer en temps réel la m-trace première Moodle. Dans un travail précédent, nous avons proposé une trace première basée sur une ontologie suite à une expérimentation que nous avons réalisée dans Moodle (Djouad, 2008). Nous avons décomposé ce sous-système en deux modules : un module pour la collecte coté serveur et un module pour la collecte coté client. Le module de collecte intégré du coté serveur génère la m-trace première selon deux formats OWL et HTML. Le format HTML permet de visualiser

directement la m-trace première mais pas de la réutiliser, alors que le format OWL enregistre la m-trace première ce qui permet de la réutiliser. Le module de collecte intégré du côté client offre aux utilisateurs la possibilité de faire une collecte paramétrée, et de générer des m-traces premières selon leur besoin. Dans ce cas, la m-trace première ne contient que les données dont l'utilisateur estime avoir besoin. Le fonctionnement des modules de collecte est expliqué dans le tableau 1.

Le *sous-système de transformation* met en œuvre les opérateurs qui transforment les m-traces. Nous proposons trois classes d'opérateurs : une classe d'opérateurs de transformation de trace qui ne modifient pas le modèle de la m-trace, une classe d'opérateurs de transformation de trace qui le modifient, et des opérateurs spécifiques pour le calcul d'indicateur. La base de m-trace, qui permet de stocker les traces, permet également de conserver les informations des transformations ayant été utilisées. Toute trace est ainsi associée aux transformations qui ont permis de la construire. L'un des avantages décisifs de cette association entre la m-trace et sa transformation, est la possibilité de réutiliser, simplement et en contexte, la séquence de transformations pour créer d'autres m-traces équivalentes.

Le sous-système qui calcule l'indicateur associe à chaque indicateur à calculer une ou plusieurs m-traces et une règle de calcul. La m-trace utilisée pour calculer la valeur de l'indicateur est le résultat d'une séquence de transformations de la m-trace première. Pour obtenir la valeur de l'indicateur, on applique une statistique sur les observés utilisés par la règle de calcul, et enfin on calcule la valeur de l'indicateur. La figure 8 présente l'architecture générale de notre SBT-IM.



**Figure 8.** Architecture du SBT-IM

Le tableau suivant explique le rôle de chaque module de cette architecture

Sous système	Module	Rôle
	Module de filtrage et de sélection des données à collecter	Ce module sélectionne les différentes classes et attributs de classes à collecter. Ce module regroupe deux filtres : « Filtrer classes » sélectionne les classes à collecter dans la m-trace première, et « Filtrer attributs » sélectionne les attributs à collecter dans chaque classe de la m-trace première.
	Module qui génère la trace première coté client	Ce module génère la m-trace première en format OWL et l'enregistre en format interne dans une base de données, pour être ré-utilisée par la suite par le système de transformation.

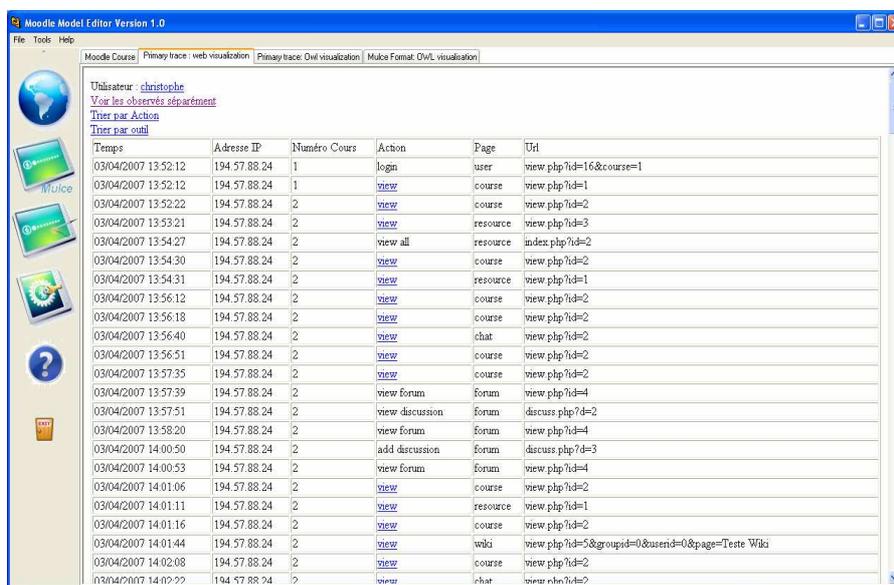
Sous-système de collecte	Générateur de m-trace première en format OWL coté serveur	Ce module est écrit entièrement en PHP, et les scripts s'exécutent sur le serveur où la plateforme Moodle est installée. Ce module interroge les sources de traçage Moodle, construit la m-trace première Moodle en format OWL, et envoie le résultat sur le poste client.
	Générateur de m-trace première en format HTML.	Ce module fonctionne de la même façon que le module précédant, et génère la m-trace première Moodle mais sous format graphique : plusieurs pages html reliées entre elles par des liens hypertextes. La m-trace première générée par ce module est plus lisible par l'observateur que celle en format OWL.
Sous-système de transformation.	Opérateurs de transformation qui ne changent pas le modèle	L'opérateur de <i>sélection</i> sélectionne un ensemble d'observés selon un critère de sélection comme le type d'observé, l'acteur associé à l'observé, le type de l'outil d'activité, etc. L'opérateur de <i>concaténation</i> regroupe les observés de deux m-traces différentes dans une nouvelle m-trace. L'opérateur de <i>matching</i> sélectionne une séquence d'observés à partir d'un motif
	Opérateurs de transformation qui modifient la structure du modèle	Nous proposons deux types d'opérateurs qui modifient le modèle de la m-trace : l'opérateur d' <i>élagage</i> qui supprime dans un modèle de m-trace une classe ou un attribut, et l'opérateur de <i>réécriture</i> qui réécrit une classe ou un attribut d'un modèle de m-trace.
	Opérateurs spécifiques aux statistiques / indicateurs	L'opérateur <i>Compter</i> qui compte le nombre des observés, et l'opérateur <i>tri</i> qui trie les observés selon des critères comme par exemple le temps.
Le Sous-système qui calcule l'indicateur <i>Indicator Calculus (IC)</i>	Ce module charge les m-traces, compte les instances, et calcule la valeur de l'indicateur avec des opérateurs arithmétiques (addition, soustraction, multiplication et division).	

**Tableau1.** Les différents sous modules de l'architecture SBT-IM

### 3.3. Interface graphique

L'interface graphique offre aux utilisateurs du SBT-IM une interaction efficace avec les services de collecte des données, de transformation des m-traces, et du calcul de l'indicateur.

L'interface (voir figure 9) communique avec le sous-système de collecte des données côté serveur. Chaque module possède son interface graphique accessible par onglet. L'onglet « *Primary trace : web visualization* » constitue l'interface du collecteur de m-trace première en format HTML et affiche la m-trace première dans cet onglet. Des options sont disponibles pour afficher les observés comme : voir les observés par utilisateur, par type d'observé, etc. Nous pouvons voir comme exemple dans la figure 9 la liste des observés liés à l'acteur *Christophe*. Ces observés font partie de la m-trace première. L'onglet : « *Primary trace : OWL visualization* » constitue une variante de l'interface du collecteur de m-trace première en format OWL et affiche le résultat selon ce point de vue. La trace première est représentée en OWL, ce qui permet de faire des requêtes sur cette m-trace première avec une sémantique précise. Le SBT-IM se connecte automatiquement à la plateforme Moodle grâce à un paramétrage accessible par menus graphiques et interactifs.

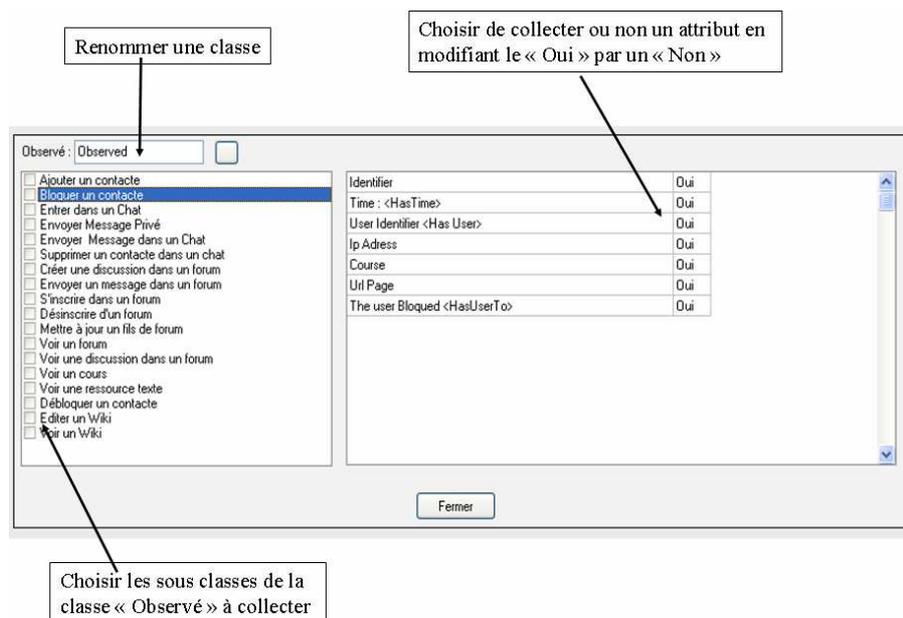


Temps	Adresse IP	Numéro Cours	Action	Page	Uri
03/04/2007 13:52:12	194.57.88.24	1	login	user	view.php?id=16&course=1
03/04/2007 13:52:12	194.57.88.24	1	view	course	view.php?id=1
03/04/2007 13:52:22	194.57.88.24	2	view	course	view.php?id=2
03/04/2007 13:53:21	194.57.88.24	2	view	resource	view.php?id=3
03/04/2007 13:54:27	194.57.88.24	2	view all	resource	index.php?id=2
03/04/2007 13:54:30	194.57.88.24	2	view	course	view.php?id=2
03/04/2007 13:54:31	194.57.88.24	2	view	resource	view.php?id=1
03/04/2007 13:56:12	194.57.88.24	2	view	course	view.php?id=2
03/04/2007 13:56:18	194.57.88.24	2	view	course	view.php?id=2
03/04/2007 13:56:40	194.57.88.24	2	view	chat	view.php?id=2
03/04/2007 13:56:51	194.57.88.24	2	view	course	view.php?id=2
03/04/2007 13:57:35	194.57.88.24	2	view	course	view.php?id=2
03/04/2007 13:57:39	194.57.88.24	2	view forum	forum	view.php?id=4
03/04/2007 13:57:51	194.57.88.24	2	view discussion	forum	discuss.php?id=2
03/04/2007 13:58:20	194.57.88.24	2	view forum	forum	view.php?id=4
03/04/2007 14:00:50	194.57.88.24	2	add discussion	forum	discuss.php?id=3
03/04/2007 14:00:53	194.57.88.24	2	view forum	forum	view.php?id=4
03/04/2007 14:01:06	194.57.88.24	2	view	course	view.php?id=2
03/04/2007 14:01:11	194.57.88.24	2	view	resource	view.php?id=1
03/04/2007 14:01:16	194.57.88.24	2	view	course	view.php?id=2
03/04/2007 14:01:44	194.57.88.24	2	view	wiki	view.php?id=5&groupid=0&userid=0&page=Texte Wiki
03/04/2007 14:02:08	194.57.88.24	2	view	course	view.php?id=2
03/04/2007 14:02:22	194.57.88.24	2	view	chat	view.php?id=2

Figure 9. Copie d'écran pour la visualisation au format html

Nous proposons pour la collecte côté Client un atelier qui permet de collecter les données selon les cinq classes par défaut décrites dans le modèle de la m-trace

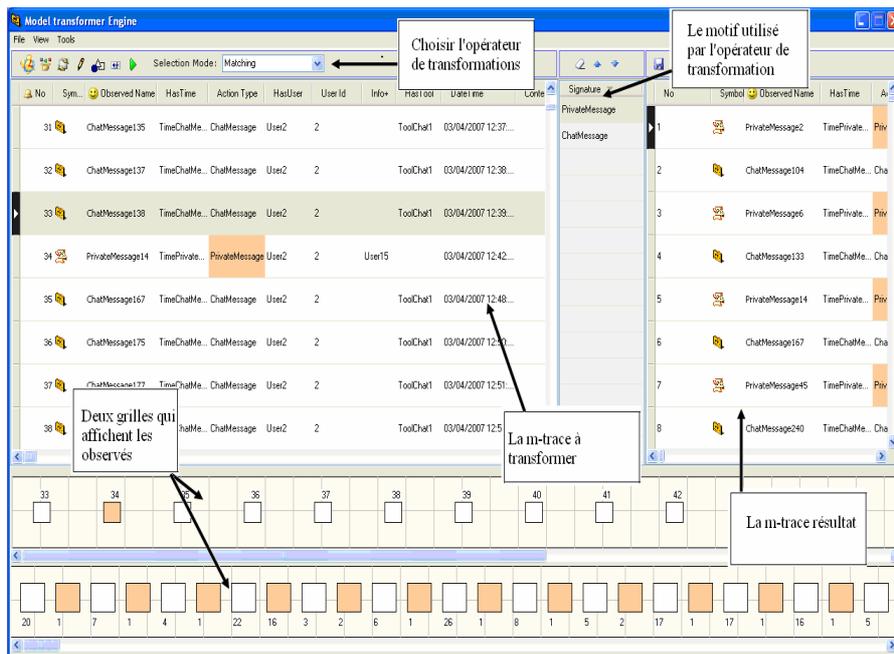
première (Djouad, 2008) et qui sont : Course, User, Time, Tools et Observed. Le module de filtrage (FM) propose plusieurs fenêtres de filtrage : il y a au total cinq boîtes de configurations, une pour chaque classe. La figure 10 présente la fenêtre associée à la classe *Observed*, et permet de lancer les différentes opérations de filtrage et de réécriture sur cette classe. Par exemple : renommer la classe *Observed* par « *Observés* » et ses attributs, comme « *Observed\_Id* par *Identifiant\_Observé* » la classe *Observed* contient plusieurs sous classes comme Entrer dans un Chat, Ecrire dans un Chat... ou chaque sous classe possède une liste d'attributs, sélectionner les différentes sous-classes de cette classe à collecter, sélectionner les différents attributs pour chaque sous-classe à collecter, etc. Nous pouvons par exemple collecter seulement les observés de type : « Envoyer un message privé » et « Envoyer un message dans un chat ». Et nous pouvons par la suite préciser pour l'observé « Envoyer un message privé » de collecter seulement les attributs : identifiant message et le message, et de ne pas collecter le reste.



**Figure 10.** Copie d'écran : sélectionner les différentes classes d'observés à afficher

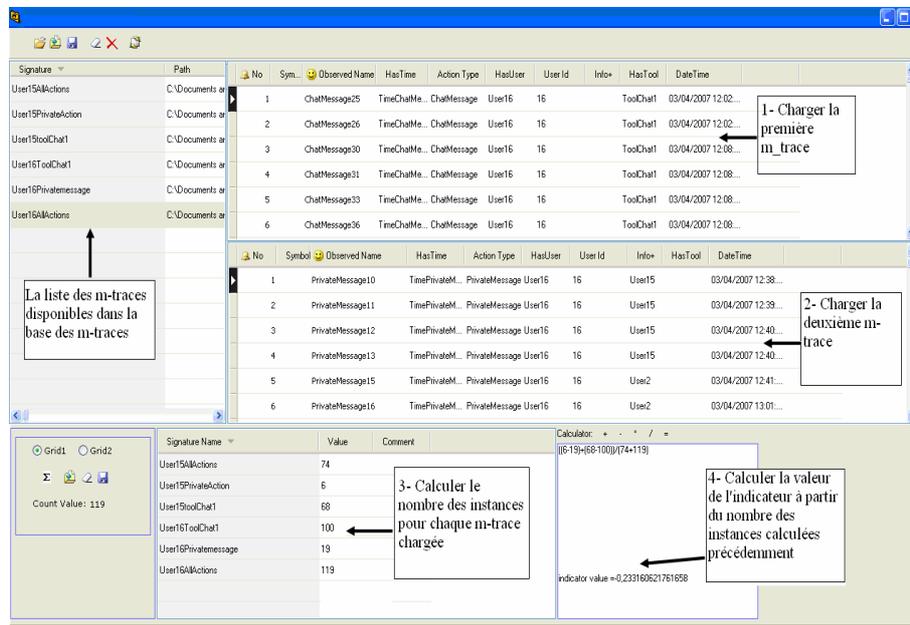
La figure 11 présente un atelier de transformation de m-traces et illustre l'utilisation de l'opérateur *Matching*. Une boîte à outil (en haut à gauche de la fenêtre) permet de sélectionner le type d'opérateur à utiliser (ici *Matching*). Trois zones permettent d'afficher : la m-trace à transformer (à gauche), la signature de tâche (la signature des tâches est un motif décrivant une séquence significative dans

la trace. C'est ce motif qui sera utilisé par le matching), et la m-trace résultat de la transformation (à droite). Les deux grilles en bas affichent les observés : la première donne un extrait de séquence d'observés de la m-trace actuelle, et la seconde tous les observés de la m-trace actuelle (regroupant et dénombrant des observés consécutifs du même type).



**Figure 11.** Le sous-système de transformation, fenêtre principale. Dans cet exemple, l'utilisateur charge la m-trace à transformer, définit un motif, et choisit un opérateur de transformations (ex ; le Matching) qui utilise ce motif. L'outil donne la m-trace résultat de la transformation.

La figure 12 illustre le calcul d'un l'indicateur. L'interface communique avec la base des m-traces et appelle : la fonction *Compter (Count)*, le module qui charge les m-traces, ainsi que le module de calcul. Nous pouvons voir dans la figure en haut à gauche, la liste des différentes m-traces disponibles dans la base des m-traces. Cette liste est générée par le sous-système de transformation. On trouve aussi à droite de la figure, les m-traces qu'on charge pour calculer les statistiques d'observés pour chaque m-trace. Le module de calcul est quant à lui disponible (en bas de la figure 12) pour définir et évaluer la fonction associée à l'indicateur. Nous chargeons alors les m-traces nécessaires pour calculer l'indicateur, et nous utilisons par la suite le module de calcul pour calculer l'équation associée l'indicateur à partir de ces m-traces.



**Figure 12.** Interface graphique pour le sous-système qui calcule la valeur de l'indicateur : un exemple de calcul de proportion entre deux types d'observés

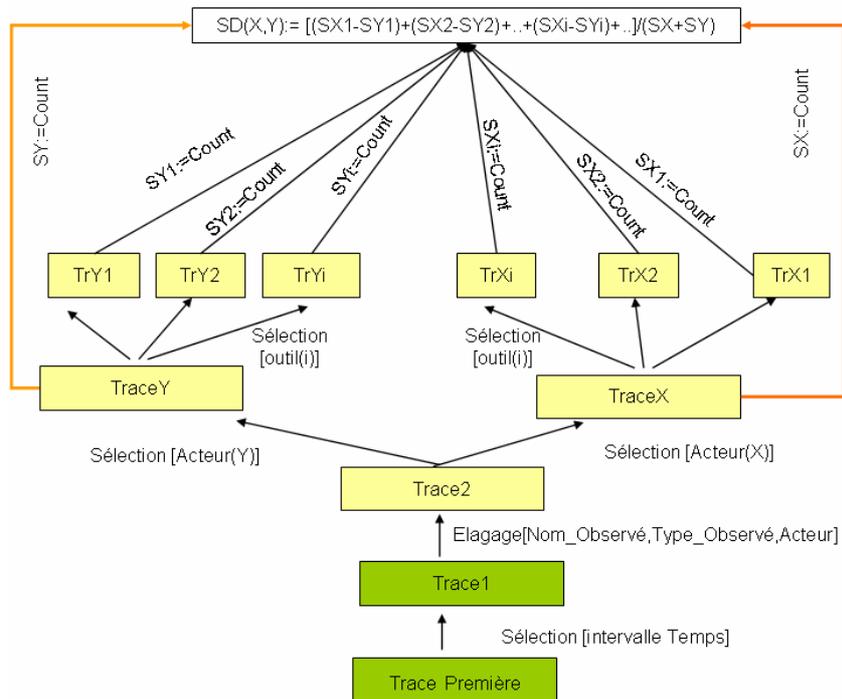
### 3.4. Illustration des transformations dans le SBT-IM

Nous montrons dans cette partie comment utiliser les opérateurs sur les m-traces pour construire les séquences de transformations, et générer par la suite des m-traces évoluées nécessaires aux indicateurs collaboratifs. Pour montrer que notre SBT-IM permet de construire aisément des indicateurs évolués, et pour illustrer l'utilisation des transformations des m-traces, nous donnons ici des exemples concrets sur le calcul de quelques indicateurs issus de la littérature du domaine des EIAH.

#### 3.4.1. Exemple 1 : La division du travail

Cet indicateur défini et implémenté dans la plateforme COTRAS « COLlaborative TRAffic Simulator » (Jermann, 2004) identifie la division du travail adoptée par deux utilisateurs qui agissent sur un ensemble de ressources. Cet indicateur destiné précisément aux chercheurs, permet d'identifier le rôle pris par chaque participant dans le processus de collaboration pour l'apprentissage. On peut distinguer trois types de division du travail : la division basée sur la tâche où chaque acteur agit séparément sur des ressources différentes. Une division basée sur le rôle, où un seul des deux acteurs agit sur toutes les ressources. Une division du travail concourant où les deux acteurs agissent (en même temps) sur l'ensemble des ressources. Le calcul de la division du travail est basé sur la somme des différences de deux

instances d'observés liées à chaque utilisateur sur chaque outil. Il n'y a pas une règle précise dans la littérature EIAH pour calculer cet indicateur, nous proposons la séquence de transformations suivante pour calculer notre indicateur. Nous utilisons les opérateurs de sélection et d'élagage pour construire la séquence (Figure 13).



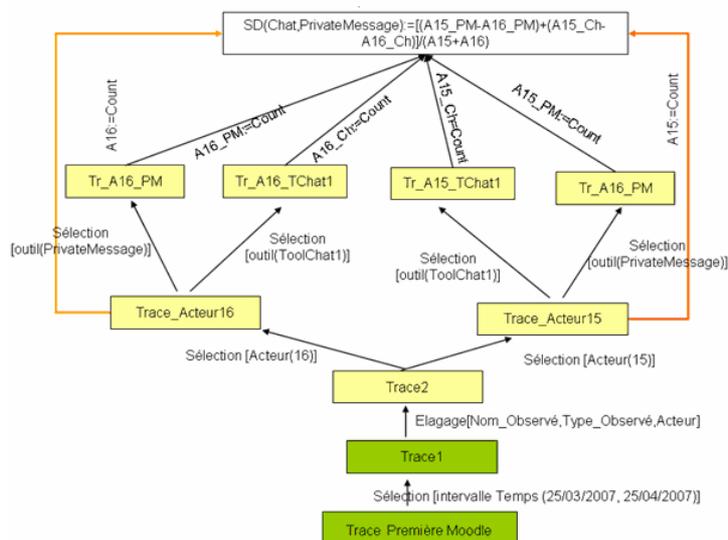
**Figure 13.** Séquence de transformations générique pour calculer l'indicateur « Division du travail » entre deux acteurs ActeurX et ActeurY

Nous pouvons remarquer que la séquence de transformations pour l'ActeurX est la même pour l'ActeurY, ce qui nous permettra de réutiliser un morceau de la même séquence pour construire la séquence globale, ce qui montre la possibilité de définir et de réutiliser les morceaux ou des briques de transformations dans d'autres séquences de transformations.

Le calcul de la division du travail entre les acteurs « 15 » et « 16 » sur les outils « ToolChat1 » et « PrivateMessage », illustré par la figure 14 est une application de la séquence générique (Figure 13).

Action	Param
TimeSelection	From 25/03/2007 12:52:18 To 25/04/2007 12:52:18
Filter Options	Tool: ToolChat1
SaveMain Instance	C:\Documents and Settings\Tarek\Bureau\Calcul SD\
By user	15,
By Param	PrivateMessage,
Save Instance	C:\Documents and Settings\Tarek\Bureau\Calcul SD\
By user	16,
SaveMain Instance	C:\Documents and Settings\Tarek\Bureau\Calcul SD\
By Param	PrivateMessage,

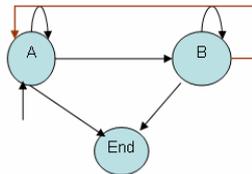
La boîte à outil contient la liste des transformations à faire par ordre pour arriver à l'équation qui calcule l'indicateur : sélection sur une intervalle de temps, ensuite une sélection par acteur... etc. en suivant le même ordre sur le schéma en bas



**Figure 14.** Exemple d'une séquence de transformations pour calculer la division du travail (acteurs 15 et 16) sur les deux outils : ToolChat1 et PrivateMessage

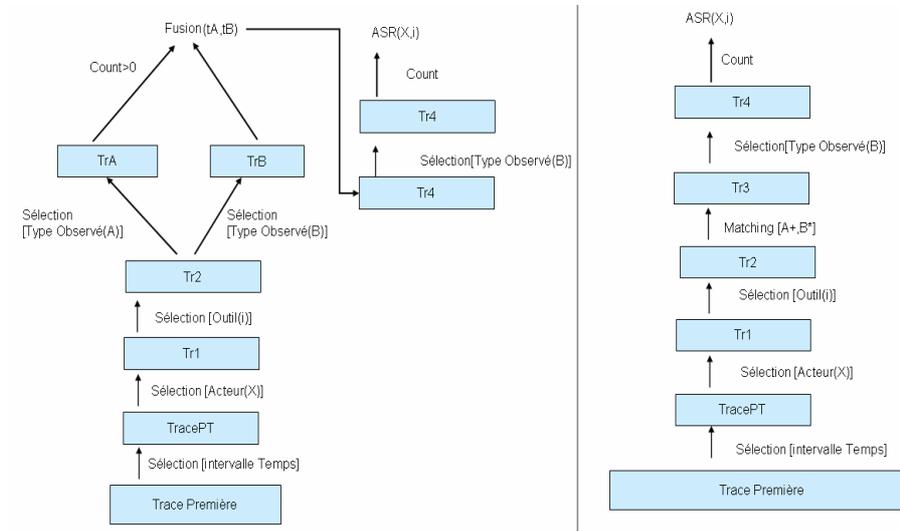
3.4.2. *Exemple2 : Un acteur qui lit ses propres contributions (auto self read) sur une ressource (i)*

Cet indicateur destiné aux enseignants et aux étudiants est cité dans (Dimitracopoulou, 2004). Nous proposons pour le calculer, d'utiliser un motif avec l'opérateur Matching. Le motif est de type  $[A^+, B^*]$  où le A représente les observés de type : « Modifier l'outil » et le B ceux de type « Lire dans le même l'outil », le + signifie que l'acteur a édité au moins une fois sa contribution, alors que le \* signifie que l'acteur a lu zéro ou plusieurs fois cette contribution. Par exemple, la séquence ABB signifie qu'un acteur a édité une fois sa contribution, et l'a lu deux fois : toutes les séquences commencent par un A. Le motif (signature de tâche) est présenté sous forme d'automate sur la figure 15 :



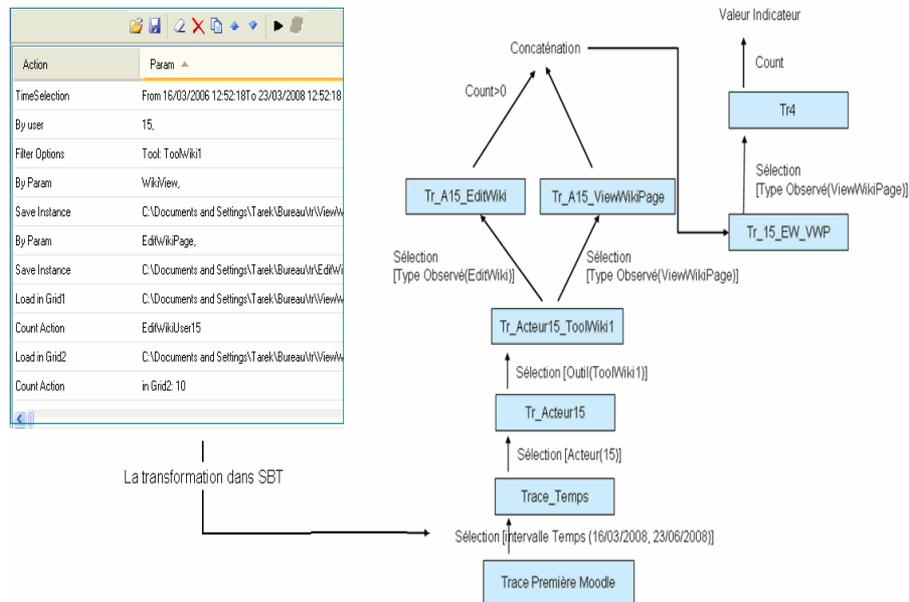
**Figure 15.** Motif pour l'indicateur auto lecture de ses propres contributions

La séquence de transformations associée à cet indicateur et à base de ce motif est présentée figure 16 (à droite). Nous proposons en parallèle le calcul du même indicateur avec une autre transformation basée sur la sélection et la fusion (Figure 16 à gauche). Notons que la fusion impose que le nombre des instances de type « Ecrire » soit supérieur à 0. Cet exemple montre l'existence d'une certaine équivalence fonctionnelle entre les séquences de transformations. Nous pouvons construire deux séquences totalement différentes, mais calculant le même indicateur ce qui donne aussi la richesse de notre méthode.



**Figure 16.** Deux transformations équivalentes pour l'indicateur ASR (Auto-Self-read). À gauche celle utilisant la fusion, à droite le Matching

En illustration (voir figure 17), nous avons choisi d'implémenter la deuxième séquence de transformations pour l'acteur « 15 » avec comme outil le Wiki « ToolWiki1 ». Le « A » dans la séquence de transformations générique devient : « EditWiki », et le « B » devient « ViewWiki ».



**Figure 17.** Séquence de transformations pour l'acteur 15 sur l'outil ToolWiki, dans le but de calculer les lectures de ses propres contributions sur ce wiki

#### 4. Conclusion

Nous avons présenté dans cet article un système à base de trace pour modéliser et transformer les traces d'interactions. Le cadre théorique comporte les notions de trace modélisée (m-trace), de transformation d'une m-trace, le principe de fonctionnement d'un système à base de trace, etc. Nous avons implémenté un système à base de trace spécialisé dans le calcul des indicateurs en EIAH SBT-IM, et nous avons choisi comme étude de cas le calcul des indicateurs de collaboration à partir des m-traces dans la plateforme Moodle. Enfin nous avons illustré avec quelques exemples concrets la facilité de calculer les indicateurs dans notre SBT-IM grâce aux transformations des m-traces.

L'architecture du SBT-IM implémenté que nous proposons est ouverte, et indépendante des plateformes d'apprentissage ainsi que du calcul spécial d'un indicateur. Le système implémenté offre un calcul d'indicateur à un niveau abstrait sans passer par un calcul spécial d'un indicateur et dans une plateforme d'apprentissage ciblée. Cette abstraction dans le calcul des indicateurs est possible grâce aux opérateurs de transformations qui transforment toute m-trace première venant de toute plateforme d'apprentissage. La séquence de transformation générée

est orientée vers le calcul d'un indicateur. Notons qu'elle est réutilisable par parties, comme briques dans la transformation elle-même, ou dans d'autres transformations. On peut bien sûr trouver des équivalences entre séquences de transformations différentes, aboutissant au calcul du même indicateur.

Les principales contributions concernent la proposition d'un cadre théorique pour le traitement des traces d'interactions à des fins de compréhension d'une activité, la proposition d'une architecture logicielle permettant de sous-traiter les traitements de traces à un système à base de traces modélisées et enfin, un environnement concret développé pour les EIAH basés sur Moodle et orienté vers l'élaboration d'indicateurs d'activités d'apprentissage.

La manière d'exploiter l'indicateur dans un EIAH (visualisation, utilisation pour l'adaptation, etc.) sort du cadre du travail de recherche actuel bien qu'il s'agisse d'une question ouverte et importante. Notre travail se focalise sur l'ingénierie des indicateurs à partir de traces modélisées.

Par la suite, nous comptons tester le SBT-IM avec d'autres plateformes d'apprentissage afin d'en démontrer l'indépendance. Pour élargir et amplifier l'expressivité et la généricité de notre mécanisme de transformation des m-traces, nous souhaitons également ajouter d'autres opérateurs de transformations plus génériques.

## 5. Bibliographie

- Baker M J., « Forms of cooperation in dyadic problem solving », *Intelligence Artificielle*, vol. 16, 2002, p. 587-620.
- Baker M J., Lund K., « Promoting reflective interactions in a computer-supported collaborative learning environment », *Journal of Computer Assisted Learning*, vol. 13, 1997, p. 175-193.
- Barros B., Verdejo M F., « Analyzing student interaction processes in order to improve collaboration. The DEGREE approach », *International Journal of Artificial Intelligence in Education*, vol. 11, 2000, p. 221-241.
- Beldame M., « Inscribing the user's experience to enact development », *In Proceeding of the third international conference on enactive interfaces*, Montpellier, France, 2006, p. 193-194.
- Betbeder M L., Reffay C., Chanier T., « Environnement audiographique synchrone : recueil et transcription pour l'analyse des interactions multimodales », *Dans les actes de JOCAIR, premières journées communication et apprentissage instrumentés en réseau*, Amiens, France, 2006, p. 406-420.
- Bouhineau D., Nicaud J F., Pavard X., Sander E., « Un micromonde pour aider les élèves à apprendre l'algèbre », *Sciences et Techniques Educatives, Numéro spécial : Environnements Interactifs d'Apprentissage avec Ordinateur*, vol. 8, 2001, p. 33-47.

- Bratitsis T., Dimitracopoulou A., « Indicators for measuring quality in asynchronous discussion forums », *Communication in the International conference on Cognition and Exploratory Learning in Digital Era*, 2006, Barcelona, Spain.
- Charlier B., Deschryver N., Peraya D., « Apprendre en présence et à distance : Une définition des dispositifs hybrides », *Revue Distances et Savoirs*, 2007, vol. 4, n° 4.
- Cram D., Visualisation de Traces : Application aux Traces Réflexives d'eLycée. Rapport de Master de recherche, Laboratoire Liris, Université Claude Bernard Lyon1, 2007.
- Corbel A., Girardot J J., Jaillon P., « DREW: A Dialogical Reasoning Web Tool », *In proceeding of the International Conference on Information and Communication technologies in Education*, 13-16 Novembre 2002, Badajoz, Spain.
- Després C., Coffinet T., « Reflet, un miroir sur la formation », *Communication dans la conférence internationale sur les Technologies de l'Information et de la Connaissance dans l'Enseignement Supérieur et l'Industrie*, 20-22 octobre 2004, Compiègne, France, p. 19-24.
- Dimitracopoulou A., State of the art on Interaction and Collaboration Analysis (D26.1.1). EU Sixth Framework programme priority 2, Information society technology, Network of Excellence Kaleidoscope, (contract NoE IST-507838), project ICALTS: Interaction and Collaboration Analysis, 2004.
- Djouad T., « Analyser l'activité d'apprentissage collaboratif : Une approche par transformations spécialisées de traces d'interactions », *dans les actes de la 2ième rencontre des jeunes chercheurs RJC-EIAH08*, 15-16 mai 2008, Lille, France, p. 93-98.
- Dyke G., Girardot J J., Lund K., Corbel A. « Analysing face to face computer-mediated interactions. EARLI (European Association for Research, Learning and Instruction) », *12th Biennial International Conference, University of Szeged, Eötvös Lorand University, Hungarian Academy of Sciences*, August 28-September 1, 2007, Budapest, Hungary.
- France L., Heraud J M., Marty J C., Carron T., « Visualisation et régulation de l'activité des apprenants dans un EIAH tracé », *dans les Actes de la conférence EIAH 2007*, 27-29 Juin 2007, Lausanne, Suisse.
- Gassner K. Diskussionen als Szenario zur Ko-Konstruktion von Wissen. [Discussion as a scenario of knowledge co-construction]. Dissertation. Faculty of Engineering Sciences, University Duisburg-Essen, 2003.
- Greffier F., Reffay C., « Les échos du forum de discussion en FAD. Dans les actes de JOCAIR, premières journées communication et apprentissage instrumentés en réseau, 2006, p. 130-144. Amiens, France.
- Guéraud V., Adam J M., Pernin J P., Calvary G., David J P., « L'exploitation d'Objets Pédagogiques Interactifs à distance : le projet FORMID », *Revue STICEF*, vol. 11, 2004, p. 109-164.
- Jaillet A., « Peut-on repérer les effets de l'apprentissage collaboratif à distance », *Distances et savoirs*, vol. 3, n°1, 2005, p. 49-66.
- Jermann P R., Computer Support for Interaction Regulation in Collaborative Problem-Solving. Thèse de doctorat, Genève, 2004.

- Laperrousaz C., Le suivi individuel d'apprenants engagés dans une activité collective à distance, TACSI : un environnement informatique support aux activités du tuteur, thèse de doctorat, Université de Maine, 2006.
- Luengo V., Vadcard L., Dubois M., Mufti-Alchawafa D., « TELEOS : de l'analyse de l'activité professionnelle à la formalisation des connaissances pour un environnement d'apprentissage », dans *les Actes de la conférence Ingénierie de Connaissances, IC 2006*, Juin 2006, Nantes, France.
- Lund K., Mille A., Traces, traces d'interaction, traces d'apprentissages : définition, modèles informatiques, structurations, traitements et usage, Chapitre 1, Analyse de traces et personnalisation des Environnements Informatiques d'Apprentissage Humain, Hermes, IC2 - Série Informatique et Systèmes d'Information, à paraître
- Martínez A., Dimitriadis Y., Gómez E., Rubia B., de la Fuente P., « Combining qualitative and social network analysis for the study of classroom social interactions », *Computers and Education*, vol. 41, n°4, 2003, p. 353-368.
- Mazza R., Milani C., « Exploring Usage Analysis in Learning Systems: Gaining Insights From Visualisations », *Communication in the Workshop on Usage analysis in learning systems, the twelfth International Conference on Artificial Intelligence in Education. Amsterdam*, 2005, the Netherlands, p. 65-72.
- Mazza R., Dimitrova V., « Visualising Student Tracking Data to Support Instructors in Web-Based Distance Education », *In proceeding of the thirteenth International World Wide Web Conference-Educational Track*, 2004, New York, USA, p. 154-161.
- Meunier J P., Peraya D., « Introduction aux théories de la communication. Analyse sémiopragmatique de la communication médiatique », *2ème édition revue et augmentée*, 2004, De Boeck, Bruxelles.
- May M., George S., Prévôt P., « Keeping Tracks of Users' Activities on Discussion Forum: Reflection and Proposal », *National Conference on Learning Through Online Exchanges: Tools, Tasks, Interactions, Multimodality, Corpora (EPAL 2007)*, 7-9 Juin 2007, Grenoble, France.
- Peraya D., (2003). « De la correspondance au campus virtuel : formation à distance et dispositifs médiatiques », *In Charlier, B. & Peraya, D. (Ed.). Technologie et innovation en pédagogie. Dispositifs innovants de formation pour l'enseignement supérieur*, 2007, De Boeck, Bruxelles, p. 79-92.
- Reffay C., Lancieri L., « Quand l'analyse quantitative fait parler les forums de discussion », *STICEF, numéro spécial forum de discussion en éducation*, 2006, p. 255-288.
- Renie D., « Apport d'une trace informatique dans l'analyse du processus d'apprentissage d'une langue seconde ou étrangère », *In Duquette L., Laurier M. (dirs) Apprendre une langue dans un environnement multimédia*. Outremont, Canada. 2000, p. 281-301.
- Rezeau J., « L'apport du concordancier à l'analyse et à la remédiation des erreurs des apprenants dans les forums de discussion en ligne », *Revue Aslic*, vol. 10, n° 2, 2007, p. 27-43.
- Santos O C., Rodríguez A., Gaudioso E., Boticario J G., « Helping the tutor to manage a collaborative task in a web-based learning environment », *Communication in the*

*Workshop Towards Intelligent Learning Management Systems*, 2003, Sydney, Australia, p. 72-81.

Settouti L S., Prié Y., Mille A., Martu J C., « Systèmes à base de traces pour l'apprentissage humain », *Communication in the international TICE, Technologies de l'Information et de la Communication dans l'Enseignement Supérieur et l'Entreprise*, 2006, Toulouse, France.

Soury-Lavergne S., « Connaissances et mise en œuvre d'un micromonde dans les interactions de préceptorat distant, le cas de la géométrie dynamique », *Sciences et techniques éducatives*, vol.8, 2001, p. 321-345.

Tedesco P A., « MArCo: Building an Artificial Conflict mediator to Support Group Planning Interactions », *International Journal of Artificial Intelligence in Education*, vol. 13, 2003, p. 117-155.