
Formalisation d'annotations produites par des apprenants

Etape préalable à la réalisation d'un annotateur

Dominique Mille

*CLIPS-IMAG
BP 53
38041 Grenoble cedex 9
Dominique.Mille@imag.fr*

RÉSUMÉ. L'objet de cet article est la description d'une formalisation computable des annotations produites par des apprenants, représentée par une ontologie. Cette formalisation explicite la sémantique des annotations grâce à des attributs auxquels le lecteur devrait donner une valeur. Elle contient également les valeurs possibles de ces attributs. L'intérêt d'une telle formalisation est de couvrir toutes les annotations produites par des apprenants, et d'explicitier tout ce qui constitue leur sens, afin de favoriser une réutilisation sans perte d'information. Nous souhaitons, à partir de notre formalisation, concevoir un annotateur (logiciel d'annotation), ce qui nous amène à alléger la tâche du lecteur en réduisant la valuation manuelle. Pour cela, nous identifions les dépendances entre les attributs, ce qui permettrait d'automatiser l'affectation d'une valeur aux divers attributs.

MOTS-CLÉS : annotation sémantique, ontologie

1. Introduction

Nous nous intéressons ici aux *annotations*, que nous définissons comme les marques inscrites par un individu sur un document dans un but spécifique. Elles ne peuvent être dissociées du document, sur lequel elles sont visuellement perceptibles. Ce travail s'inscrit dans un projet qui considère que les dispositifs actuels d'annotation sont mal adaptés, et qui souhaite proposer des formalismes et des outils efficaces pour l'annotation électronique de ressources pédagogiques par des apprenants. Nous présentons une formalisation portant sur la structure de l'annotation, les valeurs de ses attributs, et leurs liens. La finalité est la production de formalismes et d'un *annoteur* (outil d'annotation) efficace.

2. Intérêt et utilité des annotations électroniques

L'annotation d'un document papier par un apprenant doit être envisagée comme objet et comme activité. En tant qu'objet, c'est un élément de la mémoire externe de formation des apprenants, qui permet de soulager leur mémoire humaine. En tant qu'activité, c'est l'une de la lecture active. Elle sert à s'approprier et à personnaliser des documents. Elle aide à comprendre, et à retrouver aussi bien les informations contenues dans un document, que la démarche de compréhension déjà effectuée.

Bien que les documents pédagogiques soient souvent disponibles au format électronique, ils sont généralement imprimés pour lire (manque de mobilité et de confort des machines actuelles) et pour annoter (difficulté d'usage et limites des annoteurs) [ROBERT 01]. Pour une institution disposant d'un grand nombre de documents, conserver et mettre ces documents à disposition des apprenants au format électronique offre plusieurs avantages : ils sont moins encombrants et plusieurs apprenants peuvent les manipuler simultanément. De plus, grâce au format électronique, chaque apprenant peut personnaliser ses documents. Un cycle de vie tout électronique offre également des avantages dans un contexte d'usage individuel. Une étude [MARSHALL 00] montre ainsi une volonté des lecteurs de conserver les annotations sous forme électronique.

Au-delà de leur forme, les annotations sont porteuses d'un sens pour le lecteur : par exemple, les raisons pour lesquelles il a annoté, ou encore le contexte dans lequel il a annoté. Le lecteur connaît la sémantique de ses annotations. Il l'explique parfois dans des légendes, mais elle reste souvent implicite. Elle est contenue dans l'aspect visuel des annotations, mais le codage n'est conservé que dans la mémoire humaine du lecteur [WOLFE & NEUWIRTH 01]. En raison de cette sémantique implicite, une partie de l'information est perdue lors de la réutilisation, que ce soit par soi-même (oubli des légendes) ou par autrui (utilisation de légendes différentes).

L'annotation s'avérant utile aux apprenants, il est souhaitable de leur proposer un annoteur. Pour être aussi efficace que celle sur papier, l'annotation électronique nécessite des dispositifs adaptés, offrant assez d'avantages pour dispenser les lecteurs d'imprimer les documents. Nous estimons qu'il est nécessaire de conserver la sémantique des annotations, et donc de l'expliquer tant au niveau d'une structure formelle qu'au niveau d'un annoteur.

3. Formalisation de l'annotation

Notre structure d'annotation est formée d'attributs catégorisés selon leur fonction et selon ce qu'ils représentent. Les **attributs fixes** représentent l'évènement d'annotation. Ils contiennent les attributs **physiques** (ce qui est visible), **circonstanciels** (les circonstances spatiales et temporelles), et **sémantiques** (le sens). Les **attributs évolutifs** représentent l'état de l'annotation quand elle est donnée à autrui, et leurs valeurs sont modifiées à chaque nouveau don. Ils contiennent les attributs **de transmission** (liés à l'évènement de don et de réception) et les attributs **personnels** (l'apport subjectif du nouveau possesseur).

Attributs fixes	physiques	<ul style="list-style-type: none"> • Ancre physique (lieu où apparaît l'annotation) • Forme visuelle (aspect visuel de l'annotation) • Ancre syntaxique (contenu annoté)
	circonstanciels	<ul style="list-style-type: none"> • Auteur (de l'annotation) • Date (de création de l'annotation) • Lieu (lieu où se trouvait l'auteur) • Situation d'enseignement • Situation de réutilisation (réutilisation envisagée)
	sémantiques	<ul style="list-style-type: none"> • Contenu libre • Destinataire(s) (pour qui est crée l'annotation) • Objectifs d'annotation (l'objectif d'une annotation) • Objectif de lecture (pourquoi lire le document) • Importance initiale (donnée par l'auteur) • Confiance initiale (accordée par l'auteur)
Attributs évolutifs	de transmission	<ul style="list-style-type: none"> • Donneur (qui transmet l'annotation) • Possesseur (qui reçoit l'annotation) • Effet (ce que produit l'annotation, ex : sonnerie)
	personnels	<ul style="list-style-type: none"> • Droits (à qui le possesseur peu donner l'annotation) • Importance (donnée par le nouveau possesseur) • Confiance (accordée par le nouveau possesseur)

Tableau 1. *Structure d'annotation*

La **situation d'enseignement** représente le domaine du document, le niveau et le type d'étude du lecteur. Le **contenu libre** contient un texte associé à l'annotation. L'**objectif d'annotation** est composé de l'**objectif de lecture active** (but de l'annotation en tant qu'activité de lecture active) et de l'**objectif de mémorisation** (le but de l'annotation en tant que mémoire externe). Ce dernier est le but de l'objectif de lecture active (« donner un titre » pour « mémoriser des repères »).

3.1. Valeurs des formes visuelles

La **mise en évidence** consiste à modifier l'apparence d'une information afin qu'elle soit plus facilement visible. Les formes visuelles ne contiennent aucun objectif, bien qu'il soit intuitif d'en associer un dans certain cas.

1. Ajout 1.1. De texte 1.2. D'un dessin explicatif 1.3. D'une marque non explicative 1.3.1. Unaire 1.3.2. Binaire	2. Mise en évidence 2.1. En surlignant 2.2. En soulignant 2.3. En entourant 2.4. En barrant 2.5. En changeant la couleur
--	---

Tableau 2. *Formes visuelles*

3.2. Valeurs des objectifs de lecture active

1. Restructurer 1.1. Donner un titre 1.2. Hiérarchiser 1.3. Synthétiser 1.4. Reformuler 2. Catégoriser 2.1. Par valeur d'importance 2.2. Par type prédéfini 2.2.1. Théorème 2.2.2. Définition 2.2.3. <i>Autres...</i> 2.3. Par type personnel 2.4. Par similarité de contenu 3. Créer une relation 3.1. Présentation/développement 3.2. Présentation/explication 3.3. Définition/explication 3.4. <i>Autres...</i>	4. Ajouter une remarque personnelle 4.1. Critiquer 4.1.1. Positivement 4.1.2. Négativement 4.2. Exprimer une idée connexe 4.3. Développer 4.3.1. Expliquer textuellement/graphiquement 4.3.2. Ajouter un exemple 4.3.3. Résoudre un problème 4.4. Faire référence à un autre document 5. Planifier une action 6. Réviser 6.1. Supprimer 6.2. Ajouter 6.3. Reformuler 6.4. Déplacer 7. Soutenir l'attention
---	---

Tableau 3. *Objectifs de lecture active*

(1) Le lecteur **restructure**, selon une présentation personnelle et ciblée, sans interpréter le contenu. (2) Il **catégorise** des passages, selon leur type ou leur contenu (grâce à des symboles, couleurs,...). (3) Il **crée une relation** entre deux passages selon leur type ou leur contenu, pour réorganiser le contenu. (4) Il **ajoute des remarques personnelles** qui interprètent le contenu. (5) Il **planifie** des tâches (ex : relire ce paragraphe). (6) Il **révise**, en vue d'une modification du contenu. (7) Il inscrit des marques pour **soutenir son attention** ou encore pour indiquer où reprendre une lecture momentanément interrompue.

3.3. Valeurs des objectifs de mémorisation

A partir des études de la littérature [SCHILIT & al. 98] [MARSHALL 00] et de celle que nous avons menée [MILLE 01], nous déduisons que le but général d'une annotation est de *mémoriser des repères dans le document*, de *mémoriser sa propre compréhension*, ou de *mémoriser des corrections*. Nous définissons ainsi trois valeurs pour l'objectif de mémorisation : **Comprendre**, **Se repérer** et **Corriger**.

3.4. *Valuation des attributs*

Notre formalisation doit servir de base à la réalisation d'un annotateur dont l'utilisation doit perturber le moins possible l'activité de lecture. Or la valuation est une tâche lourde, ce qui nous amène à réduire le nombre d'attributs à renseigner.

Nous diminuons la tâche de l'annotateur en déterminant les attributs que la machine peut valuer, c'est le cas pour *l'ancre physique*, *l'ancre syntaxique*, *la date*, *l'auteur* (si l'annotateur reconnaît l'utilisateur), *la situation d'enseignement* (selon ce qui est connu du document). Le système peut aussi valuer *le donneur* et *le possesseur* lors d'un don, et modifier *l'effet* lorsqu'il a eu lieu.

On peut aussi exploiter les dépendances entre attributs et les contraintes sur les valeurs. Ainsi, **ancre physique** et **forme visuelle** sont liées (un surlignage se situe forcément sur une partie non vierge), de même **ancre syntaxique** et **forme visuelle** (un soulignage n'annote pas tout un document); **ancre physique** et **objectif de lecture active** (une annotation de reformulation se situe forcément sur une partie vierge); **ancre syntaxique** et **objectif de lecture active** (une annotation catégorisant un passage ne peut pas annoter le document entier). Les *ajouts* et les annotations de *planification* ne sont placés que sur une partie vierge. Les *marques binaires* ont deux ancres syntaxiques et n'annotent donc pas le document entier. Les *misés en évidence* sont placées sur du contenu et ne peuvent donc pas avoir pour ancre physique une partie vierge; elles annotent une partie du document mais pas sa totalité. Les annotations de *restructuration* et *d'ajout de remarques* personnelles sont placées dans une partie vierge. Elles annotent tout ou partie du document, sauf pour celles de hiérarchisation qui sont relatives à d'autres passages. Les annotations de *catégorisation* ou de *mise en relation* n'annotent pas le document entier.

Il est fréquent que le lecteur construise un système de couleurs et/ou de formes correspondant à l'objectif de lecture active [MARSHALL & al. 99]. Des patrons personnalisables pourraient lier automatiquement la forme d'une annotation aux valeurs de ses attributs sémantiques. L'activité d'annotation électronique serait alors très proche de celle sur papier, tout en conservant la totalité de la sémantique. L'intérêt de ces patrons, outre le fait d'alléger la tâche d'annotation, serait de transmettre la sémantique et d'adapter la forme visuelle des annotations selon les patrons du nouveau possesseur. La construction de patrons nécessite d'étudier les pratiques des lecteurs, pour en déterminer les différents types de base. Par la suite, les lecteurs affinaient leurs patrons, selon leur système de représentation.

4. Conclusion

Notre formalisation permet d'explicitier la sémantique des annotations, grâce à une ontologie contenant à une structure d'attributs, leur valeurs, leurs dépendances et restrictions. Nous pouvons à présent aborder la réalisation d'un annotateur efficace, répondant aux besoins des apprenants. Au niveau interne, nous souhaitons utiliser les normes et standards en vigueur : représentation des ontologies en OWL et des annotations en RDF. Ainsi, des utilisateurs utilisant différents annotateurs pourront s'échanger leurs annotations. De plus, les ontologies construites pourront être partagées, sans nécessiter une traduction. Au niveau externe, nous souhaitons que

notre annotateur soit *un outil de transition*. Il doit permettre au lecteur de conserver ses habitudes papier tout en offrant de nouvelles fonctionnalités [SCHILIT & al. 98], afin de satisfaire les lecteurs qui préfèrent transposer leurs habitudes papier et non pas s'adapter à de nouvelles procédures d'annotation [MARSHALL & al. 99]. Pour cela, nous envisageons d'utiliser une métaphore du papier. Nous souhaitons une interaction Outil/Objet, puisque que, sur papier, le lecteur saisit son stylo avant de marquer le document. Les outils (stylos) seront regroupés dans une trousse, en laissant sur l'espace de lecture les derniers utilisés pour qu'ils soient plus rapides d'accès. Enfin, l'annotation sera accessible à l'utilisateur par une forme visuelle associée à une fenêtre d'attributs contenant la sémantique. Nous espérons alléger considérablement les manipulations grâce aux automatisations proposées.

5. Bibliographie

- [MARSHALL et al. 99] Marshall, C. C., Price, M. N., Golovchinsky, G., Schilit, B. N., « Introducing a digital library reading appliance into a reading group », *Proceedings of the fourth ACM conference on Digital libraries DL'99*, Berkeley, California, USA, 11-14 août 1999, ACM Press, New York, USA, p. 77-84.
- [MARSHALL 00] Marshall, C. C. « The Future of Annotation in a Digital (Paper) World », *Successes and Failures of Digital Libraries*, Harum & Twidale eds. Urbana-Champaign, University of Illinois, 2000, p. 97-117.
- [MILLE 01] Mille, D., Vers une mémoire individuelle de formation : annotation didactique de ressources pédagogiques. DEA, Université J. Fourier, Grenoble, 2001, 78p.
- [ROBERT 01] Robert, L. Annotation et visualisation interactives de documents hypermédias, Thèse de doctorat, ENST de Paris, 2001, 224 p.
- [SCHILIT et al. 98] Schilit, B. N., Golovchinsky, G., Price, M. N., « Beyond Paper: Supporting Active Reading with Free Form Digital Ink Annotations », *Proceedings of the ACM Conference on Human Factors in Computing Systems, CHI'98*, Los Angeles, USA, 18-23 avril, ACM Press, New York, USA, p. 249-256.
- [WOLFE & NEUWIRTH 01] Wolfe, J. L., Neuwirth, C. M., « From the margins to the center: the future of annotation », *Journal of Business and Technical Communication*, vol. 15, 2001, n°3, p. 333-371.