



HAL
open science

Ontologies et relations spatiales dans la lecture d'une bande dessinée

Norbert Tsopze, Clément Guérin, Karell Bertet, Arnaud Revel

► **To cite this version:**

Norbert Tsopze, Clément Guérin, Karell Bertet, Arnaud Revel. Ontologies et relations spatiales dans la lecture d'une bande dessinée. Ingénierie des Connaissances, Jun 2012, Paris, France. pp.175-182. hal-00713720

HAL Id: hal-00713720

<https://hal.science/hal-00713720>

Submitted on 6 Dec 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Ontologies et relations spatiales dans la lecture d'une bande dessinée

N. Tsopze, C. Guérin, K. Bertet, A. Revel

LABORATOIRE L3I, Université de La Rochelle, France
ntsopze, cguerin, kbertet, arevel@univ-lr.fr

Résumé : Nous présentons dans cet article une ontologie permettant d'associer de la sémantique à une bande dessinée lors de sa numérisation. Nous nous intéressons particulièrement ici au calcul de l'ordre de lecture des cases, basé sur leur relations spatiales.

Mots-clés : Ontologies, Images, Bandes dessinées, Relations spatiales.

1 Introduction

La bande dessinée (BD) représente, au même titre que le cinéma, la musique ou la littérature, un patrimoine culturel important pour beaucoup de pays. Elle est aujourd'hui numérisée en masse pour des besoins d'archivage, de conservation et de valorisation. Cette dématérialisation à échelle industrielle permet actuellement l'indexation des pages mais pas de leur contenu. Si cette "limite à la page" peut être franchie, alors de nouvelles utilisations du 9ème art sont envisageables, telles qu'une lecture adaptée aux appareils mobiles ou encore la recherche d'informations dans un grand nombre d'albums. Lorsque l'on parle de bande dessinée sur support informatique, il faut bien faire la distinction entre la BD *numérique*, créée et pensée pour une lecture sur dispositif nomade, et la BD *numérisée* qui reste, quoi qu'il arrive, une image fixe détachée de toute sémantique. Nous nous intéressons dans cet article au passage du numérisé vers le numérique, basé sur l'extraction de la connaissance scénaristique portée par un document physique, proposant ainsi une approche de résolution du problème du

fossé sémantique particulièrement important dans le domaine de l'image.

Des travaux, basés sur le traitement d'images, sont menés depuis quelques années afin d'extraire le contenu des BD numérisées, plus particulièrement les cases et les bulles de dialogues. Peu d'études ont en revanche été consacrées à l'extraction de la sémantique associée à ces éléments. Une page de bande dessinée regorge pourtant d'informations scénaristiques, portées bien sûr par les dessins à l'intérieur des cases, mais également par la structure géométrique et l'agencement topologique de ces dernières. Cette connaissance peut être formalisée sous la forme d'une ontologie afin, par exemple, de servir de support de navigation dans une BD.

En effet, les ontologies (Uschold & King, 1996) utilisées, entre autres, dans le cadre du web sémantique (Berners-Lee *et al.*, 2001) permettent d'organiser et de décrire les connaissances d'un domaine dans le but de les partager, les diffuser et les actualiser. Dans le secteur de l'image, les ontologies semblent être une solution prometteuse pour définir une sémantique explicite : elles permettent de surmonter le problème de la représentation implicite et cachée d'une connaissance visuelle et peuvent être utilisées comme une référence commune pour échanger des contenus sémantiques. Le travail présenté s'intéresse à la construction d'une telle ontologie et à son utilisation pour rechercher et visualiser certains éléments dans une BD.

La deuxième section de cet article s'intéressera à la structure et au vocabulaire de la bande dessinée. La troisième partie décrira le modèle que nous proposons et présentera une première application sur le calcul des relations d'ordre entre les cases. La dernière partie conclura et présentera les perspectives de nos travaux.

2 Bande dessinée : structure, vocabulaire et traitement

En BD, une histoire est mise en scène à travers une succession de cases et peut être racontée en un ou plusieurs épisodes. Les épisodes de l'histoire sont rassemblés en albums, ou tomes. La succession des planches dans l'album définit leur ordre de lecture. Une planche peut s'étendre sur une ou deux pages d'un album. Formellement, comme défini dans (Yamada *et al.*, 2004), si nous notons *Comic* un album, alors $Comic = \{p_1, p_2, \dots, p_n\}$ avec n le nombre de pages, p_i une page numéro i de l'album et $p_1 < p_2 < \dots < p_n$. Cette dernière relation définit l'ordre de lecture des pages.

Chaque case décrit une scène à travers des personnages, des bulles de dialogue, un décor, du texte et d'autres objets présents dans la case. Ces

cases sont généralement de forme rectangulaire entourées ou non, et séparées ou non les unes des autres par une bande appelée *espace inter-iconique*. Dans certaines bandes dessinées, la géométrie des cases est plus complexe et des chevauchements et imbrications entre elles sont possibles. Formellement, on note une page $p_i = \{f_1, f_2, \dots, f_m\}$ où m est le nombre de cases dans la page, f_k est la case k de la page. L'ordre de lecture $f_1 < f_2 < \dots < f_m$ dépend à la fois des relations spatiales entre les cases, mais également de leur interprétation scénaristique. En effet, l'auteur peut choisir d'introduire délibérément une ambiguïté dans la séquence de lecture afin de renforcer sa mise en scène. Incruster une case dans une plus grande permet par exemple d'introduire un effet de simultanéité dans une scène (Lainé & Delzant, 2007). Plusieurs ordres coexistent alors et c'est au lecteur de faire un choix. Par ailleurs, l'ordre de lecture peut varier selon le type de BD, les mangas japonais se lisant par exemple de droite à gauche. D'une manière plus générale, les auteurs sont libres quant à la séquence de lecture qu'ils souhaitent adopter et, si l'ordre consensuel induit une lecture de gauche à droite et de haut en bas, il existe dans la littérature de nombreux contre-exemples. C'est pourquoi nous avons choisi de formaliser la disposition des cases par un ordre partiel (ie. graphe sans cycles), l'ordre de lecture correspondant alors à un tri topologique de cet ordre partiel. La même approche est utilisée concernant les phylactères.

Certains travaux récents (Rigaud *et al.*, 2012) proposent des algorithmes de traitement de la bande dessinée. Le but de ces travaux est de segmenter une page de bande dessinée pour en extraire le texte, les cases ainsi que leur contenu. Ces travaux utilisent des techniques basées sur un découpage par des lignes tracées selon la *transformée de Hough*, l'algorithme *X-Y récursif* ou encore à partir du *gradient*. Plusieurs paramètres sont nécessaires au bon fonctionnement de ces algorithmes : la couleur du contour des cases, la couleur de l'espace séparant des cases, la couleur du fond de la bulle, etc. Ces éléments sont décrits par les caractéristiques de bas niveau telles que les coordonnées et les valeurs des pixels. Notre travail s'intéresse à l'annotation et à l'interprétation des résultats de ces travaux.

3 Ontologie de la bande dessinée

3.1 Etat de l'art ontologie image et bande dessinée

Plusieurs travaux proposent une étude des techniques de résolution du problème du fossé sémantique : (Zhang *et al.*, 2012) se sont penchés sur

l'utilisation des méthodes d'apprentissage automatique (SVM, arbres de décision, réseaux de neurones, réseaux bayésiens) permettant de classer les images (ou les régions) suivant des concepts préalablement appris, ainsi que sur la définition de métadonnées dépendantes (ou non) du contenu de l'image ou bien liées au contenu sémantique de cette dernière. (Bannour & Hudelot, 2012) ont travaillé sur la combinaison des informations visuelles, conceptuelles et contextuelles dans la construction de la hiérarchie sémantique, tandis que (Neumann & Möller, 2008) et (Tousch *et al.*, 2012) ont respectivement étudié l'interprétation des scènes par une logique de description et l'introduction de l'utilisateur dans l'annotation.

Parmi les rares études du domaine appliquées à la bande dessinée, on peut citer (Yamada *et al.*, 2004) et (Han *et al.*, 2007). Les premiers proposent une ontologie dont le but est de détecter les parties de l'image les moins importantes, afin de les éliminer lors de l'affichage sur un écran de téléphone. L'ontologie de (Han *et al.*, 2007) permet quant à elle de modéliser précisément le processus d'extraction du contenu, mais écarte totalement l'aspect scénaristique de la BD.

3.2 Modèle général

Notre ontologie a pour ambition de fournir un cadre adapté à :

- La description d'une bande dessinée numérisée au niveau de l'image. La connaissance portera alors aussi bien sur les caractéristiques visuelles des régions de l'image, que sur les caractéristiques de l'image en elle-même (résolution, dimension, etc.). On nommera ce niveau de description "bas niveau".
- La description d'une bande dessinée numérisée au niveau "scénaristique", ou "haut niveau". L'accent étant alors mis sur le sens propre du contenu de la bande dessinée. Sera représentée ici toute connaissance déductible des informations bas niveau.
- L'évaluation d'algorithmes d'extraction automatique de contenu.

Afin de se conformer à la philosophie du modèle d'extraction du contenu, notre ontologie est centrée sur la notion de *Region Of Interest*, ou *ROI*. Une *ROI* représente une zone de l'image identifiée comme porteuse d'information, elle peut se spécialiser en *Panel* ou en *TextRegion*. Chaque *ROI* est liée à la page d'où elle est extraite, elle-même associée à la bande dessinée de laquelle elle est issue par la relation transitive *hasPart*. Chaque page est également reliée à une *Image* (au sens "fichier informatique") par la relation *hasImage*.

Une région d'intérêt est calculée par et liée à un extracteur via la relation *hasExtractor*. L'extracteur est soit automatique, résultat d'un algorithme d'extraction, soit humain dans le cas d'une extraction des *ROI* par l'humain. Le concept *Validation* et ses spécialisations, liés à *ROI* par les propriétés *hasValidation* et *hasReference*, nous permet de comparer entre eux les résultats de différents extracteurs.

La Fig. 1 propose un extrait de notre modèle comportant les concepts et les rôles nécessaires à la description des travaux présentés dans cet article.

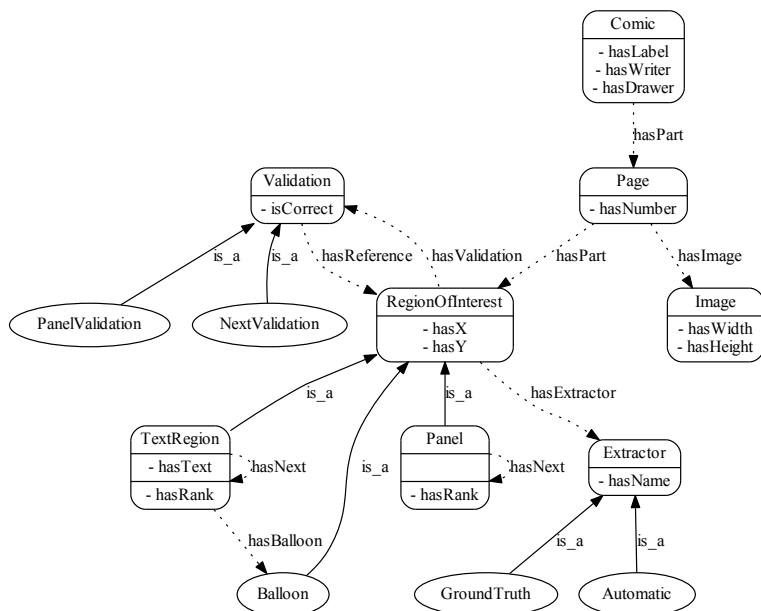


FIGURE 1 – Extrait de la hiérarchie de concepts (en flèches pleines) de notre ontologie et leurs relations (en flèches pointillées).

3.3 Règles de peuplement

Les informations bas niveau sont directement injectées dans la A-Box (base de faits) du modèle, à partir des données fournies par les extracteurs. Elles servent de base à l'élaboration de règles de peuplement de la connaissance scénaristique. La première notion déduite est la relation d'ordre entre les cases, i.e. l'ordre de lecture, modélisée par la relation *hasNext*.

Chaque case est caractérisée par un polygone dont les sommets prennent leurs coordonnées dans le repère de la page, l'origine de cette dernière étant son coin supérieur gauche. Deux cases *A* et *B* d'une même planche

sont liées par des relations spatiales topologiques et d'orientation cardinale. Concernant les premières, nous avons choisi d'utiliser trois des relations décrites par RCC8 (Cohn *et al.*, 1997), à savoir la disjonction DC , le chevauchement PO et la contenance, qu'elle soit tangentielle ou non, PP . Ces dernières n'étant pas suffisantes pour calculer une séquence de lecture, et considérant l'environnement très cadré d'une page de bande dessinée, nous intégrons en complément les relations d'orientation cardinales Up , $Down$, $Right$ et $Left$.

Avant de pouvoir calculer les relations $hasNext$, il nous faut formaliser les relations d'orientation cardinale entre les cases. Il existe plusieurs façons de calculer de telles relations entre des polygones. Dans les cas des bandes dessinées, nous voulons garantir l'absence de cycles dans ces relations car cela est nécessaire à la définition d'un ordre de lecture. Pour cela, nous décomposons le calcul en deux étapes selon l'équation (1) qui décrit la relation $Down$ – la relation $Right$ est définie de manière similaire à partir des abscisses des points. Lorsque les cases sont disjointes ($DC(A, B)$), nous considérons les points des polygones $\{(x_{A1}, y_{A1}); \dots; (x_{An}, y_{An})\}$, $\{(x_{B1}, y_{B1}); \dots; (x_{Bm}, y_{Bm})\}$. Lorsqu'elles se chevauchent ($PO(A, B)$), nous utilisons leurs centres de gravité $\{(x_{AC}, y_{AC})\}$, $\{(x_{BC}, y_{BC})\}$. Lorsque qu'une case est incluse dans une autre ($PP(A, B)$), nous considérons que la relation d'orientation est indéterminable. Il ne peut donc pas y avoir de cycles et les relations $Down$ et $Right$ sont des ordres partiels. De plus, on peut noter $Up = Down^{-1}$ et $Left = Right^{-1}$.

$$Down(A, B) \Leftrightarrow (DC(A, B) \wedge \forall (y_{Ai}, y_{Bi}), y_{Bi} > y_{Ai}) \vee (PO(A, B) \wedge y_{BC} > y_{AC}) \quad (1)$$

Une relation d'ordre est ensuite établie en deux étapes à partir des relations $Down$ et $Right$. Deux cases n'étant pas toujours liées par une relation d'orientation déterminée (lorsque l'une est incluse dans l'autre par exemple), on calcule d'abord l'ordre partiel (i.e. relation transitive, antisymétrique et réflexive) du séquençement topologique des cases selon (2).

$$StructuralNext(A, B) \Leftrightarrow Down(A, B) \vee (\neg Down(A, B) \wedge \neg Down(B, A) \wedge Right(A, B)) \quad (2)$$

Nous effectuons ensuite un tri topologique de la relation $StructuralNext$ afin d'obtenir un ordre de lecture total $f_1 < f_2 < \dots < f_m$, modélisé par la

relation *hasNext*. Le relation *StructuralNext* ne représentant pas un ordre total, il peut exister plusieurs ordres de lecture (ie. plusieurs tris) pour un même séquençement topologique de cases. A noter qu'il est possible de travailler sur un sens de lecture oriental (de droite à gauche) en remplaçant simplement $Right(A, B)$ par $Right^{-1}(A, B)$ dans l'équation 2.

3.4 Expérimentation

La A-Box de notre ontologie a été peuplée à partir de bandes dessinées provenant d'auteurs tels que Cyb (Cyb, 2009) ou Trébla (Trébla, 2008). Sur les 360 planches à notre disposition, 22 ont été choisies pour l'élaboration d'une vérité terrain portant à la fois sur le périmètre des cases et sur leur ordre de lecture dans la page. Les cases ont d'abord été détournées manuellement, une à une, produisant ainsi un ensemble de polygones dont les coordonnées peuvent être passées en entrée de notre algorithme. Chaque case s'est ensuite vue attribuée un numéro correspondant à son ordre de lecture dans la page. Notre algorithme de calcul des relations *hasNext* a ensuite été appliqué sur la base des cases délimitées. Les deux jeux de données ont été injectés dans notre ontologie, associés à deux extracteurs *GroundTruth* et *Automatic*. Pour chaque axiome $hasNext(A, B)$ de l'extracteur *Automatic* on recherche l'existence de l'atome $hasNext(A, B)$ associé à *GroundTruth*. Les polygones utilisés dans les deux jeux de données étant les mêmes, il ne peut pas y avoir d'ambiguïté dans l'appariement des cases A et B des deux extracteurs. Sur les 22 planches annotées, 338 relations *hasNext* ont été établies et toutes ont été validées par la vérité terrain. Nous avons mis en pratique ces relations en chargeant notre ontologie dans Sewelis (Ferré & Hermann, 2011), le logiciel nous permettant alors de naviguer à travers les pages et les cases de notre base de BD.

4 Conclusions et perspectives

Cet article présente notre contribution dans la résolution du problème du fossé sémantique dans le cas de la bande dessinée. Nous avons présenté une ontologie permettant une annotation des éléments extraits des images de la BD et facilitant la navigation dans ces images en formalisant l'ordre de lecture des cases. Cette dernière fournit un contexte à la formalisation d'informations extraites d'images de BD, qu'elles soient de haut ou de bas niveau. Elle est amenée à évoluer afin d'inclure les objets contenus dans les cases et leurs relations (scénaristiques et spatiales). Notre vérité terrain sera prochainement étendue à l'ensemble de nos 360 planches et complé-

tée par les concepts de *zone de texte* et de *personnages*. Ce corpus est lui-même susceptible de s'enrichir afin de couvrir un panel d'une diversité plus importante. Notre algorithme de calcul des relations *hasNext* pourra être appliqué aux zones de texte et à des cases issues de mécanismes d'extraction. Un système d'interaction avec l'utilisateur est également à l'étude dans le but d'apporter des corrections et de créer un bouclage de pertinence entre le modèle et les extracteurs.

Références

- BANNOUR H. & HUDELLOT C. (2012). Combinaison d'information visuelle, conceptuelle, et contextuelle pour la construction automatique de hiérarchies sémantiques adaptées à l'annotation d'images. In *Actes de RFIA'12*, p. 460–469.
- BERNERS-LEE T., HENDLERS J. & LASSILA O. (2001). The semantic web. *Scientific American*, **284**(5), 34–43.
- COHN A., BENNET B., GOODAY J. & GOTTS N. (1997). Representing and reasoning with qualitative spatial relations about regions. In *Spatial and temporal reasoning*, p. 97–134. Kluwer.
- CYB (2009). Cyborg 07's website. <http://www.cosmozone.fr/>.
- FERRÉ S. & HERMANN A. (2011). Semantic search : reconciling expressive querying and exploratory search. In *proc. of the ISWC'11*.
- HAN E., YANG J., YANG H. & JUNG K. (2007). Automatic mobile content conversion using semantic image analysis. In *Proc. of the HCI'07*, p. 298–307, Berlin, Heidelberg : Springer-Verlag.
- LAINÉ J. & DELZANT S. (2007). *Le dessin des planches*. Eyrolles.
- NEUMANN B. & MÖLLER R. (2008). On scene interpretation with description logics. *Image Vision Comput.*, **26**, 82–101.
- RIGAUD C., TSOPZE N., BURIE J. & OGIER J. (2012). Extraction robuste des cases et du texte de bandes dessinées. In *Proc of CIFED'12*.
- TOUSCH A.-M., HERBIN S. & AUDIBERT J.-Y. (2012). Semantic hierarchies for image annotation : A survey. *Pattern Recognition*, **45**, 333–345.
- TRÉBLA (2008). Trebla's blog. <http://trebla.over-blog.com/>.
- USCHOLD M. & KING M. (1996). Ontologies : principles, methods and applications. *Knowledge Engineering Review*, **11**(2), 93–155.
- YAMADA M., BUDIARTO R., ENDO M. & MIYAZAKI S. (2004). Comic image decomposition for reading comics on cellular phones. *IEICE Transactions*, **87-D**(6), 1370–1376.
- ZHANG D., ISLAM M. M. & LU G. (2012). A review on automatic image annotation techniques. *Pattern Recognition*, **45**, 346–362.