

PRISE EN COMPTE DE L'IMPERFECTION DES CONNAISSANCES DEPUIS LA SAISIE DES DONNÉES JUSQU'À LA RESTITUTION 3D

1. INTRODUCTION

En interrogeant le passé, l'information archéologique est naturellement sujette à de l'imperfection. Nous en présenterons succinctement différents aspects (imprécision, incertitude, vague, conflit, lacune) qui peuvent intervenir sur les trois composantes que sont le temps, l'espace et la fonction. Dans un premier temps, elle doit donc pouvoir être identifiée, caractérisée et mémorisée dans le système d'information archéologique (qui peut-être basé sur un système d'information géographique comme l'indique James CONOLLY et Mark LAKE 2006). En second lieu, se pose alors la question de la prise en compte de cette imperfection des connaissances dans les hypothèses archéologiques d'analyse, de production et de restitution. Face à ce besoin, nous avons fait le choix dans le projet SIGRem de travailler dans le cadre de la théorie des ensembles flous. L'étape finale de médiatisation des résultats peut de nos jours s'effectuer par une modélisation en 3 dimensions. Mais lors de cette dernière étape, la reconstruction réaliste étant très souvent perçue comme la vérité, nous perdons fréquemment la richesse des niveaux de confiance que nous avons en nos connaissances. Nous présenterons alors comment certains paradigmes visuels peuvent être mobilisés pour en permettre la perception dynamique dans des environnements virtuels 3D dédiés.

2. DIFFÉRENTS ASPECTS DE L'IMPARFAIT

L'imperfection des connaissances est inhérente à toute appréhension du monde. Ne pas la prendre en compte peut être source d'erreurs d'interprétation et sa visualisation peut aider à la compréhension ou la prise de décision (DETRICK 2007). Elle doit donc être intégrée tout au long du processus, de l'acquisition des données à la restitution des hypothèses. Cela nécessite de pouvoir l'identifier, la modéliser et la quantifier.

Bien que de nombreux travaux aient été réalisés ces dernières décennies, il n'y a pas réellement à ce jour de terminologies et de définitions définitivement établies.

Il est possible toutefois, en s'appuyant par exemple sur les travaux de Peter FISHER (1999, 191-205) et de DE RUNZ *et al.* (2008), d'identifier les quatre grandes catégories suivantes rappelées dans le Schéma 1:

Incertitude:

Il y a un doute sur la validité de la connaissance.

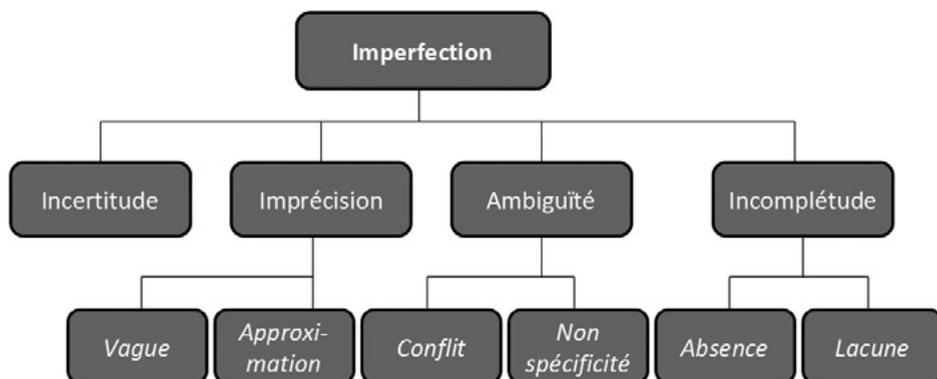


Schéma 1 – Une typologie de l'imperfection des données.

L'objet est bien défini mais sa réalisation est incertaine. Cette situation est souvent liée à l'aspect aléatoire de la mesure d'un phénomène physique.

Imprécision:

Il y a une difficulté à exprimer clairement les connaissances.

Cette fois, l'objet n'est pas suffisamment clairement défini. Il y a un manque de précision dans la définition de l'objet. Cela peut être dû à l'aspect *vague* de la sémantique du concept et/ou de ces limites. Il en est de même lorsqu'il est *approximé* par analogie avec des exemples avoisinants.

Ambiguïté:

Il y a une difficulté à être d'accord.

L'ambiguïté survient lorsqu'il y a un doute sur la manière de définir un objet ou un phénomène, c'est-à-dire quand un élément peut appartenir à plusieurs catégories disjointes ou d'échelles différentes, ou encore quand la description de l'élément peut donner lieu à plusieurs sens. D'une manière plus spécifique, il y a *conflit* si au minimum deux classifications contradictoires pour un unique objet sont possibles. Lorsqu'une définition d'une relation ou d'un objet peut amener à plusieurs sens, ou lorsque l'échelle de l'analyse est susceptible d'amener à de multiples interprétations, on parle alors de *non-spécificité*.

Incomplétude:

Il y a des connaissances manquantes ou partielles.

Les incomplétudes sont des absences de connaissances ou des connaissances lacunaires. L'*absence* est le phénomène qui survient quand dans une base de données, des valeurs manquent à la description de certains objets. La

lacune est le fait qu'un ou plusieurs objets de la base de données ne décrivent que de manière partielle une structure les englobant.

Les catégories que nous venons de présenter ne sont nullement exclusives. D'une manière générale, les données sont sujettes simultanément à des imperfections de plusieurs natures.

3. DIVERSES THÉORIES DE REPRÉSENTATION DE L'IMPERFECTION

Afin de manipuler l'incertitude, de nombreux cadres théoriques ont été formalisés. Nous en présentons ici les principaux.

Probabilité:

Elle est liée à la mesure ou à l'estimation d'événements bien définis mais dont la réalisation est caractérisée par le hasard. La probabilité qu'un tel événement A survienne est ainsi définie par une fonction $P(A)$ à valeur dans $[0;1]$. Elle est très bien adaptée au traitement de l'incertitude. Toutefois, la contrainte d'additivité des événements disjoints est une hypothèse forte qui ne peut pas toujours être validée.

(Sous-)Ensembles flous:

Dans ce cadre introduit par Lotfi A. ZADEH (1965, 338-353), l'appartenance d'un objet x à un ensemble A n'est plus booléenne (vrai/faux) mais est caractérisée par un degré d'appartenance $\mu_A(x)$ pouvant varier dans l'intervalle $[0;1]$ où les valeurs extrêmes correspondent aux valeurs booléennes classiques avec 0 pour faux et 1 pour vrai. De nombreuses extensions au "flou" de la représentation booléenne (tout ou rien) du monde ont été développées: logique floue, arithmétique floue... Ce formalisme conceptuel est bien adapté au traitement de l'imprécis.

Théorie des possibilités:

Introduite par Lotfi A. Zadeh en 1978 (ZADEH 1978, 3-28), les fondements de la théorie des possibilités ont été édités par Didier Dubois et Henri Prade dans *Théorie des possibilités* (DUBOIS, PRADE 1985) et *Possibility Theory* (DUBOIS, PRADE 1988). L'incertitude d'un événement A , au contraire des probabilités, est caractérisée par deux valeurs: sa possibilité (A) et sa nécessité $N(A)$. Cette approche est bien adaptée à la gestion d'information imprécise dans un environnement incertain

Théorie de l'évidence/des croyances/de Dempster-Shafer:

Dans ce cadre plus général introduit par Arthur DEMPSTER (1968, 205-247) puis Glenn SHAFER (1976) et étendu par Philippe SMETS et Robert KENNES (1994, 191-243), une connaissance imparfaite sur l'univers est représentée par une masse de croyance définie comme une fonction de l'ensemble

des parties de cadre de discernement dans $[0;1]$. Elle permet aussi la construction de croyance et de plausibilité qui encadrent la probabilité. Pouvant être perçue comme une généralisation des théories précédentes, elle est un cadre idéal pour le traitement différencié de l'incertain et de l'imprécis, ainsi que pour la fusion d'informations multi-sources.

Logique bipolaire

Lorsque l'on dispose en plus de l'information sur la non existence (information négative), il est intéressant de pouvoir gérer un degré pour l'appartenance à A et un autre pour $non-A$ comme le précise Isabelle BLOCH (2010, 91-112).

Ensembles approximatifs (Rough Sets):

Une autre approche peut consister en une approximation formelle d'un ensemble classique (ensemble approximation haute/possible et ensemble approximation basse/nécessaire).

Il n'était pas ici question d'être exhaustif mais de présenter les approches fondatrices de la manipulation de l'imparfait formant une branche de ce que l'on nomme maintenant le *Soft computing*.

4. REPRÉSENTATION DANS LE SYSTÈME D'INFORMATION

Ces diverses formes d'imperfection de l'information s'appliquent à toutes les caractéristiques propres des entités manipulées dont, selon Xavier RODIER et Laure SALIGNY (2007), les composantes sont le temps, l'espace et la fonction/sémantique. Il en est de même dans la gestion des relations entre entités.

Issus des travaux de normalisation en géomatique pour la gestion et l'échange des données (normes ISO19113, ISO19115), la qualité de l'information géographique peut être mémorisée dans le système d'information par le biais des métadonnées. Toutefois cette connaissance ne correspond pas directement à une représentation de l'imparfait et ne peut pas être directement utilisée.

Les modalités de stockage interne les plus classiques sont l'association d'intervalles de confiance, de degrés de croyance, de fonctions d'appartenance... Le choix que nous avons mis en œuvre actuellement dans GISSAR est celui des fonctions d'appartenance des ensembles flous.

5. EXEMPLES DE RECONSTRUCTION À PARTIR DE CONNAISSANCES IMPARFAITES: 2D

Dans le cadre du projet GISSAR relatif à la mise en œuvre d'un Système d'Information Géographique en Archéologique dédié à la cité des Rems, nous nous sommes intéressés à l'analyse du maillage des rues en fonction des dif-

férentes époques. Chaque donnée est issue de résultats de fouilles et correspond à l'identification ponctuelle d'un segment de rues. Elles sont caractérisées par trois attributs: un positionnement, une estimation de l'orientation et de la période d'activité.

Au niveau de la position, les sources d'imperfection peuvent être multiples: précision des instruments de mesure, aspect approximatif du recalage par rapport à des repères visuels pris dans le voisinage du site de fouille... Pour l'orientation, en fonction de la qualité des objets trouvés, celle-ci est objectivement estimée avec une imprécision plus au moins grande. Quant à la datation, celle-ci est fournie par un label identifiant une période temporelle dont les limites sont floues.

Pour exemple, un segment de rue peut être caractérisé par les valeurs d'attributs:

- Localisation: 723325 ± 2 m, 174361 ± 2 m;
- Orientation: $210^\circ \pm 5^\circ$;
- Datation: Gallo-romain.

Ces imperfections sont modélisées par l'utilisation d'objets flous. Pour cela, nous définissons des fonctions d'appartenance pour chacun des attributs (composantes de l'information) dont les valeurs sont maximales (vrai=1) pour les valeurs caractéristiques et décroissantes (jusqu'à faux=0) plus on s'en éloigne. Des exemples sont présentés dans la Fig. 1, Pl. XV, a.

Ces informations sont alors utilisées dans une adaptation floue, proposé par Joon HAN *et al.* en 1994 (HAN, KOCZY, POSTON 1994, 649-648), de la transformée de Hough. Cyril de RUNZ et Eric DESJARDIN (DE RUNZ *et al.* 2010, 341-356) ont publié en 2010 la méthode de construction de cette requête. Nous obtenons alors un ensemble d'hypothèses de présence de rues en fonction des époques, hypothèses portant une estimation de l'assurance que nous avons de la réalisation de cet événement (Fig. 2, Pl. XV, b).

6. EXEMPLES DE RECONSTRUCTION À PARTIR DE CONNAISSANCES IMPARFAITES: 3D

Un autre exemple porte sur la reconstruction virtuelle en 3D d'un bâtiment. Construite au 13^{ème} siècle, l'abbaye Saint-Symphorien de Reims fût définitivement détruite au XVIII^e siècle. Elle n'est plus perceptible qu'à travers les traces qu'elle a laissées dans le passé: fouilles de l'INRAP, croquis anciens, documents sur des travaux médiévaux qui peuvent être enrichis par des connaissances historiques complémentaires telles les hypothèses fortes d'une très grande similarité avec la basilique St-Remi (Reims).

Il est bien évident dans ce genre de travaux que les connaissances que nous pouvons recueillir sont partielles, lacunaires, approximatives, imprécises...

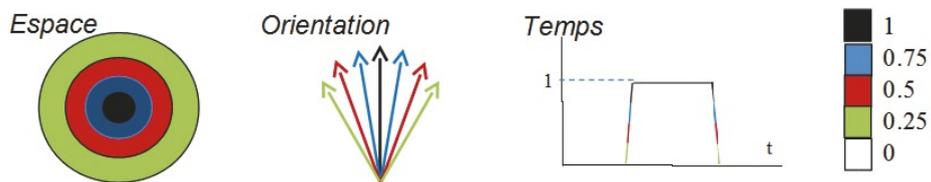


Fig. 1 – Exemples de modélisation par des objets flous.



Fig. 2 – Hypothèses évaluées de présence de rues de *Durocortorum* au III^e siècle ap. JC: les hypothèses sont en niveau de gris superposées sur un plan des voiries issu d'expertises (en noir).

Les questions qui se posent sont alors: une reconstruction est-elle possible à partir des informations partielles retrouvées? Comment formaliser les hypothèses? Comment les valider interactivement? La structure imaginée par l'historien est-elle consistante en 3D?

Les recherches que nous menons portent alors sur la mise à disposition efficace auprès de l'utilisateur de l'évaluation de la certitude que nous avons sur les connaissances visualisées dans une reconstruction.

7. MODALITÉS DE RESTITUTION

Un certain nombre de travaux ont déjà été réalisés dans le cadre de la visualisation 2D: cartographie, Systèmes d'Information Géographique, visualisation de données scientifiques, urbanisme, archéologie... La visualisation de l'imparfait dans le monde de l'information géospatiale est un problème important comme l'indique l'article d'Alan MACEACHREN *et al.* (2005, 139-160). Un des axes de la recherche dans ce domaine concerne le passage à la prise en compte de la troisième dimension et du temps.

L'utilisation d'un environnement immersif 3D stéréo permet-elle l'activation de nouveaux paradigmes de représentation de l'imparfait? Afin de répondre à cette question, nous mettons en place quelques modalités de perception permettant la prise en compte de l'imperfection. Celles-ci s'appuient sur les techniques:

Virtualité augmentée

Dans ce cas précis, Il s'agit d'ajouter des informations non visuelles (degré d'appartenance ou de croyance) au sein de l'environnement virtuel dans lequel l'utilisateur est immergé:

Représentation interne

Les attributs non visuels d'un objet peuvent être représentés par une couleur ou une info bulle accrochée à l'objet. Cette représentation assez courante en cartographie n'est pas appropriée aux environnements virtuels immersifs où la lecture de texte est peu confortable (dans le cas des infos bulles) et où l'utilisation de couleurs peut interférer avec les textures et l'éclairage du monde virtuel et nécessite l'affichage en surimpression de l'échelle des couleurs.

Représentation dynamique

Plutôt que d'ajouter de l'information visuelle statique en superposition de l'environnement virtuel, on peut altérer les propriétés visuelles intrinsèques des objets pour renseigner l'utilisateur sur la valeur de leurs attributs non visuels. Par exemple, faire scintiller un objet 3D avec une fréquence proportionnelle à son degré d'appartenance. Ou animer un objet pour que ses

dimensions géométriques (longueur, largeur, hauteur) parcourent toutes les valeurs de l'intervalle de confiance.

Représentation interactive

Dans ce mode de représentation, le spectateur du monde virtuel devient acteur dans la mesure où l'on peut solliciter chez lui d'autres sens que la vision. A l'aide d'un dispositif haptique à retour d'effort, l'utilisateur peut glisser sa main "virtuelle" sur la surface d'un mur et "ressentir" le degré de croyance en fonction de la réponse, dure ou molle, du dispositif haptique.

7.1 Lunette de réalité dans les mondes virtuels

Alors que la tendance a été ces dernières années d'aller vers de plus en plus de réalisme visuel, force est de constater qu'elle nous éloigne très souvent de la réalité de la connaissance que nous avons en ne permettant pas la perception par l'utilisateur de son imperfection. Il nous faut alors dans un environnement interactif inverser ce paradigme et "chausser des lunettes de réalité". Le système de restitution fournit alors, en fonction des réglages du filtrage perceptif du monde, des modalités différentes permettant de percevoir les zones où notre connaissance est plus ou moins sûre.

Pour exposer notre propos, nous nous appuyerons sur un exemple de reconstruction d'un site à partir de données de fouilles. L'emplacement au sol est bien défini. Par contre, l'architecture des bâtiments est assez sûre sur le devant et hypothétique sur l'arrière. Trois niveaux de perception sont présentés dans les Figs. 3, 4 et 5. Dans le premier, nous souhaitons percevoir le monde avec un niveau élevé de réalité quant à la justesse de la connaissance que nous avons. Comme nous connaissons en fait peu de chose, nous utilisons un paradigme de représentation de type *cartoon*. A l'opposé, dans le dernier niveau, nous souhaitons une visualisation dite *réaliste* en synthèse d'image dans le sens où elle se rapproche d'une photo que l'on aurait pu prendre d'un bâtiment *réel*. Mais dans ce cas, au regard de la réalité de nos connaissances, il s'agit bien d'une construction purement virtuelle. Les niveaux intermédiaires peuvent utiliser d'autres paradigmes de représentation comme la transparence, le floutage, la superposition, le niveau de détail... ZUK *et al.* (2005, 99-106) ou encore Maria RIVEIRO (2007, 1-8) ont testé plusieurs d'entre-elles en 2005. Dans l'exemple intermédiaire que nous avons appelé semi-réel, des techniques différentes sont utilisées selon les degrés de confiance: réaliste si la connaissance est sûre (partie avant), *cartoon* si elle est hypothétique (partie arrière) et réaliste avec superposition de traits gras sur les contours si elle est assez probable (partie intermédiaire).

La possibilité de changer dynamiquement de *point de vue* dans la dimension *vérité de l'information construite* est un apport nouveau dans les mondes virtuels.



Fig. 3 – Modalité de représentation de la réalité de la connaissance.



Fig. 4 – Modalité de représentation semi-réelle de la réalité.



Fig. 5 – Modalité de représentation virtuelle des connaissances.

8. CONCLUSION

L'imperfection des connaissances est un phénomène naturel auquel l'archéologie est tout particulièrement sujette. Nous avons présenté une typologie des sources d'imperfection puis des principales théories. Puis dans un premier exemple, nous en avons exposé une prise en compte de bout en bout de l'acquisition à la restitution d'hypothèses spatiales et temporelles. L'exemple suivant a permis d'exprimer nombre d'interrogations quant à la validité de représentation pseudo-réaliste de bâtiments en 3D. Dans ce cadre il semble pertinent de fournir à l'utilisateur plusieurs modalités de représentation permettant de rendre compte dynamiquement de l'assurance que nous avons dans les connaissances projetées en lui permettant de chausser des *lunettes de réalité* dans les environnements virtuels immersifs 3D stéréo. L'extension de cette perception de l'imperfection par le biais des capteurs haptiques est l'un des axes de nos travaux à venir.

ERIC DESJARDIN
GEGENAA – URCA
Université de Reims Champagne-Ardenne
OLIVIER NOCENT, CYRIL DE RUNZ
CRESTIC SIC
Université de Reims Champagne-Ardenne

Remerciements

Ce travail a été financé par la Région Champagne-Ardenne dans le cadre des programmes SIGRem puis ArchéoChamp et s'est nourri des collaborations avec le laboratoire de recherche GEGENAA, le SRA, l'INRAP et le Service Archéologique de Reims-Métropole que nous tenons à remercier. Le modèle 3D a été conçu par le cabinet rémois d'architectes "BLP Architecte".

BIBLIOGRAPHIE

- BLOCH I. 2010, *Bipolar fuzzy spatial information: First operations in the mathematical morphology setting*, in R.K. DE, D.P. MANDAL, A. GHOSH (eds.), *Machine Interpretation of Patterns: Image Analysis and Data Mining*, World Scientific Press, 91-112.
- CONOLLY J., LAKE M. 2006, *Geographical Information Systems in Archaeology*, Cambridge, Cambridge University Press.
- DE RUNZ C., DESJARDIN E. 2010, *Imperfect spatiotemporal information modeling and its analysis in a generalization process in a GIS: Application to archaeological information*, in R. JEANSOULIN, O. PAPINI, H. PRADE, S. SCHOCKAERT (eds.), *Methods for Handling Imperfect Spatial Information*, Studies in Fuzziness and Soft Computing, 256, Berlin, Springer Verlag, 341-356.
- DE RUNZ C., DESJARDIN E., PIANTONI F., HERBIN M. 2008, *Toward handling uncertainty of excavation data into a GIS*, in *36th Annual Conference on Computer Applications and Quantitative Methods in Archaeology (CAA)*, Budapest, Hongrie.

- DEITRICK S.A. 2007, *Uncertainty visualization and decision making: does visualizing uncertain information change decisions*, in XXIII International Cartographic Conference, Moscou.
- DEMPSTER A. 1968, *Generalization of Bayesian inference*, «Journal of the Royal Statistical Society», 30, 2, 205-247.
- DUBOIS D., PRADE H. 1985, *Théorie des Possibilités. Applications à la Représentation des Connaissances en Informatique*, Paris, Masson (Collection Méthode + Programmes).
- DUBOIS D., PRADE H. 1988, *Possibility Theory*, New York, Plenum Press.
- FISHER P. 1999, *Models of uncertainty in spatial data*, in P.A. LONGLEY, M.F. GOODCHILD, D.J. MAGUIRE, D. RHIND (eds.), *Geographic Information Systems: Principles and Technical Issues*, New-York, Wiley, 191-205.
- HAN J., KOCZY L., POSTON T. 1994, *Fuzzy Hough Transform*, «Pattern Recognition Letters», 15, 649-648.
- MACEachREN A., ROBINSON A., HOPPER S., GARDNER R.M., GAHEGAN M., HETZLER E. 2005, *Visualizing geospatial information uncertainty: What we know and what we need to know*, «Cartography and Geographic Information Science», 32:3, 139-160.
- RIVEIRO M. 2007, *Evaluation of uncertainty visualization techniques for information fusion*, in 10th International Conference on Information Fusion (ICIF), 1-8.
- RODIER X., SALIGNY L. 2007, *Modélisation des objets urbains pour l'étude des dynamiques urbaines dans la longue durée*, in M. BATTON-HUBERT, T. JOLIVEAU, S. LARDON (eds.), SAGEO'07, *Rencontres internationales Géomatique et territoire*, <http://www.emse.fr/site/SAGEO2007/CDROM/index.html>.
- SHAFER G. 1976, *A Mathematical Theory of Evidence*, Princeton, N.J., Princeton University Press.
- SMETS P., KENNES R. 1994, *The transferable belief model*, «Artificial Intelligence», 66, 191-243.
- ZADEH L.A. 1965, *Fuzzy sets*, «Information and control», 8, 338-353.
- ZADEH L.A. 1978, *Fuzzy sets as a basis for a theory of possibility*, «Fuzzy Sets and Systems», 1, 3-28.
- ZUK T., SHEELAGH M., CARPENDALE T., GLANZMAN W.D. 2005, *Visualizing temporal uncertainty in 3D virtual reconstructions*, in *International Symposium on Virtual Reality, Archaeology and Cultural Heritage (VAST)*, The Eurographics Association, 99-106.

ABSTRACT

By questioning the past, archaeological information is naturally prone to imperfection. Through a series of examples, we will briefly present its various aspects (inaccuracy, uncertainty, vagueness, conflict, lack) which can apply to time, space and function. The first stage consists in the identification, characterization and recording of imperfection in the archaeological information system. At the second stage, the question arises of how the imperfection of knowledge in archaeological hypotheses should be taken into account in terms of analysis, production and restitution. In the SIGRem project, we have chosen to resort to the Fuzzy sets theory. At a final stage, although the promotion through the media of results can nowadays be carried out by a 3D modeling, realistic reconstruct being very often confusingly perceived as truth, we tend to lose the richness of confidence levels we have in our knowledge. Therefore, we will also describe how visual paradigms can be used to enable dynamic perception of uncertainty in dedicated 3D virtual environments.

