

Compte-rendu de la conférence

17^e journées Francophones d'Ingénierie des Connaissances 2006

Coordonné par Myriam Lewkowicz (Tech-CICO, UTT, Troyes), présidente du comité de programme de IC 2006

L'édition 2006 de la conférence Ingénierie des Connaissances s'est tenue dans le cadre de la Semaine de la Connaissance, à Nantes, du 26 au 30 juin 2006.

40 propositions d'articles longs ont été évaluées par le comité de programme. 18 articles longs ont été retenus pour les actes et présentés lors de 8 sessions dont le compte-rendu est présenté ci-dessous. 15 posters ont été retenus comme articles courts présents sur le CD des actes, et ont été présentés lors d'une session dédiée.

Modèles de connaissance

Session présidée par Chantal Reynaud

Cette première session d'IC'2006 a porté sur les modèles de connaissance. Le premier papier concernait l'extraction de connaissances à partir de données sous forme de règles d'association. Son objectif était d'aider à analyser les résultats de l'extraction en confrontant les règles à un modèle de connaissance de l'expert. Celui-ci contient au départ les dépendances les plus évidentes entre les données. Il est ensuite complété à partir des résultats issus de la confrontation des règles au modèle et permet ainsi d'identifier trois sortes de règles : les connues mais inintéressantes, les règles non valides, les règles valides et surprenantes. Le second papier a porté sur la transformation d'un modèle exprimé dans le langage des Topic Maps en graphes conceptuels de façon à pouvoir mettre en œuvre différents mécanismes de raisonnement via l'API CoGITaNT.

Langages de modélisation

Session présidée par Marie-Christine Jaulent

La session sur les langages de modélisation a donné lieu à deux présentations. La première portait sur l'association d'une sémantique formelle aux descripteurs UCD (Unified Content Descriptors) en astronomie afin de construire une ontologie de ce domaine. Une méthodologie pour construire une ontologie des UCD normalisés en OWL a d'abord été présentée, suivie d'une méthode d'annotation dans laquelle l'ontologie est exploitée pour l'attribution semi-automatique d'UCD à des descriptions de propriétés d'objets célestes issues de catalogues d'astronomie. La seconde présentation a porté sur les apports de la modélisation algébrique pour la représentation de connaissances par objets. Elle a d'abord permis de rappeler les principes généraux de la représentation de connaissances par objets, et des langages de modélisation algébrique. En prenant l'exemple d'AROM, cette présentation a montré comment la présence du LMA augmente le pouvoir expressif (ou déclarativité) d'AROM en donnant la possibilité à un concepteur de formaliser, dès la construction, une base des connaissances opératoires, et, à un utilisateur, d'interroger la base de connaissances à l'aide d'un langage de requêtes formel basé sur le calcul ensembliste.

Raisonnement à partir de cas

Session présidée par Yannick Prié

Au cours de la session consacrée au Raisonnement à Partir de Cas (RàPC), Amélie Cordier nous a présenté un article écrit avec Béatrice Fuchs et Alain Mille, intitulé "De l'Ingénierie à l'Apprentissage des connaissances d'adaptation en Raisonnement à Partir de Cas". La présentation - dense - a traité de la problématique de l'Ingénierie des Connaissances en RàPC. La première partie a permis de faire l'inventaire des connaissances mobilisées dans le cycle du RàPC, la seconde s'est intéressée au type particulier de connaissances que sont les connaissances d'adaptation. Des stratégies d'apprentissage de ces connaissances ont été proposées, à partir d'une formalisation basée sur les notions de dépendance et d'influence. L'un des points importants de cette approche est que toutes les stratégies donnent un rôle prépondérant à l'utilisateur. La discussion qui a suivi la présentation a porté sur la façon dont il était possible de sélectionner les connaissances d'adaptation à appliquer et si un système de pondération (cf. Évaluation de similarité de cas) pourrait être utilisé.

Annotations

Session présidée par Rose Dieng

La problématique émergente des annotations fait son apparition à IC 2006 avec cette session, rassemblant deux présentations, représentatives de deux usages différents des techniques du Web Sémantique pour la gestion des annotations. La première présentation porte sur une approche basée sur l'annotation sémantique des CVs et des offres d'emploi pour automatiser le processus de recrutement électronique. Le principe de base consiste à modéliser formellement le contenu de ces documents en termes de leurs acquis (cas des CVs) ou requis (cas des offres d'emploi), à l'aide d'un référentiel commun entre les recruteurs et les chercheurs d'emploi. Les annotations ne sont pas ici à l'usage d'agents humains, mais ont pour objectif d'optimiser l'appariement sémantique entre les CVs et les offres d'emploi. La seconde présentation, à l'inverse, propose d'assister la coopération par l'indexation semi-automatique d'annotations. Les annotations sont donc ici un moyen de faciliter l'interaction autour de documents au sein d'une équipe projet. Ici, les techniques d'IC sont utilisées pour faciliter l'indexation des annotations textuelles par l'utilisateur et les outils et protocoles de communication proposés par le Web Sémantique (Annozilla, Annotea) sont réutilisés afin de concevoir une solution robuste.

Analyse de corpus

Session présidée par Sylvie Szulman

Cette session comportait deux articles, mais O. Lartillot n'ayant pas pu venir, seul l'article de Axel Reymonet, Nathalie Aussenac-Gilles et Jérôme Thomas, intitulé « tâche, domaine et application : influences sur le processus de modélisation de connaissances » a fait l'objet d'une présentation par son premier auteur. Partant du constat selon lequel un nombre croissant d'outils de gestion de documents et de connaissances a désormais recours à des ressources terminologiques et/ou ontologiques (RTO) pour répondre à leurs besoins applicatifs, les auteurs montrent que le processus de modélisation de telles ressources passe par la prise en compte de la nature du domaine, de la tâche et de l'application visés. Le propos est illustré par une étude de cas de construction de RTO à partir de textes dans le domaine du diagnostic automobile. Cette étude de cas souligne notamment la nécessité d'une analyse préliminaire de l'application pour laquelle une ontologie est élaborée afin d'assurer la bonne adéquation entre la ressource termino-ontologique construite et les besoins applicatifs.

Conception d'outils supports à l'enseignement

Session présidée par Jean-François Boujut

Cette session comportait deux présentations qui ont abordé des thèmes importants du domaine. La première présentation abordait la question du support à l'activité individuelle du professionnel, les auteurs s'inscrivent dans le domaine du PMK (personal knowledge management). Le parti pris consiste à privilégier une approche à partir d'annotation sur les documents de travail et le domaine applicatif est celui de l'enseignement. L'outil proposé s'appuie sur le concept d'annotation sémantique support à la création, la mémorisation et au classement d'informations. Cette présentation a soulevé un certain nombre d'interrogations notamment par rapport à l'activité des enseignants, ici réduite à une activité individuelle alors que celle-ci a une dimension collective très forte. De même la finalité des annotations n'est pas vraiment caractérisée. Un parallèle avec le dossier patient a été soulevé, particulièrement autour des travaux sur les outils d'annotation du dossier patient dans le cadre du partage d'information. La seconde présentation abordait un domaine de l'assistance à la formation dans le domaine professionnel, ici la formation aux chirurgiens. L'originalité du positionnement se situe dans le caractère intermédiaire du type connaissance visé : des connaissances qualifiées d'empiriques (accessibles à travers une mise en scène dans un contexte donné), par opposition à des connaissances déclaratives ou gestuelles (habiletés). Les questions soulevées portaient sur le modèle sous jacent, sous forme de réseau Bayésien, et notamment si celui-ci était utilisé par les auteurs comme outil de travail (d'apprentissage) avec les chirurgiens. Quelles sont les limites de cette approche et donc quelles voies d'amélioration apparaissent ? Enfin, y a-t-il exploitation dans l'outil des logiques de choix et des décisions prises par les chirurgiens en situation d'apprentissage ?

Résolution de problèmes

Session présidée par Nathalie Aussenac-Gilles

La vocation initiale de modélisation des connaissances en IC était de développer des systèmes de résolution de problème. Après avoir presque disparu des contributions présentées à la conférence IC, cette notion est revenue à IC 2006 au cours d'une session rassemblant deux communications. La première, plus fondamentale, précise la notion de rôle pour la modélisation du raisonnement, alors que la seconde illustre l'intérêt de la modélisation du raisonnement pour mieux comprendre la nature d'une démarche méthodologique en Conception Innovante. S. Bruaux (LIA, Amiens) a d'abord présenté "OntoKADS : une ontologie générale de la résolution de problèmes". OntoKADS est une méthode de modélisation s'appuyant sur des ontologies pour décrire une méthode de résolution de problème au sens de CommonKADS, distinguant le domaine du raisonnement mis en oeuvre. Le raisonnement manipule des rôles, représentés par des méta-classes par rapport aux classes représentant les concepts du domaine. En typant les modes d'utilisation des connaissances dans le raisonnement, cette approche guide le cognicien dans la structuration des connaissances. F. Rousselot (LGeCo, INSA Strasbourg) s'intéresse à la Conception Innovante, et s'interroge sur la nature réelle du processus mis en oeuvre pour réaliser cette tâche : s'agit-il de la synthèse de systèmes ou d'une résolution de problèmes ? Cette question se pose car l'équipe cherche à développer et numériser l'approche TRIZ, une méthode de conception inventive, et cherche pour cela à standardiser sa mise en oeuvre. La démarche d'analyse proposée consiste à décrire la démarche préconisée par TRIZ à l'aide d'un modèle de raisonnement de CommonKADS.

Le Web Sémantique peut-il être social ?

Session présidée par Manuel Zacklad

Cette session a été organisée en deux temps autour de la présentation par F. Gandon de son article « Le Web Sémantique n'est pas anti-social ». Dans cet article F. Gandon (Inria Sophia / Accacia) faisait une revue critique de l'ensemble des articles présentés à IC par l'équipe Tech-CICO promotrice du Web Socio Sémantique. Bien que n'étant pas en désaccord avec la vision du Web promue par cette équipe, F. Gandon considère que celle-ci ne relève pas d'un courant spécifique mais intrinsèque au Web Sémantique. Selon lui, RDF est le véritable langage pivot du Web Sémantique et il permet toutes les applications qu'elle soit sociales ou computationnelles. Dans un deuxième temps, M. Zacklad (Tech-CICO / UTT) lui a apporté la contradiction en soutenant qu'il existait bien deux courants distincts dans le Web Sémantique, l'un relevant d'une forte tendance à la formalisation (le Web Computationnellement Sémantique) l'autre donnant la primauté à la médiation des activités humaines sans accorder d'importance particulière à la formalisation et notamment à celle des ontologies (le Web Socio Sémantique). Selon lui ces deux courants ne partagent pas la même vision de la signification (cf. la conférence invitée de la SDC par M. Slodzian). La suite de la session a été consacrée à des échanges avec la salle. Pour schématiser, trois tendances sont apparues : des chercheurs soutenant la thèse de F. Gandon et la vision « centrée RDF » du Web Sémantique, d'autres d'accord avec M. Zacklad pour considérer qu'il existait bien un fort courant formel du Web Sémantique assez distinct dans ses sous-basements épistémologiques des approches donnant la primauté aux langues naturelles et à la médiation de la coopération (W2S), et enfin une troisième tendance soutenant que le processus de formalisation est une tendance irréversible de l'informatisation qui finirait de toute manière par s'imposer dans l'ensemble des applications du Web. Cette session, fort riche, appelle à une prolongation de ces débats essentiels pour l'IC !

Cartographie et visualisation des connaissances

Session présidée par Nada Matta

Quatre présentations ont été présentées dans cette session. Les deux premières présentations : « Environnement d'ingénierie des connaissances : intégration et visualisation dans le domaine du vivant » par Fabien Jalabert et « Cartes conceptuelles pour l'ingénierie de cartes de connaissances adaptatives » par Michel Crampes ont montré les avantages d'une visualisation basée sur un modèle de force. Cette visualisation permet une navigation aisée et une focalisation suivant des vues spécifiques. L'approche a été présentée sur deux domaines d'applications différentes : la biologie et la musique. La troisième présentation « Techniques de gestion et d'ingénierie des connaissances pour la conception des dispositifs de transfert de savoir-faire dans les métiers pétroliers » par Djilali Benmahamed a mis l'accent sur l'importance de la définition d'une cartographie de connaissances pour le transfert et le partage des connaissances dans une entreprise. Ce travail a été illustré sur une application dans le domaine pétrolier. Enfin, la dernière présentation « Modélisation des méta-données bibliographiques : une contribution instrumentée destinée à améliorer la circulation des usagers dans les connaissances de la bibliothèque » par Fabrice Papy montre les problèmes de gestion de bibliothèque. Fabrice Papy propose une organisation des documents selon trois vues : classification de thèmes, notices bibliographiques et matières RAMEAU. Cette organisation permet de répondre aux différents besoins d'utilisation des documents.

Conférencier invité : Michael Baker

Michael Baker, fondateur en 2000 de l'équipe « Interaction & Cognition », au sein de l'UMR 5612 GRIC, a présenté le rôle et l'usage des modèles pour la conception de systèmes, et en particulier pour les Environnements Informatiques pour l'Apprentissage Humain. Sa proposition consiste en premier lieu à voir les EIAH comme étant issus d'une combinaison de trois types de modèles : modèles des situations, modèles informatiques, et modèles de conception. Selon lui, c'est ce qui distingue l'EIAH de recherches plus descriptives ou normatives en sciences de l'éducation. Plus généralement, Michael Baker propose de distinguer trois types de modèles :

- (1) Les modèles *composants logiciels*, qui sont les modèles théoriques qui ont inspiré la réalisation du logiciel, les modèles effectivement sous-jacents au logiciel.
- (2) Les modèles pour *l'évaluation*, qui sont les modèles de la cognition d'un individu, ou, au sens large, le modèle d'une cognition située, distribuée, ou socialement partagée.
- (3) Les modèles de la conception, la conception étant vue comme un processus collectif (vision descriptive).

Michael Baker souligne l'intérêt de disposer de modèles qui soient à la fois divers et imprécis, flous, ... et qui ne représentent pas forcément la sémantique d'un système formel. C'est à cette condition que les modèles peuvent contribuer à un processus situé de conception faisant intervenir des disciplines différentes, et permettre la « production des idées dans l'interaction chaotique ». Le modèle peut être ainsi appréhendé comme un objet intermédiaire entre concepteur, utilisateur et chercheur. Il peut être le moyen d'établir de nouvelles relations entre l'informatique et les sciences humaines dans l'élaboration des systèmes supports à l'activité.