



Extraction de la structure de la sémantique dans les modèles de système d'information d'entreprises collaboratives

Esma Yahia, Mario Lezoche, Alexis Aubry, Hervé Panetto

► To cite this version:

Esma Yahia, Mario Lezoche, Alexis Aubry, Hervé Panetto. Extraction de la structure de la sémantique dans les modèles de système d'information d'entreprises collaboratives. *Revue des Sciences et Technologies de l'Information - Série ISI: Ingénierie des Systèmes d'Information*, 2012, 17 (4), pp.49-77. 10.3166/ISI.17.4.49-77 . hal-00722247

HAL Id: hal-00722247

<https://hal.science/hal-00722247v1>

Submitted on 1 Aug 2012

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.

Extraction de la structure de la sémantique dans les modèles de systèmes d'information d'entreprises collaboratives

Esma YAHIA^{1,2}, Mario LEZOCHÉ^{1,2}, Alexis AUBRY^{1,2}, Hervé PANETTO^{1,2}

¹Université de Lorraine, CRAN, UMR 7039
Campus Scientifique, Faculté des Sciences et Technologies
BP 70239, 54506 Vandoeuvre-lès-Nancy Cedex, France

²CNRS, CRAN, UMR 7039, France

{Esma.Yahia, Mario.Lezoche, Alexis.Aubry, Herve.Panetto}@univ-lorraine.fr

RÉSUMÉ. L'interopérabilité d'entreprise est désormais considérée comme une performance clé pour la réalisation de la mission d'un réseau d'entreprises collaboratives. Il faut garantir que l'information est partagée, comprise par rapport à une sémantique commune et utilisée par les différents systèmes d'entreprise pour réaliser une mission globale.

Dans ce contexte, évaluer l'interopérabilité sémantique des modèles de données associés aux systèmes d'information d'entreprises est important. Nous nous intéressons dans ce travail à l'extraction de la sémantique dans un modèle de système d'information. Le but est de faire émerger la sémantique nécessaire et suffisante pour que le système d'information opère correctement. Pour cela nous définissons ce que nous appelons des blocs sémantiques qui permettent de mettre en évidence, à l'intérieur d'un modèle conceptuel, des ensembles indépendants de concepts qui définissent la sémantique minimale nécessaire pour l'opération d'un sous-système du système d'information considéré. Ainsi les blocs sémantiques doivent permettre à terme l'évaluation de la relation d'interopérabilité entre des sous-systèmes de deux systèmes d'information donnés.

ABSTRACT. Enterprise interoperability is considered as a key performance when dealing with the mission satisfaction of a collaborative enterprise network. It is necessary to guarantee that the shared information is understood with regard to a common semantics and is used by the various enterprise systems for reaching a global mission.

In this context, assessing the semantic interoperability of the information systems conceptual models is a major issue. This paper deals with the semantics extraction of any information system conceptual model. The purpose is to bring out the necessary and sufficient semantics needed by the information system for operating correctly. We define the concept of so-called “semantic block” in order to enact the minimal necessary semantics for sub-systems operation. Those semantics blocks are facilitating the assessment of the interoperability relationships between different information systems.

MOTS-CLÉS : Systèmes d'information d'entreprises, extraction de la sémantique, évaluation de l'interopérabilité.

KEYWORDS: Enterprise Information systems, semantics enactment, interoperability assessment.

1. Introduction

De nos jours, les entreprises sont amenées à devenir de plus en plus agiles et flexibles pour pouvoir gérer au mieux les changements rapides de leurs activités dans un marché toujours plus compétitif. Les défis scientifiques auxquels elles doivent faire face concernent notamment l'interopérabilité d'entreprise. En effet, l'interopérabilité d'entreprise doit permettre d'une part de surmonter les barrières organisationnelles et d'autre part d'améliorer les interactions entre les systèmes, le personnel, les applications, les départements et les entreprises (en termes de flux de matière, flux informationnels et flux décisionnels) (Vernadat 2009).

La modélisation d'entreprise joue un rôle critique dans cette intégration, et ceci en capturant les informations ainsi que les connaissances les plus significatives pour assurer la réalisation des processus d'entreprise et de leur organisation (Vernadat 1996) (Panetto et Molina 2008). Les modèles d'entreprise réalisés sont principalement liés aux artefacts de type processus, comportements, activités, informations, ressources, flux d'objets/matières, objectifs, infrastructures et architectures de systèmes. Ces modèles doivent formaliser la sémantique nécessaire et suffisante pour qu'ils soient intelligibles et qu'ils permettent ainsi l'intégration globale de l'entreprise. Par exemple, si nous considérons un modèle de processus, sa sémantique métier peut être modélisée en utilisant des langages de modélisation comme BPMN (Business Process Modelling Notation). De plus, l'enrichissement de cette sémantique est désormais une problématique scientifique faisant intervenir différentes communautés; nous pouvons par exemple citer les recherches menées par (Boudjlida et Panetto 2008) concernant les annotations sémantiques de modèles de processus.

Les Systèmes d'Information (SI) se trouvent au cœur des principaux modèles d'entreprises. Concrètement, la complexité de l'interopérabilité d'entreprise réside dans le fait qu'une entreprise donnée (seule ou en réseau d'entreprises) comprend une multitude de systèmes d'information, implémentés dans diverses applications, hétérogènes de par leur mission, comme par exemple les ERP (progiciel de gestion intégré), les MES (système de pilotage de la fabrication), les SCM (progiciel de gestion de la chaîne logistique), les PDM (progiciel de gestion de données techniques) et les CRM (progiciel de gestion de la relation client) pour ne citer que les plus usuels. Ces systèmes d'information ont besoin i) de partager des informations spécifiques et ii) de traiter des informations selon une sémantique opérationnelle partagée iii) pour réaliser une mission spécifiée dans un contexte donné. La réalisation de ces actions est alors appelée « interopération » (Whitman, Santanu et Panetto 2006).

Assurer l'interopérabilité sémantique entre différents systèmes d'informations dans le contexte des entreprises collaboratives, nécessite ainsi l'analyse et la manipulation de

leurs modèles conceptuels. Un constat général montre que ces modèles conceptuels, dans leur état actuel, sont non exploitables pour évaluer l'interopération. En effet, la sémantique de ces modèles conceptuels, de par leur hétérogénéité (outils, méthodes), est souvent implicite, voire bruitée par des exigences liées à l'implémentation ou non exprimée car elle est dépendante du métier et des expériences de modélisation (Castano et De Antonellis 1998). Extraire la sémantique implicite est ainsi une étape nécessaire dans la mesure où nous cherchons à étudier l'interopérabilité à travers la formalisation des relations sémantiques entre les concepts des systèmes d'informations mis en œuvre.

La prochaine section met en évidence les défis associés à l'extraction de la sémantique dans les modèles de systèmes d'information. L'un de ces défis est l'extraction de la structure de cette sémantique.

Nous proposons dans un premier temps de structurer la sémantique d'un modèle conceptuel pour faciliter la compréhension de la connaissance incorporée dans le cœur des modèles d'entreprise. L'approche adoptée propose tout d'abord d'identifier les concepts « cœur de métier » (core knowledge) puis fait émerger ensuite ce que nous appelons des blocs sémantiques (des agrégats de concepts). Nous avons également développé un algorithme qui permet de construire automatiquement ces blocs sémantiques (Yahia, Aubry et Panetto 2011).

Dans un deuxième temps, nous nous servons d'une étude de cas pour illustrer l'approche proposée précédemment. Cette étude de cas traite des exigences d'interopérabilité entre deux progiciels d'une entreprise à savoir un ERP et un MES, et consiste à appliquer notre approche pour extraire la structure de la sémantique véhiculée par le MES.

Finalement, nous concluons ce papier avec quelques remarques et perspectives concernant nos recherches en cours.

2. Modèles conceptuels de systèmes d'information d'entreprises

Les Systèmes d'Information sont des outils qui participent à la gestion, au traitement, au transport et à la diffusion de l'information au sein d'une organisation ou de plusieurs organisations en réseau. Ils sont au cœur du processus de la création de la valeur d'autant plus qu'ils contribuent à l'atteinte des objectifs de l'organisation. Le contexte collaboratif a fait émerger la nécessité de faire interopérer une multitude de systèmes d'information, autonomes et hétérogènes (architectures, fonctions, modèles...), pour assurer l'objectif désiré de la collaboration.

Ce besoin en interopération entre les différents Systèmes d'information requiert que les informations soient physiquement échangées (l'interopérabilité technique), soient

comprises (l'interopérabilité sémantique) et soient utilisées pour atteindre le but pour lequel elles ont été produites (l'interopérabilité sémantique et organisationnelle).

Il a été constaté par (Castano et De Antonellis 1998) et (Chapurlat 2007) que les systèmes d'information sont pour la plupart définis et conçus par différents experts ayant chacun leur propre expérience de modélisation. Ils font appel, par la suite, à une multitude de modèles d'entreprise, nécessairement hétérogènes de par leur représentation.

L'hétérogénéité des modèles conceptuels d'entreprise se situe alors à deux niveaux : syntaxique et sémantique.

- L'hétérogénéité syntaxique prend place selon (Benhlilima et Chiadmi 2006) au niveau des formats de stockage des données (XML, etc.), dans les langages d'interrogation (XQuery, SQL, OQL, etc.), dans les protocoles d'accès (HTTP, etc.), dans les interfaces, etc.
- L'hétérogénéité sémantique résulte selon (Park et Ram 2004) des différences entre les interprétations du monde réel par les modélisateurs durant le processus de conception des systèmes d'information. Ces différences concernent le contexte et l'utilisation des données et engendrent des conflits sémantiques qui peuvent survenir au niveau des schémas et/ou des données (schema-level, data-level).

(Kavouras 2003) attribue, par exemple, les causes qui peuvent générer des hétérogénéités sémantiques à :

- la différence de couverture (niveau de détail)
- la différence de classification (conceptualisation)

Des standards d'échanges d'informations (STEP, ISO...) ont contribué à la résolution des problèmes liés à l'hétérogénéité syntaxique. Cependant, la question de la résolution des problèmes liés aux conflits sémantiques reste ouverte et difficile de par la nature même de ces conflits. Nous caractérisons, dans la section qui suit, l'ensemble des conflits issus de l'hétérogénéité sémantique de systèmes d'information.

2.1. Les conflits sémantiques des modèles d'entreprises

Les conflits sémantiques rencontrés dans les modèles conceptuels d'entreprises indiquent des problèmes d'interprétation des modèles souvent dus à la difficulté d'explicitation la sémantique implicite.

(Zougggar 2009) distingue deux types de problèmes sémantiques dans le domaine de la modélisation d'entreprise :

- Problèmes sémantiques concernant les langages de modélisation d'entreprise et portant sur l'ensemble des artefacts (syntaxe concrète) qui définissent la sémantique de chacun des langages. Dans ce cadre, par exemple, le concept de « fonction » n'a pas la même sémantique dans les langages fonctionnels définis par CIMOSA¹ que dans IDEF0². En effet, le concept « fonction » dans le contexte de CIMOSA décrit l'enchaînement des opérations de la gestion de production alors que, dans IDEF0, il traite les aspects statiques d'un système.
- Problèmes sémantiques portant sur les informations conceptualisées au niveau d'un modèle d'entreprise spécifique relatif à une entreprise particulière. Dans ce contexte, par exemple, le concept « employé » n'a pas la même sémantique dans toutes les entreprises. En effet, pour une entreprise donnée, un « employé » désigne toute personne travaillant dans cette entreprise. Alors que pour une autre entreprise, un employé est toute personne n'ayant pas le statut cadre.

Nous constatons ainsi que la sémantique véhiculée par les différents modèles est plus ou moins explicite. Dans ce contexte, la difficulté d'opérer avec les divers modèles d'entreprise réside dans le fait que la majorité de ces modèles ont été élaborés par différents experts ayant chacun leur propre expérience de modélisation. Par conséquent, un processus de conceptualisation peut produire plusieurs représentations syntaxiquement différentes mais sémantiquement équivalentes.

De plus, la plupart de ces modèles proviennent d'une retro-ingénierie d'applications existantes, qui incluent alors diverses fonctionnalités liées à l'implémentation puisqu'ils ont été réalisés *a posteriori* et non *a priori* (Lezoche, Panetto et Aubry 2011). Ce processus de modélisation induit alors des interférences néfastes pour la compréhension de la sémantique des modèles produits.

Considérons, par exemple, l'extrait (très succinct) de deux modèles conceptuels différents représenté dans la Figure 1, selon la notation des diagrammes de classes UML. Intuitivement, nous remarquons que ces classes véhiculent une sémantique équivalente bien que celle-ci soit spécifiée différemment dans chacun des modèles. Il s'agit, dans ce simple exemple, d'une modélisation des concepts de produit et de sa propriété de poids, de deux manières syntaxiquement distinctes. Sur la droite de la Figure 1, on met en évidence deux classes associées *PRODUCT* et *WEIGHT*. Ce choix de modélisation peut être justifié par les contraintes d'implémentation visant à faciliter la recherche des valeurs de poids par exemple, ou par la possibilité de partager ce concept avec d'autres entités. Alors que, sur la gauche de la Figure 1, le poids associé au produit est modélisé par deux attributs (valeur et unité) de la classe *PRODUCT*.

¹ Computer Integrated Manufacturing Open System Architecture (<http://www.cimosa.de/>)

² Integration Definition for Function Modeling (<http://www.idef.com/idef0.htm>)

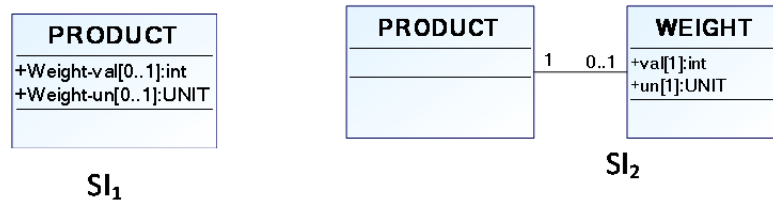


Figure 1. Deux extraits de modèles conceptuels

2.2. Problématique : Extraction de la sémantique

La difficulté de faire interopérer les divers systèmes d'information d'entreprises collaboratives, à travers l'interopération de leurs modèles conceptuels, réside alors dans l'explicitation de leurs sémantiques. En ce sens, (Lezoche, Panetto et Aubry 2011) argumente que *le pré-requis principal pour garantir l'interopérabilité des systèmes d'information est de maximiser la quantité de la sémantique utilisée et de la rendre de plus en plus explicite* (Obrst 2003), ce qui permet, par conséquent, de rendre les systèmes sémantiquement interopérables.

Extraire la sémantique implicite est ainsi une étape nécessaire avant l'étude de l'interopérabilité des systèmes concernés. Cette extraction peut se faire en deux étapes. La première étape consiste à enrichir et expliciter la sémantique implicite des systèmes d'entreprise, dans leur contexte, bruitée par les exigences d'implémentation de ces modèles d'entreprise (cf. section 3). La deuxième étape consiste à structurer cette sémantique pour mettre en évidence la sémantique nécessaire au système d'information étudié pour opérer (cf section 4).

3. Enrichissement et explicitation de la sémantique implicite

La première étape se traduit par l'extraction de chaque « atome » de sémantique incorporée dans l'application d'entreprise étudiée. Dans (Lezoche, Panetto et Aubry 2011), nous avons proposé une méthode (présentée dans la Figure 2) pour réaliser ce travail à travers l'enrichissement des modèles conceptuels par l'intervention d'experts métier et par l'explicitation de la sémantique implicite dans ces modèles. Cette méthode consiste en :

1. la construction d'un premier modèle conceptuel par une rétro-ingénierie du modèle implémenté de l'application d'entreprise (étape 1 de la Figure 2),
2. l'enrichissement de la sémantique de ce modèle conceptuel en injectant les connaissances des experts du domaine. (étape 2 de la Figure 2).

3. l'explicitation de la sémantique implicite en procédant à une transformation « fact-oriented modelling » (Yahia, Lezoche, et al. 2011) et (Yahia, Aubry et Panetto 2011). (étape 3 de la Figure 2).

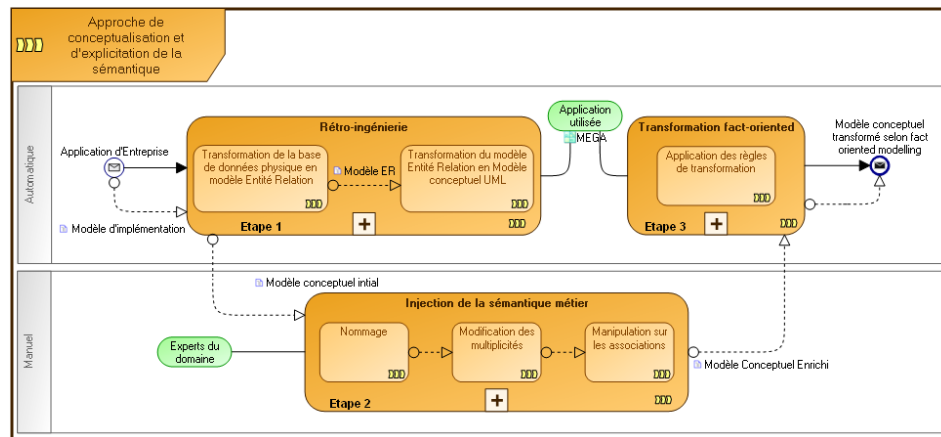


Figure 2. L'approche de conceptualisation et d'explicitation de la sémantique (Lezoche, Panetto et Aubry 2011)

Dans cette section, nous détaillons ces trois étapes, dont le résultat est le point de départ de notre proposition pour structurer la sémantique d'un modèle conceptuel de système d'information en vue de mettre en évidence la sémantique minimale et nécessaire pour opérer.

Etape 1 : rétro-ingénierie

Développer des modèles conceptuels consiste, d'abord, à spécifier les composants essentiels du système étudié ainsi que les relations entre ces composants, et à reconnaître, ensuite, les changements possibles au niveau des composants et leurs relations qui affectent le fonctionnement du système (Engelbart 1962).

De plus, le processus de conceptualisation a été caractérisé par (Guarino 1998) comme étant un processus décisionnel, il fournit d'après (Lezoche, Panetto et Aubry 2011) *une vue, dans laquelle la partie étudiée de la connaissance qui est souvent implicite et complexe, est réorganisée sous forme d'agrégats plus simples à représenter.*

Nous pouvons disposer, en entrée du processus de conceptualisation, d'une application avec un modèle physique déjà implémenté, c'est-à-dire d'un modèle de données. Diverses méthodes de rétro-ingénierie parmi lesquelles celle de (Fonkam et Gray 1992) et (Chiang, Barron et Storey 1994) permettent de retro-dériver un modèle conceptuel à partir de ce type d'entrées.

Ces méthodes sont, pour la plupart, implémentées dans des outils de modélisation comme MEGA Suite³ qui fournit un environnement permettant notamment l'extraction d'un modèle de données à partir d'une application existante vers un modèle relationnel, puis sa rétro-dérivation en un modèle conceptuel.

Etape 2 : injection de la sémantique par l'expert

(Zouggar 2009) considère que le langage de modélisation permet d'assurer une certaine rigueur syntaxique au modèle conceptuel du système d'information réalisé, alors que la sémantique est principalement injectée par les experts du domaine. D'ailleurs, la sémantique des modèles obtenus repose uniquement sur les compétences de ces experts. Injecter les bonnes pratiques d'une entreprise donnée dans les modèles conceptuels obtenus après l'étape 1 permet de les enrichir sémantiquement. Ces bonnes pratiques sont exprimées par les usages des applications, les traitements qui y sont implémentés et les contraintes dictées par les experts du domaine concerné.

Considérons l'exemple illustratif formulé par (Lezoche, Panetto et Aubry 2011). Soit a_1 un attribut particulier, deux cas de figure se présentent :

- a_1 est un attribut non-obligatoire dans le modèle conceptuel, mais les utilisateurs (dans l'entreprise) lui affectent toujours une valeur spécifique. Le modèle enrichi doit donc prendre en considération que cet attribut a_1 doit être traité comme étant obligatoire ;
- a_1 est implémenté comme attribut obligatoire dans le modèle conceptuel, mais en pratique, les utilisateurs ne se soucient guère de renseigner sa valeur concrète et lui attribuent une valeur factice. Dans ce cas, le modèle enrichi peut formaliser que cet attribut n'est pas contraint.

De la même manière, on peut ajouter des associations qui ne sont pas définies au niveau du modèle conceptuel initial mais seulement détectées à travers les bonnes pratiques des utilisateurs ou les traitements internes effectués par l'application. Le modèle conceptuel enrichi doit formaliser ainsi toute la sémantique du système d'information, en prenant en considération les connaissances des parties prenantes et les bonnes pratiques de l'entreprise.

Etape 3 : normation du modèle à travers une transformation « Fact oriented »

Pour éviter que la sémantique des modèles conceptuels étudiés soit dépendante des différentes pratiques de modélisation utilisées, nous proposons de normer ces modèles en faisant abstraction de leurs différences syntaxiques pour se concentrer, uniquement, sur la sémantique qu'ils véhiculent, avec un niveau de granularité le plus fin possible.

³ MEGA International, <http://www.mega.fr/fr/c/product>

Nous proposons ainsi de mettre en application les préceptes définis par l'approche « fact-oriented modelling » (Halpin 2007). En effet, le paradigme de modélisation « fact-oriented modelling » est une approche visant à éviter les représentations conflictuelles comme celles définies dans l'exemple de la Figure 1. Il remet en cause la sémantique implicite sous-jacente de n'importe quel domaine et ceci en verbalisant les faits et les règles (Halpin 2007). Les modèles résultants sont sans attribut et repose sur l'utilisation de faits (présentant des choses réelles ou présentant les caractéristiques de ces choses réelles), ainsi que les relations entre ces faits.

Le langage « NIAM (Natural-language Information Analysis Method) » proposé par (Nijssen et Halpin 1989), et sa variante « ORM (Object-Role Modelling) » sont les langages les plus significatifs basés sur le paradigme de modélisation « fact-oriented modelling ». Les atouts de cette approche sont qu'elle favorise l'expressivité de la sémantique et sa compréhension univoque par une verbalisation naturelle (Halpin 2007). D'ailleurs, cette approche n'utilise pas explicitement les attributs mais fait intervenir, au lieu de cela, des concepts lexicaux (Lexical Object Type - LOT) et non-lexicaux (NON-Lexical Object Type - NOLOT) pour spécifier les modèles produits (Halpin 2006).

(De Bo, Spyns et Meersman 2003) définissent ces concepts comme suit: *a lexical object (LOT), a term, is an object in a certain reality that can be written down. LOTs always consist of letters, numbers, symbols or other characters. They can be used as names for or references to other objects. A non-lexical object (NOLOT), a concept, is an object in a certain reality that cannot be written down. Non-lexical objects must be named by lexical objects or referred to by means of lexical objects.*

Nous proposons ainsi d'utiliser cette approche mais en gardant une notation de type diagramme de classes UML car les modèles conceptuels d'applications d'entreprise existants sont très souvent représentés avec cette notation standardisée. Par ailleurs, et en prenant en compte les définitions de NIAM, nous considéreront les classes et les attributs d'UML comme respectivement des concepts non-lexicaux et des concepts lexicaux.

(Lezoche, Panetto et Aubry 2011) identifient, dans la Figure 3, l'ensemble des règles de transformation d'un diagramme UML (classe, attribut, classe d'association, agrégation, composition, généralisation) en « Fact-Oriented Modelling » (noté FOM).

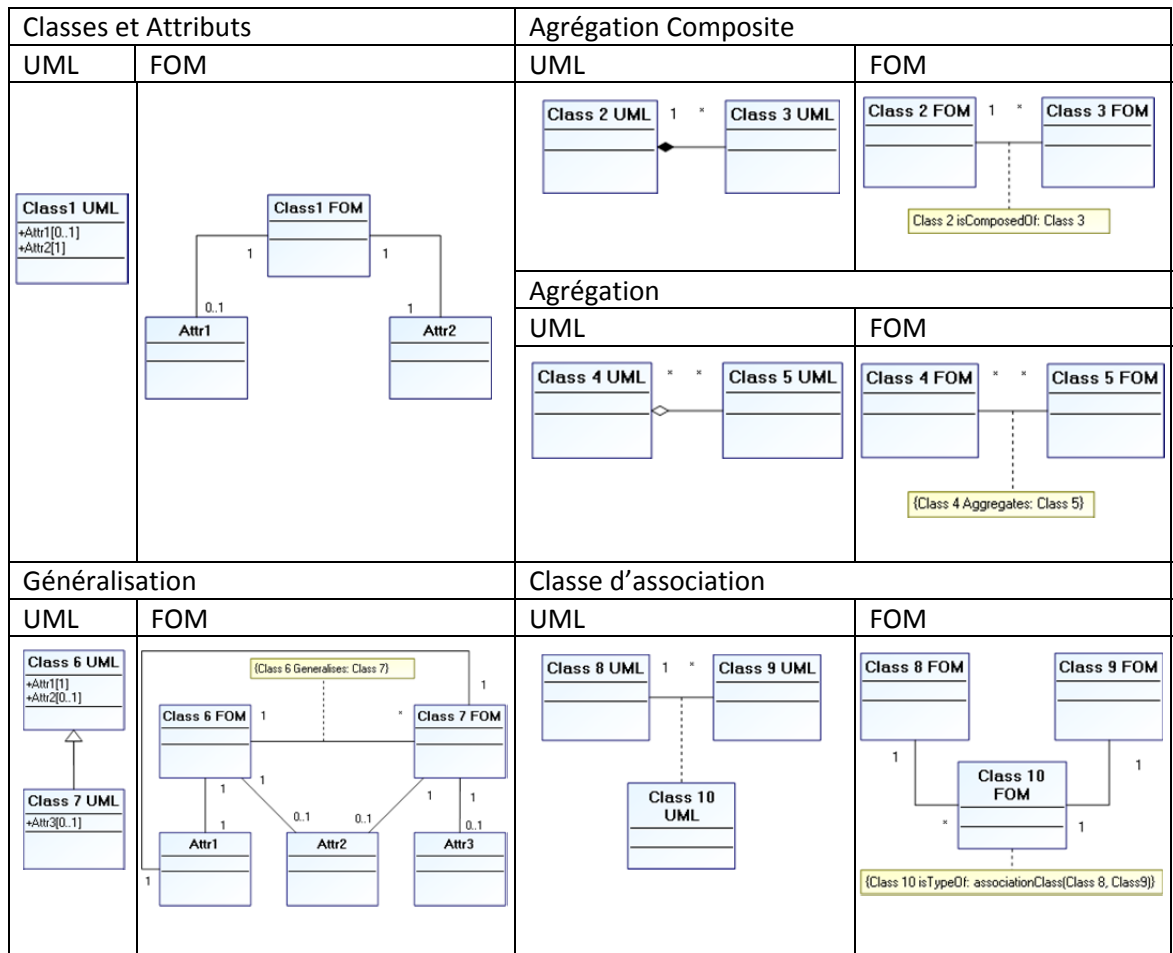


Figure 3. Les règles de transformation utilisant la syntaxe d'UML (Lezoche, Panetto et Aubry 2011)

En appliquant le paradigme « fact-oriented modelling » sur l'exemple de la Figure 1, nous obtenons alors les modèles conceptuels de la Figure 4, caractérisés par leur sémantique explicitée.

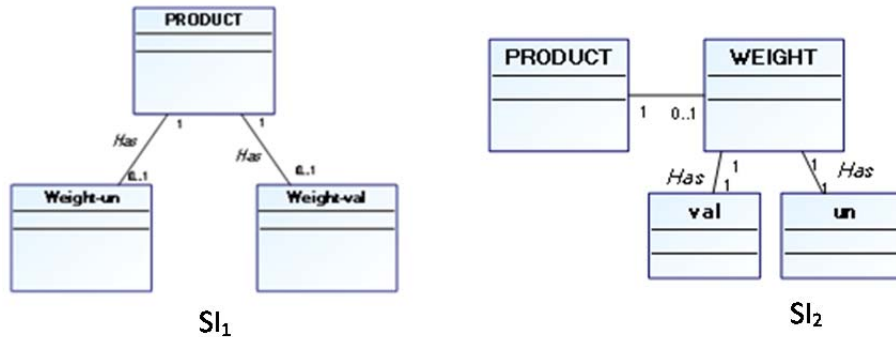


Figure 4. Les modèles conceptuels de la Figure 1 transformés selon le paradigme « fact-oriented modelling »

Dans la suite du papier, nous ferons l'hypothèse que l'application de la méthode présentée dans cette section permet d'obtenir un modèle conceptuel où chaque atome de sémantique du système d'information est représenté sans aucune influence des contraintes liées à l'implémentation. La deuxième étape dans l'extraction de la sémantique consiste alors à mettre en évidence, par l'analyse du modèle conceptuel obtenu, la structure de cette sémantique. C'est l'objet de la section suivante. On propose ainsi de définir ce que nous appelons des « blocs sémantiques ». Chacun de ces agrégats de concepts exprime alors la sémantique minimale et nécessaire liée à un ou plusieurs concepts, mettant ainsi en évidence des sous-systèmes sémantiquement indépendants pour opérer.

4. Structuration de la sémantique

4.1. Définition de la sémantique centrale et étendue

L'étude d'un modèle conceptuel d'un système d'information élaboré selon l'approche « fact-oriented modelling » met en évidence des concepts obligatoires et non-obligatoires qui expriment la sémantique du modèle, et ceci à travers l'identification des associations obligatoires (rôles contraints) et non-obligatoires (rôles non contraints) entre concepts.

Les concepts obligatoires sont ceux qui amènent la sémantique minimale et nécessaire à la cohérence et la compréhension du modèle conceptuel. Ils comprennent tous les concepts lexicaux et non-lexicaux, associés entre eux par des rôles d'association ayant une multiplicité contrainte (m ou $p..*$) avec $m, p \geq 1$. De manière duale, les concepts non-obligatoires correspondent aux concepts ayant uniquement des rôles

d'association avec une multiplicité non-contrainte ($0..n$ ou $*$) avec $n \geq 1$. Ces derniers ne font qu'enrichir la sémantique des modèles conceptuels produits.

La sémantique dite « centrale » d'un modèle correspond alors à la sémantique minimale contenue dans chaque modèle conceptuel d'un Système d'Information (SI) et formalisée par l'ensemble des concepts obligatoires et leurs associations contraintes. Cette sémantique représente l'essence du SI. La sémantique dite « étendue » englobe non seulement la sémantique « centrale » mais aussi tous les concepts et associations qui lui sont faiblement couplés (multiplicité non-contrainte).

4.2. Quelques définitions mathématiques

Nous définissons, pour chaque modèle conceptuel d'un système d'information SI donné, les notations suivantes :

Définition 1. C_{SI} est l'ensemble des concepts lexicaux ou non lexicaux identifiés, il est formellement défini par :

$$C_{SI} = \{c_i | c_i \text{ est un concept lexical ou non lexical du modèle conceptuel de } SI\}$$

Nous définissons également deux sous-ensembles comme suit :

- NLC_{SI} est le sous-ensemble de C_{SI} contenant tous les concepts non lexicaux et,
- LC_{SI} est le sous-ensemble de C_{SI} contenant tous les concepts lexicaux.

Nous noterons que :

$$C_{SI} = NLC_{SI} \cup LC_{SI}$$

$$NLC_{SI} \cap LC_{SI} = \emptyset$$

Définition 2. Rel_{SI} est l'ensemble des relations binaires (rel) identifiées entre les concepts. Il est formellement défini par :

$$Rel_{SI} = \{rel(c_i, c_j) | (c_i, c_j) \in C_{SI} \times C_{SI} \wedge c_i \text{ est associé à } c_j\}$$

Définition 3. $Mult(rel(c_i, c_j))$ représente la multiplicité du rôle de c_j dans l'association entre les deux concepts c_i et c_j .

Pour chaque $(c_i, c_j) \in C_{SI} \times C_{SI}$, si $rel(c_i, c_j)$ existe alors $Mult(rel(c_i, c_j)) \in \{*, 0..n, m, p..*\}$ avec $n, m, p \geq 1$. Ceci se lit c_i est associé à c_j dans la relation $rel(c_i, c_j)$ avec une multiplicité égale à $Mult(rel(c_i, c_j))$.

Définition 4. MC_{SI} est le sous-ensemble de C_{SI} contenant tous les concepts obligatoires (la sémantique centrale). Il est formellement défini par :

$$MC_{SI} = \{c_i \mid \exists (c_j, rel(c_i, c_j)) \in C_{SI} \times Rel_{SI} \wedge Mult(rel(c_i, c_j)) \in \{m, p, .*\} ; m, p \geq 1\}$$

Nous définissons également deux sous-ensembles comme suit :

- $MNLC_{SI}$ est le sous-ensemble de C_{SI} contenant tous les concepts non lexicaux obligatoires et,
- MLC_{SI} est le sous-ensemble de C_{SI} contenant tous les concepts lexicaux obligatoires.

Nous noterons que :

$$MC_{SI} = MNLC_{SI} \cup MLC_{SI}$$

$$MNLC_{SI} \cap MLC_{SI} = \emptyset$$

Définition 5 NMC_{SI} est le sous-ensemble de C_{SI} contenant tous les concepts non obligatoires (la sémantique centrale). Il est formellement défini par :

$$NMC_{SI} = \{c_i \mid \exists (c_j, rel(c_i, c_j)) \in C_{SI} \times Rel_{SI} \wedge Mult(rel(c_i, c_j)) \in \{*, 0, .n\} ; n \geq 1\}$$

Nous définissons également deux sous-ensembles comme suit :

- $NMNLC_{SI}$ est le sous-ensemble de C_{SI} contenant tous les concepts non lexicaux non obligatoires et,
- $NMLC_{SI}$ est le sous-ensemble de C_{SI} contenant tous les concepts lexicaux non obligatoires.

Nous noterons que :

$$NMC_{SI} = NMNLC_{SI} \cup NMLC_{SI}$$

$$NMNLC_{SI} \cap NMLC_{SI} = \emptyset$$

Définition 6. Pour chaque concept non lexical c_j , nous définissons l'ensemble des concepts lexicaux et obligatoires qui lui sont associés par :

$$MLC(c_j) = \{c_i \in LC_{SI} \mid (\exists rel(c_j, c_i) \in Rel_{SI} \mid Mult(rel(c_i, c_j)) \in \{m, p, .*\} ; m, p \geq 1)\}$$

Définition 7. Pour chaque concept non lexical c_j , nous définissons l'ensemble des concepts non lexicaux et obligatoires qui lui sont associés par :

$$MNLC(c_j) = \{c_i \in NLC_{SI} \mid (\exists rel(c_j, c_i) \in Rel_{SI} \mid Mult(rel(c_i, c_j)) \in \{m, p, .*\} ; m, p \geq 1)\}$$

4.3. Identification des blocs sémantiques

4.3.1. Objectifs

(Evans 2006) caractérise les différents concepts par leurs propres **potentiels sémantiques** et postule que *la sémantique des concepts est souvent liée au contexte et est déterminée à travers la connaissance « potentielle » propre à chaque concept*⁴.

Ce potentiel sémantique, propre à chacun des concepts, met en évidence une structure sémantique spécifique de la connaissance.

De plus, si nous considérons la distinction entre les concepts lexicaux et non-lexicaux, nous remarquons que les concepts non-lexicaux possèdent une valeur sémantique propre, tandis que les concepts lexicaux fournissent principalement une caractérisation informationnelle au concept non-lexical auquel ils sont associés. C'est pour cette raison que nous déterminons le potentiel sémantique pour chaque concept non-lexical à travers la structure du modèle conceptuel. De plus, ces concepts non-lexicaux pourraient être incorporés au niveau de plus grandes structures appelées structure composite de concepts non lexicaux (Evans 2006).

Cette structure composite définit la sémantique minimale et nécessaire pour un concept non lexical c_i , formulée à travers :

- La sémantique intrinsèque centrale définie par la sémantique de l'ensemble $MLC(c_i)$ (Définition 6), et,
- La sémantique extrinsèque centrale définie par la sémantique de l'ensemble $MNLC(c_i)$ (Définition 7)

Nous assignons à la structure composite d'un concept non lexical (obligatoire ou non obligatoire) donné c_i , la notion de Bloc Sémantique, dénoté $SB(c_i)$ et définissant la sémantique minimale et nécessaire de c_i . Formellement, $SB(c_i)$ est défini comme suit :

$$SB(c_i) = \{c_i\} \cup MLC(c_i) \bigcup_{c_j \in MNLC(c_i)} SB(c_j)$$

À travers cette définition, nous notons que la notion de bloc sémantique est récursive. En ce sens, un bloc sémantique fournit, par exemple, toute la sémantique obligatoire d'un sous-système du système d'information. Nous présentons dans les sections qui suivent, le méta-modèle du bloc sémantique ainsi que la procédure de construction du bloc sémantique illustrée par un exemple explicatif.

⁴«The meaning is always contextually determined, and is selected among the knowledge 'potential' that a concept provides access to».

4.3.2. Construction des blocs sémantiques

Considérons le modèle conceptuel de la Figure 5 et sa transformation selon l'approche « fact-oriented modelling » (Figure 6).

Nous cherchons à construire le bloc sémantique du concept *C2*.

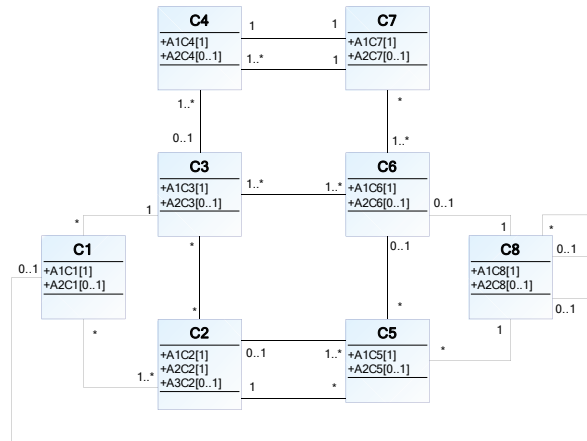


Figure 5. Un modèle conceptuel

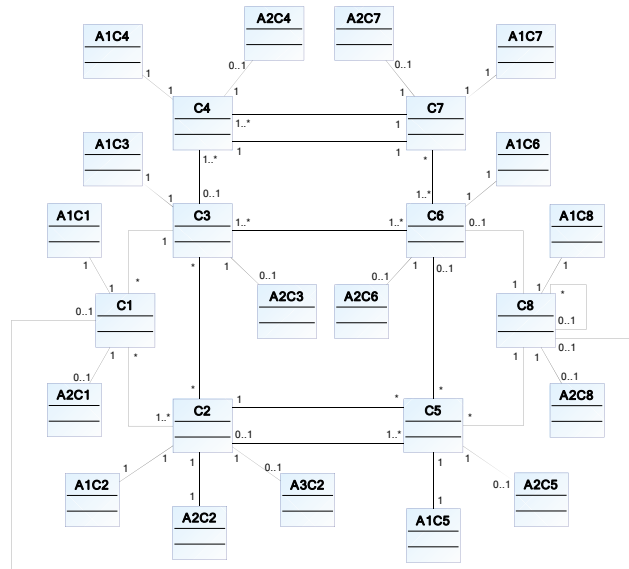


Figure 6. Le modèle conceptuel de la Figure 5 transformé selon le paradigme « fact-oriented modelling »

La sémantique intrinsèque centrale de $C2$ est définie par la sémantique des concepts lexicaux obligatoires qui lui sont associés $A1C2$ et $A2C2$. De plus, une instance donnée du concept non lexical $C2$ existe si seulement si elle est associée à au moins une instance du concept non lexical $C5$. Cela signifie que $C5$ est obligatoire pour exprimer la sémantique de $C2$. Par ailleurs, étant donné que les rôles de $C1$ et $C3$ dans leur association avec $C2$ ont une multiplicité minimale égale à 0, l'existence de n'importe quelle instance de $C2$ n'est pas conditionnée par l'existence d'une instance de $C1$ ou $C3$. Finalement, $SB(C2) = \{C2\} \cup \{A1C2, A2C2\} \cup SB(C5)$. (Définition de la section 4.3.1)

De manière itérative, nous pouvons montrer que la sémantique intrinsèque centrale de $C5$ est définie par la sémantique de $A1C5$ et une instance donnée de $C5$ existe si seulement si elle est associée à exactement une instance de $C8$ et exactement une instance de $C2$. Donc $SB(C5) = \{C5\} \cup \{A1C5\} \cup SB(C2) \cup SB(C8)$.

En procédant au même raisonnement, nous pouvons démontrer que $SB(C8) = \{C8\} \cup \{A1C8\}$. Ce qui implique que $SB(C2) = \{C2, C5, C8\} \cup \{A1C2, A2C2\} \cup \{A1C5\} \cup \{A1C8\}$.

Si on considère un concept non lexical c_i , son Bloc Sémantique va effectivement contenir tous les concepts non-lexicaux qui doivent être instanciés pour assurer l'existence d'une de ces instances, en plus des concepts lexicaux obligatoires qui lui sont associés.

Pour simplifier l'identification du Bloc Sémantique d'un concept particulier c_i , nous proposons de procéder, en premier lieu, à la détermination de l'ensemble des concepts non-lexicaux qui lui sont associés. Les concepts lexicaux obligatoires sont ajoutés, en deuxième lieu. Ce qui implique que l'identification du bloc sémantique du concept c_i consiste à déterminer $SB(c_i) = SB_c(c_i) \cup SB_a(c_i)$ tel que :

$$SB_c(c_i) = \{c_i\} \cup_{c_j \in MNL(c_i)} SB_c(c_j) \text{ et } SB_a(c_i) = \{MLC(c_j) \mid c_j \in SB_c(c_i)\}$$

Par exemple, $SB_c(C2) = \{C2, C5, C8\}$ et $SB_a(C2) = \{A1C2, A2C2, A1C5, A1C8\}$.

4.3.3. Le graphe des dépendances sémantiques

Pour faciliter la construction des blocs sémantiques, nous proposons, pour chaque $c_i \in NLC_{SI}$, d'identifier l'ensemble associé $SB_c(c_i)$ en utilisant la théorie des graphes et ses outils mathématiques.

Nous proposons tout d'abord de définir ce que nous appelons un « graphe des dépendances sémantiques » associé à un modèle conceptuel donné. Ce graphe des dépendances sémantiques est un digraphe $G = (V, E)$ où V est l'ensemble des sommets et E est l'ensemble des paires ordonnées de sommets, appelées aussi des arcs. Chaque

sommet de V représente un concept non-lexical du modèle conceptuel. Chaque arc de E est construit au niveau du modèle conceptuel comme suit : l'arc (c_i, c_j) existe si (i) il y a une association entre c_i et c_j dans le modèle conceptuel et (ii) tout en considérant l'association entre c_i et c_j , la multiplicité minimale du rôle de c_j est supérieure ou égale à 1 ($c_j \in MNL C(c_i)$). Cela signifie que l'arc (c_i, c_j) tel qu'il est défini représente le fait que c_j est obligatoire pour exprimer la sémantique de c_i . La Figure 7 montre le graphe des dépendances sémantiques associé au modèle conceptuel de la Figure 6.

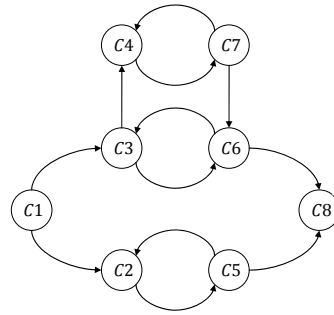


Figure 7. Le graphe des dépendances sémantiques associé au modèle conceptuel de la Figure 6

4.3.4. Quelques propriétés

Théorème 1. Étant donné deux concepts c_i et c_j , c_j appartient à $SB_c(c_i)$ si et seulement s'il existe un chemin de c_i à c_j .

Preuve. Pour construire l'ensemble $SB_c(c_i)$ associé au concept c_i , nous considérons, tout d'abord, ce concept comme étant le point de départ et la racine du graphe des relations sémantiques. Dans un premier temps, nous pouvons ajouter dans $SB_c(c_i)$ tous les concepts c_{1k} qui doivent être instanciés pour assurer l'existence de n'importe quelle instance de c_i , c'est-à-dire tous les concepts c_{1k} tel qu'il y a une association entre c_i et c_{1k} dans le modèle conceptuel et tel que la multiplicité minimale du rôle de c_{1k} est supérieure ou égale à 1. C'est la définition exacte de tous les successeurs de c_i dans le graphe des dépendances sémantiques. Notons aussi que, par définition, il existe un chemin du concept c_i aux concepts c_{1k} . Ensuite par itération, les nouveaux concepts c_{2k} qui peuvent être rajoutés à $SB_c(c_i)$, sont les successeurs des premiers concepts c_{1k} . Pour les successeurs des concepts c_{1k} , il existe là aussi un chemin du concept c_i aux concepts c_{2k} (le chemin de c_i à c_{1k} plus l'arc (c_{1k}, c_{2k})). Finalement l'ensemble $SB_c(c_i)$ associé à c_i contient exactement tous les concepts c_j tel qu'il existe un chemin de c_i à c_j . ■

Théorème 2. Étant donné deux concepts c_1 et c_2 , si c_2 appartient à $SB_c(c_1)$ alors $SB_c(c_2)$ est inclus dans $SB_c(c_1)$.

Preuve. Si c_2 appartient à $SB_c(c_1)$ alors ceci signifie qu'il existe un chemin de c_1 à c_2 (voir théorème 1). Considérons maintenant un concept non lexical particulier de $SB_c(c_2)$ noté c . Par définition de $SB_c(c_2)$, il existe un chemin de c_2 à c donc, il existe un chemin de c_1 à c (le chemin de c_1 à c_2 plus le chemin de c_2 à c). Cela signifie que c est aussi dans $SB_c(c_1)$. Finalement $SB_c(c_2) \subseteq SB_c(c_1)$. ■

Théorème 3. Tous les concepts qui sont dans le même cycle du graphe des dépendances sémantiques sont associés au même et unique ensemble de type SB_c .

Preuve. Un cycle est un chemin fermé. Considérons deux concepts particuliers, notés c_i et c_j , et appartenant à un cycle donné. Il existe alors un chemin de c_i à c_j . Cela signifie que c_j est contenu dans $SB_c(c_i)$. D'après le théorème 2, nous pouvons aussi démontrer que $SB_c(c_j) \subseteq SB_c(c_i)$. De plus, il y a un chemin de c_j à c_i . Cela signifie que c_i est contenu dans $SB_c(c_j)$. D'après le théorème 2, ceci signifie que $SB_c(c_j) \supseteq SB_c(c_i)$. Finalement, nous trouvons que $SB_c(c_j) = SB_c(c_i)$. ■

Pour chaque cycle du graphe des dépendances sémantiques, le théorème 3 implique qu'il existe un seul ensemble $SB_c(c_i)$ associé à tous les concepts non lexicaux qui sont dans le même cycle. Ces concepts appartiennent alors à des composantes fortement connexes⁵ du graphe des relations sémantiques. Ainsi il y a un seul ensemble $SB_c(c_i)$ par composante fortement connexe du graphe des relations sémantiques.

4.3.5. Algorithme de construction des blocs sémantiques

L'application des théorèmes 1, 2 et 3, nous permet d'établir une procédure automatisable d'élaboration de tous les Blocs Sémantiques d'un modèle conceptuel donné :

- Élaboration du graphe des dépendances sémantiques
- Élaboration du graphe des composantes fortement connexes en se basant sur le graphe des dépendances sémantiques
- Élaboration de l'ensemble SB_c associé à chacune des composantes fortement connexes
- Élaboration du Bloc Sémantique SB_a en rajoutant tous les concepts lexicaux obligatoires associés à tous les concepts non lexicaux de SB_c .
- Élaboration du Bloc Sémantique $SB = SB_c \cup SB_a$

Ces étapes sont détaillées comme suit.

⁵ En théorie des graphes, une composante fortement connexe d'un graphe orienté G est un sous-graphe maximal de G tel que pour toute paire de sommets u et v dans ce sous-graphe, il existe un chemin de u à v et un chemin de v à u (Aho, Hopcroft et Ullman 1974).

1) *Élaboration du graphe des dépendances sémantiques*

Ce graphe peut être facilement obtenu en considérant chaque association entre les concepts c_i et c_j et en construisant, par la suite, les arcs de c_i à c_j si la multiplicité minimale du rôle de c_j dans cette association est supérieure ou égale à 1.

2) *Élaboration du graphe des composantes fortement connexes*

Le théorème 3 implique, que pour construire les Blocs Sémantiques, nous pouvons considérer un seul concept de chaque composante fortement connexe (étant donné que les autres concepts de la même composante fortement connexe ont le même Bloc Sémantique). Ainsi nous pouvons simplifier le graphe des dépendances sémantiques en ne préservant que les composantes fortement connexes. Dans ce graphe, il existe un arc d'une composante fortement connexe $SCC1$ à une seconde composante fortement connexe $SCC2$, s'il existe un arc allant d'un concept de $SCC1$ à un autre concept de $SCC2$.

L'identification de toutes les composantes fortement connexes d'un graphe est un problème facile qui peut être résolu avec l'algorithme polynômial de Kosaraju-Sharir présenté dans (Sharir 1981). On donne, dans la Figure 8, le graphe des composantes fortement connexes associé au graphe des dépendances sémantiques de la Figure 7. Sur ce graphe, les composantes fortement connexes sont définies comme suit : $SCC1 = \{C1\}$, $SCC2 = \{C2, C5\}$, $SCC3 = \{C3, C4, C6, C7\}$ et $SCC4 = \{C8\}$.

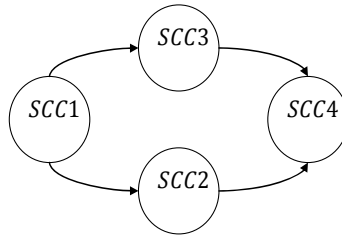


Figure 8. Le graphe des composantes fortement connexes associé au graphe des dépendances sémantiques de la Figure 7

3) *Élaboration de l'ensemble SB_c associé à chacune des composantes fortement connexes*

L'application d'un algorithme que nous proposons dans (Yahia, 2011) nous permet de déterminer, automatiquement, l'ensemble des Blocs Sémantiques SB_c associés à chacune des composantes fortement connexes sont définis comme suit :

$$- \quad SB_c(SCC1) = SCC1 \cup SCC2 \cup SCC3 \cup SCC4,$$

- $SB_c(SCC2) = SCC2 \cup SCC4$,
- $SB_c(SCC3) = SCC3 \cup SCC4$ et,
- $SB_c(SCC4) = SCC4$.

Finalement, en remplaçant les composantes fortement connexes par leur contenu, nous obtenons les blocs sémantiques suivants :

- $SB_c(C1) = \{C1, C2, C3, C4, C5, C6, C7, C8\}$,
- $SB_c(C2, C5) = \{C2, C5, C8\}$,
- $SB_c(C3, C4, C6, C7) = \{C3, C4, C6, C7, C8\}$ et
- $SB_c(C8) = \{C8\}$.

4) *Élaboration du Bloc Sémantique SB_a associé à chacune des composantes fortement connexes*

Le Bloc Sémantique SB_a contient l'ensemble des concepts lexicaux obligatoires associés aux concepts contenus dans SB_c . En appliquant cette définition sur l'exemple de la Figure 6, nous obtenons :

- $SB_a(C1) = \{A1C1, A1C2, A2C2, A1C3, A1C4, A1C5, A1C6, A1C7, A1C8\}$,
- $SB_a(C2, C5) = \{A1C2, A2C2, A1C5, A1C8\}$,
- $SB_a(C3, C4, C6, C7) = \{A1C3, A1C4, A1C6, A1C7, A1C8\}$ and
- $SB_a(C8) = \{A1C8\}$.

5) *Élaboration du Bloc Sémantique SB*

Le Bloc Sémantique SB agrège les éléments de SB_a et SB_c . Les Blocs sémantiques du modèle conceptuel de la Figure 6 sont alors définis comme suit :

- $SB(C1) = \{C1, C2, C3, C4, C5, C6, C7, C8\} \cup \{A1C1, A1C2, A2C2, A1C3, A1C4, A1C5, A1C6, A1C7, A1C8\}$
- $SB(C2, C5) = \{C2, C5, C8\} \cup \{A1C2, A2C2, A1C5, A1C8\}$,
- $SB(C3, C4, C6, C7) = \{C3, C4, C6, C7, C8\} \cup \{A1C3, A1C4, A1C6, A1C7, A1C8\}$ et,
- $SB(C8) = \{C8\} \cup \{A1C8\}$.

4.3.6. *Méta-modèle de l'architecture du bloc sémantique*

Dans cette section, nous proposons de formaliser l'architecture du bloc sémantique à travers le méta-modèle représenté sur la Figure 9. Ce méta-modèle est créé sur la base

du pattern composite (Gamma, et al. 1995). L'objectif de ce méta-modèle est d'intégrer tous ses constituants au niveau d'une arborescence permettant de présenter une ou plusieurs hiérarchies d'objets.

Un bloc sémantique définit la sémantique nécessaire et minimale d'un ou plusieurs concepts non-lexicaux tout en respectant la contrainte qui impose que ces concepts non lexicaux appartiennent forcément à la même composante fortement connexe.

De plus, la sémantique d'un ou plusieurs « concepts », qu'ils soient des concepts non-lexicaux ou des concepts lexicaux, est agrégée dans un ou plusieurs blocs sémantiques.

Cette contrainte traduit la définition formelle que nous avons donnée pour les blocs sémantiques $(SB(c_i) = \{c_i\} \cup MLC(c_i) \cup_{c_j \in MNL(c_i)} SB(c_j))$.

Un Bloc Sémantique est aussi une spécialisation du concept abstrait « Concept ». Il résulte qu'un Bloc Sémantique, étant lui-même un « Concept », subsume la sémantique d'un ou plusieurs « concept » de niveau inférieur. Le Bloc Système (BS) représente le dernier niveau issu de l'agrégation de l'ensemble des Blocs Sémantiques.

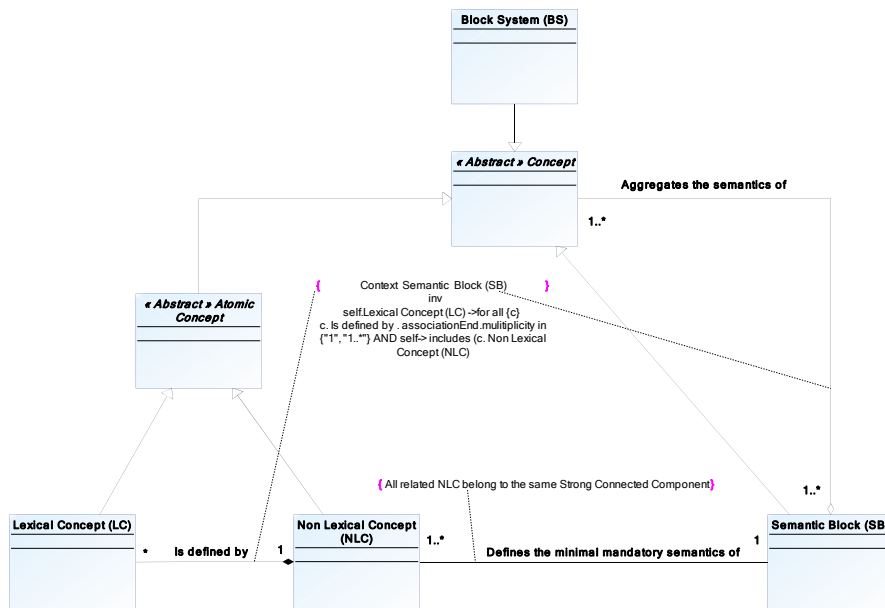


Figure 9. Méta-modèle de l'architecture du bloc sémantique

4.3.7. Élaboration automatique des blocs sémantiques

Dans cette partie, nous présentons l'outil permettant d'automatiser l'élaboration de l'ensemble des Blocs Sémantiques pour n'importe quel modèle conceptuel d'un SI. En effet, les procédures présentées au niveau de la section précédente ont été mise en œuvre à l'aide de l'outil *MEGA Suite*⁶. Cet outil est organisé autour d'un référentiel commun, permettant de conduire des projets variés tels que l'analyse des processus, la cartographie applicative et la conception des systèmes.

Par ailleurs, *MEGA Suite* supporte les spécifications d'UML, ce qui permet de représenter les modèles conceptuels étudiés. De plus, *MEGA Suite* offre un outillage permettant d'implémenter l'algorithme d'élaboration des blocs sémantiques par la programmation de scripts ayant directement accès aux concepts modélisés par l'intermédiaire de leur meta-modèle. Le bloc sémantique est alors conceptualisé comme un paquetage UML tandis que les concepts lexicaux et non-lexicaux sont conceptualisés comme des classes UML.

La Figure 10 présente un extrait du résultat de l'application de l'algorithme sur le modèle conceptuel de la Figure 6.

Cette figure montre le Bloc Sémantique $SB_c(C5, C2)$ (représenté comme un paquetage UML) et associé à $C2$ et $C5$ et incluant le concept $C8$. La Figure 11 montre le sous-modèle formalisant la sémantique entière de $SB(C5, C2)$. Ce sous-modèle est extrait du modèle conceptuel complet de la Figure 6, obtenu en conservant, uniquement, les concepts contenus dans $SB(C5, C2)$ ainsi que leurs associations.

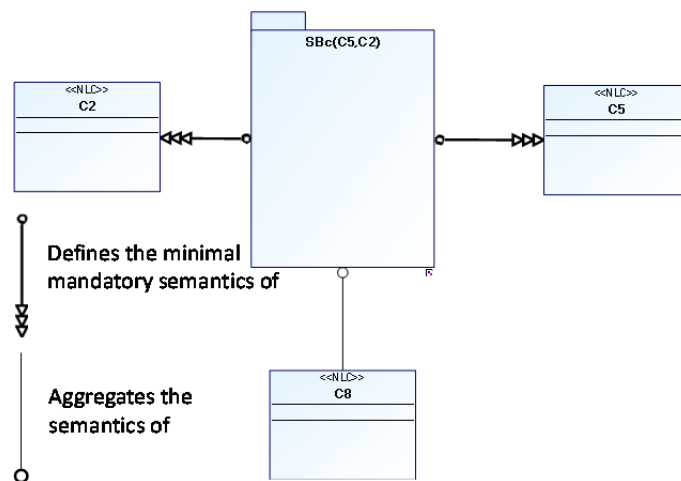


Figure 10. Le Bloc Sémantique $SB_c(C5, C2)$ du modèle conceptuel de la Figure 6

⁶ MEGA International, <http://www.mega.fr/fr/c/product>

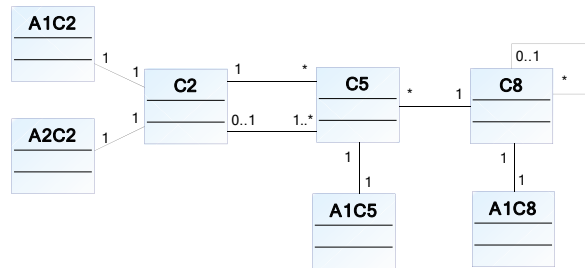


Figure 11. Extrait du modèle conceptuel formalisant la sémantique de SB(C5, C2)

5. Étude de cas : Processus d'achat de la matière première à l'AIPL-PRIMECA

5.1. Présentation de l'étude de cas

Pour illustrer l'approche d'explicitation de la sémantique des modèles conceptuels d'un système d'information donné, dans un contexte industriel, nous avons choisi une étude de cas qui met en évidence deux systèmes d'information différents qui doivent interopérer pour assurer le processus d'approvisionnement en matière première à l'Atelier Inter-établissements de Productique Lorrain (AIP-PRIMECA Lorraine). Il s'agit de l'ERP Sage X3⁷ qui coopère avec le MES Flexnet Apriso⁸ pour assurer la fabrication d'une famille de produits.

De tels systèmes d'information industriels sont architecturés autour d'un système de gestion de base de données relationnelle. Le modèle logique, extrait de ces applications par rétro-ingénierie, fait émerger une multitude de tables et de relations. Flexnet, par exemple, implémente environ 800 tables et 300 relations, qui, après rétro-dérivation, donne un modèle conceptuel brut comportant environ 600 concepts et 500 associations. Sage X3 assure la planification à long terme en spécifiant, pendant une période donnée, les besoins en termes de matières premières, pour ensuite lancer les bons de commande associés. Ces derniers sont transférés de l'ERP au MES pour exécution par le service achat. Un rapport contenant l'état des stocks ainsi que le statut des ordres d'achat est ensuite envoyé à l'ERP par le MES. Pour la suite de notre étude, nous considérons uniquement les sous-systèmes du MES Flexnet (SI_1) et de l'ERP Sage X3 (SI_2) concernés par l'approvisionnement en matière première. Cependant, pour faciliter la lecture, nous continuerons à utiliser les termes Flexnet et Sage X3 pour parler de ces sous-systèmes.

⁷ <http://www.sage.fr/mge/logiciels-erp>

⁸ <http://www.aprison.com/products/overview.php>

5.2. Explicitation et structuration de la sémantique de Flexnet et Sage X3

L'approche d'explicitation de la sémantique des systèmes d'information présentée dans la section 3 permet d'obtenir un modèle conceptuel normé et enrichi. Ce modèle est par la suite analysé pour permettre la structuration de la sémantique sous formes de blocs sémantiques grâce à l'application des algorithmes présentés dans la section 4 et implémentés dans *MEGA Suite*.

5.2.1. Explicitation de la sémantique

Nous présentons dans cette section du papier, l'application de notre approche pour la mise en évidence des modèles conceptuels normés et enrichis relatifs à l'ERP Sage X3 et au MES Flexnet.

Une première étape semi-automatique de rétro-ingénierie (dans *MEGA Suite*) a permis d'obtenir les deux modèles conceptuels. La Figure 12 représente le modèle conceptuel du processus de l'émission du bon de commande déclenché au niveau de Figure 13 conceptuel correspondant de Sage X3.

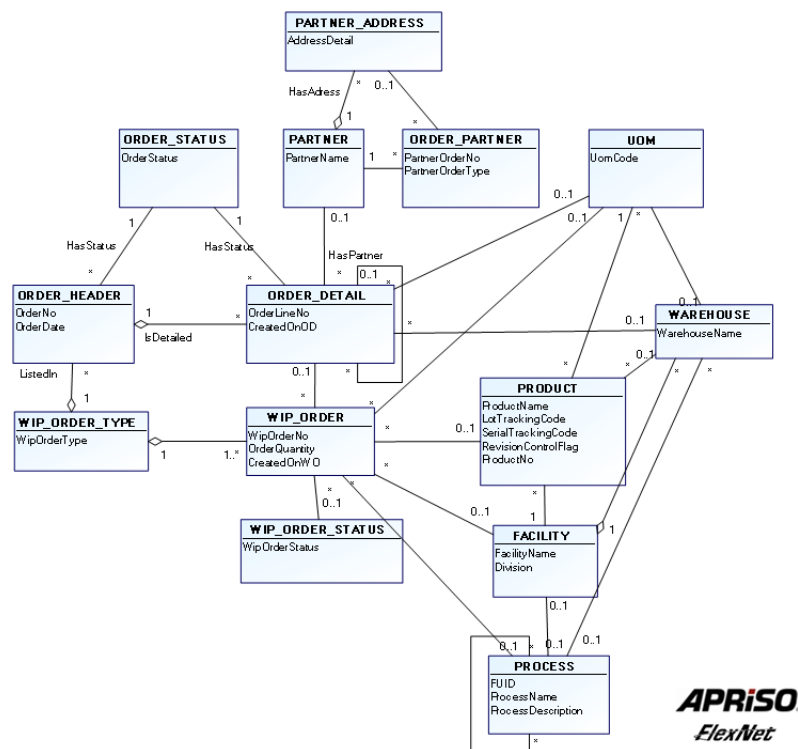


Figure 12. Modèle conceptuel de Flexnet pour l'enregistrement, l'exécution et le suivi du bon de commande

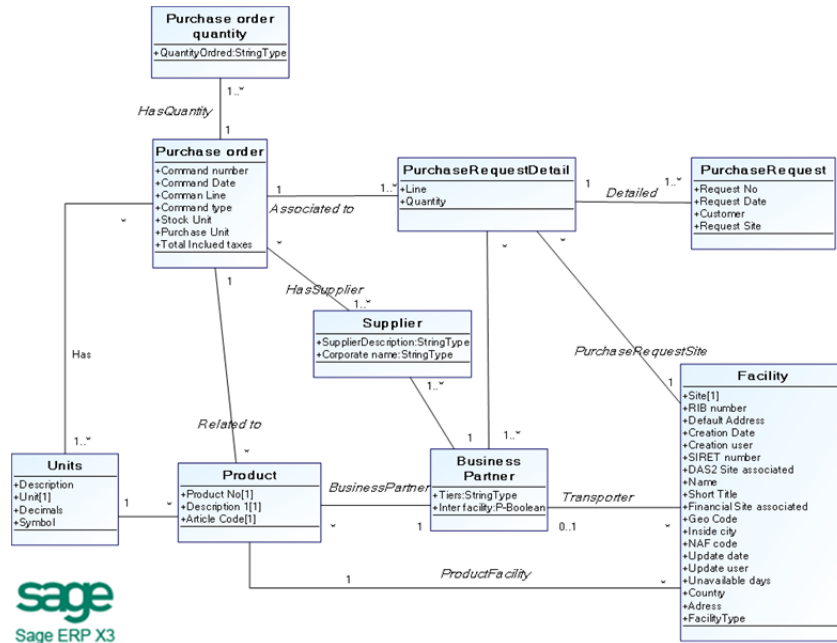


Figure 13. Modèle conceptuel de Sage X3 pour l'élaboration du bon de commande

Ces modèles sont ensuite complétés (en particulier par des contraintes) en y injectant la sémantique apportée par les parties prenantes (utilisateurs et experts) à travers les bonnes pratiques de l'entreprise ainsi que l'analyse des traitements effectués par les applications support.

Les modèles conceptuels enrichis de Flexnet et Sage X3 sont ensuite transformés selon la perspective « fact-oriented modelling » tout en utilisant les notations UML. L'approche de normation des modèles conceptuels permet de produire des modèles conceptuels enrichis et normés ayant une sémantique bien explicitée. Le modèle conceptuel enrichi et normé de Flexnet est donné dans la Figure 14 et celui de Sage X3 est donné dans la Figure 15. La génération automatique de ces modèles est élaborée avec l'environnement *MEGA Suite* par l'exécution des règles de transformation proposées dans la section 3.

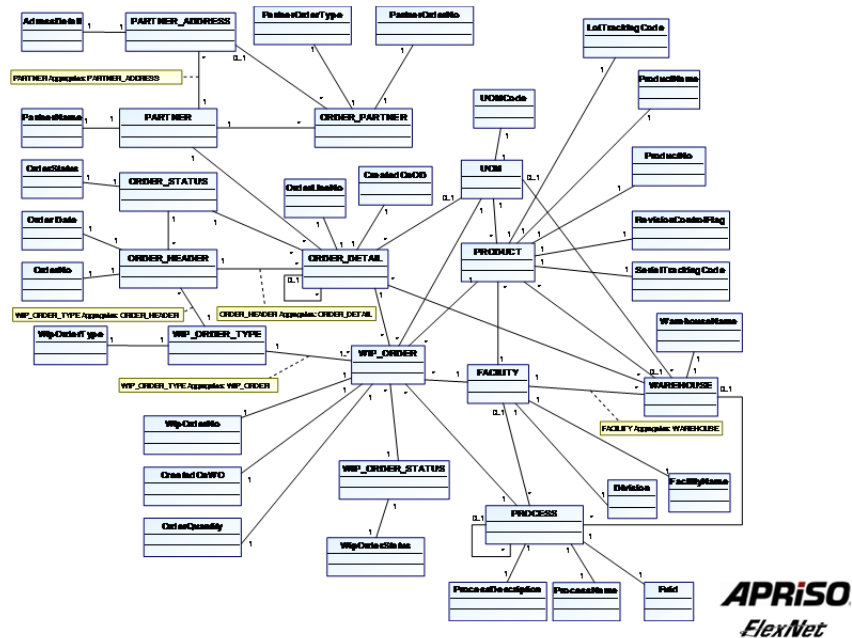


Figure 14. Modèle conceptuel de Flexnet enrichi et normé selon la perspective « fact-oriented modelling »

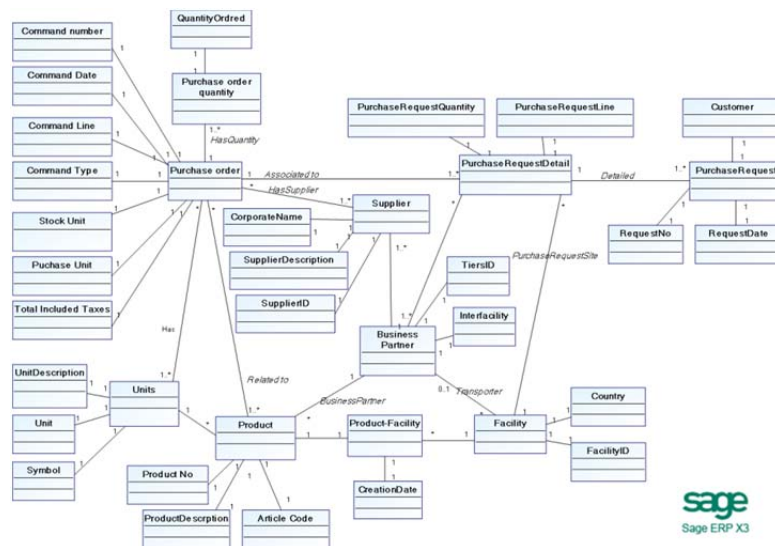


Figure 15. Modèle conceptuel de SAGE X3 enrichi et normé selon la perspective « fact-oriented modelling »

5.2.2. Structuration de la sémantique

Pour structurer la sémantique des modèles conceptuels de l'ERP et du MES sous forme de blocs sémantiques, nous appliquons la procédure présentée dans la section 4.3.5. Cette procédure a été implémentée dans *MEGA Suite* à l'aide de scripts VB, et nous générons alors automatiquement l'ensemble des blocs sémantiques correspondants à chacun des modèles conceptuels enrichis et normés.

Par ailleurs, un diagramme de paquetage UML, généré automatiquement dans *MEGA Suite*, permet de représenter l'ensemble des blocs sémantiques du modèle conceptuel du MES Flexnet (Figure 16). Ce diagramme permet de mettre en évidence les relations entre blocs sémantiques. De plus, un diagramme de classe UML est également généré automatiquement pour chaque bloc sémantique : il représente la restriction du modèle conceptuel complet aux concepts contenus dans le bloc sémantique. Par exemple, nous représentons dans la Figure 17, le diagramme de classes associé au bloc sémantique *SB(PRODUCT)* du concept *PRODUCT*. Ce modèle inclut ainsi tous les concepts exigés pour exprimer la sémantique minimale obligatoire du concept *PRODUCT*.

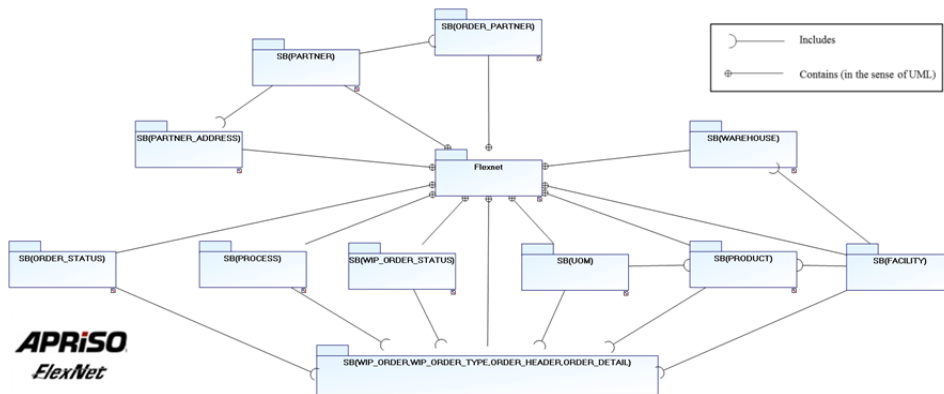


Figure 16. Les blocs sémantiques identifiés dans le modèle conceptuel de Flexnet de la Figure 14.

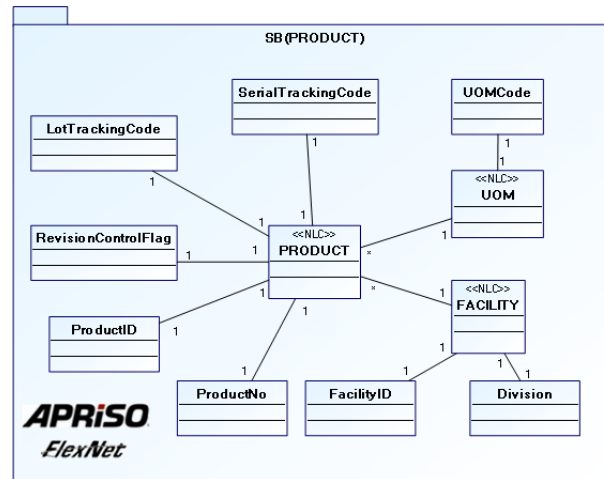


Figure 17. Le modèle conceptuel associé au bloc sémantique $SB(PRODUCT)$

L'ensemble des blocs sémantiques liés respectivement à Flexnet et Sage X3 sont identifiés automatiquement par l'exécution de l'algorithme implémenté dans *Mega Suite*.

5.3. Discussion

L'identification des blocs sémantiques respectifs à chacun des systèmes d'information étudiés permet de mettre en évidence des sous-systèmes indépendants relativement à leur sémantique minimale et nécessaire pour que les applications puissent fonctionner correctement. Ces informations sont très utiles pour l'analyse a priori de l'interopérabilité (capacité à interopérer) entre les applications en présence.

Considérons par exemple le sous-système associé au bloc sémantique $SB_7^1(PRODUCT)$ de Flexnet. Ce sous-système, sémantiquement autonome, implique que le processus de création d'un produit dans Flexnet nécessite, au minimum, de spécifier toutes les informations contenues dans ce bloc sémantique pour pouvoir opérer. Notre analyse a ainsi montré que la sémantique minimale et nécessaire de ce bloc est agrégée par l'ensemble des concepts $\{PRODUCT, LotTrackingCode, ProductId, ProductNo, RevisionControlFlag, SerialTrackingCode, UOM, UOMCode, FACILITY, FacilityId, Division\}$. La spécification de l'interopérabilité de Sage X3 vers Flexnet, pour ce processus particulier, doit alors évaluer le niveau de complétude de l'ensemble des informations échangées en référence au besoin identifié par ce bloc

sémantique. La même analyse peut être réalisée pour chacun des blocs sémantiques et dans les deux sens d'interopération.

En pratique, ce type d'analyse montre que l'interopérabilité est le plus souvent partielle. Il est alors nécessaire d'évaluer le gap sémantique pour identifier les axes prioritaires d'amélioration devant donner lieu à des développements spécifiques. Nous avons proposé dans (Yahia, 2011) un système de mesure quantitatif et qualitatif associé à une cartographie contribuant à cet objectif. Ce système de mesure s'appuie sur les résultats présentés dans ce papier.

6. Conclusion

L'explicitation de la sémantique au niveau des modèles conceptuels des différents systèmes d'information dans une entreprise donnée (seule ou appartenant à un réseau d'entreprises), est de nos jours une problématique critique s'inscrivant dans le contexte de modèles d'entreprises. L'explicitation de la sémantique a pour avantage de faciliter la compréhension et par la suite l'utilisation des informations échangées entre des systèmes d'information hétérogènes.

Dans ce papier, nous avons proposé d'appliquer une approche « fact-oriented modelling » afin d'obtenir une normation des différents modèles conceptuels de systèmes d'information. Cela nous a permis d'identifier les connaissances centrales et étendues respectivement formalisées par les concepts obligatoires et non obligatoires des modèles sous-jacents.

L'originalité de ce papier concerne l'élaboration de blocs sémantiques pour expliciter la sémantique minimale et nécessaire enfouie dans les modèles d'entreprise.

La détermination des blocs sémantiques permet de spécifier des frontières entre les sous-systèmes et ainsi de faire émerger leur propre sémantique minimale et nécessaire pour opérer. En d'autres termes, nous mettons en évidence « ce qui est essentiel » dans le système sans prendre en considération les artefacts de mise en œuvre.

Nous avons validé notre proposition par un outillage informatique implémenté dans MEGA Suite et appliqué sur un cas d'étude. Cette étude de cas est relative à l'interopérabilité entre les applications Sage X3 et Flexnet dans le cadre d'un processus particulier. Il a permis de montrer comment les blocs sémantiques pouvaient être utilisés pour évaluer et améliorer l'interopérabilité entre ces systèmes.

Les outils et l'approche présentés dans ce papier ont servi de base aux travaux sur l'évaluation formelle de l'interopérabilité sémantique entre systèmes d'information présentés dans (Yahia, 2011).

7. Bibliographie

Aho, A.V., J.E. Hopcroft, et J.D. Ullman. *The design and analysis of computer algorithms*. Addison Wesley, 1974.

Boudjlida, N., et H. Panetto. «Annotation of enterprise models for interoperability purposes .» *Proceedings of the International Workshop on Advanced Information Systems for Enterprises IWAISE'08*. Constantine, Algeria, 2008.

Castano, S., et V. De Antonellis. «A framework for expressing semantic relationships between multiple information systems for cooperation.» *Information Systems* (Elsevier) 23, n° 3-4 (May-June 1998): 253-277.

Chapurlat, V. «Vérification et validation de modèles de systèmes complexe : application à la modélisation d'entreprise.» Habilitation à Diriger les Recherches de Université de Montpellier 2, 2007.

Chiang, R.H.L., T.M. Barron, et V.C. Storey. «Reverse engineering of relational databases: Extraction of an EER model from a relational database.» *Data and Knowledge Engineering* 12, n° 2 (1994): 107--142.

De Bo, J., P. Spyns, et R. Meersman. «Creating a “DOGMAtic” multilingual ontology infrastructure to support a semantic portal.» *In proceedings of the OTM confederated international workshops HCI-SWWA, IPW, JTRES, WORM, WMS, and MRS*. Catania , Italie: Springer-Verlag New York Inc, 2003. 253-266.

Engelbart, D.C. *Augmenting human intellect: A conceptual framework*. Summary Report AFOSR-3233, Menlo Park, CA, USA: Standord Reasearch Institute , 1962.

Evans, V. «Lexical concepts, cognitive models and meaning-construction.» Édité par Walter de Gruyter. *Cognitive Linguistics* 17, n° 4 (2006): 491-534.

Fonkam, M., et W. Gray. «An Approach to Eliciting the Semantics of Relational Databases.» *In proceeding of CAiSE'92, Lecture Notes in Computer Science 593 Springer*. Manchester, UK, 1992. 463-480.

Gamma, E., R. Helm, R. Johnson, et J. Vlissides. *Design patterns: elements of reusable object-oriented software*. Vol. 206. Addison-wesley Reading, MA, 1995.

Guarino, N. «Formal Ontology and Information Systems.» *Proceedings of FOIS'98*. Trento, Italy, 1998. 3-15.

Halpin, T. «Fact-oriented modeling: Past, present and future.» Édité par A. Opdahl & S. Brinkkemper J. Krogstie. *Conceptual Modelling in Information Systems Engineering* (Springer Berlin Heidelberg), 2007: 19-38.

Halpin, T. «Object-role modeling (ORM/NIAM).» Édité par K. Mertins & G. Schmidt P. Bernus. *Handbook on Architectures of Information Systems* (Springer Heidelberg), 2006: 81-103.

Kavouras, M. «A unified ontological framework for semantic integration.» *International Workshop on Next Generation Geospatial Information*. Boston MA, USA, 2003.

Lezoche, M., H. Panetto, et A. Aubry. «Conceptualisation approach for cooperative information systems interoperability.» *In Proceedings of 13th International Conference on Enterprise Information Systems*. Beijing, China, 2011.

Nijssen, G.M., et T.A Halpin. *Conceptual Schema and Relational Database Design: a fact oriented approach*. Upper Saddle River, NJ, USA: Prentice-Hall, 1989.

Obrst, L. «Ontologies for semantically interoperable systems.» *Proceedings of the 12 th international conference on Information and knowledge management*. New Orleans, USA, 2003. 366-369.

Panetto, H., et A. Molina. «Enterprise Integration and Interoperability in Manufacturing Systems: trends and issues .» *Enterprise Integration and Interoperability in Manufacturing Systems. Computers In Industry* 59, n° 7 (2008): 641-646.

Sharir, M. «A strong-connectivity algorithm and its applications in data flow analysis.» *Computers and Mathematics with Applications* (Elsevier) 7 (1981): 67-72.

Vernadat, F.B. *Enterprise Modelling and Integration: principles and applications*. Chapman and Hall, 1996.

Vernadat, F.B. «Enterprise Integration and Interoperability.» *Springer Handbook of Automation*, 2009: 1529-1538.

Whitman, L., D. Santanu, et H. Panetto. «An Enterprise Model of Interoperability .» *In Proceeding of the 12th IFAC Symposium on Information Control Problems in Manufacturing (INCOM'2006*. Saint Etienne, France,; Elsevier, 2006.

Yahia, E., A. Aubry, et H. Panetto. «Explicitation de la sémantique de modèles de systèmes d'information d'entreprise.» *4èmes Journées Doctorales , Journées Nationales MACS, JD-JN-MACS*. Marseille, France, 2011.

Yahia, E., M. Lezoche, A. Aubry, et H. Panetto. «Semantics enactment in Enterprise Information Systems.» *Proceedings of the 18th World Congress: The International Federation of Automatic Control*. Milan, Italy, 2011.

Yahia, E. *Contribution à l'évaluation de l'interopérabilité sémantique entre systèmes d'information d'entreprises : application aux systèmes d'information de pilotage de la production*. Thèse de doctorat, Nancy, France: Université Henri Poincaré – Nancy I, 2011.

Zouggar, N. *Contribution au développement de la sémantique en modélisation d'entreprise*. Thèse de doctorat, Bordeaux, France: Université Bordeaux 1, 2009.