Système d'information multi-agent pour une mémoire organisationnelle annotée en RDF

Fabien Gandon — Rose Dieng — Alain Giboin — Olivier Corby

Projet ACACIA – INRIA Sophia Antipolis

{Fabien.Gandon|Rose.Dieng|Alain.Giboin|Olivier.Corby}@sophia.inria.fr

RÉSUMÉ. *La distribution, au sens large, d'une mémoire organisationnelle est une caractéristique fondamentale dans le développement et le déploiement de systèmes de gestion des connaissances. Le web sémantique offre de nouvelles technologies pour matérialiser cette mémoire et les ontologies jouent un rôle important dans cette nouvelle génération de systèmes d'information à base de connaissances. Parallèlement, les systèmes multi-agents proposent un nouveau paradigme pour concevoir une plate-forme logicielle permettant de s'adapter à la topographie interne des organisations, et de capitaliser, en les intégrant dans une vue globale, les connaissances de l'organisation. Nous présentons ici un premier retour d'expérience sur notre activité de conception d'une telle architecture multi-agent et d'une ontologie pour la mémoire d'entreprise.*

ABSTRACT. *Distribution, in the broad sense, of an organizational memory, is a fundamental characteristic for developing and deploying knowledge management systems. The semantic web offers new technologies to materialize this memory and ontologies play a significant role in this new generation of knowledge-based information systems. In parallel, multi-agents systems propose a new paradigm to design software platform able to adapt to the internal topography of organizations, and to capitalize the corporate knowledge through integrated views. In this article, we present the first experience feedback on our design activity of such a multi-agent architecture and its underlying ontology for the corporate memory.*

MOTS-CLÉS : *mémoire organisationnelle, systèmes multi-agents, XML, RDF, ontologies.*

KEYWORDS: *organizational memory, multi-agent systems, XML, RDF, ontologies.*

1. Introduction

Avec l'avènement de la société de l'information et la modification consécutive des règles du jeu économique, les entreprises ont été amenées à s'intéresser davantage à la gestion de leurs connaissances ou de leur « mémoire », et les organismes de recherche à développer de nouvelles méthodes et de nouveaux outils. CoMMA – Corporate Memory Management through Agents [COM 00] – est un projet européen IST rassemblant des industriels et des organismes de recherche pour développer et tester un environnement de gestion de la mémoire d'entreprise en s'intéressant en particulier à deux scénarios : l'aide à l'insertion d'un nouvel employé et le support à la veille technologique. Pour atteindre ses objectifs, CoMMA a entrepris de conjuguer au sein d'un système multi-agent (SMA) des techniques de l'ingénierie des connaissances, de la galaxie XML, de la recherche d'information et de l'apprentissage symbolique. Après avoir justifié ces orientations techniques, nous présenterons notre retour d'expérience sur deux aspects du travail effectué par notre équipe : la conception de l'architecture multi-agent et la construction de l'ontologie O'CoMMA qui joue un rôle-clef dans le système.

2. L'organisation : un monde hétérogène et distribué

Nous définissons la mémoire organisationnelle comme la représentation persistante, explicite, désincarnée des connaissances et des informations dans une organisation, afin de faciliter leur accès, leur partage et leur réutilisation par les membres adéquats de l'organisation, dans le cadre de leurs tâches individuelles et collectives [DIE 00]. La mémoire d'entreprise correspond au cas particulier où l'organisation considérée est une entreprise. L'enjeu de la construction d'un système de gestion de mémoire organisationnelle est donc de réussir l'intégration cohérente de la connaissance dispersée dans l'organisation afin d'en promouvoir la croissance, la communication et la préservation [STE 93]. *La mémoire organisationnelle est un système distribué au sens large du terme.* Elle n'est pas seulement distribuée entre des systèmes informatiques, sa distribution inclut d'autres artefacts (livres, panneaux d'affichage...), mais aussi, et surtout, les membres de l'organisation. Les solutions de gestion de la mémoire, doivent prendre en compte cette « super-distribution » et les modèles de la distribution doivent aller au-delà de la dimension technique et rendre compte de l'organisation, de son infrastructure, de son environnement et de ses caractéristiques psychosociologiques.

Une mémoire distribuée et hétérogène. La mémoire organisationnelle a cela en commun avec le web qu'elle est un paysage d'informations hétérogènes et distribuées. De ce fait, tous les deux partagent les problèmes de bruit et d'imprécision lors de la recherche d'information. Pour reprendre la formule de John Naisbitt, « les utilisateurs du web se noient dans l'information alors qu'ils ont soif de connaissance ». Parmi les initiatives visant à résoudre ces problèmes, le web sémantique est une approche prometteuse [BER 01] ; l'idée est de rendre explicite la sémantique des documents au travers de métadonnées ou d'annotations. Les intranets reposant sur la technologie internet, peuvent bénéficier des progrès du web sémantique, et la mémoire d'entreprise peut alors être étudiée comme un « web sémantique d'entreprise ».

Une population d'utilisateurs distribuée et hétérogène. La population des utilisateurs d'une mémoire organisationnelle est hétérogène car il existe plusieurs profils de membres concernés par la gestion et l'exploitation de la mémoire. Chacun a ses particularités, ses centres d'intérêts et son rôle dans l'organisation. Cette population est aussi distribuée car les utilisateurs sont dispersés dans l'infrastructure de l'organisation. Ils ne sont pas forcément conscients de l'existence des autres, de leur fonction ou de leurs capacités. Ils ignorent parfois l'existence, l'emplacement et la disponibilité d'artefacts et de gens avec lesquels ils partagent des centres d'intérêts. Pour s'adapter aux profils des utilisateurs et aux contextes d'utilisation de la mémoire, l'apprentissage symbolique propose des techniques permettant d'apprendre les particularités individuelles ou de faciliter l'apparition de communautés d'intérêt et la diffusion proactive d'information par une exploitation collective des profils.

Une gestion distribuée et hétérogène. La mémoire organisationnelle est distribuée et hétérogène. La population de ses utilisateurs est distribuée et hétérogène. Il semble donc intéressant que l'interface entre ces deux mondes soit elle-même de nature distribuée et hétérogène pour pouvoir d'une part, se calquer sur le paysage d'information, et, d'autre part, s'adapter aux utilisateurs. Comme le note [WOO 00], les progrès de la programmation se sont faits à travers le développement d'abstractions de plus en plus puissantes permettant de modéliser et de développer des systèmes de plus en plus complexes. L'abstraction procédurale, les types de données abstraits, et plus récemment les objets et les composants sont tous des exemples de telles abstractions. Le paradigme des SMA [WOO 00] représente une nouvelle étape dans l'abstraction et il peut être employé pour comprendre, modéliser, et développer des systèmes distribués complexes. Ce paradigme apparaît particulièrement adapté au déploiement d'une architecture logicielle au-dessus de ce paysage d'information distribué.

La figure 1 présente une vue globale du système CoMMA. D'un côté, les agents de CoMMA sont les acteurs d'une société qui, par leur collaboration et leur coopération basées sur l'échange de messages au niveau sémantique, résolvent des problèmes complexes nécessitant une vue globale et intégrée de la mémoire d'entreprise. *La collaboration intelligente des agents du SMA permet de réaliser une capitalisation globale de la connaissance de l'entreprise.*

D'un autre côté, ces mêmes agents sont des entités localisées et affectées à un utilisateur pour l'assister dans son interaction avec la mémoire ou dédiées à une ressource pour en permettre l'exploitation au bénéfice de l'ensemble de la société des agents. Pour ce faire, les agents possèdent des capacités d'inférence, voire d'apprentissage utilisant les annotations, les modèles, l'ontologie et les profils

utilisateurs ainsi que des interfaces graphiques indispensables pour gérer les interactions entre ces objets conceptuels et les préoccupations des utilisateurs finaux. Les agents intègrent ces capacités à leur comportement afin de remplir leur rôle social. *L'autonomie et l'individualité de chaque agent lui permet de s'adapter localement à la spécificité des ressources et des utilisateurs individuels, tout en en faisant bénéficier l'ensemble de la société.*

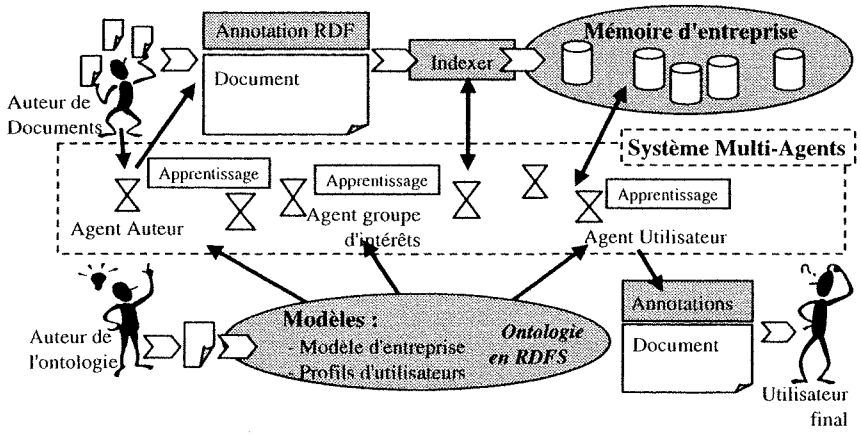

Figure 1. *Schéma général du système CoMMA*

Les SMA proposent une approche sous la forme d'un réseau de composants logiciels faiblement couplés, les agents, qui fonctionnent ensemble comme une société visant à résoudre des problèmes qui seraient hors d'atteinte pour n'importe quel agent pris individuellement. Un SMA est hétérogène s'il inclut des agents d'au moins deux types différents. Un système d'information multi-agent (SIMA) est un SMA visant à fournir tout ou partie des fonctionnalités permettant de contrôler et d'exploiter un ensemble d'informations. L'architecture logicielle de CoMMA est un SIMA hétérogène. Nous étudions la conception d'une telle architecture pour assister la gestion de la distribution de l'information. La dualité de la définition du mot « distribution » indique deux problèmes importants à traiter :

– distribution signifie « dispersion » : la propriété spatiale de ce qui est épars dans un espace. Le problème associé est de pouvoir gérer les données, l'information ou la connaissance naturellement distribuées dans l'organisation.

– distribution signifie également « propager » et « répartir ». Le problème associé est de s'assurer que l'information appropriée arrive à l'agent intéressé.

Ces deux aspects de la distribution trouvent leur réponse technique dans l'utilisation de XML pour la matérialisation de la mémoire (voir section 3) et

l'utilisation d'une architecture multi-agent pour sa gestion et son exploitation (voir section 4).

Les ontologies fournissent, d'une part, des ressources conceptuelles et notionnelles pour formuler et expliciter un savoir ; d'autre part, elles constituent un cadre partagé par les différents acteurs ; enfin, elles représentent le sens de différents contenus échangés dans les systèmes d'informations [BAC 01]. Ainsi, dans CoMMA, nous pouvons distinguer plusieurs raisons pour lesquelles nous avons besoin d'une ontologie : une venant de la nature même des SMA où les agents ont besoin d'un vocabulaire conceptuel consensuel pour exprimer les messages qu'ils échangent ; une autre venant de l'aspect système d'information utilisant des requêtes et des annotations basées sur un vocabulaire conceptuel consensuel ; enfin l'ontologie est une composante de la mémoire d'entreprise qui capture une connaissance potentiellement intéressante en elle-même pour l'utilisateur membre de cette communauté d'entreprise. Nous ne développons qu'une seule ontologie (O'CoMMA) mais les différents agents et les différents utilisateurs n'exploiteront pas les mêmes aspects de cette ontologie ; notre retour d'expérience sur la conception de cette ontologie sera présenté dans la section 5.

3. La technologie XML pour matérialiser la mémoire

XML, *Extensible Markup Language*, est en passe de devenir un standard pour l'échange et le stockage d'informations dans l'entreprise. C'est aussi dans la galaxie XML qu'a été proposé le langage RDF utilisé pour le web sémantique et présentant de nombreux avantages pour l'annotation puis la fouille sémantique d'une mémoire.

3.1. *XML : la technique de structuration*

XML est un langage de description de documents ou de données recommandé par le consortium du World Wide web. La syntaxe de XML utilise des balises avec des attributs ; elles permettent de structurer des documents et des données dans des fichiers au format texte facilement échangeables et utilisables sur des réseaux basés sur la technologie internet (figure 2). Le standard rend les données compréhensibles à l'homme et à la machine, même si à terme l'aspect XML est caché derrière les logiciels applicatifs et les interfaces utilisateur. XML est un format indépendant d'une technologie propriétaire et permet ainsi une mémoire pérenne. *Dans CoMMA, en utilisant XML on transmet aux agents logiciels une information structurée qu'ils peuvent ensuite traiter. On peut ainsi distribuer la charge de travail du système, en implantant des mécanismes locaux de manipulation des données ; par exemple, un agent d'interface peut trier, filtrer ou synthétiser des données qu'il a reçues sans refaire appel à ses sources.*

```
<details_contact>
  <nom>INRIA-Sophia</nom>
  <adresse pays="FR">
<rue>2004 Route des Lucioles</rue>
  <ville>Sophia Antipolis</ville>
  <code_postal>06902</code_postal>
  </adresse>
  <tel>04 92 38 77 00</tel>
</details_contact>
```

Figure 2. *Exemple de document XML*

XML permet de définir de nouvelles balises et des attributs pour les paramétrer. Les structures sont emboîtables jusqu'à un niveau arbitrairement grand, ce qui permet de représenter, par exemple, des schémas de base de données ou des hiérarchies d'objets. La structure des documents et des données échangés peut, au besoin, être définie formellement dans une DTD *Document Type Definition*. La DTD donne les noms des éléments et des attributs autorisés à être utilisés dans un document, la structure et l'emboîtement permis pour les balises, les valeurs des attributs, leurs types et leurs valeurs par défaut, etc. (figure 3). L'intérêt de pouvoir explicitement décrire la structure type d'un document est que l'on peut alors s'assurer que les instances des documents s'y conforment et garantir ainsi une certaine cohérence parmi les éléments de la mémoire. Malheureusement, la sémantique des balises ne peut pas être décrite dans une DTD car elle se limite à la syntaxe. *XML Schema*, une nouvelle recommandation qui doit remplacer les DTD, ne permet pas non plus de décrire la sémantique, mais permet de typer les documents. *Dans CoMMA, la structure aide un agent à trouver l'information appropriée et permet de répondre à des requêtes de façon plus précise que la recherche en texte intégral basée sur des mots.*

```
<!DOCTYPE details_contact [
 <!ELEMENT details_contact (nom,adresse,tel)>
 <!ELEMENT nom (#PCDATA)>
 <!ELEMENT adresse (rue,ville,code_postal)>
 <!ELEMENT tel (#PCDATA)>
 <!ELEMENT rue (#PCDATA)>
 <!ELEMENT ville (#PCDATA)>
 <!ELEMENT code_postal (#PCDATA)>
 <!ATTLIST adresse pays CDATA #REQUIRED >
]>
```

Figure 3. *Exemple de DTD*

A la différence de HTML, les balises de XML visent à permettre de décrire la structure des documents, plutôt que leur présentation. La présentation et la

structuration du contenu sont complètement indépendantes. Le langage XSLT, *Extensible Stylesheet Language Transformation,* est utilisé pour les feuilles de style (figure 4), il fournit des outils de manipulation de la structure des documents et des données qui vont au-delà du simple formatage. Ainsi un document de la mémoire peut être visualisé différemment (figure 5) et transformé pour s'adapter aux besoins des agents, aux profils des utilisateurs, aux contextes d'utilisation, aux logiciels avec lesquels le système est interfacé, etc., tout en étant stocké et communiqué dans un format unique. *Dans CoMMA les feuilles de style nous permettent notamment de présenter des résultats et de naviguer dans l'ontologie.*

```
<xsl:template match="/">
<HTML>
 <HEAD>
  <TITLE>Telephones</TITLE>
 </HEAD>
 <BODY>
   <xsl:apply-templates />
 </BODY>
</HTML>
</xsl:template>
<xsl:template match="details_contact">
 <xsl:value-of select='nom'>
 <xsl:text> : </xsl:text>
 <xsl:value-of select='tel'>
 <BR/>
</xsl:template>
```

```
<HTML>
 <HEAD>
  <TITLE>Telephones</TITLE>
 </HEAD>
 <BODY>
INRIA-Sophia : 04 92 38 77 00<BR>
 </BODY>
</HTML>
```

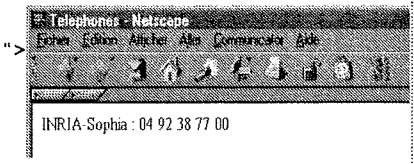

Figure 4. *Exemple de XSL*

Figure 5. *Exemple de rendu HTML*

3.2. *RDF : la technique d'annotation*

Selon [DOY 98], les agents logiciels doivent avoir la capacité d'acquérir des informations sémantiques utiles à partir du contexte du monde dans lequel ils évoluent et des environnements annotés permettent aux agents de devenir rapidement des acteurs intelligents dans des espaces virtuels. Bien que les auteurs aient choisi, pour domaine d'application celui des agents guidant des enfants dans des mondes virtuels, leur remarque est transposable aux agents évoluant dans le monde complexe des informations. De ce fait, les paysages d'information annotés sont, en l'état de l'art actuel, une façon rapide de rendre les agents d'information plus « intelligents ». Si la mémoire d'entreprise devient un monde annoté, les agents peuvent utiliser la sémantique des annotations et à l'aide d'inférences assister les utilisateurs dans leurs activités.

RDF, *Resource Description Framework*, fournit « la technologie pour exprimer la signification de termes et de concepts dans un format qu'une machine peut traiter » [BER 01]. Le langage RDF, utilise une syntaxe XML pour décrire des

triplets représentant les propriétés de ressources telles que des images, des pages web, etc., disponibles sur le réseau et les relations qui existent entre elles. On peut ainsi décrire les documents par des annotations sémantiques puis utiliser ces annotations pour fouiller plus efficacement la masse d'informations contenues dans la mémoire. RDF définit un mécanisme pour décrire des ressources par des annotations internes ou externes à celles-ci, et ne fait aucune hypothèse sur le domaine d'application, ni ne définit de sémantique *a priori*. Une caractéristique importante des nouveaux logiciels est la capacité à s'interfacer avec un système patrimonial ; de même une caractéristique importante d'une plate-forme de gestion d'une mémoire d'entreprise est sa capacité à intégrer des archives patrimoniales. Puisque RDF permet les annotations externes, des documents existants de la mémoire peuvent être gardés intacts et annotés extérieurement. Les annotations sont basées sur une ontologie décrite et partagée dans le langage RDF Schéma (RDFS). RDFS est proche des langages de représentation par objets avec la particularité que les propriétés sont définies en dehors des classes (figure 6).

```
<Class ID='Document'/>

<Class ID='Article'>
  <subClassOf resource='#Document'/>
</Class>

<Property ID='relecteur'>
  <domain resource='#Document'/>
  <range resource='Literal'/>
</Property>

<Article about='MyArticle.ps'>
  <relecteur>Fabien Gandon</relecteur>
</Article>
```

Figure 6. *Exemple simplifié de schéma et d'annotation*

Afin de manipuler annotations et ontologies, et d'effectuer des inférences, notre équipe (ACACIA) a développé CORESE [COR 00], un prototype de moteur de recherche permettant des inférences sur des annotations RDF en traduisant les triplets RDF en graphes conceptuels (GC) [SOW 84] et vice versa. CORESE combine les avantages d'utiliser, d'une part, le langage standard RDF pour exprimer et échanger les annotations, et de l'autre, les mécanismes de requête et d'inférence disponibles dans le formalisme des GC. En particulier l'opérateur de projection permet maintenant de fouiller une base d'annotations RDF en exploitant les liens de spécialisation définis dans le schéma. *Dans CoMMA, des modules de CORESE sont intégrés aux agents des sociétés dédiées à l'ontologie et aux documents, afin de leur fournir les capacités techniques nécessaires pour remplir leur rôle.*

3.3. *Une mémoire basée sur un modèle*

A la différence du web ouvert, le web d'une mémoire organisationnelle a un contexte délimité et mieux défini : l'organisation. Nous pouvons envisager d'identifier les différents profils d'utilisateurs, et comme la communauté de l'entreprise partage souvent un certain nombre de points de vue (politique d'entreprise, pratiques instituées...), un consensus ontologique est possible dans une certaine mesure, d'où notre cadre (voir figure 7) et notre démarche de modélisation :

1. Application des techniques de l'ingénierie des connaissances pour le recueil de données afin de fournir le vocabulaire conceptuel identifié comme nécessaire dans les scénarios d'application. Nous spécifions les concepts de la mémoire d'entreprise et leurs relations dans une ontologie que nous formalisons en RDFS.

2. L'analyse des données recueillies permet de décrire les profils utilisateurs et le modèle d'entreprise capturant la situation actuelle dans l'organisation. Ces modèles sont implantés en RDF en instanciant la description RDFS de l'ontologie.

3. Les agents assistent la structuration de la mémoire basée sur des annotations RDF à propos des documents. Ces annotations instancient la description RDFS de l'ontologie et font référence aux modèles. Par exemple: « le rapport d'analyse de tendance "nouvelles normes de la téléphonie mobile" a été écrit par la cellule de veille technologique du département D12 et concerne les sujets : UMTS, WAP ».

4. Les annotations, la situation actuelle et l'ontologie sont exploitées par les inférences des agents pour rechercher et naviguer dans la mémoire et pour la gérer.

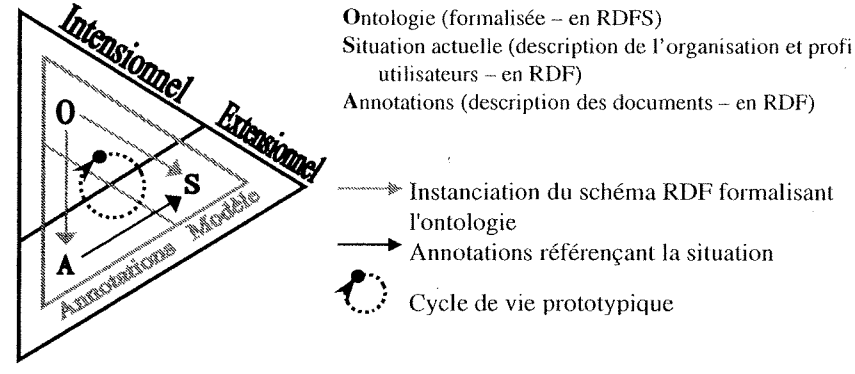

Figure 7. *Schéma O.S.A*

L'ontologie et la description de la situation actuelle forment le modèle de l'organisation. La structuration de la mémoire dépend des deux et c'est pourquoi la mémoire est dite basée sur un modèle. Le modèle d'entreprise est une représentation explicite, motivée et focalisée de l'organisation. Il vise à supporter les activités sur la

mémoire d'entreprise en permettant aux agents de CoMMA d'obtenir un aperçu partiel mais focalisé de leur contexte d'exécution et de leur environnement organisationnel, et d'exploiter intelligemment les aspects décrits dans ce modèle pour leurs interactions entre agents et avec les utilisateurs. Ce modèle rend compte des structures organisationnelles (hiérarchies, groupes) et des activités et centres d'intérêts qui leur sont attachés (veille technologique, R&D...).

Les modèles d'utilisateur sont des représentations explicites des caractéristiques des utilisateurs qui ont été jugées pertinentes pour l'adaptation du comportement du système. L'identification et les informations administratives permettent de distinguer un utilisateur des autres et de le positionner dans le modèle de l'organisation en explicitant son rôle, sa position, suggérant un réseau potentiel de connaissances, etc. Le profil contient aussi des préférences directement explicitées qui vont de la personnalisation d'interfaces aux centres d'intérêts. En plus de l'information explicitement indiquée, le système déduit des informations de l'utilisation passée. Il maintient un historique des documents demandés ou visités et des éventuelles réactions de l'utilisateur lors de sa visite. De ceci le système peut déduire et apprendre les habitudes et les préférences de l'utilisateur. Ces informations dérivées sont utilisées à des fins d'interface et pourront être exploitées pour la suggestion proactive d'informations. Enfin, les modèles des utilisateurs peuvent être comparés pour regrouper les utilisateurs en se basant sur des critères de similarité dans leurs profils et utiliser le profil des utilisateurs semblables pour faire des suggestions. Des techniques d'apprentissage symboliques sont étudiées par le LIRMM [KIS 01], en particulier pour prédire les préférences de l'utilisateur.

4. Conception d'une société d'agents pour la mémoire

Lorsqu'il envisage une solution logicielle dans une perspective multi-agent, le concepteur doit gérer la relation entre :

– le *niveau macroscopique du SMA (la société des agents)*, où se posent les problèmes d'ingénierie des interactions et d'organisation de la société du SMA afin d'obtenir, du point de vue global du système, les fonctionnalités correspondant aux exigences de l'utilisateur ;

– le *niveau microscopique du SMA (les agents en eux-mêmes)* où se posent les problèmes d'identification des rôles nécessaires, d'ingénierie des comportements tenant compte des interactions qui se produiront et fournissant les différentes compétences recherchées.

Notre approche partage avec AALAADIN [FER 97] et GAIA [WOO 00] le souci de chercher à concevoir un SMA comme une organisation humaine, en identifiant les rôles nécessaires au fonctionnement des sociétés d'agents, les relations qui existent entre ces rôles et les interactions systématiques auxquelles ils participent suivant des protocoles institutionnalisés. Cette approche est séduisante dans le contexte de la gestion d'une mémoire organisationnelle et nous présentons

dans cette section notre retour d'expérience sur la conception d'un prototype de SIMA appliqué à une telle mémoire. Nous partirons des fonctionnalités du système exprimées au niveau social pour aboutir au comportement interne des agents. La plate-forme JADE [BEL 01], développée par l'université de Parme, membre du consortium CoMMA, est utilisée pour l'implantation.

4.1. *Du macroscopique au microscopique*

Les fonctionnalités souhaitées pour le système ne se transfèrent pas directement en fonctionnalités d'agents, mais influencent la conception et sont finalement distribuées dans les interactions sociales des agents et l'ensemble des capacités, des rôles et des comportements qui leur sont associés. Dans cette section, nous partons des contraintes de CoMMA exprimées au niveau macroscopique et nous irons jusqu'au point où les rôles nécessaires peuvent être identifiés.

4.1.1. *Architecture et configuration*

L'architecture d'un système multi-agent est une structure qui dépeint les différentes familles d'agents possibles dans le système et les rapports qu'ils entretiennent. Une configuration est une instance d'une architecture avec un agencement choisi et un nombre approprié d'agents de chaque type. A une architecture donnée peuvent correspondre plusieurs configurations.

Dans le cas d'une mémoire d'entreprise, la configuration choisie dépend étroitement de l'environnement de déploiement du système (agencement de l'organisation, topographie et contraintes réseau, localisation des zones d'intérêt et des personnes concernées par le système). Elle doit s'adapter au paysage d'informations et changer avec lui pour permettre une intégration symbiotique du système de gestion de la mémoire à son organisation commanditaire. L'architecture multi-agent doit donc permettre plusieurs configurations de déploiement couvrant les différentes structures organisationnelles envisageables. La description d'une configuration peut être étudiée et documentée lors du déploiement en utilisant des diagrammes de déploiement UML pour représenter les machines hôtes (serveurs, machines personnelles...), les conteneurs de la plate-forme SMA, les instances d'agents et leur réseau d'accointances.

L'architecture, quant à elle, est étudiée et fixée lors de la conception du SMA. L'analyse architecturale part du plus haut niveau d'abstraction du système (*i.e.* la société) et par raffinements et décompositions successives (*i.e.* sous-sociétés gigognes) elle descend jusqu'au point où les rôles d'agents et les interactions peuvent être identifiés et spécifiés. Nous ne développons pas, comme dans GAIA, la formalisation des rôles, ou, comme dans AALAADIN, la notion de réflexivité, ces aspects n'étant par ailleurs pas implantés dans la plate-forme JADE utilisée pour CoMMA. Nous allons par contre détailler la conception de l'architecture. En considérant les fonctionnalités du système CoMMA nous avons identifié quatre sous-sociétés : 1) une sous-société dédiée au modèle d'entreprise et à l'ontologie ;

2) une sous-société dédiée aux documents ; 3) une sous-société dédiée aux utilisateurs ; 4) une sous-société dédiée à l'interconnexion. Nous décrivons chacune d'entre elles dans les parties suivantes.

4.1.2. *Organisation des sous-sociétés*

Lors de l'analyse des sous-sociétés dédiées aux ressources (modèles, documents et pages jaunes), nous avons constaté la récurrence de trois types de sociétés.

La société hiérarchique (figure 8) distingue deux types d'agents :

1. *Les représentants* : médiateurs entre leur société et le reste du SMA. Ils sont responsables du traitement des requêtes externes, de leur décomposition, du dialogue avec les exploitants et de l'intégration des résultats intermédiaires pour fournir une réponse appropriée au commanditaire externe ;

2. *Les exploitants* : assignés à une ressource locale, ils contribuent autant que possible à la résolution des requêtes reçues, avec cette ressource locale. Dans cette société, l'information est répartie entre les agents assignés aux ressources. Ceci permet de conserver la répartition initiale des ressources et de distribuer et équilibrer la charge de travail : le travail de recherche locale est réparti entre les exploitants et le travail de fusion est effectué par les représentants. La spécialisation des agents et la distribution des rôles permettent la distribution de la charge de travail mais requièrent en contrepartie un volume plus important d'échanges sur le réseau.

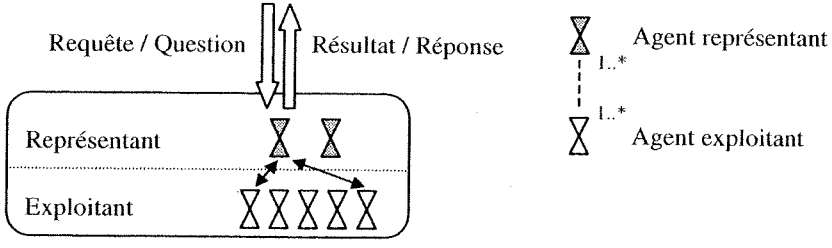

Figure 8. *Société hiérarchique*

La société égalitaire (figure 9) repose sur des relations égalitaires entre des rôles complètement redondants. N'importe quel agent peut être contacté de l'extérieur de la société pour répondre à une requête concernant le type de ressources auquel sa société est dédiée. Il doit alors coopérer avec ses pairs pour résoudre efficacement la requête. Dans cette société, les agents sont uniquement spécialisés par le contenu de la ressource locale à laquelle ils sont assignés. La charge de travail est moins distribuée que dans l'exemple précédent mais la charge réseau peut être diminuée. Il n'y a qu'un seul type d'agent et il joue les deux rôles précédents. Chaque agent est capable de former une coalition avec d'autres agents pour résoudre les requêtes externes.

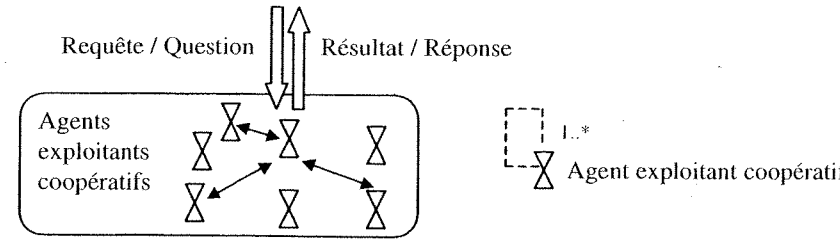

Figure 9. *Société égalitaire*

La société de duplication (figure 10) est un cas particulier du cas précédent : ni les rôles ni le contenu ne sont distribués. Chaque agent maintient à jour une copie complète de toute l'information et peut résoudre les requêtes seul. Par conséquent, les seules interactions sociales qui existent concernent les mises à jour du contenu. La charge de travail est bien moins distribuée que dans le cas précédent et le contenu doit être dupliqué auprès de chaque agent de la société, ce qui peut être une contrainte inacceptable. En revanche, il n'y a aucune spécialisation, le système est fortement redondant, et par conséquent fortement résistant aux pannes. L'utilisation du réseau est minimale lors du traitement d'une requête. Le seul rôle en commun avec les sociétés précédentes, est celui de l'exploitant.

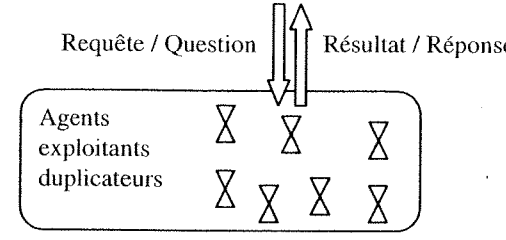

Figure 10. *Société de duplication*

Selon les tâches à exécuter, la taille des données, et plus généralement selon les contraintes imposées, une sous-société dédiée sera organisée selon l'un ou l'autre des modèles précédents, et les interactions exigeront différents protocoles (requête, question, appel à proposition, enchères, etc.).

4.1.3. *Sous-société dédiée au modèle et à l'ontologie*

Les agents de cette sous-société fournissent des services de diffusion, de mise à jour et de requêtes sur l'ontologie et sur le modèle décrivant la situation actuelle de l'organisation. Si le système admet plusieurs ontologies, les agents peuvent assurer

la correspondance d'une ontologie vers une autre en utilisant, par exemple, des règles de mise en correspondance exprimées par rapport à une ontologie globale. Les agents en charge de la description de la situation organisationnelle, maintiennent un modèle de l'organisation décrit en utilisant les primitives de l'ontologie et le rendent accessible aux autres agents. Si le système utilise une ontologie structurée en points de vue, les agents de cette sous-société peuvent permettre d'utiliser ces points de vue pour filtrer l'accès aux données. S'il existe un niveau terminologique dans l'ontologie, les agents peuvent fournir des services de recherche de termes et de synonymes définis sur les concepts.

Pour la sous-société dédiée à l'ontologie ou au modèle d'entreprise, les trois types d'organisation sont envisageables :

1. Dans le cas d'une hiérarchie, nous aurions un rôle de *Médiateur* en charge de la résolution des requêtes externes et un rôle d'*Archiviste* en charge d'une partie ou d'une vue sur l'ontologie ou le modèle.

2. Dans une société égalitaire, nous aurions un rôle d'*Archiviste* coopératif en charge d'une partie du modèle ou de l'ontologie et susceptible de créer des coalitions temporaires pour répondre à une requête.

3. Dans une société de duplication, nous aurions un rôle d'*Archiviste* ayant en permanence une copie complète de l'ontologie ou du modèle.

Le dernier choix est acceptable si l'ontologie ou le modèle ne sont pas trop volumineux, si le processus de consensus ontologique est centralisé et si les mises à jour sont propagées dans la société des agents. C'est cette organisation qui est actuellement implantée dans le premier prototype de CoMMA. Elle est acceptable uniquement parce que la gestion du consensus ontologique n'est pas à la charge du système et que la propagation des changements dans l'ontologie se fait par rechargement complet. Si le processus d'édition de l'ontologie et la maintenance du consensus devaient être à la charge du système, les autres options deviendraient intéressantes, en particulier si l'ontologie ou les modèles sont assez volumineux et/ou partitionnables ; la société maintiendrait alors la répartition et la cohérence.

4.1.4. *Sous-société dédiée aux documents*

Les agents de cette sous-société sont en charge de l'exploitation des documents et des annotations distribués dans des bases de l'entreprise. Ces agents recherchent et extraient les références satisfaisant la requête de l'utilisateur avec l'aide des agents dédiés à l'ontologie et aux modèles. Pour cette société, seules les deux premières options sont envisageables : 1) Dans une société hiérarchique, nous aurions un rôle de *Médiateur* en charge de gérer les requêtes posées sur la base documentaire de la mémoire, et un rôle d'*Archiviste* associé à la gestion d'une base d'annotations. 2) Dans une société égalitaire nous aurions un rôle d'*Archiviste* coopératif. 3) Une société de duplication n'est pas réaliste, car cela signifierait la duplication d'une image complète de toute la mémoire pour chaque instance d'un agent de cette

société. Ceci n'est pas concevable dans le cadre d'une importante mémoire d'entreprise distribuée sur un intranet.

Pour le système CoMMA nous avons opté pour la première solution en l'appliquant à la gestion de bases d'annotations RDF. Le rôle de *Médiateur* d'Annotations est de proposer ses services aux agents des autres sociétés pour résoudre leurs requêtes et archiver leurs annotations ; le médiateur engage les services des *Archivistes* d'Annotations pour effectivement archiver une annotation et résoudre une requête : 1) Il décompose et distribue les requêtes aux *Archivistes* 2) Il compile les réponses partielles pour construire le résultat final. Le rôle de l'*Archiviste* d'Annotations est affecté à l'exploitation d'une archive locale. Quand cet agent reçoit une requête, il essaie de la résoudre même partiellement pour permettre au *Médiateur* de retrouver des réponses distribuées. Le premier prototype de CoMMA utilise une implantation simplifiée de la distribution des annotations et des requêtes permettant de tester l'architecture et l'intégration. Des algorithmes plus poussés d'archivage et de résolution distribués sont en cours de développement. Ils utilisent des statistiques sur les bases d'annotations exploitant l'ontologie pour l'allocation d'une nouvelle annotation et la répartition des tâches entre agents lors de la résolution d'une requête.

4.1.5. *Sous-société dédiée aux interconnexions*

Les agents de cette sous-société sont responsables de la connexion, par appariement, des autres agents en se basant sur l'expression de leurs besoins et la description de leurs capacités. Chaque agent fournisseur d'un service doit enregistrer sa description auprès d'un agent d'interconnexion pour que les attentes des agents demandeurs puissent y être comparées. La littérature donne plusieurs types d'agents d'interconnexion et la terminologie n'est pas consensuelle : la même désignation peut dénoter des rôles différents. Nous adoptons les définitions de [KLU 99] :

– un agent courtier (*broker*) identifie les fournisseurs potentiels, leur transmet la requête, récupère les résultats et renvoie une réponse complète au demandeur.

– un agent apparieur (*matchmaker*) identifie les fournisseurs potentiels, fournit cette liste de candidats au demandeur et le laisse prendre contact avec eux.

Dans JADE les agents responsables des pages jaunes (*Directory Facilitators*) sont des apparieurs que l'on peut fédérer en une société égalitaire.

4.1.6. *Sous-société des agents utilisateurs*

Les agents de la sous-société dédiée aux utilisateurs sont en charge des aspects d'interface, d'observation, d'aide et d'adaptation à l'utilisateur. Ces agents sont en général demandeurs de services. Etant donné qu'ils ne sont pas liés à une ressource comme dans les cas précédents, ces agents ne suivent pas la typologie des sous-sociétés que nous venons d'exposer. Les rôles définis dans cette sous-société et leur distribution dépendent de caractéristiques fonctionnelles supplémentaires du système telles que l'apprentissage des profils, le filtrage collaboratif, etc. Nous nous

sommes intéressés en priorité à deux rôles récurrents dans les SIMA : 1) Le rôle de gestion de l'interface utilisateur : en charge du dialogue avec l'utilisateur, il doit permettre l'expression de requêtes par l'utilisateur, leur raffinement et la présentation des résultats dans un format approprié. 2) Le rôle de gestion du profil utilisateur : premier point d'intégration des techniques d'apprentissage symbolique, il exploite le profil pour s'adapter et apprendre les préférences de l'utilisateur non seulement pour des aspects d'interface et de présentation mais aussi pour de la diffusion proactive. Un autre aspect de ce deuxième rôle concerne l'archivage et l'accès aux profils des utilisateurs. D'autres rôles sont décrits dans [GAN 00].

4.2. *Des rôles et interactions aux comportements*

La deuxième étape de notre approche est de dériver de l'analyse de l'architecture les caractéristiques des rôles, les protocoles de leurs interactions, afin de choisir une implantation des comportements de chaque type d'agents.

4.2.1. *Les rôles*

Un rôle représente la position d'un agent dans une société et les responsabilités et les activités, assignées à cette position, que les autres agents s'attendent à voir correctement remplies. L'analyse des rôles est à la charnière entre le niveau microscopique des agents et le niveau macroscopique du SMA. Dans les sous-sociétés, nous avons identifié les rôles suivants ; ils ont motivé l'implantation de comportements d'agents remplissant ces rôles dans notre premier prototype :

- *archiviste d'Ontologie* : gestion et accès à l'ontologie ;
- *archiviste Modèle d'Entreprise* : gestion et accès au modèle d'entreprise ;
- *archiviste d'Annotations* : gestion et accès à une base d'annotations ;
- *médiateur d'Annotations* : médiation entre les archivistes et les demandeurs ;
- *apparieur* : maintenance et accès aux pages jaunes ;
- *contrôleur d'Interface* : contrôle de l'interface graphique utilisateur ;
- *gestionnaire de profils* : modification des profils d'utilisateurs connectés ;
- *archiviste de profils* : archivage et accès aux profils utilisateurs ;

La liste des caractéristiques qui sont débattues dans la communauté SMA donne une excellente vision des capacités actuellement envisageables pour un agent. La diversité des caractéristiques révèle l'influence et l'intégration dans les SMA de résultats issus des multiples domaines de recherche. Le tableau 1 compile les caractéristiques extraites de [ETZ 95, FRA 96, NWA 96] et [WOO 95] et donne pour chaque rôle les caractéristiques identifiées. Nos partenaires de l'Université de Parme ont aussi documenté chaque rôle en utilisant les cartes de descriptions pour les facettes d'un rôle d'agent, définies dans [KEN 99]. L'article [WOO 00] propose aussi une formalisation de la définition des rôles mais cette formalisation n'a pas été jugée nécessaire pour le premier prototype de CoMMA.

	Archiviste d'ontologie	Archiviste Modèle d'Entreprise	Archiviste Annotations	Médiateur d'Annotations	Apparieur	Contrôleur d'Interface	Gestionnaire de profils	Archiviste de profils
Réactif	Non	Non	Non	Non	Non	Non	Non	Non
Etat Mental Complexe	Non	Non	Non	Non	Non	Non	Non	Non
Dégradation Elégante	Oui	Oui	Oui	Oui	Oui	Oui	Oui	Oui
Continuité Temporelle	Oui	Oui	Oui	Oui	Oui	Non	Oui	Oui
Autonomie								
Orienté but	Non	Non	Non	Oui	Non	Non	Oui	Non
Collaboratif	Oui	Oui	Oui	Oui	Oui	Oui	Oui	Oui
Flexible	Non	Non	Non	Oui	Non	Non	Oui	Non
Proactif	Non	Non	Non	Non	Non	Non	Oui	Non
Personnalité	Non	Non	Non	Non	Non	Non	Non	Non
Communication	Oui	Oui	Oui	Oui	Oui	Oui	Oui	Oui
Adaptabilité								
Apprentissage	Non	Non	Non	Non	Non	Non	Oui	Non
Personnalisable	Non	Non	Non	Non	Non	Oui	Oui	Non
Mobilité	Non	Non	Non	Non	Non	Non	Non	Non
Représentation visuelle	Non	Non	Non	Non	Non	Oui	Non	Non
Véracité	Oui	Oui	Oui	Oui	Oui	Oui	Oui	Oui
Volontariat	Oui	Oui	Oui	Oui	Oui	Oui	Oui	Oui
Rationalité	Oui	Oui	Oui	Oui	Oui	Oui	Oui	Oui

Tableau 1. *Analyse des rôles d'identifiés en utilisant les caractéristiques extraites de [ETZ 95] [FRA 96] [NWA 96] et [WOO 95]*

4.2.2. *Interactions sociales*

Pour répondre aux fonctionnalités globales du système, les agents doivent pouvoir communiquer les uns avec les autres pour déléguer des tâches, échanger des informations, coopérer. Après la spécification des rôles, nous nous intéressons donc aux interactions.

L'interaction entre agents est plus que le simple envoi d'un message isolé. Le modèle d'une conversation doit être spécifié par des protocoles que les agents doivent respecter pour que le SMA fonctionne effectivement. Les protocoles sont des codes de conduite sociale pour l'interaction ; ils décrivent des procédures standards régulant l'échange d'informations entre agents en institutionnalisant des modèles de communications pouvant survenir entre des rôles identifiés. La spécification d'un protocole part d'un graphe d'accointances au niveau des rôles, qui représente par un graphe orienté les voies de communication existant entre les agents jouant ces rôles. A partir de ce graphe, on spécifie la séquence possible des messages échangés pendant l'interaction. La figure 11 montre, par exemple, une sous-partie d'un diagramme d'interaction documentant les échanges ayant lieu lors de l'allocation d'une nouvelle annotation ; on pourrait décrire ce protocole comme un *contract-net* à deux niveaux.

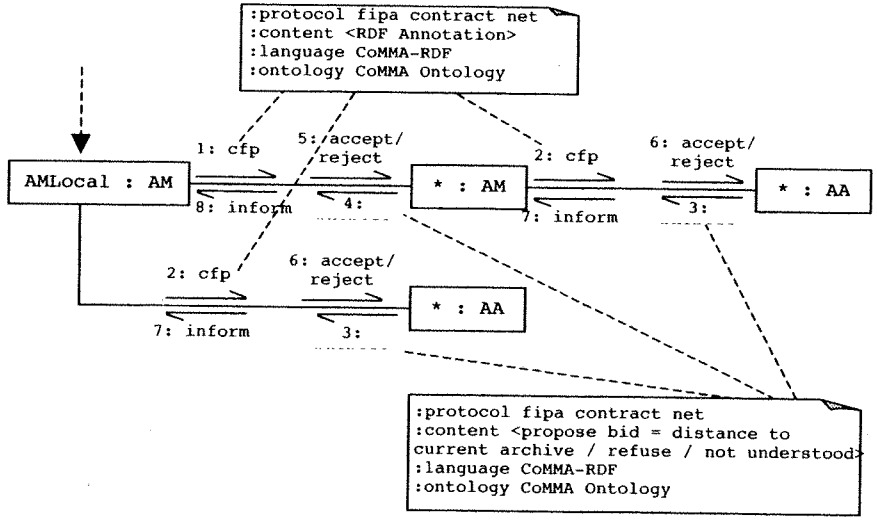

Figure 11. *Sous-partie d'un diagramme d'interaction pour l'allocation d'une annotation*

Le graphe d'accointances se déduit à la fois de l'analyse organisationnelle précédente et des *use cases* du système, décrits à partir des scénarios d'application

proposés dans le cahier des charges. Le graphe d'accointances et les messages sont décrits dans des diagrammes de protocoles, une restriction des diagrammes de collaboration UML [AUM 00] [BER 00].

La plate-forme JADE offre un environnement de développement en langage Java conforme à la norme FIPA [FIP 01]. Son langage de communication FIPA ACL (*Agent Communication Language*) est basé, comme son homologue KQML, sur la théorie des actes de langage et dispose d'une liste de protocoles élémentaires déjà spécifiés avec leur ontologie. Le prototype actuel de CoMMA utilise le langage SL spécifié par FIPA pour l'enveloppe des messages (figure 12-a) et les actes de langage spécifiques à CoMMA (figure 12-b), et RDF pour l'expression d'annotations et de requêtes sur le contenu de la mémoire (figure 12-c). Il est également intéressant de noter ici que le besoin d'un langage commun dans les SMA rejoint l'initiative du W3C sur les langages XML et RDF et, compte tenu des évolutions prochaines de JADE, nous espérons pouvoir intégrer entièrement XML en tant que langage de contenu et de communication.

```
(a) (QUERY-REF
     :sender   ( agent-identifier
                                  :name localUPM@fapollo:1099/JADE)
     :receiver  (set ( agent-identifier
                                  :name AM@fapollo:1099/JADE) )
     :content
(b)   ((all ?x (is-answer-for
                 (query
      :pattern
(c)     <?xml version ="1.0"?> <rdf:RDF xml:lang="en"
        xmlns:rdf= "http://www.w3.org/1999/02/22-rdf-syntax-ns#"
        xmlns:comma="http://www.inria.fr/acacia/comma#">
        <comma:Memo><comma:Designation>?</comma:Designation>
        </comma:Memo>
        </rdf:RDF>
(b)              ) ?x ) ) )
     :reply-with  QuerylocalUPM987683105872
(a)  :language  CoMMA-RDF
     :ontology  CoMMA-annotation-ontology
     :protocol  FIPA-Query
     :conversation-id  QuerylocalUPM987683105872 )
```

(a) Enveloppe FIPA ACL - **(b)**Requête SL0 CoMMA - **(c)**Pattern RDF

Figure 12. *Exemple de message échangé demandant le titre des mémos annotés*

4.2.3. *Implantation : le comportement et les capacités techniques*

A partir de la description des rôles et des interactions, nous avons proposé et implanté des types d'agents remplissant un ou plusieurs rôles. Le comportement d'un type d'agent combine les différents comportements implantés pour remplir les activités correspondant aux rôles qui lui sont assignés. Le comportement est fixé par

les choix d'implantation déterminant les actions et réactions de l'agent ; ces choix sont libres dans la limite du respect des contraintes données par les rôles et les protocoles d'interaction. Il existe plusieurs types de programmation de comportement. Ils reposent, par exemple, sur des automates finis ou des réseaux de Petri. L'implantation d'un comportement repose aussi sur la palette de capacités techniques disponibles pour les concepteurs et c'est pourquoi les SMA tendent à intégrer des résultats de plusieurs autres domaines de recherche. Les comportements intègrent donc des outils de différents domaines comme des briques de base pour ajouter aux différents types d'agents les capacités techniques nécessaires pour remplir leurs rôles.

Dans JADE, les comportements sont décrits par composition de sous-comportements soit directement issus des comportements de base (séquentiel, non déterministe, cyclique, ponctuel...), soit implantés pour le besoin. Nous avons intégré des modules du moteur de recherche CORESE dans les agents des sociétés dédiées à l'ontologie et aux documents. Les API Xerces et Xalan du consortium Apache permettent de traiter XML et les feuilles de style. L'API Sirparc du W3C permet de traiter les annotations RDF et le schéma RDFS. Les techniques d'apprentissage symbolique étudiées par le LIRMM utilisent l'API WEKA de l'Université de Waikato, Nouvelle-Zélande.

```
<cos:rule>
  <cos:if>
    <rdf:RDF>
      <CoMMA:OrganizationalEntity>
        <CoMMA:Include> <CoMMA:Person rdf:about="?x"/> </CoMMA:Include>
        <CoMMA:Include> <CoMMA:Person rdf:about="?y"/> </CoMMA:Include>
    </CoMMA:OrganizationalEntity>
    </rdf:RDF>
  </cos:if>
  <cos:then>
    <rdf:RDF>
      <CoMMA:Person rdf:about="?x">
        <CoMMA:Colleague>
          <CoMMA:Person rdf:about="?y"/>
        </CoMMA:Colleague>
      </CoMMA:Person>
    </rdf:RDF>
  </cos:then>
</cos:rule>
```

Figure 13. *Exemple de règle pour une définition formelle de « collègue »*

A titre d'exemple, un Agent Archiviste d'Annotations, va dans son comportement inclure la participation à la résolution d'une requête distribuée. Une sous-tâche de cette activité consiste à projeter un sous-graphe de la requête globale pour proposer des solutions partielles à l'Agent Médiateur. Lors de la projection,

l'Agent Archiviste d'Annotations peut être amené à inférer des connaissances qui ne sont pas explicites dans sa base. Par exemple le fait que « si deux personnes travaillent ensemble, elles sont collègues ». Le comportement de cet agent inclut donc des appels à CORESE pour appliquer des règles propres à sa tâche comme celle que nous venons d'énoncer et donnée en figure 13 sous sa forme XML.

Les messages d'interaction, les annotations et les requêtes sur la mémoire, les profils des utilisateurs, et le modèle de l'entreprise peuvent être communiqués et traités par les agents car ils reposent sur une ontologie partagée ; L'ontologie est la clef de voûte de notre système d'information multi-agents : elle assure la cohérence dans les échanges entre les composants distribués. La conception de cette ontologie a été étudiée avec attention, et la section suivante présente les étapes de sa construction ainsi que nos premiers retours d'expérience.

5. Conception de l'ontologie O'CoMMA

CoMMA exploite un modèle de la situation actuelle de l'organisation décrivant les aspects pertinents de l'environnement organisationnel, des profils d'utilisateur et des annotations structurant la mémoire. Pour les formaliser, nous avons besoin d'une ontologie définissant les primitives nécessaires à leur représentation et la sémantique qui leur est associée. Dans la suite, nous précisons notre position et les définitions du domaine ontologique puis nous présentons en détail chacune des phases de notre méthode de conception de l'ontologie O'CoMMA sous l'angle, en particulier, de la distribution des connaissances.

5.1. *Position et définitions*

L'ontologie est une théorie logique qui rend compte partiellement mais explicitement d'une conceptualisation [GUA 95], *la conceptualisation étant une structure sémantique intensionnelle qui capture les règles implicites contraignant la structure d'un morceau de réalité.* L'ontologie est une représentation explicite partielle parce qu'elle se concentre sur les aspects de la conceptualisation nécessaires à l'application envisagée.

Un *concept* est un constituant de la pensée (un principe, une idée, une notion abstraite) sémantiquement évaluable et communicable. L'ensemble des propriétés d'un concept s'appelle sa compréhension ou son *intension* et l'ensemble des êtres qu'il englobe, son *extension*. Il existe une dualité entre l'intension et l'extension : à des intensions incluses $I_1 \supset I_2$ correspondent des extensions incluses $E_1 \subset E_2$. L'intension est déterminée par l'identification de propriétés qualitatives ou fonctionnelles communes aux individus auxquels le concept s'applique ainsi que ses relations avec les autres individus, cet ensemble de caractères permettant de définir le concept. Pour exprimer et communiquer l'intension, nous choisissons une représentation symbolique, souvent verbale, ex : les définitions des différentes

notions associées à un terme et données dans les dictionnaires. Dans notre contexte d'un système d'information distribué dans une société humaine, une notion, pour être communicable, doit par conséquent être dotée d'un libellé. Il est donc vital d'*expliciter et de maintenir les liens entre les intensions capturées dans l'ontologie, les termes employés et les définitions associées en langue naturelle*. C'est dans ces liens que se détectent l'éventualité des ambiguïtés en particulier dès lors qu'un terme est susceptible de dénoter plusieurs notions. Notons que les exemples et les illustrations utilisés dans un dictionnaire montrent, respectivement, qu'il est parfois nécessaire pour clarifier une définition en langue naturelle d'exhiber un échantillon représentatif de son extension ou d'utiliser d'autres formes de représentations.

Les représentations des intensions sont organisées, structurées et contraintes pour exprimer une théorie logique rendant compte des relations qui existent entre les concepts. *L'ontologie est l'objet capturant les expressions des intensions et la théorie logique visant à rendre compte des aspects de la réalité choisis pour leur pertinence dans les scénarios d'application considérés*. La représentation des intensions et de la structure ontologique peut faire appel à des langages plus ou moins formels selon l'opérationalisation envisagée pour l'ontologie. La construction formelle de l'intension donne une représentation précise et non ambiguë de la manière dont on peut concevoir son sens, ce qui permet sa manipulation logicielle et son utilisation comme une primitive de représentation de connaissances pour les modèles (ex : modèle d'entreprise, profils utilisateurs) et les annotations. *Les langages de formalisation ont tous des limites et des biais* dans leur expressivité qui font que la formalisation ne peut pas être une traduction parfaite de la conceptualisation. Le simple fait d'adopter une vue objet amène la dichotomie classe/instance qui n'était peut-être pas naturelle dans la conceptualisation du monde faite à l'origine.

L'expression de l'intension partant presque systématiquement d'une définition en langue naturelle, définir une ontologie est une tâche de modélisation menée à partir de l'expression linguistique des connaissances [BAC 00]. Par raffinements successifs, nous augmentons l'ontologie en développant les pendants formels des aspects sémantiques pertinents pour les scénarios d'application envisagés. Dans l'ontologie, les concepts en intension sont habituellement organisés en *taxinomie*, c'est-à-dire une classification basée sur leurs similitudes.

Quand des personnes s'accordent sur l'utilisation et la théorie spécifiée par l'ontologie, elles prennent *un engagement ontologique*. L'*ingénierie ontologique* s'occupe des aspects pratiques, essentiellement les méthodologies et les outils de l'application des résultats de la théorie des ontologies à la construction d'ontologies, pour un cas précis et avec un but spécifique.

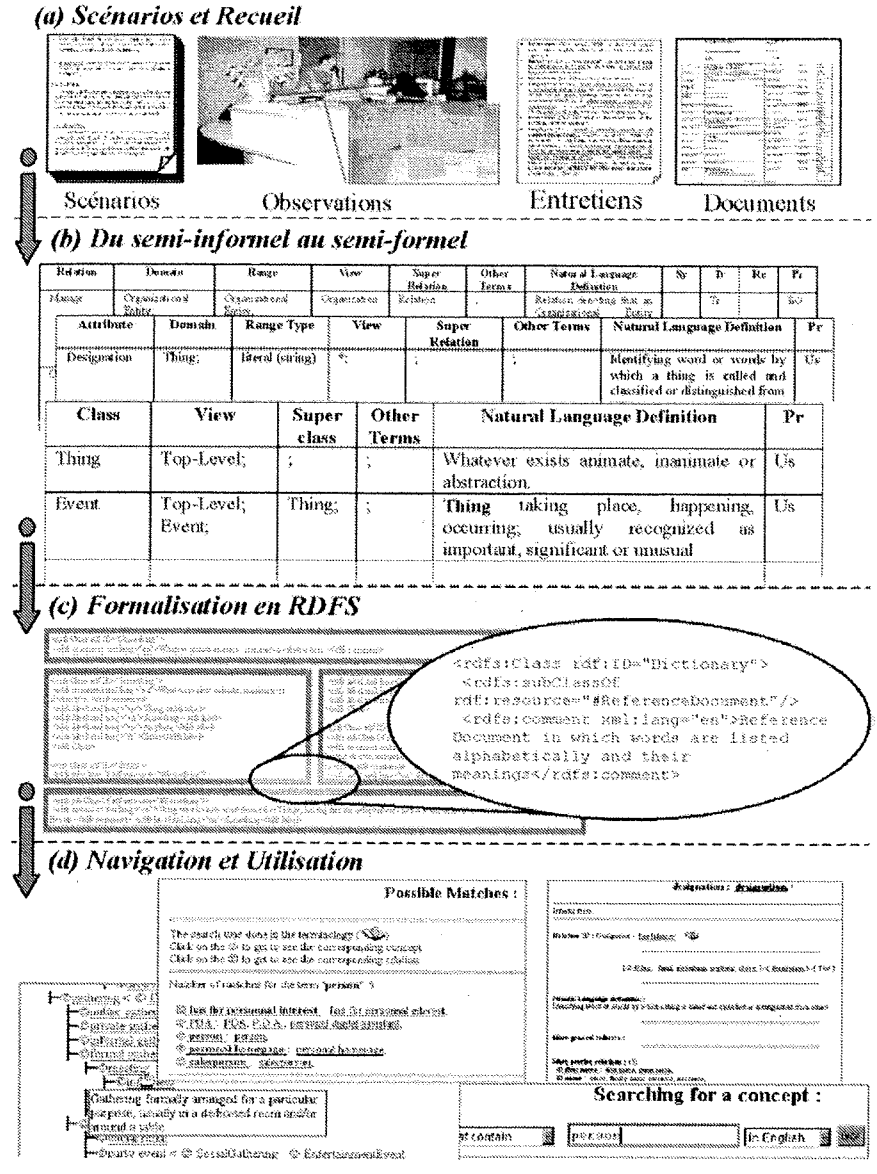

Figure 14. *Etapes de conception de l'ontologie O'CoMMA*

5.2. *Scénarios et recueil*

Construire l'ontologie, c'est identifier les concepts du monde que l'on veut représenter ainsi que leurs relations. Concepts et relations étant accessibles en particulier au travers du langage, il s'agit de recueillir et d'analyser des corpus langagiers représentatifs du monde de l'entreprise. Les corpus que nous avons recueillis et analysés sont de deux types : des corpus propres à l'entreprise et des corpus non spécifiques à l'entreprise. Le recueil et l'analyse de ces corpus ont été guidés par des scénarios d'utilisation.

5.2.1. *Scénarios : guides pour le recueil et l'analyse de corpus*

Les scénarios d'utilisation sont des descriptions textuelles de l'activité organisationnelle. Ils rendent compte de la distribution des rôles et des connaissances dans l'entreprise. C'est une façon naturelle et efficace de capturer les besoins des utilisateurs dans leur contexte [CAR 97]. Les scénarios ont déjà été utilisés dans la construction d'ontologies [GRU 94] [FER 99]. Les principaux avantages identifiés lors de l'utilisation des scénarios pour CoMMA sont :

– se focaliser sur les aspects de gestion des connaissances spécifiques au cas d'application considéré, capturant ainsi la portée du travail effectué ;

– capturer et présenter une image globale permettant de voir le système comme un composant d'une solution possible pour la gestion de connaissances dans une entreprise ;

- rassembler un ensemble concret d'interactions avec la mémoire d'entreprise, compréhensible pour tous les intéressés et fournir un excellent point de départ pour construire des « use cases »
- et autres diagrammes d'interactions ;

– fournir un cadre pour évaluer et contrôler chaque nouvelle idée.

L'analyse par scénario nous a amenés à définir une grille suggérant les principaux aspects à considérer lors de la description d'un scénario. Durant la phase de recueil, cette grille rappelle les informations à rechercher et les points qui pourraient être affinés. Les scénarios aident à définir le domaine de l'ontologie : l'ontologie doit fournir l'expressivité nécessaire aux interactions décrites dans les scénarios. Les rapports de scénarios sont extrêmement riches et donc de très bons candidats à inclure dans le corpus de l'étude terminologique effectuée pour construire l'ontologie. Ils représentent notre première étape de recueil.

5.2.2. *Recueil de corpus propres à l'entreprise*

Trois grandes catégories de corpus d'entreprise ont été recueillis (figure 14-a) :

– documents d'entreprise : documents textuels (lexiques, terminologies, rapports, etc.) et documents graphiques (formulaires, organigrammes...) ;

– transcriptions d'entretiens avec des membres de l'entreprise sur leur rôle ;

– des comptes rendus d'observations menées dans l'entreprise : observations de membres d'entreprise effectuant une tâche particulière dans leur environnement de travail.

Nos techniques de recueil sont détaillées dans [GAN 01]. Nous ne ferons ressortir ici que quelques aspects de distribution des connaissances dans l'entreprise, dont il est nécessaire de rendre compte dans l'ontologie.

Les *documents* officiels d'entreprise reflètent la distribution supposée des connaissances dans l'entreprise. Un exemple pour CoMMA est « la fiche d'itinéraire du nouvel employé » d'un de nos partenaires industriels : elle décrit, à l'aide d'un certain vocabulaire, ce qu'il faut faire, où aller, qui contacter, comment contacter et l'ordre dans lequel ces tâches doivent être effectuées par un nouvel employé. Cependant, le reflet de la distribution des connaissances fourni par les documents officiels n'est pas toujours exact : il ne tient pas compte par exemple des évolutions incessantes du « paysage cognitif ». On peut alors chercher à obtenir un reflet plus réel à partir de l'analyse de l'usage de ces documents (cf. les annotations portées par les employés sur la fiche d'itinéraire).

On peut aussi obtenir un reflet plus exact de la distribution des connaissances à partir des *entretiens* menés auprès des employés. Ainsi l'entretien réalisé auprès d'une jeune secrétaire nouvellement arrivée chez un autre de nos partenaires industriels nous a permis de mettre en évidence le réseau réel de connaissances de cette employée. Ce réseau inclut non seulement les interlocuteurs désignés officiellement par l'entreprise, mais aussi les interlocuteurs officieux que l'employée consulte pour s'intégrer plus rapidement. Le même entretien a permis de mettre en évidence que la perception qu'une personne a de son rôle dans l'entreprise peut être différente de la définition officielle de ce rôle telle qu'on peut la trouver dans une fiche de poste. Ontologiquement, cela conduit à introduire une distinction entre rôle prescrit et rôle réel et à fournir le vocabulaire permettant aux différents intéressés de personnaliser leur profil.

Les *observations* des membres de l'entreprise dans leur environnement quotidien permet d'obtenir un reflet encore plus exact de la distribution des connaissances. En observant ainsi la jeune secrétaire, on a pu voir que celle-ci, en s'appropriant son nouvel espace de travail, avait réorganisé tous les documents qui lui avaient été remis lors de son arrivée et modifié les classifications qui avaient été faites de ces documents. A titre d'exemple, la secrétaire utilise deux critères pour étiqueter son « bac de documents en cours » : le type des documents (ex : « formulaires de congés ») et la tâche à réaliser sur ces documents (ex : « à faire signer par M. X »). Il est important de rendre compte dans l'ontologie des concepts associés à ces pratiques de classification.

D'une manière générale, on pourrait dire que les documents d'entreprise sont déjà le résultat d'un consensus. Les transcriptions d'entretiens et les comptes rendus d'observation reflètent les écarts possibles au consensus. Il nous paraît nécessaire de tenir compte de ces écarts, parce qu'une fois que l'on aura

« distribué » l'ontologie avec les système, il faudra faire face aux écarts qui ne manqueront pas de se produire lors de l'utilisation du système.

Les différents types de corpus que l'on vient de décrire aident à comprendre quels types d'annotations peuvent exister et révèlent tout un vocabulaire lié à l'utilisation et l'organisation de la mémoire documentaire. Le recueil de ces corpus est cependant très consommateur en temps, ce qui empêche de le répéter sur un grand nombre de cas pour permettre une véritable généralisation des résultats.

5.2.3. *Recueil de corpus non spécifiques à l'entreprise*

Pour construire l'ontologie O'CoMMA, nous avons également exploité des sources externes à l'entreprise, en les adaptant en fonction de nos scénarios (ainsi la relation de « propriété » présente dans l'Ontologie d'Entreprise de AIAI [USC 98] n'a pas été car elle n'était jamais exploitée dans nos scénarios).

Nous avons partiellement réutilisé des *ontologies existantes*. Nous avons en particulier étudié l'Ontologie d'Entreprise [USC 98] et celle de TOVE [TOV 00] spécifiques au monde de l'entreprise, ainsi que l'Ontologie Supérieure de Cyc [CYC 00] celle de PME [KAS 00] celle de CGKAT [MAR 96] et WebKB [MAR 00] plus générales, rappelant en cela que l'entreprise de par sa distribution s'inscrit dans le reste du monde. La réutilisation d'ontologies est à la fois séduisante (elle devrait permettre d'économiser du temps et des efforts et favoriser la normalisation) et difficile (les engagements et les conceptualisations doivent être réajustés entre l'ontologie réutilisée et l'ontologie désirée).

Certaines sources très générales, comme des ouvrages sur l'usage de la langue (ex. : le livre « Using Language » de Herbert H. Clark) nous ont aidés à structurer les parties supérieures de certaines branches de notre ontologie. D'autres sources très spécifiques (normes, taxinomies documentaires) nous ont permis de recenser plus rapidement certaines feuilles de l'arbre taxinomique. Par exemple la norme MIME était une excellente source pour la description des formats électroniques et a permis d'énumérer les feuilles descendantes de la notion « format de fichier de données ». Le Dublin Core a permis d'amorcer l'énumération des propriétés utilisées pour l'annotation d'un document dans sa globalité.

Des *lexiques et dictionnaires* existants, en particulier les méta dictionnaires, se sont révélés extrêmement utiles. Les méta dictionnaires permettent un accès simultané à différentes définitions pour identifier ou construire celle qui correspond à l'intention voulue.

L'usage de corpus externes à l'entreprise ne reflète pas uniquement un souci d'efficacité ou de commodité de la part des constructeurs de l'ontologie. Cet usage reflète également certains aspects de la distribution des connaissances dans l'entreprise. Cet usage indique par exemple que certaines connaissances dépassent le cadre de l'entreprise et sont partagées par une communauté plus large ou par des communautés en relation avec la communauté d'entreprise.

5.3. *Phase terminologique*

Chaque terme dénotant un concept jugé utile pour le déroulement d'un scénario est candidat à l'analyse terminologique. Il entraîne avec lui tous les termes dénotant des notions proches et ses synonymes connus susceptibles de participer aux scénarios applicatifs. Si nous considérons ainsi les termes *document*, *rapport* et *compte rendu d'analyse de tendance technologique* impliqués dans les scénarios de veille technologique, leurs candidatures sont très liées. Le point de départ peut être le terme *document* du fait d'une étude de la partie haute d'ontologies existantes, ou le terme *rapport* apparu lors de l'entretien avec la jeune assistante ou enfin le terme *compte rendu d'analyse de tendance technologique* à la suite de l'observation d'une instance particulière de ce document (ex : « compte-rendu d'analyse de tendance technologique intitulé "Capitalisation des expériences du WAP et passage à la génération UMTS" »). Les termes candidats sont rassemblés dans un jeu de tableaux, passant ainsi à une structure de recueil semi-informelle. ACACIA a proposé des définitions en langage naturel pour chacun des termes candidats. La première version a été présentée aux membres du consortium et des extensions de bas niveau ont été faites par les partenaires industriels, par exemple :

*« **Référent de Domaine** : Observateur responsable de l'expertise d'un domaine et qui est chargé de gérer un groupe d'observateurs contributeurs pour ce domaine. »*

L'intérêt d'avoir d'un côté la contribution d'un ingénieur de la connaissance pour la méthodologie et les concepts de haut niveau, et de l'autre celle des industriels pour les concepts spécifiques est flagrant. L'étude terminologique est au cœur de l'ingénierie ontologique et le principal travail sur les termes identifiés est la production des définitions consensuelles qui expriment l'intension de chaque concept et qui assureront la cohérence dans la distribution.

Si à *un terme* correspond une et *une seule notion*, ce terme non ambigu devient le libellé de la notion. Si *plusieurs termes* dénotent *une notion*, les termes sont des synonymes et on garde la liste des synonymes en choisissant le terme le plus communément utilisé comme libellé de la notion. Si à *un terme* correspondent *plusieurs notions*, il y a ambiguïté et différentes expressions d'intension sont définies avec des termes non ambigus (souvent des termes composés) choisis comme libellés internes. Etiqueter des concepts avec des termes est à la fois indispensable et délicat. C'est une source importante de « rechute d'ambiguïté » : nous retombons rapidement et souvent inconsciemment dans l'ambiguïté en utilisant les libellés selon la définition que nous leurs associons naturellement et non pas selon la définition qui leur a été réellement associée au cours de l'engagement sémantique. Cependant, si du point de vue du système, les concepts tirent leur signification de l'expression formelle de leur intension et de leur position dans l'ontologie, il n'en est pas de même du point de vue utilisateur pour qui le libellé linguistique est la représentation privilégiée. La représentation explicite du niveau terminologique et l'existence de fonctionnalités de gestion de ce niveau dans les

outils assistant l'ontologiste et l'utilisateur final sont donc nécessaires pour assurer un déploiement symbiotique du système distribué.

5.4. *Structuration de l'ontologie*

Les concepts obtenus sont organisés en taxinomie. Nous avons commencé à regrouper des concepts premièrement de façon intuitive, puis en raffinant itérativement la structure suivant les principes aristotéliciens étendus [BAC 00]. Ces principes poussent à éliminer le multi-héritage au profit de la multi-instanciation. Ceci nous pose un problème avec l'introduction de « rôles » dans l'ontologie car ces « rôles » ont tendance, à cause de RDFS, à rendre le multi-héritage nécessaire. Une solution serait d'expliciter des points de vue dans la taxinomie [RIB 99] et de limiter l'application de ces principes étendus à l'intérieur d'un point de vue. Une approche serait d'introduire la notion d'axes sémantiques [KAS 00] pour regrouper sous un « genus » les différentes dimensions ou natures de critères suivant lesquelles le « différentia » s'énonce. De même, [GUA 92] et [GUA 00] contribuent à poser les bases fondamentales de l'ingénierie ontologique fournissant un cadre théorique, des définitions et des contraintes basées sur ces dernières permettant de vérifier la taxonomie. Le problème est qu'aucun outil n'est disponible pour aider un ontologiste à effectuer ce travail facilement et indépendamment d'un langage de formalisation ; cela peut devenir un travail titanesque d'appliquer ces théories à de grandes ontologies et ces travaux semblent plus adaptés à la validation de la partie haute de l'ontologie qui, par extension, assure une cohérence minimale dans le reste de l'ontologie.

La façon de concevoir une ontologie est encore sujette à beaucoup de discussions dans la communauté. D'après notre compréhension des différentes contributions, il y a une tendance à distinguer trois options de construction :

– *approche ascendante* : L'ontologie est construite par généralisation en partant des concepts des basses couches taxinomiques. Cette approche encourage la création d'ontologies spécifiques et adaptées ;

– *approche descendante* : L'ontologie est construite par spécialisation en partant de concepts des hautes couches taxinomiques. Cette approche encourage la réutilisation d'ontologies ;

– *approche centrifuge* : La priorité est donnée à l'identification de concepts centraux à l'application que l'on va ensuite généraliser et spécialiser pour compléter l'ontologie. Cette approche encourage l'émergence de domaines thématiques dans l'ontologie et favorise la modularité.

Après expérience, nous ne sommes pas convaincus qu'il existe réellement une approche purement ascendante, descendante ou centrifuge. Pour nous, il s'agit de trois perspectives complémentaires d'une méthodologie complète. Lors de la structuration taxinomique, la spécialisation ou la généralisation d'un concept sont

présentes de façon concourante à chaque niveau de profondeur taxinomique et de granularité (au niveau des concepts ou des groupes de concepts). Quand un ontologiste conçoit une ontologie il doit effectuer les tâches définies dans ces trois perspectives en parallèle.

Pour CoMMA, nous avons effectivement effectué certaines tâches en parallèle dans différentes perspectives :

– nous avons utilisé l'approche descendante quand nous avons étudié des ontologies de haut niveau et les couches hautes d'ontologies existantes pour structurer notre couche supérieure et réutiliser des distinctions déjà existantes, par exemple la distinction chose / entité / situation ;

– nous avons utilisé l'approche centrifuge quand nous avons étudié différentes branches d'ontologies existantes ainsi que les thèmes principaux détectés pendant le recueil pour identifier les domaines de l'ontologie et regrouper les concepts ;

– nous avons utilisé l'approche ascendante quand nous avons exploité les rapports de l'analyse par scénarios et les traces du recueil pour identifier les concepts spécifiques et les regrouper par généralisation.

Les différents candidats dans chaque perspective (concept de haut niveau, concept central, concept spécifique) sont à l'origine de sous-taxinomies partielles. L'objectif est alors de faire se rejoindre ces différentes sous-taxinomies partielles ; chaque approche influence les autres, un événement dans une perspective déclenche des vérifications et des tâches dans les autres.

5.5. *Du semi-informel au semi-formel*

A partir de la terminologie informelle, nous avons séparé les attributs et les relations des concepts et nous avons obtenu trois tableaux (figure 14-b). Ces tableaux ont évolué depuis une simple représentation semi-informelle (tableaux terminologiques terme/notion) vers une représentation semi-formelle (relations taxinomiques, signatures des relations). La logique de conception de l'ontologie devrait être capturée parce qu'elle explique ce qui a motivé sa forme actuelle et ceci peut aider la compréhension, l'adaptation de l'ontologie, voire l'adhésion à celle-ci lors de sa distribution.

De cette approche résulte une première ontologie contenant 420 concepts organisés approximativement en trois couches :

– une partie supérieure très générale qui ressemble plus ou moins à la partie supérieure d'autres ontologies ;

– une couche centrale qui tend à se diviser en deux zones distinctes : 1) une liée au domaine de la mémoire d'entreprise en général (documents, organisation, personnes...) 2) l'autre liée aux sujets du domaine d'application (par exemple pour les télécommunications : technologies sans fil, technologies réseau...) ;

– une couche d'extension qui tend à être spécifique aux scénarios et à l'entreprise, avec des concepts complexes propres à l'organisation (rapport d'analyse de tendances, fiche d'itinéraire du nouvel employé...).

5.6. *Formalisation de l'ontologie*

Le travail de formalisation ne consiste pas à substituer une version formelle à une version informelle. Il s'agit d'augmenter une version informelle avec la correspondance formelle des aspects sémantiques intéressants et pertinents de l'ontologie informelle, afin d'obtenir une ontologie documentée (description en langage naturel éventuellement augmentée par des capacités de navigation introduites par la description formelle) et opérationnelle (description formelle des aspects sémantiques appropriés nécessaires aux opérations du système envisagé).

La version informelle de l'ontologie n'est donc pas une trace intermédiaire qui disparaît après la formalisation. L'ontologie formelle doit inclure les définitions en langue naturelle, les commentaires, les remarques, qui seront exploitées par les personnes essayant de s'approprier l'ontologie. Les ontologies doivent être intelligibles aux ordinateurs comme aux humains [MIZ 97]. Ceci joue un rôle important dans la réutilisabilité des ontologies, la re-conception et la rétro-conception. Une fois la taxonomie soit suffisamment explicite pour être traduite en RDFS (figure 14-c), un jeu de scripts utilisant les libellés non-ambigus nous a permis de basculer de la représentation semi-formelle des tableaux vers une représentation formelle en RDFS incluant tous les aspects de la représentation semi-informelle mais reposant sur la structure forte et explicite donnée par les relations taxinomiques. Ce point de basculement ne change rien à l'ontologie en elle-même, il change uniquement sa représentation interne. En RDFS, les concepts sont représentés comme des classes de ressources et leurs relations et leurs attributs comme des classes de propriétés. Un individu (*i.e.* la réalisation d'un concept) est représenté comme l'instance d'une ou plusieurs classes puisque RDF(S) autorise la multi-instanciation.

La figure 15 montre comment le langage RDFS est utilisé pour implanter les différents niveaux introduits précédemment :

– le *niveau terminologique* où s'organisent les termes recueillis. Les relations entre le niveau intensionnel et le niveau terminologique dénotent les libellés possibles pour chaque intension (propriété `rdfs:label`). Une intension ayant des liens avec différents termes (ex. dans la figure 15 : C_4) est caractéristique de la synonymie de ces termes. Un terme ayant des liens avec différentes intensions (ex. dans la figure 15 : T_3) est caractéristique de l'ambiguïté de ce terme ;

– le *niveau intensionnel* où s'inscrit la structure intensionnelle formalisant l'ontologie. Les relations entre le niveau intensionnel et le niveau extensionnel représentent l'instanciation de chaque classe de concept. Les faisceaux de liens relient une intension à son extension (ex. dans la figure 15 : C_8) ;

– le *niveau extensionnel* où s'organise la mémoire factuelle (les annotations des documents, la description de la situation organisationnelle et les profils utilisateurs). Une extension reliée à plusieurs intensions (ex. dans la figure 15 : C_6 et C_7) est caractéristique de la multi-instanciation.

Grâce à la technologie XML nous pouvons régénérer en permanence les vues informelles ou semi-formelles précédentes en passant par des feuilles de style XSLT (figure 14-d). On peut recréer le tableau terminologique initial, présentant un lexique de la mémoire. On peut rechercher des termes au niveau terminologique et explorer le niveau intensionnel en proposant en permanence de passer de l'un à l'autre. On peut naviguer dans la taxonomie des intensions de concepts ou de relations. Les points d'entrée de l'ontologie peuvent être filtrés par un profil utilisateur pour s'adapter aux intéressés. Un échantillon d'extensions peut être généré, jouant le rôle d'exemplification automatique et améliorant la compréhension du concept. Une dernière feuille propose un arbre indenté des intensions. L'aspect terminologique permet une meilleure navigation ainsi qu'un accès multilingue à la même structure des intensions et la même mémoire des extensions. Les feuilles de style sont aussi utilisées pour documenter le résultat des requêtes en utilisant les définitions et les termes associés aux intensions.

Enfin, comme nous le notions au début de cette section, la formalisation a ses limites, et celles de RDF(S) sont très vites atteintes. La principale limitation est l'impossibilité d'associer une définition formelle à une notion. Or dans un contexte distribué, une définition formelle peut permettre d'adapter le comportement du système au contexte d'utilisation. Reprenons la notion de « collègue ». A l'origine « un collègue » était un concept. Puis en regardant son utilisation ce concept est devenu une relation « x est collègue de y ». Enfin en considérant le scénario applicatif, on s'aperçoit que cette relation n'est pas utilisée pour l'annotation (personne n'envisagerait d'énumérer dans une annotation tous les collègues d'une personne). Cette relation est en fait systématiquement inférée du modèle d'entreprise : « si x et y sont dans la même équipe alors ils sont collègues ». RDF ne permettant pas la définition formelle de cette relation, nous avons dû introduire une règle donnée en figure 13 dans un nouveau langage XML créé pour RDF et CORESE. A terme il faudrait envisager l'implantation d'une extension à RDF(S) comme proposé par [DEL 01] afin de permettre la factorisation des connaissances dans l'ontologie sous forme de règles de définition ce qui serait très utiles dans notre cas.

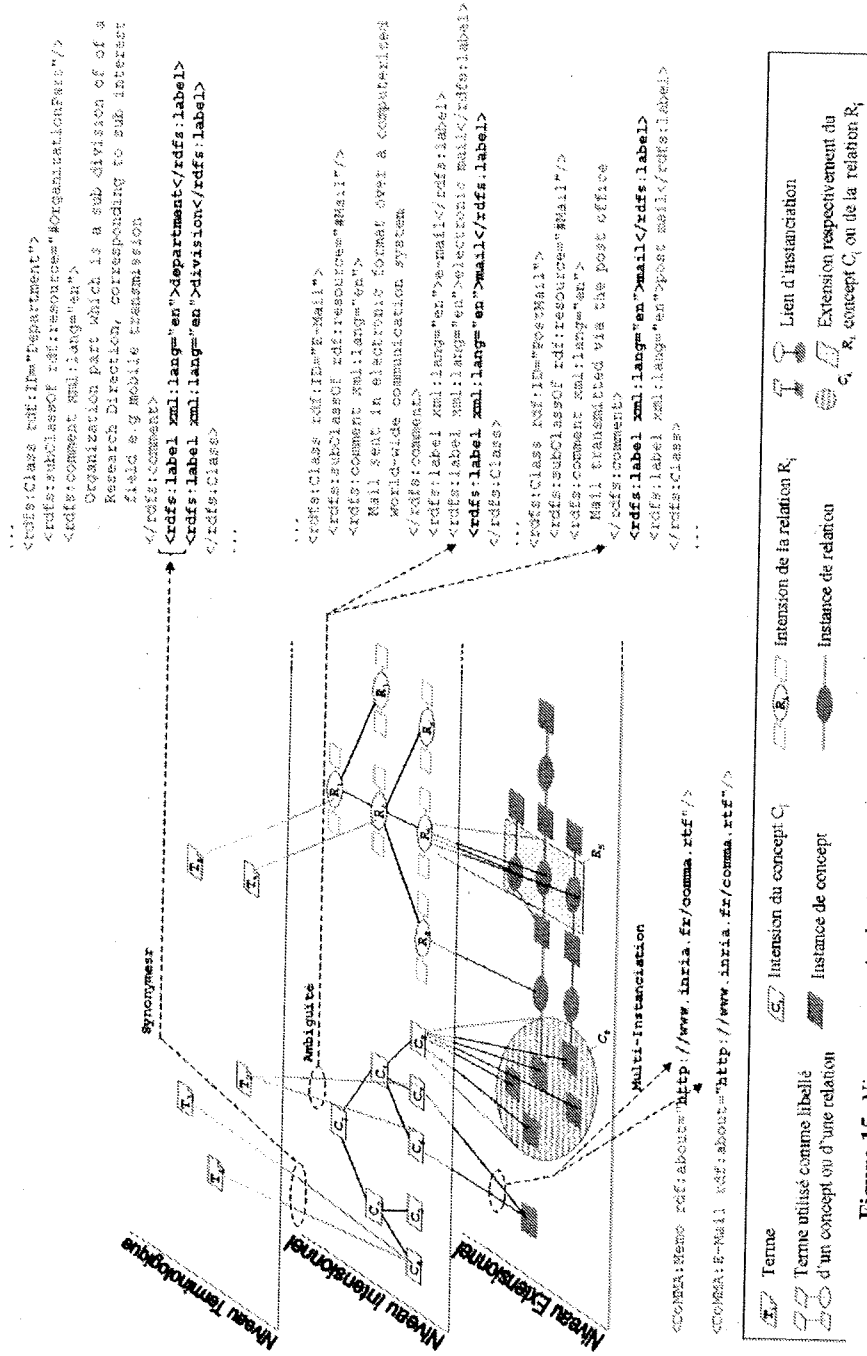

Figure 15. *Niveaux terminologique, intensionnel et extensionnel dans la représentation d'une ontologie*

6. Conclusion

Dans cet article, nous avons présenté une vision de la mémoire d'entreprise ainsi que l'approche évaluée dans le projet CoMMA basée sur les systèmes multi-agents, l'apprentissage symbolique et le web sémantique. Nous avons expliqué notre approche pour la conception de l'architecture du système d'information multi-agents en soulignant l'intérêt de multiples configurations à partir d'une même architecture. Nous avons tout d'abord présenté notre décomposition de la société d'agents de CoMMA en différentes sous-sociétés et les différentes organisations possibles pour ces sous-sociétés. A partir de cette analyse, nous avons montré que nous pouvions identifier les rôles des agents pour ensuite analyser leurs caractéristiques et leurs interactions. A ce stade l'implantation des comportements est possible en intégrant des solutions techniques issues de plusieurs autres domaines de recherche. Nous avons montré que l'ontologie est une clef de voûte d'un tel système d'information et nous avons détaillé chaque étape de la conception de l'ontologie O'CoMMA. Nous avons notamment expliqué notre démarche de recueil et l'utilité de la définition de scénarios d'utilisation. Nous avons ensuite décrit comment d'autres ontologies et d'autres sources d'expertise ont influencé notre travail en soulignant qu'une réutilisation immédiate de ces autres ontologies n'était pas possible. Nous avons détaillé le travail sur la terminologie en insistant sur la nécessité de préserver cet aspect dans l'ontologie. Enfin nous avons exposé comment se sont déroulés le développement et la formalisation de l'ontologie par structuration itérative et comment RDF et les feuilles de style nous permettent de représenter et de naviguer dans l'ontologie au travers des niveaux terminologique, intensionnel et extensionnel.

Il reste beaucoup à faire pour proposer des méthodes et des modèles permettant de passer des représentations des aspects pertinents de l'organisation et des activités de l'entreprise aux spécifications techniques (architecture et configuration) d'un SIMA de mémoire d'entreprise ; ces travaux sont notre première contribution à une méthodologie complète. De plus, tout au long de notre travail sur l'ontologie, nous avons fait appel à de multiples outils informatiques ou conceptuels ; il ne fait aucun doute que l'ingénierie ontologique exige une convergence de toutes ces contributions vers un environnement intégré permettant de suivre tout le cycle de vie de l'ontologie.

Remerciements

Nous remercions le consortium CoMMA (Atos-Origin, CSELT/Telecom Italia, CSTB, INRIA, LIRMM, T-Nova/Deutsche Telekom, l'Université de Parme) pour les discussions fructueuses que nous avons eues. Nous remercions la Commission Européenne qui subventionne ce projet IST (1999-12217).

7. Bibliographie

[AUM 00] Agent Unified Modelling Language http://www.auml.org.

[AUS 00] AUSSENAC-GILLES N., BIEBOW B., SZULMAN S., Revisiting Ontology Design: a Method Based on Corpus Analysis, *Actes EKAW'2000*, Juan-les-Pins, Springer LNAI, 1937, p. 172-188, 2000.

[BAC 00] BACHIMONT B., Engagement sémantique et engagement ontologique: conception et réalisation d'ontologies en ingénierie des connaissances, *in* Charlet J., Zacklad M., Kassel G., Bourigault D., *Ingénierie des connaissances Evolutions récentes et nouveaux défis*, Eyrolles, ISBN 2-212-09110-9, 2000.

[BAC 01] BACHIMONT B., « Modélisation linguistique et modélisation logique des ontologies » l'apport de l'ontologie formelle, *Actes IC'01* p. 349-368, 2001.

[BEL 01] BELLIFEMINE F., POGGI A., RIMASSA G., Developing multi agent systems with a FIPA-compliant agent framework, *Software Practice et Experience*, 31:103-128, 2001.

[BER 00] BERGENTI F., POGGI A., « Exploiting UML in the Design of Multi-Agent Systems », in Omicidi A., Tolksdorf R., Zambonelli F., *Engineering Societies in the Agents World*, Springer LNAI, vol. 1972, p. 106-113, 2000.

[BER 01] BERNERS-LEE T., HENDLER J., LASSILA O., « The Semantic Web », in *Scientific American*, p. 35-43, May 2001.

[CAR 97] CAROLL J. M., Scenario-Based Design, *in* M. Helander, T.K. Landauer, P. Prabhu, *Handbook of Human-Computer Interaction*, 2nd edition, Ch. 17, Elsevier Science, 1997.

[COM 00] CoMMA, « Corporate Memory Management through Agents », *Actes de la Conférence E-Work & E-Business*, Madrid, Octobre 2000, p. 383-406, 2000.

[COR 00] CORBY O., DIENG R., HEBERT C., « A Conceptual Graph Model for W3C Resource Description Framework », *Actes ICCS'2000*, Springer LNAI 1927, p. 468-482, 2000.

[CYC 00] Ontology http ://www.cyc.com/cyc-2-1/cover,html Cycorp Inc. Cyc®.

[DEL 01] DELTEIL A., FARON C., DIENG R., « Extension of RDFS based on the CG formalism », in *Proc. ICCS'01*, p. 275-289.

[DIE 00] DIENG R., CORBY O., GIBOIN A., GOLEBIOWSKA J., MATTA N., RIBIERE M., *Méthodes et outils pour la gestion des connaissances*, Dunod, 2000.

[DOY 98] DOYLE P., HAYES-ROTH B., « Agents in Annotated Worlds », *Actes 2nd Annual Conference on Autonomous Agents*, Minneapolis, ACM, p. 173-180, 1998.

[ETZ 95] ETZIONI O., WELD D., Intelligent Agents on the Internet: Fact, Fiction and Forecast, University of Washington, *IEEE Expert/Intelligent Systems & Their Applications* vol. 10, No. 4, 1995.

[FER 97] FERBER J., GUTKNECHT O., Aalaadin : a meta-model for the analysis and design of organizations in multi-agent systems, *Rapport de Recherche LIRMM 97189*, 1997.

[FER 99] Fernandez-Lopez M., Overview of Methodologies for Building Ontologies, *Actes IJCAI-99 Workshop on Ontologies and Problem-Solving Methods: Lessons Learned and Future Trends*, 1999.

[FIP 01] FIPA -Foundation for Intelligent Physical Agents - http://www.fipa.org.

[FRA 96] FRANKLIN S., GRAESSER A., Is it an Agent, or just a Program ? : A Taxonomy for Autonomous Agents, *Actes 3rd International Workshop on Agent Theories, Architectures and Languages*, Springer LNAI, 1996.

[GAN 00] GANDON F., DIENG R., CORBY O., GIBOIN A., « A Multi-Agents System to Support Exploiting an XML-based Corporate Memory », *Actes 3rd International Conference on Practical Aspects of Knowledge Management*, Basel, p. 10-1 to 10-12, 2000.

[GAN 01] GANDON F., DIENG R., « Ontologie pour un système multi-agents dédié à une mémoire d'entreprise », *Actes IC'01*, p. 1-20, 2001.

[GUA 92] GUARINO N., « Concepts, Attributes, and Arbitrary Relations: Some Linguistic and Ontological Criteria for Structuring Knowledge Bases », in *Data and Knowledge Engineering* 8: p. 249-261, 1992.

[GUA 95] GUARINO N., GIARETTA P., Ontologies and Knowledge Bases: Towards a Terminological Clarification, in Mars, *Towards very large knowledge bases*, IOS Press, 1995.

[GUA 00] GUARINO N, WELTY C, Towards a methodology for ontology-based model engineering, *Actes ECOOP-2000 Workshop on Model Engineering*, Cannes, 2000.

[GRU 94] GRUNINGER M., FOX M.S., The Role of Competency Questions in Enterprise Engineering , *Actes IFIP WG5.7 Workshop on Benchmarking - Theory and Practice*, Trondheim, Norway, 1994.

[KAS 00] KASSEL G., ABEL M.-H., BARRY C., BOULITREAU P., IRASTORZA C., PERPETTE S., « Construction et exploitation d'une ontologie pour la gestion des connaissances d'une équipe de recherche », *Actes IC'2000*, Toulouse, p. 251-260, 2000

[KEN 99] KENDALL E., Role Modeling for Agent System Analysis, Design, and Implementation, *Actes ASA'99*, Symposium on Mobile Agents (MA'99), Palm Springs. 1999.

[KIS 01] KISS A., QUINQUETON J., Multiagent Cooperative Learning of User Preferences, *Actes European Conference on Machine Learning Principles and Practice of Knowledge Discovery in Databases*, Freiburg, Germany, 2001.

[KLU 99] KLUSH M., Introduction of Part 1, *Intelligent Information Agent: Agent-Based Information Discovery and Management on the Internet*, Springer, p. 3-9, 1999.

[MAR 96] MARTIN P., Exploitation de Graphes Conceptuels et de documents structurés et Hypertextes pour l'acquisition de connaissances et la recherche d'informations, Thèse en informatique, Univ. Nice Sophia-Antipolis, 1996.

[MAR 00] MARTIN P., EKLUND P., « Knowledge Retrieval and the Word Wide Web » *IEEE Intelligent Systems on Knowledge Management and the Internet*, p. 18-25, 2000.

[MIZ 97] MIZOGUCHI RIICHIRO, IKEDA MITSURU, *Towards Ontology Engineering*, in *Actes Joint 1997 Pacific Asian Conference on Expert systems / Singapore International Conference on Intelligent Systems*, p. 259-266, 1997.

[NWA 96] NWANA H.S., « Software agents: an overview », *The Knowledge Engineering Review*, vol. 11:3, p. 205-244, 1996.

[RIB 99] RIBIERE M., Représentation et gestion de multiples points de vue dans le formalisme des graphes conceptuels, Thèse en informatique Univ. Nice Sophia-Antipolis, 1999.

[STE 93] STEELS L., « Corporate Knowledge Management », in Barthès J. P. (ed.), *Proc. of ISMICK'93*, p. 9-30, Compiègne, 1993.

[SOW 84] Sowa, J.F. *Conceptual Structures – Information Processing in Mind and Machine*, Addison-Wesley Readings, Massachusetts, 1984.

[TOV 00] http://www.eil.utoronto.ca/tove/ontoTOC.html Enterprise Integration Laboratory (EIL), Department of Industrial Engineering University of Toronto.

[USC 96] USCHOLD M. ET GRUNINGER M. Ontologies: Principles, methods and applications, *Knowledge Engineering Review*, 11(2):93-136, 1996.

[USC 98] USCHOLD M., KING M., MORALEE S., ZORGIOS Y., « The Enterprise Ontology », in Uschold, Tate, *The knowledge engineering review special issue* vol. 13, 1998.

[WOO 95] WOOLDRIDGE M., JENNINGS N.R. Intelligent Agents: Theory and Practice, *in The Knowledge Engineering Review*, 10 (2): 115-152, 1995.

[WOO 00] WOOLDRIDGE M., JENNINGS N., KINNY D., The Gaia Methodology for Agent-Oriented Analysis and Design, in *Autonomous Agents and Multi-Agent Systems*, 3 (3): 285-312, 2000.