

Analyse des Dépendances entre Fragments Documentaires pour le Management des Communautés d'Action Médiatisées

N. Prat, M. Zacklad

Département
Institution
Adresse
Code_Postal Lieu
adresse-email@auteur1
adresse-email@auteur2

Résumé: Nous nous intéressons aux situations dans lesquelles les documents sont des supports à la coordination d'une Communauté d'Action répartie engagée dans une activité commune finalisée. Dans ces contextes, le document apparaît comme un ensemble de fragments portés par des auteurs divers et il est possible de recourir à une analyse des traces de l'activité collective laissées sur les documents, pour comprendre notamment l'évolution de la dynamique cognitive et sociale de la communauté. Après avoir justifié le choix du cadre conceptuel des Documents pour l'Action (DopA), nous en proposerons une formalisation inédite représentée en UML qui nous fournira une structure pour la représentation conceptuelle des traces d'activités collectives médiatisées. Nous présenterons ensuite la problématique de l'analyse et de la mesure des dépendances que nous utiliserons pour proposer des dépendances particulières adaptées aux DopA. Nous commencerons à l'appliquer à un scénario de rédaction collective et nous discuterons des perspectives ouvertes par notre démarche d'analyse et de mesure des dépendances à partir des traces d'activités collectives médiatisées.

1 Introduction

L'essor du Web dans un contexte Internet ou Intranet a entraîné une multiplication des pratiques collectives centrées sur le document : rédaction collective de contrats et de plans, prise en charge médicale ou médico-sociale par un réseau de santé, conception Open Source pour ne citer que quelques exemples. Toute une série de nouvelles pratiques centrées sur le document sont ainsi en train d'émerger, allant de la rédaction collective d'un document échangé par la messagerie électronique à l'aide d'annotations à la contribution à des forums de discussion en passant par l'usage des Wiki et Blogs.

Nous nous intéresserons particulièrement ici aux situations dans lesquelles les documents sont des supports à la coordination d'une communauté répartie engagée dans une activité commune finalisée, correspondant aux Communautés d'Action [Za03]¹. Dans ces contextes, le document apparaît comme un ensemble de fragments portés par des auteurs divers et dont le contenu final reste largement indéterminé alors même que sa circulation rapide lui fait déjà jouer un rôle majeur d'information, d'aide à la décision et de preuve. L'animation de ces communautés dont les membres sont souvent dispersés géographiquement passe d'abord par une analyse de leur dynamique à la fois sur un plan cognitif (les problèmes métier qu'elles résolvent dans le contexte de processus souvent peu structurés) et social (structuration parfois informelle en terme d'affinités, de compétences, de préoccupations communes).

Or, du fait même de la médiatisation, il est possible de recourir à une analyse des traces de l'activité collective laissées sur les documents, pour comprendre notamment l'évolution de la dynamique cognitive et sociale de la communauté. Ces traces peuvent être issues de sources très diverses et sont d'autant plus faciles à appréhender que le système conserve l'historique des modifications effectuées comme dans les systèmes de Wiki mais également dans les systèmes de traitement de textes qui offrent un mode « révision ». Par exemple, dans le cas d'école que nous considérerons, nous travaillerons sur la rédaction collective d'un fichier faisant l'objet de divers échanges d'e-mail dont on considérera que les versions successives ont été archivées.

Notre objectif est double :

1. Définir un modèle formel des traces documentaires. Ce modèle s'inspire des recherches effectuées dans le domaine de la traçabilité des systèmes d'information [ZP01], et se fonde sur le cadre conceptuel des Document pour l'Action [Za04].
2. Ensuite, sur la base de ce modèle, expliciter des dépendances entre les fragments du document et entre auteurs et lecteurs de ces fragments. Ces résultats permettront d'analyser et mesurer ces dépendances, cette analyse pouvant déboucher sur de nombreuses applications pratiques. L'analyse des dépendances peut notamment permettre de restituer la dynamique collective au sein de la communauté afin de faciliter l'analyse de son activité et sa régulation, par un manager, un animateur ou par les participants eux-mêmes, une perspective qui sera développée dans un deuxième temps de la recherche.

¹ Différents types de communautés sont mises en évidence dans la littérature en CSCW, en sociologie et en management : communauté de pratique, communauté d'intérêt, communauté épistémique, etc.

Nous suivrons le plan suivant. Après une justification du choix du cadre conceptuel des Documents pour l'Action (DopA), nous en proposerons une formalisation inédite représentée en UML [Ob06] qui nous fournira une structure pour la représentation conceptuelle des traces d'activités collectives médiatisées. Nous présenterons ensuite la problématique de l'analyse des dépendances que nous utiliserons pour proposer des dépendances particulières adaptées aux DopA. Nous commencerons à l'appliquer à un scénario de rédaction collective et nous discuterons des perspectives ouvertes par notre démarche d'analyse et de mesure des dépendances à partir des traces d'activités collectives médiatisées.

2 Vers une modélisation des Documents pour l'Action

2.1. Le DopA : production sémiotique associée à une transaction communicationnelle

Pour analyser les systèmes d'information de nature documentaire qui permettent la coordination des communautés médiatisées dans les phases de conception (par ex. rédaction) et d'exploitation (par ex. lecture) associés aux usages très divers d'Internet (circulation accélérée des fichiers via la messagerie, gestionnaire de contenu pour les sites web, forum de discussion, annotation en ligne, weblogs...), le cadre conceptuel des documents pour l'action (DopA) offre plusieurs avantages. Alors que la conceptualisation en termes d'hypermédia visait surtout à rendre compte des nouvelles pratiques de lecture associées aux hypertextes, le DopA vise à rendre compte des processus « d'hyper-rédaction » associés aux documents numérisés qui apparaissent eux-mêmes comme constitués d'une multiplicité de fragments documentaires mouvants réalisés par différents auteurs.

Le DopA permet ainsi de redéfinir le concept de document et d'appréhender autrement ses contenus qui relèvent de moins en moins de la catégorie du texte classique pour inclure des images et des sons accessibles à partir de systèmes de navigation de nature de plus en plus variée. En insistant sur la dimension collective de l'activité rédactionnelle, il permet d'analyser les documents comme relevant de processus de communication pour partie différés, au sens des processus asynchrones décrits dans le champ du CSCW, entre des producteurs et des récepteurs liés par des intérêts communs. Cette vision communicationnelle du document s'inspire elle-même de l'approche en termes de transactions communicationnelles symboliques dans laquelle le document est analysé comme étant l'objet d'une transaction entre des acteurs impliqués dans un processus d'échange visant à la fois des engagements mettant en jeu leur « self » et des connaissances liées à la production d'une « œuvre » au moins pour partie commune [Za04].

Dans la description d'un échange en termes de transactions communicationnelles, on considère qu'un self « réalisateur » et un self « bénéficiaire » ont une relation qui est médiatisée par des productions sémiotiques qu'ils coproduisent. La décomposition de la production sémiotique amène à établir une distinction entre, d'une part, la forme et le support de cette production ou médium de la production sémiotique et d'autre part, le contenu sémiotique « véhiculé » par ce médium. Une production sémiotique est donc l'association d'un contenu sémiotique, par exemple l'histoire de Blanche Neige et d'un médium, qui se décompose lui-même en une modalité d'expression (récit oral, texte écrit, film) et un type de support (présence du narrateur ou enregistrement magnéto pour le récit oral, papier ou électronique pour le texte, DVD ou VHS pour le film, etc.). Bien sur le choix du médium aura une influence sur le « pouvoir d'évocation » associé au contenu sémiotique.

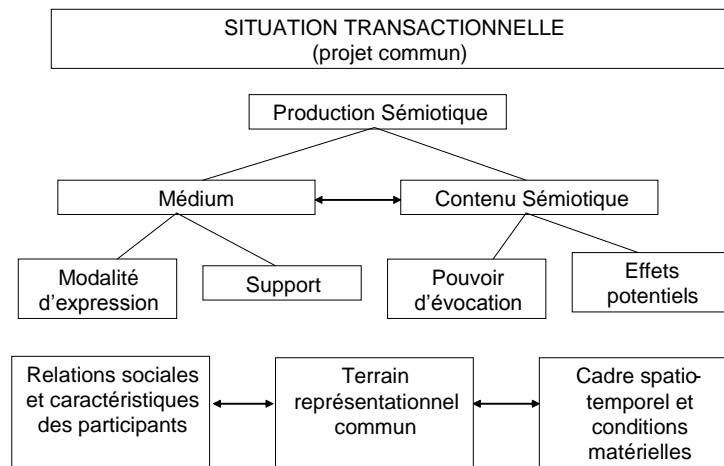


Figure 1: Analyse de la production sémiotique [Za04]

Mais, du fait de la distribution des activités collectives, réalisateurs et bénéficiaires des transactions communicationnelles peuvent ne pas être présents dans le même cadre spatio-temporel. Cela implique de donner une forme pérenne aux objets de la transaction de manière à ce qu'elle puisse être initialisée, interrompue, réactualisée, répétée dans toutes les configurations de présence/absence du bénéficiaire et du réalisateur. Une des stratégies permettant de pallier à la présence physique est la stratégie de documentarisation qui s'appuie sur la transcription ou l'enregistrement de la production sémiotique sur un support pérenne. Mais la transcription ou l'enregistrement sont des conditions nécessaires mais non suffisantes de la documentarisation. Une prise de note personnelle ou un enregistrement ponctuel visant à faciliter une activité sémiotique dans une situation transactionnelle donnée pourra avoir une utilité locale certaine. Cependant, s'ils ne font pas l'objet d'un investissement documentaire, ces supports ne pourront pas être ré-exploités dans des situations transactionnelles différées et distribuées.

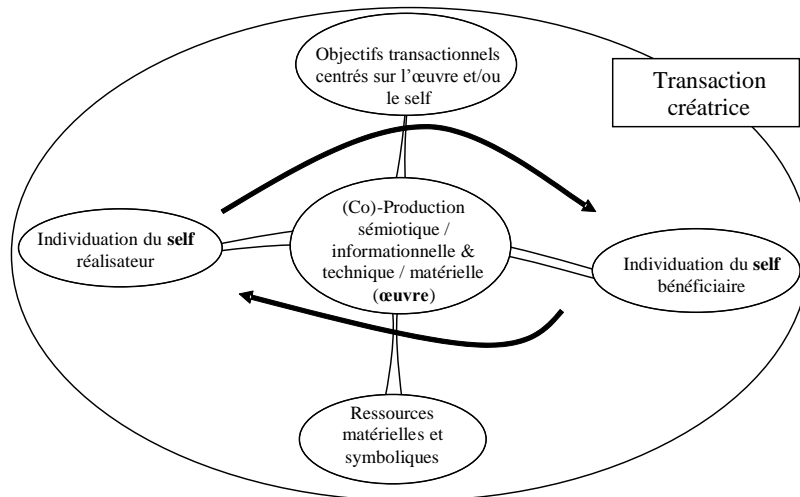


Figure 2: Composantes d'une transaction créatrice dans la théorie transactionnelle de l'action [Za06b]

L'investissement documentaire, ou documentarisation de la production sémiotique, consiste à équiper le supports d'attributs spécifiques visant à faciliter les pratiques liées à son exploitation ultérieure dans le cadre de la préservation de transactions communicationnelles réparties. Ces attributs (nom de l'auteur, date, titre, sous-titres, version...) doivent permettre au document de circuler à travers l'espace, le temps, les communautés d'interprétation, pour tenter de prolonger les transactions communicationnelles initiées par ses réalisateurs. Quand le document est écrit par un collectif réparti dans l'espace et le temps, il est constitué de multiples fragments eux-mêmes dotés d'attributs permettant leur articulation. Ces documents fragmentaires constituent des DopA.

Selon [Za04], les exemples de DopA sont très divers et correspondent à de nombreux contextes professionnels. Citons, sans être exhaustif, les documents de conception en ingénierie (mécanique, logiciel...), les dossiers patients en médecine, les documents contractuels dans un contexte d'affaires qui passent du stade de proposition commerciale au stade de contrat en bonne et due forme, les dossiers qualité de plus en plus souvent numérisés, les rapports d'étude dans le conseil en management, les forum d'échange dans le domaine, par exemple, du logiciel libre [RS06].

Les caractéristiques principales des DopA sont les suivantes :

1. Inachèvement prolongé : ils possèdent un statut d'inachèvement prolongé pendant la phase active de la production sémiotique collective, phase durant laquelle ils sont qualifiés de DopA en évolution (vs DopA stabilisés) ;
2. Pérennité : ils possèdent une certaine pérennité associée aux engagements des acteurs à l'égard des contenus sémiotiques véhiculés et à la distribution des transactions, d'où des enjeux liés à leur documentarisation et corollairement à leur stockage et à leur indexation ;

3. Fragmentation : au moins pendant leur phase évolutive, ils articulent plusieurs fragments plus ou moins liés sémantiquement (notamment des annotations) qui ne peuvent être intégrés mécaniquement ou de manière organique à l'ensemble du document (cf. infra) ;
4. Rapport non trivial aux réalisateurs : les différentes parties du DopA sont le plus souvent produites par différents réalisateurs (plurivocité, pluri-textualité...) qui peuvent avoir des statuts variables dans la situation transactionnelle et donc vis-à-vis de la production sémiotique (droits, engagement, responsabilités...) ;
5. Rapport argumentatif non trivial aux autres parties du document : chaque partie possède une relation potentiellement complexe aux autres – modalisation, incertitude, relation argumentative vis-à-vis des autres fragments...

2.2. Relation des fragments avec le reste du DopA

Dans le contexte des transactions communicationnelles réparties médiatisées par des DopA, les contributions sont médiatisées par des fragments qui seront liés avec plus ou moins de succès à la production sémiotique principale (par ex. le texte principal). Le fragment passe par différents états (statuts) qui correspondent à son degré d'intégration vis-à-vis du « corps » du document. Si la liaison du fragment avec le reste du contenu sémiotique est inexistante ou faible celui-ci constitue un *fragment libre*. Quand le fragment est doté d'attributs qui explicitent ses liens avec d'autres fragments ou avec la production sémiotique principale, celui-ci est un *fragment documentarisé* ou annotation contributive. Quand le fragment est complètement intégré à la production sémiotique principale, par des liaisons sémantiques explicites (utilisation d'annotations index, cf. infra) et implicites (relations de présuppositions internes au « texte »), celui-ci constitue une partie intégrante du contenu sémiotique principal du DopA et à ce titre est désigné comme partie intégrée du DopA (*partie de document*).

Les attributs ou liens qui permettent la documentarisation du fragment libre, son ancrage sur le DopA, peuvent également être considérées comme de petits fragments auxiliaires et à ce titre comme des annotations. On les qualifie d'annotations index [Za06a], afin de les différencier des annotations contributives. L'annotation index ne contribue pas directement au contenu sémiotique principal mais vise à compléter un document ou un fragment pour faciliter son articulation avec d'autres documents (externe) ou avec d'autres parties ou fragments du document (interne). Les annotations index, qui correspondent aux attributs ou liens permettant la documentarisation, sont de nature très diverse : nom de l'auteur (réalisateur) du fragment, date, relation graphique matérialisant une relation sémantique, lieu de réalisation, mots-clés issus d'un thésaurus, concept appartenant à une ontologie, etc. Le schéma UML suivant correspond au modèle formel des transactions médiatisées par des DopA basé sur ce cadre conceptuel.

L'analyse des dépendances et des mécanismes de coordination est une problématique cruciale dans de nombreuses disciplines, parmi lesquels le management [Mi82] et l'informatique. A partir d'une analyse pluridisciplinaire, T.Malone et K.Crowston [MC94] [Cr03] proposent une typologie des dépendances et des mécanismes de coordination. La coordination est définie elle-même comme le management des dépendances. Les auteurs identifient notamment les types de dépendances suivants :

1. la dépendance de flux (une tâche utilise une ressource créée par une autre tâche)
2. la dépendance de partage (plusieurs tâches utilisent la même ressource)
3. la dépendance d'ajustement (plusieurs tâches créent la même ressource)
4. la dépendance de composition entre tâches (décomposition d'une tâche en sous-tâches).

La caractérisation des dépendances permet notamment de choisir les mécanismes de coordination et les outils (de groupware en particulier) les plus adaptés, en fonction du type de dépendance.

En systèmes d'information, [YM93] propose un modèle des dépendances entre acteurs. Ce modèle fait partie d'un cadre plus global visant à assister le processus d'ingénierie des besoins ; il s'applique également à la reconfiguration des processus. Une dépendance est un lien orienté entre deux acteurs. Elle concerne un objet particulier (but, tâche ou ressource), ce qui permet de typer plus finement les dépendances. Le modèle permet de caractériser la force des dépendances, cependant les auteurs ne définissent pas de métriques pour quantifier celle-ci. La dépendance entre agents est également centrale dans [Su00], qui propose une méthode d'ingénierie des besoins pour la conception des systèmes sociotechniques. Les dépendances entre agents sont étudiées afin de déterminer le degré de couplage entre agents. Des métriques de couplage sont appliquées pour évaluer le degré de dépendance entre les systèmes et les agents (utilisateurs). Ces métriques sont utilisées pour arbitrer entre différents scénarios de systèmes.

En génie logiciel, de nombreuses métriques ont été proposées afin d'évaluer la cohésion des modules. La cohésion (modularité) favorise la réutilisation et facilite la maintenance. Dans le cas de l'approche objet, de nombreuses métriques ont été proposées pour mesurer la cohésion des classes [Et04]. [Wa05] propose une telle métrique, qui s'appuie sur une typologie des dépendances entre éléments (méthodes et attributs d'une classe). L'approche consiste à construire le graphe des dépendances entre éléments d'une classe, à calculer la matrice carrée des dépendances (directes et indirectes) entre chaque couple d'éléments de la classe, puis à en déduire la mesure de cohésion globale de la classe. Enfin, dans le domaine des bases de données, les données peuvent être stockées en mémoire en fonction de l'analyse des dépendances entre celles-ci. Cela permet d'améliorer les performances d'accès aux données, en particulier dans les applications navigationnelles et multimédia [Va03].

Dans notre approche, l'analyse et la mesure des dépendances se fonde sur les DoPA, représentés à l'aide du modèle décrit à la section précédente. Une dépendance est un lien orienté d'un objet A vers un objet B dont il dépend. Pour le management des communautés médiatisées, les dépendances les plus pertinentes sont celles qui concernent les fragments et les selfs. Il existe ainsi trois types fondamentaux de dépendances : (1) les dépendances entre deux fragments, (2) les dépendances entre un fragment et un self, et enfin (3) les dépendances entre deux selfs.

Notre approche s'appuie sur une typologie des dépendances, définie à partir du modèle de DoPA. La démarche d'analyse et mesure des dépendances pour un DoPA consiste alors en les étapes suivantes :

1. Explicitation dans le DoPA de tous les liens (annotations index) demeurés implicites (jusqu'alors). Cette étape concerne essentiellement l'explicitation des liens de traçabilité, si ceux-ci n'ont pas pu être déterminés de manière automatique.
2. Génération du graphe de dépendances pour le DoPA. A partir du moment où tous les liens ont été explicités, cette génération peut s'effectuer de manière automatique. Dans ce graphe, les nœuds sont les fragments et les selfs individuels ; les arcs représentent les dépendances directes (canoniques) entre deux fragments, ou entre un fragment et un self individuel.
3. A partir de ce graphe, calcul de deux matrices de dépendances : (1) une matrice de dépendances entre selfs individuels et (2) une matrice de dépendances entre fragments (dans le cas de plusieurs versions successives, liées par un lien de traçabilité, seule la version actuelle des fragments est considérée). Les matrices de dépendances sont construites en tenant compte, dans le graphe de dépendances, des dépendances directes et indirectes.

Dans la suite de cette section, nous présentons plus particulièrement la typologie des dépendances à la base de notre approche. Nous présentons un exemple d'application simplifié de cette approche (rédaction collaborative d'un article pour un workshop).

3.1 Typologie des dépendances

Nous distinguons deux types de dépendances : (1) les dépendances canoniques (ou directes), définies à partir de modèle de DoPA présenté à la section 2, et (2) les dépendances indirectes, obtenues par combinaison des dépendances canoniques.

Dépendances canoniques

Celles-ci sont définies à partir des liens existant dans le DoPA entre deux fragments, ou entre un fragment et un self (Tableau 1). Dans ce tableau, la représentation de chaque dépendance est accompagnée de sa définition formelle au moyen du langage de contraintes OCL (Object Constraint Language) associé à UML [Ob06]. Cette expression formelle permet d'établir le lien avec le modèle de DoPA décrit précédemment.

DÉPENDANCES ENTRE DEUX FRAGMENTS	
Dépendance	Expression
F1 $\xrightarrow{\text{référence}}$ F2	dopA.annotations→exists(LR:Annotation index LR.ocIsTypeOf(Lien de référence) and LR.source=F1 and LR.cible=F2)
F1 $\xrightarrow{\text{provient de}}$ F2	dopA.annotations→exists(LT:Annotation index LT.ocIsTypeOf(Lien de traçabilité) and LT.source=F1 and LT.cible=F2)
F1 $\xrightarrow{\text{composé de}}$ F2	dopA.annotations→exists(LC:Annotation index LC.ocIsTypeOf(Lien de composition) and LC.composite=F1 and LC.composant=F2)
F1 $\xrightarrow{\text{compose}}$ F2	dopA.annotations→exists(LC:Annotation index LC.ocIsTypeOf(Lien de composition) and LC.composant=F1 and LC.composite=F2)
DÉPENDANCES ENTRE UN FRAGMENT ET UN SELF	
Dépendance	Expression
F1 $\xrightarrow{\text{réalisé par}}$ S1	dopA.réalisations→exists(Réa:Réalisation Réa.production=F1 and ((Réa.réalisateur.ocIsTypeOf(Self individuel) and Réa.réalisateur=S1) or (Réa.réalisateur.ocIsTypeOf(Self collectif) and Réa.réalisateur.constitué de→includes(S1)))
S1 $\xrightarrow{\text{récepteur de}}$ F1	dopA.réceptions→exists(Réc:Réception Réc.production=F1 and ((Réc.récepteur.ocIsTypeOf(Self individuel) and Réc.récepteur=S1) or (Réc.récepteur.ocIsTypeOf(Self collectif) and Réc.récepteur.constitué de→includes(S1)))

Tableau 1 : Dépendances canoniques

Parmi les liens entre deux fragments, seul le lien de composition se traduit par deux liens de dépendance. Le lien de référence (dont le lien de traçabilité est un cas particulier) se traduit par une dépendance de la source vers la cible.

Pour les dépendances entre un fragment et un self, deux types de dépendances sont définies, par analogie avec les dépendances entre les méthodes et les attributs en lecture ou en écriture dans DMC [Wa05].

Dépendances indirectes

Celles-ci sont définies par combinaison des dépendances canoniques, en fonction de règles de combinaison. En effet, la dépendance n'est pas toujours transitive. Les règles de combinaison dépendent notamment de la sémantique des dépendances considérées. A titre d'exemple, le Tableau 2 indique quelques règles permettant de déduire des dépendances à partir des dépendances canoniques.

DÉPENDANCES ENTRE DEUX FRAGMENTS	
Obtenues par combinaison des dépendances canoniques entre fragments.	
DÉPENDANCES ENTRE DEUX SELFS	
Si	Alors
S1 $\xrightarrow{\text{récepteur de}}$ F1 et F1 $\xrightarrow{\text{réalisé par}}$ S2	S1 $\xrightarrow{\text{dépend de}}$ S2
S1 $\xrightarrow{\text{récepteur de}}$ F1 et F1 $\xrightarrow{\text{provient de}}$ F2 et F2 $\xrightarrow{\text{réalisé par}}$ S2	S1 $\xrightarrow{\text{dépend de}}$ S2
S1 $\xrightarrow{\text{récepteur de}}$ F1 et F1 $\xrightarrow{\text{composé de}}$ F2 et F2 $\xrightarrow{\text{réalisé par}}$ S2	S1 $\xrightarrow{\text{dépend de}}$ S2
F1 $\xrightarrow{\text{réalisé par}}$ S1 et F1 $\xrightarrow{\text{réalisé par}}$ S2	S1 $\xrightarrow{\text{dépend de}}$ S2 et S2 $\xrightarrow{\text{dépend de}}$ S1
S1 $\xrightarrow{\text{dépend de}}$ S2 et S2 $\xrightarrow{\text{dépend de}}$ S3	S1 $\xrightarrow{\text{dépend de}}$ S3

Tableau 2 : Dépendances indirectes

Concernant les dépendances entre deux fragments, on peut obtenir des dépendances indirectes en combinant des dépendances canoniques, même si elles ne sont pas de même type (par exemple, si un fragment F1 référence un fragment F2 lui-même composé d'un fragment F3, alors F1 dépend de F3). L'algorithme doit cependant éviter les références circulaires, notamment en suivant les dépendances de composition dans un seul sens (composant vers composé ou l'inverse).

La première règle de dépendance indirecte entre deux selfs est obtenue par composition des deux dépendances canoniques entre fragments et selfs. Cette règle, ainsi que la seconde, peut être rapprochée de la dépendance de flux dans [MC94]. La troisième règle se rapproche de la dépendance de composition entre tâches de [MC94], et la quatrième de la dépendance d'ajustement (nécessité de coordination de deux selfs individuels dans une situation de coproduction). Enfin, la dernière règle de dépendance indirecte entre deux selfs permet de trouver de nouvelles dépendances par combinaison des 4 règles précédentes.

Pour la combinaison des dépendances, l'algorithme doit veiller à ne pas comptabiliser deux fois la même dépendance, et pour cela, considérer les dépendances au niveau le plus détaillé. Par exemple, un self individuel qui réalise un composant d'un fragment fait par définition partie des réalisateurs du fragment ; il ne faut comptabiliser qu'une fois la dépendance correspondante, en considérant uniquement la dépendance « réalisé par » entre le composant du fragment et le self individuel.

3.2 Exemple d'application de la démarche

Afin d'illustrer la démarche d'analyse et de mesure des dépendances, nous présentons un exemple simplifié de rédaction collaborative d'un article pour un workshop. L'article a trois co-auteurs et porte sur une méthode de conception de site Web. Les trois auteurs utilisent Microsoft Word et les possibilités offertes par ce logiciel en termes de structuration de documents (en particulier, définition explicite de la structure hiérarchique du document –sections et sous-sections-, utilisation des références pour renvoyer à des figures ou à des sections...). La communication s'effectue par e-mail (transmission du fichier au fur et à mesure de sa rédaction) et en présentiel.

Le plan de l'article rédigé par ces trois co-auteurs est le suivant :

1. Introduction
2. La notation graphique (section constituée d'une figure illustrant la notation graphique utilisée dans la méthode de conception de site Web, et d'un texte d'accompagnement).
3. La démarche
 - 3.1. Présentation de la démarche
 - 3.2. Exemple d'application
4. Conclusion/discussion.

Le scénario de rédaction collaborative de l'article est le suivant :

1. Auteur3 fait la figure de la section 2 et le texte explicatif.
2. A partir de cette version, les trois auteurs se réunissent pour rédiger ensemble la section 3.1.
3. Auteur1 et Auteur2 rédigent alors ensemble l'exemple (section 3.2) et envoient le fichier ainsi complété à Auteur3.
4. Auteur3 complète l'exemple, rédige la conclusion/discussion et l'introduction et envoie le fichier ainsi complété à Auteur1 et Auteur2.

5. Auteur1 et Auteur2 relisent l'ensemble, retouchent l'introduction (ajout de la présentation du plan de l'article) et envoient la version finale au responsable du workshop.

En appliquant à ce cas notre démarche d'analyse et mesure des dépendances, le contenu des étapes est le suivant :

1. Explicitation de tous les liens encore implicites. Dans ce cas, ces liens sont peu nombreux, puisque par hypothèse, toutes les possibilités de Word en termes de structuration de documents sont utilisées. D'autre part, les réalisateurs et récepteurs des différents fragments peuvent être en partie inférés à partir des propriétés des versions successives de l'article (fichier Word) et des parties « From » et « To » des mails échangés. Les liens de traçabilité doivent quant à eux être représentés de manière explicite.
2. Génération du graphe de dépendances pour le DoPA de cet exemple. Ce graphe est représenté Figure 4. Par souci de lisibilité, cette figure distingue les dépendances canoniques entre deux fragments (à gauche) et entre un self individuel et un fragment (à droite).
3. A partir de ce graphe, calcul de deux matrices de dépendances : (1) une matrice de dépendances entre les 4 selfs individuels et (2) une matrice de dépendances entre les fragments (seule la version 2 de l'introduction et de la section 3.2 sont considérées).

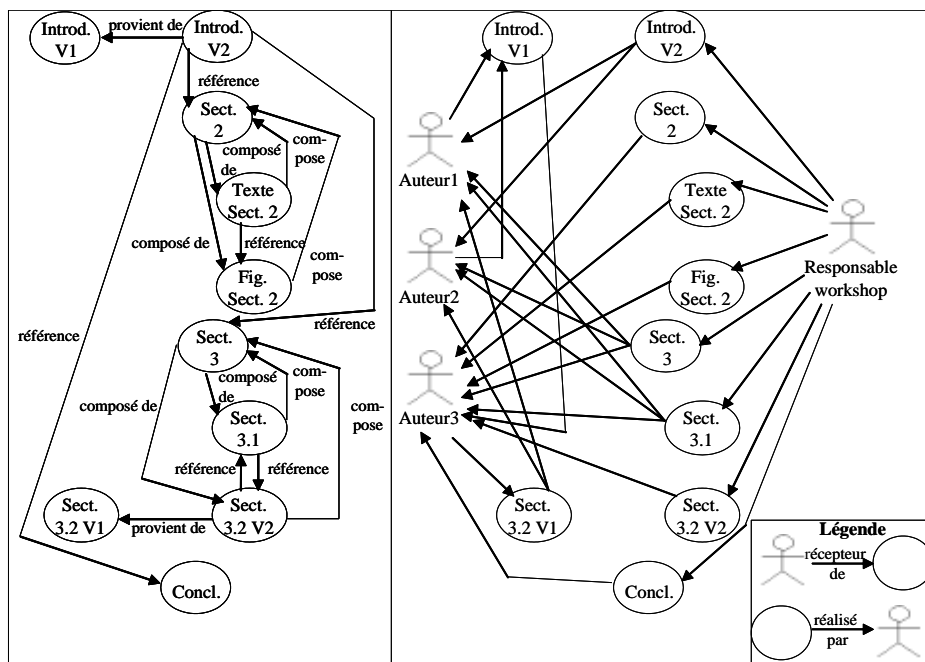


Figure 4: Graphe de dépendances de l'exemple

4 Discussion et conclusion

Le cadre conceptuel des documents pour l'action et des transactions communicationnelles présenté dans la première partie peut sembler assez complexe. Cependant, cette complexité est selon nous nécessaire pour rendre compte de la complexité corollaire des situations sociales de production documentaires médiatisées dans lesquelles les documents émergent à partir d'une multiplicité de fragments potentiellement hétérogènes produits par des communautés dont l'organisation est elle-même émergente et mouvante. Les prochaines étapes de notre recherche viseront à la fois (1) l'approfondissement de notre approche d'analyse et mesure des dépendances et (2) l'évaluation empirique de sa pertinence.

En ce qui concerne le premier point, nous poursuivons plusieurs pistes :

1. Il nous semble d'abord nécessaire de préciser encore la caractérisation des dépendances, en particulier en considérant plus finement la sémantique des liens dans le modèle de DopA. Ainsi, les liens de référence peuvent être détaillés davantage pour préciser la sémantique de la référence (pour le moment, le seul type de lien de référence représenté explicitement dans le modèle de DopA est le lien de traçabilité). De la même manière, il nous apparaît pertinent de préciser davantage la sémantique du lien de réalisation et de réception entre fragments et selfs. Il est ainsi possible de distinguer différents types/rôles de réalisateurs (par ex. : auteur principal/auteur secondaire, vérificateur, rédacteur, signataire, etc....). Ces différents types de réalisateurs contribuent chacun à leur manière à la réalisation du fragment, par conséquent la dépendance d'un fragment vis-à-vis d'un réalisateur dépend du rôle du réalisateur.
2. Afin d'affiner notre approche, il semble également pertinent d'associer des poids aux différents types de dépendances et de combiner ces poids (si A dépend de B avec un poids/coefficient p_1 et B dépend de C avec un coefficient p_2 , alors A dépend de C avec un coefficient $p_1 * p_2$). L'affectation de poids aux différents types de dépendances passe par un typage plus fin des dépendances, et peut également s'appuyer sur des heuristiques. Par exemple, il semble a priori que dans le cas d'une relation de composition entre fragments, la dépendance du composite vis-à-vis du composant soit plus forte que la dépendance inverse.

En ce qui concerne la problématique de l'évaluation empirique, nous projetons de passer en revue les principales situations transactionnelles médiatisées par des DopA ce qui nous permettra de sélectionner un terrain d'application réel mais également de fournir des préconisations pour la spécification d'environnements de production documentaire coopératifs facilitant la traçabilité. Les situations de productions collectives les plus étudiées par le biais de traces sont aujourd'hui les communautés de développement Open Source coopérant à travers des forums de discussion ([SD06] et [RS06]). Les premiers résultats de ces recherches sont prometteurs tant pour la mise en évidence de patterns de résolution de problèmes médiatisée dans le traitement de bugs logiciels [RS06] que pour une bonne compréhension des processus sociaux associés [SD06]. En nous appuyant sur le modèle formel des DopA et en l'appliquant à un scénario de rédaction médiatisée dans un petit groupe de travail, nous pensons avoir présenté de forts arguments plaidant pour une généralisation des recherches sur les traces de production documentaire permettant d'assister l'animation et le management des communautés d'action médiatisées qui concernent un nombre grandissant de situations professionnelles. Le modèle formel des DopA que nous avons présenté est générique. Il s'applique ainsi à la rédaction médiatisée comme nous l'avons illustré dans cet article, aux forums de discussion des communautés de développement Open Source... L'approche d'analyse et de mesure des dépendances, qui s'appuie sur le modèle de DopA, est par conséquent également générique, même si elle devra être affinée pour représenter plus précisément les types de dépendances, les poids associés, et les règles de composition. Une fois mesurées les dépendances entre fragments et entre selfs individuels, les informations fournies par ces mesures peuvent être utilisées pour améliorer le management des communautés d'action médiatisées : optimisation de la présentation et/ou de l'indexation des documents en fonction des dépendances entre fragments, détection de situations « anormales » de dépendance entre selfs (« anormales » car déséquilibrées, ou en contradiction avec l'organigramme...), etc. L'évaluation empirique permettra donc à la fois d'affiner/calibrer notre approche mais également de mettre en évidence sur un exemple réel l'utilité de l'analyse et de la mesure des dépendances.

Bibliographie

- [Cr03] Crowston, K.: A taxonomy of organizational dependencies and coordination mechanisms. (Malone, T.; Crowston, K.; Herman, G., eds): Tools for organizing business knowledge: the MIT process handbook. MIT Press, Cambridge, Massachusetts, 2003; p. 85-108.
- [Et04] Etkorn, L.H. et al.: A comparison of cohesion metrics for object-oriented systems. Information and Software Technology, vol. 46, n° 10, août 2004; p. 677-687.
- [MC94] Malone, T.; Crowston, K.: The interdisciplinary study of coordination. ACM Computing Surveys, vol. 26, n° 1, mars 1994; p. 87-119.
- [Mi82] Mintzberg, H.: Structure et dynamique des organisations. Les Editions d'Organisation, Paris, 1982.
- [Ob06] Object Management Group: Unified Modeling Language (UML) specification, version 2.0. <http://www.omg.org/technology/documents/formal/uml.htm>.
- [RS06] Ripoché, G., Sansonnet ; J.P., Experiences in Automating the Analysis of Linguistic Interactions for the Study of Distributed Collectives, CSCW Journal, 2006 (to appear).

- [SD06] Sack, W., Détienne, F., Ducheneaut, N., Burkhardt, J.M., Mahendran, D., Barcellini, F., A Methodological Framework for Socio-Cognitive Analyses of Collaborative Design of Open Source Software, CSCW Journal, 2006 (to appear).
- [Su00] Sutcliffe, A.: Requirements analysis for socio-technical system design. *Information Systems*, vol. 25, n° 3, mai 2000; p. 213-233.
- [Va03] Vakali, A. et al.: Hierarchical data placement for navigational multimedia applications. *Data & Knowledge Engineering*, vol. 44, n° 1, janvier 2003; p. 49-80.
- [Wa05] Wang, J. et al.: DMC: a more precise cohesion measure for classes. *Information and Software Technology*, vol. 47, n° 3, mars 2005; p. 167-180.
- [YM93] Yu, E.; Mylopoulos, J.: An actor dependency model of organizational work – with applications to business process reengineering. In *Proc. Conf. on Organizational Computing Systems (COOCS 93)*, Milpitas, Californie, USA, novembre 1993. ACM Press, New-York; p. 258-268.
- [Za03] Zacklad, M. Communities of Action: a Cognitive and Social Approach to the Design of CSCW Systems. *GROUP'2003*, 09-12 November 2003, Sanibel Island. ACM, 2003, 190-197 p.
- [Za04] Zacklad, M. Processus de documentarisation dans les Documents pour l'Action (DopA) : statut des annotations et technologies de la coopération associées. *Le numérique : Impact sur le cycle de vie du document pour une analyse interdisciplinaire* http://archivesic.ccsd.cnrs.fr/sic_00001072.html, 13-15 Octobre 2004, Montréal: 19 p.
- [Za06a] Zacklad, M., Benel, A., Barcellini, F., Barry-Greboval, C., Boujut, J.-F., Bringay, S., Burkhardt, J.-M., Charlet, J., Cure, O., Darses, F., De Saint Leger, M., Detienne, F., Guibert, S., Lewkowicz, M., Lortal, G., Pierrat, M.-J., Sack, W., Salzano, G., Todirascu, A., Treins, M., et Turner, W. Processus d'annotation dans les documents pour l'action : textualité et médiation de la coopération. In: Charlet, J. *Documents et contenu : création, indexation, navigation*. Paris: Hermès, 2006.
- [Za06b] Zacklad, M., Gestion du connaissant et du connu dans la théorie transactionnelle de l'action (TTA), soumis à la revue *Annales des Télécommunications*, 2006.
- [ZP01] Zamfiroiu, M.; Prat, N.: Traçabilité du processus de conception des systèmes d'information. (Cauvet, C.; Rosenthal-Sabroux, C., eds): *Ingénierie des systèmes d'information*. Editions Hermès, Paris, 2001; p. 245-276.