

IFT-66529 Ingénierie des connaissances

D'après Schreiber *et al.* (2000), chapitre 7

Prof. Nicole Tourigny, Ph.D.

Département d'informatique et de génie logiciel

Faculté des sciences et de génie

Université Laval

Hiver 2003

Construction du modèle des connaissances (MdC)

- Décomposition du processus de construction du MdC en étapes et activités
- Étapes :
 - Identification des sources de connaissances
 - Spécification des connaissances
 - Raffinement des connaissances
- Une approche particulière, parmi d'autres

Construction du MdC

- Contenu du MdC (ch. 5)
- Comment réaliser ce modèle ?
 - Guide de réalisation
 - Étapes
 - Activités
 - Techniques
- Recette non unique
 - Construction : spécification de besoins
 - Synthèse
- Techniques d'AC (Ch. 8)

IFT-66529 IC N. Tourigny, Ph.D.
H-2003 Schreiber et al. (2000)

3

Construction du MdC

- 3 étapes
 - Identification des sources de connaissances
 - Spécification des connaissances
 - Raffinement des connaissances

IFT-66529 IC N. Tourigny, Ph.D.
H-2003 Schreiber et al. (2000)

4

Construction du MdC

- Un survol des 3 étapes
 - Identification des sources de connaissances
 - Spécification des connaissances
 - Raffinement des connaissances

Construction du MdC : identification des sources de connaissances

- Quelles sources seront utiles pour la modélisation des connaissances ?
- Élaboration d'un lexique (glossaire) des termes du domaine
- Survol des composants des modèles existants (bibliothèque des modèles)
- Utilisation des modèles antérieurs
 - Modèle de l'organisation : description des items de connaissances
 - Modèle des tâches : caractérisation de la tâche

Construction du MdC : spécification des connaissances

- Élaboration des spécifications du MdC par le cogniticien
 - Sélection d'un modèle de base (gabarit) de tâche (ex.: évaluation, diagnostic, configuration, ...)
 - Construction d'un schéma initial du domaine
 - Emploi de la liste de composants réutilisables (étape précédente)
- Remplir les blancs
 - 2 approches
 - inférences → connaissances de la tâche et du domaine
 - connaissances de la tâche et du domaine → inférences
- Sélection basée sur la qualité et le degré de détail du modèle de base choisi (s'il y a lieu)
- Accent sur le schéma du domaine (pas sur les BC)

IFT-66529 IC N. Tourigny, Ph.D.
H-2003 Schreiber et al. (2000)

7

Construction du MdC : raffinement des connaissances

- Validation du MdC
 - Simulation de scénarios recueillis durant l'identification des sources de connaissances
 - Papier
 - Prototype
 - Le MdC permettra-t-il de simuler le comportement attendu du système ?
- Validation positive → compléter les BC
 - Instanciation
 - instances de concepts, de types de règles

IFT-66529 IC N. Tourigny, Ph.D.
H-2003 Schreiber et al. (2000)

8

Construction du MdC : trois étapes

- Un chemin non linéaire
- Étapes entrelacées
 - Boucles de rétroaction
 - Spécifications → validation → modifications (re-spécifications) → validation → ...
 - Complétion des BC : possibilité de besoin de nouvelles sources de connaissances
- Fréquence des boucles de rétroaction
 - Plus grande si problème moins structuré

Construction du MdC

- Les 3 étapes plus en détail
 - Identification des sources de connaissances
 - Spécification des connaissances
 - Raffinement des connaissances

Figure 7.1 Vue d'ensemble des 3 principales étapes de la construction du MdC (Schreiber *et al.* 2000, p. 169)

Identification des sources de connaissances

- Familiarisation avec le domaine (source de connaissances, glossaire, scénarios)
- Liste potentielle des composants pour réutilisation

• Spécification des connaissances

- Sélection d'un modèle de base
- Construction initiale de la conceptualisation du domaine
- Complétion de la spécification (MdC + BC partielle)

• Raffinement des connaissances

- Validation du MdC (simulation papier, prototype)
- Raffinement de la BC (compléter BC)

IFT-66529 IC N. Tourigny, Ph.D.
H-2003 Schreiber et al. (2000)

11

Étape 1 : identification des connaissances

- Préconditions
 - Tâche de C (k-intensive) reconnue
 - Principaux items de connaissances impliqués dans tâche ont été identifiés
 - Type de tâche identifié (évaluation, configuration,...)
- Buts
 - Recueil des items de C
 - Préparation des C pour spécification (2e étape)
- Activités
 - Familiarisation
 - Explorer et structurer les sources de connaissances pour la tâche, créer un lexique (ou glossaire) des termes du domaine
 - Listage des composants réutilisables
 - Étudier plus en détail la nature de la tâche, vérifier et réviser le type de tâche, lister tous les composants réutilisables pour cette application

IFT-66529 IC N. Tourigny, Ph.D.
H-2003 Schreiber et al. (2000)

12

Étape 1 : Identification des connaissances

Activité 1 : Familiarisation avec le domaine de C

- Point de départ
 - Liste des items de C décrits dans feuille TM-2
 - 2 facteurs importants à considérer
 - Nature des sources
 - Structure du domaine du problème
 - Diversité des sources
 - Mono-expertise, multi-expertise ?
 - Manuels, sources diverses, connaissances conflictuelles, ...
- Techniques
 - Marquage du texte dans les sources de base (manuel, volume)
 - Quelques entrevues structurées pour se faire une bonne idée du domaine d'application
 - Au besoin, explorations plus détaillées pour des zones moins comprises
 - Se familiariser sans devenir expert

IFT-66529 IC N. Tourigny, Ph.D.
H-2003 Schreiber et al. (2000)

13

Étape 1 : Identification des connaissances

Activité 1 : Familiarisation avec le domaine de C

Principes de base (P)

- P 1 : Parler aux gens de l'entreprise qui ont à parler avec des experts mais qui ne sont pas experts eux-mêmes
 - Démarche de compréhension du domaine
 - Caractéristiques du processus de résolution
- P 2 : Éviter théories détaillées compliquées, sauf si utiles
 - Théories détaillées à examiner ultérieurement
 - Existence de formules plutôt que leur expression à ce stade
- P 3: Construire quelques scénarios types compris (par IC) à un niveau global
 - Construire des scénarios avec les experts
 - Demander aux non experts impliqués leur opinion sur ces scénarios
 - Comprendre le domaine afin d'expliquer le raisonnement dans les scénarios
 - Peu de temps (2 semaines)

IFT-66529 IC N. Tourigny, Ph.D.
H-2003 Schreiber et al. (2000)

14

Étape 1 : Identification des connaissances
Activité 1 : Familiarisation avec le domaine de C
Résultats tangibles

- Liste des sources de connaissances du domaine, incluant une brève caractérisation
- Résumés de textes-clés choisis
- Description des scénarios développés
- Compréhension du domaine par l'IC

Étape 1 : Identification des connaissances
Activité 2 : Liste potentielle des composants du modèle

- Étude du potentiel de réutilisation selon 2 aspects
 - Tâche
 - Caractérisation du type de la tâche (modèle de la tâche)
 - Validité de cette caractérisation ? Construction des modèles de base de la tâche
 - Domaine
 - Établir le type du domaine : technique, heuristique, ... ?
 - Chercher descriptions standardisées du domaine ou de domaines similaires
 - Objets d'art: *Art and Architecture Thesaurus (AAT)*
 - Terminologie médicale : *Medical Subject Headings (MeSH)*
 - Bibliothèques d'ontologies
 - Modèles de référence (hôpitaux, bibliothèques de modèles de produits, ...)
- Principes pour guider la sélection du bon type de tâche
 - P4: Les tâches d'application sont souvent une combinaison de plusieurs types de tâche
 - P5 : Les noms donnés aux tâches d'application ne correspondent pas nécessairement à des noms génériques de types de tâche

Étape 2 : Spécification des connaissances

But et activités

- But
 - Obtenir une spécification complète du MdC, sauf les contenus des BC (exemples seulement)
- Activités
 - Choisir un modèle de tâche (*template*)
 - Construire un schéma du domaine initial
 - Spécifier les 3 catégories de connaissances (domaine, inférence, tâche).

Étape 2 : Spécification des connaissances

Activité 1 : Choisir un modèle de tâche (*template*)

- Bibliothèques de modèles de base (*templates*) : Ch. 6, pointeurs
- Aspects importants à considérer
 - Nature des sorties (ex.: catégorie d'erreur, catégorie de décision, plan)
 - Nature des entrées: sortes de données disponibles pour RDP ?
 - Nature du système: processus physique comme dans une usine nucléaire, un artefact fait par les humains comme un photocopieur
 - Contraintes posées par la tâche sur l'environnement : certitude de la solution requise, coûts des observations
- Principes
 - P6 : Préférer les modèles de base ayant été utilisés plusieurs fois
 - P7 : Si vous pensez avoir un modèle de base convenable, construire une structure d'inférence (SI) annotée
 - P8 : Si aucun modèle de base ne semble convenir, remettre en cause le caractère « orienté vers la connaissance » de la tâche (*k-intensive*)
 - P9 : Un mauvais modèle de base vaut mieux qu'aucun modèle

Étape 2 : Spécification des connaissances

Activité 2 : Construire un schéma du domaine initial

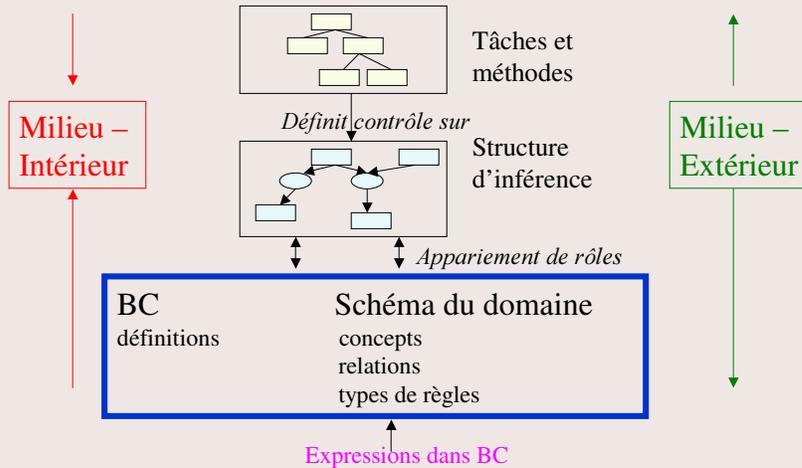
- But
 - Construire un modèle de données initial du domaine indépendant du problème d'application à résoudre ou des méthodes de tâches choisies.
 - Typiquement, le schéma du domaine contient au moins 2 parties
 - Conceptualisation spécifique au domaine
 - Conceptualisations spécifiques à la méthode
- Principes
 - P10 : Baser les conceptualisations spécifiques au domaine sur les modèles de données, si possible
 - P 11: Limiter l'utilisation des langages de modélisation de CommonKADS aux concepts, aux sous-types et aux relations

Étape 2 : Spécification des connaissances

Activité 3 : Compléter la spécification du MdC

- Prérequis
 - Modèle de base choisi
 - Schéma initial du domaine construit
- Routes (figure 7.2)
 - 1. Milieu – extérieur
 - 2. Milieu – intérieur
- Principes
 - P 12 : La SI est suffisamment détaillée ssi l'explication qu'elle fournit du processus de raisonnement est suffisamment détaillée
 - P 13: La SI est suffisamment détaillée ssi il est facile de trouver pour chaque inférence un type simple de connaissance du domaine qui peut agir comme rôle statique pour cette inférence

Deux approches pour la spécification du MdC (d'après fig. 7.2, Schreiber *et al.* 2000, p. 176)



IFT-66529 IC N. Tourigny, Ph.D.
H-2003 Schreiber *et al.* (2000)

21

Étape 2 : Spécification des connaissances
Activité 3 : Compléter la spécification du MdC

- Principes pour spécifier les tâches et les méthodes de tâche
 - P 14 : Au départ de la spécification d'une méthode de tâche, commencer avec la structure de contrôle
 - P 15 : Lors de l'écriture d'une structure de contrôle, ne pas trop s'occuper des détails de représentation
 - P 16 : Choisir un nom de rôle qui indique clairement comment cet item de données est utilisé dans la tâche
 - P 17 : Ne pas inclure des rôles de connaissance statiques comme E/S de tâche
 - P 18 : Pour les applications en temps réel, considérer une représentation autre que le pseudo-code

IFT-66529 IC N. Tourigny, Ph.D.
H-2003 Schreiber *et al.* (2000)

22

Étape 2 : Spécification des connaissances

Activité 3 : Compléter la spécification du MdC

- Principes pour spécifier la SI
 - P 19 : Commencer par représenter graphiquement la SI
 - P 20: Autant que possible, utiliser un ensemble standard d'inférences
 - P 21 : Distinguer clairement les simples rôles d'objet des ensembles
 - P 22 : Questionner les inférences sans E et celles ayant plusieurs S
 - P 23 : Choisir des noms de rôles indépendants du domaine
 - P 24 : Standardiser le plan
 - P 25 : Ne pas trop se préoccuper de la dynamique des rôles des objets dans la SI
 - P 26 : Utiliser un attribut pour une spécification claire de ce que l'inférence doit faire, et indiquer les méthodes pouvant être considérées dans la phase de conception

Étape 2 : Spécification des connaissances

Activité 3 : Compléter la spécification du MdC

- Principes pour spécifier la connaissance du domaine
 - P 27 : Un type de connaissance du domaine utilisé comme rôle statique pour une inférence n'est pas tenu d'avoir exactement la bonne représentation nécessaire pour cette inférence (disponibilité en principe des connaissances)
 - P 28 : La portée de la connaissance du domaine est typiquement plus large que ce qui est couvert par les inférences

Étape 3 : Raffinement des connaissances

- Activités

- 1. Valider le MdC, habituellement avec l'aide d'une technique de simulation
- 2. Compléter les BC en ajoutant les instances de la connaissance du domaine
 - Prérequis: MdC validé

Étape 3 : Raffinement des connaissances Activité 1 : Validation du MdC

- Activités

- Validation interne (vérification)
 - Valider la cohérence interne du modèle
 - Est-ce que le modèle est correct ?
- Validation externe (validation)
 - Validation / spécifications de l'utilisateur
 - Est-ce le bon modèle ?

Étape 3 : Raffinement des connaissances

Activité 1 : Validation du MdC

- Activités
 - Validation interne (vérification)
 - Valider la cohérence interne du modèle
 - Est-ce que le modèle est correct ?
 - Omissions : partie manquante du modèle
 - Redondances : ex.: inférence non utilisée dans aucune méthode de tâche
 - Techniques
 - « Walk-throughs »
 - Outils logiciels pour vérifier la syntaxe, complétude, ...

Étape 3 : Raffinement des connaissances

Activité 1 : Validation du MdC

- Activités
 - Validation externe (validation)
 - Validation / spécifications de l'utilisateur
 - Est-ce le bon modèle ?
 - Le modèle capture-t-il le bon comportement de RDP ?
 - Plus difficile à réaliser
 - Techniques
 - Simuler le comportement : définir des scénarios et les valider sur le modèle
 - » Papier
 - » Prototype
 - Comment se comporte le modèle ?
 - Y a-t-il des différences entre le modèle et les scénarios ?
 - Quelles adaptations doivent être apportées au modèle ?

Étape 3 : Raffinement des connaissances
Activité 2 : Compléter les BC

- Principes

- P 29 : S'il est difficile de trouver des instances d'un type de connaissances, remettre en cause ce type
- P 30 : Considérer les BC existantes du même domaine