

Analyse des diagrammes de l'apprenant dans un EIAH de la modélisation orientée objet

Le système ACDC

Ludovic Auxepaules

Thèse soutenue le 24 septembre 2009

Site web : <http://doclaux.free.fr>

Mail : ludovic.auxepaules@gmail.com

Thèse dirigée par Dominique Py

Professeure à l'Université du Maine



Contexte général

- Projet "Interaction et Connaissance" (I&C) du LIUM
- **Participants** : Dominique Py, Mathilde Alonso, Thierry Lemeunier et Ludovic Auxepaules
- **But du projet** : Elaboration de modèles, de méthodes et d'outils pour la conception d'EIAH dédiés à la modélisation
 - Conception de l'interaction
 - Diagnostic
- **Application** : l'environnement *Diagram*
 - Construction de diagrammes de classes UML de niveau analyse par des novices à partir d'un énoncé textuel



Problématique

- Comment analyser les productions de l'apprenant dans un environnement d'apprentissage de la MOO ?
- Application aux diagrammes de classes UML
- Difficultés dans le contexte de la modélisation
 - Pas de méthode de résolution définie formellement
 - Plusieurs solutions à un même problème
 - Notion d'erreur mal définie
 - Pas de résolveur pédagogique en modélisation

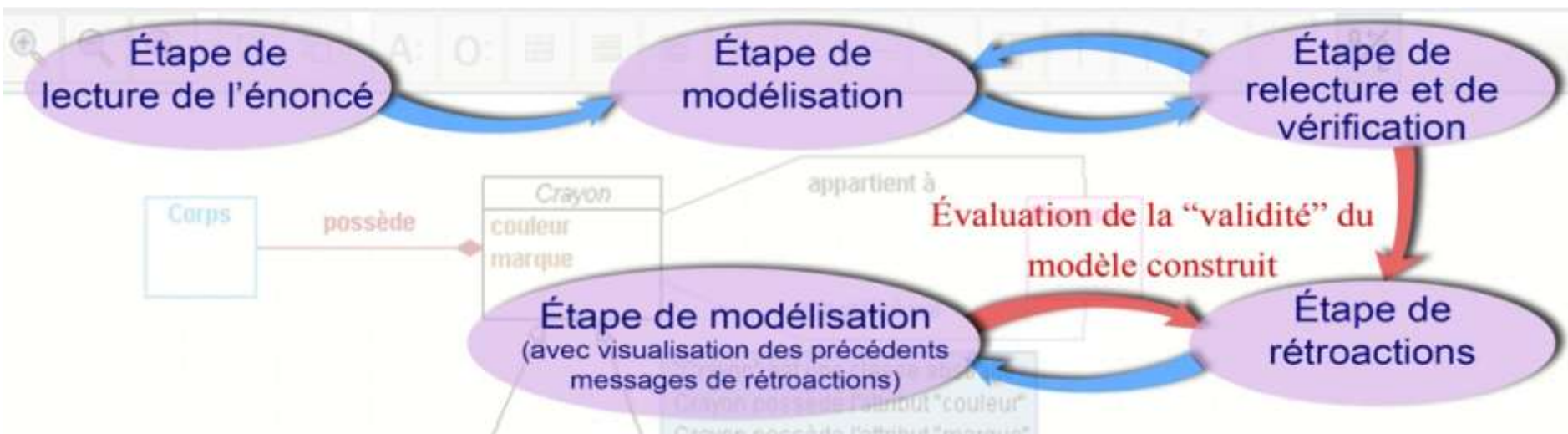


Objectifs de nos travaux de thèse

- Proposer une méthode de diagnostic des diagrammes de classes UML de l'apprenant
- Conserver un degré de généricité suffisant pour des applications à d'autres types de modèles
- Appliquer les résultats dans *Diagram* pour fournir des rétroactions pédagogiques synchrones
- Implanter et évaluer qualitativement la méthode proposée

L'environnement *Diagram*

- Éditeur de diagrammes de classes UML intégrant
 - L'énoncé textuel de l'exercice à l'interface
 - Fournissant des modalités d'interaction et des aides spécifiques aux novices
- Méthode de modélisation instanciée en cinq étapes





Notre proposition

- Diagnostic reposant sur l'appariement d'un diagramme de l'apprenant à un diagramme de référence
- Inspiration des techniques d'appariement de modèles
- Nécessité d'avoir un diagramme de référence
- Pas d'erreurs mais des différences entre les deux diagrammes
- Prise en compte de la structure des diagrammes pour déduire des constructions alternatives



Approches d'appariement de modèles

- Processus d'appariement : identification et qualification de relations entre les éléments de plusieurs modèles
- Résultat du processus : un alignement ou un *mapping* des éléments
- Grande variété de techniques individuelles [Euzenat & Shvaiko 2005]
- Combinaison de plusieurs techniques [Bernstein *et al.* 2001]
- Problème difficile à traiter de manière totalement automatique

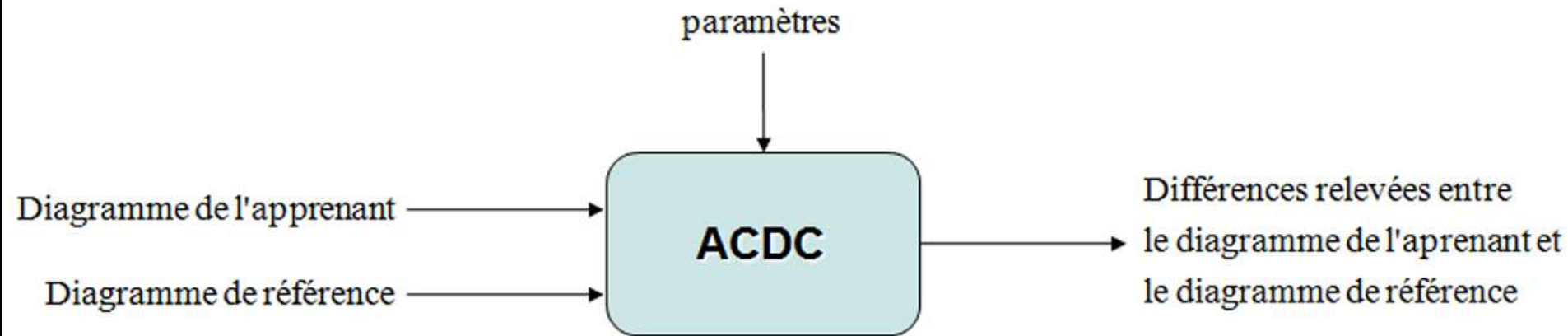


Contraintes de notre contexte

- Méthode robuste acceptant des diagrammes de l'apprenant très différents
- Pas de données auxiliaires en plus des modèles
- Pas d'intervention d'un acteur humain lors de l'activité
- Temps de calcul suffisamment bref pour les rétroactions synchrones

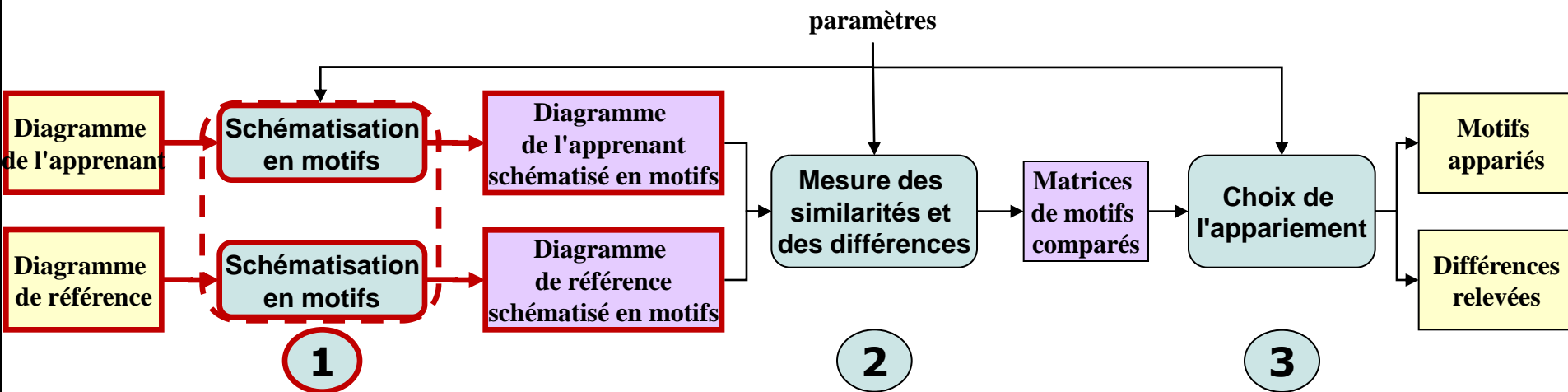
La méthode d'appariement ACDC (*Automatic Class Diagrams Comparator*)

- Méthode d'appariement hybride paramétrable

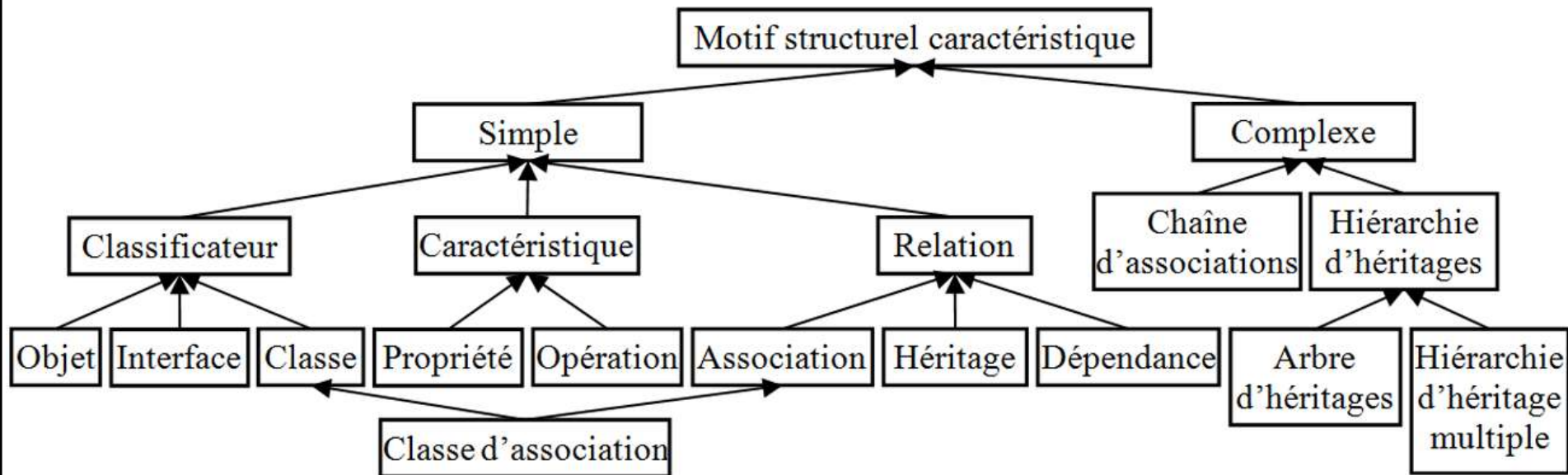


- Trois étapes séquentielles
 - ① Schématisation des diagrammes en motifs structurels caractéristiques
 - ② Évaluation des similarités et des différences locales à chaque couple de motifs
 - ③ Choix de l'appariement des motifs et des différences

Étapes de la méthode d'appariement ACDC

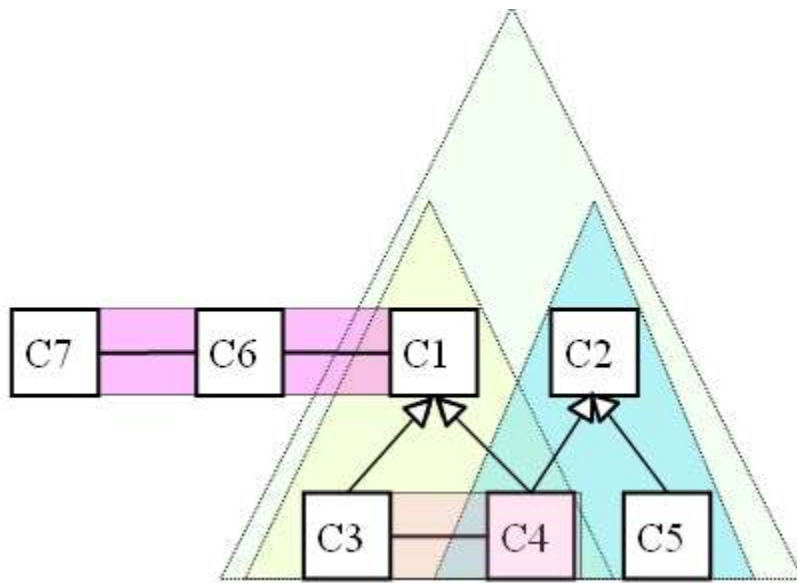


Motifs structurels caractéristiques

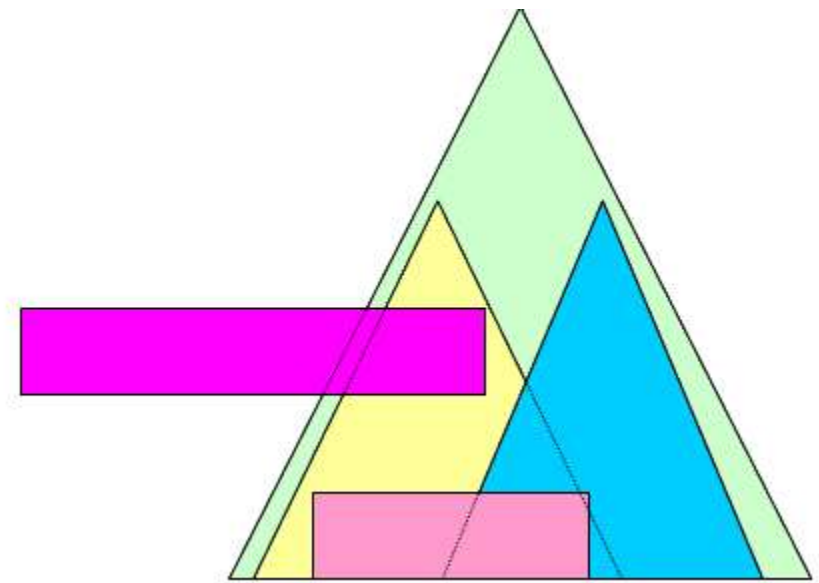


Classification des motifs structurels caractéristiques pour les diagrammes de classes UML de niveau analyse

Exemple de schématisation d'un diagramme de classes en motifs

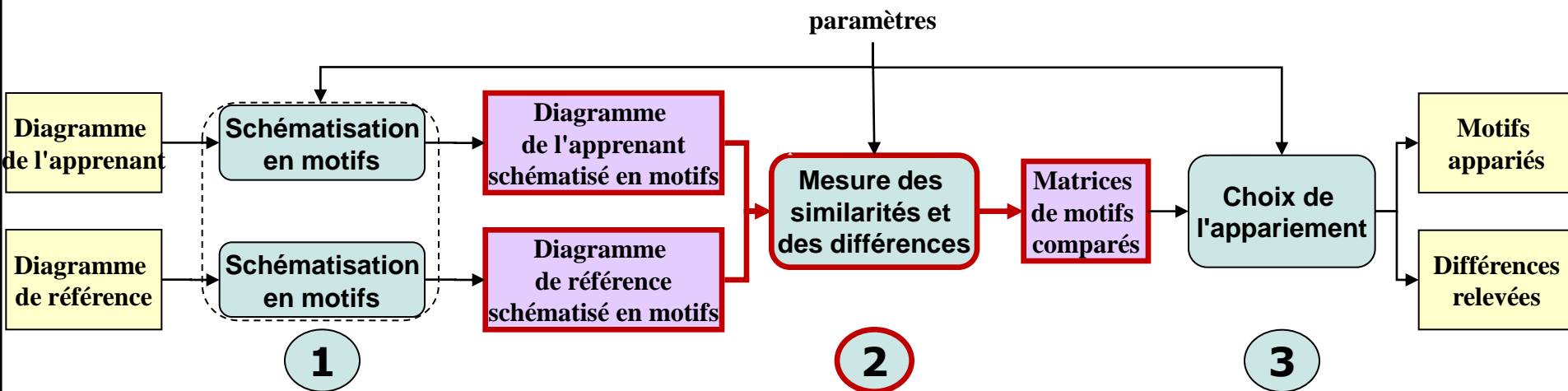


Motifs simples correspondant aux éléments du diagramme de classes



Motifs complexes structurant les motifs simples du diagramme de classes

Étapes de la méthode d'appariement ACDC





Mesure des similarités et des différences

○ Objectifs

- Affecter un score de similarité à chaque paire de motifs comparée par type
- Déterminer des différences locales relatives aux motifs en contexte
- Permettre le classement des motifs comparés par score de similarité

○ Principes

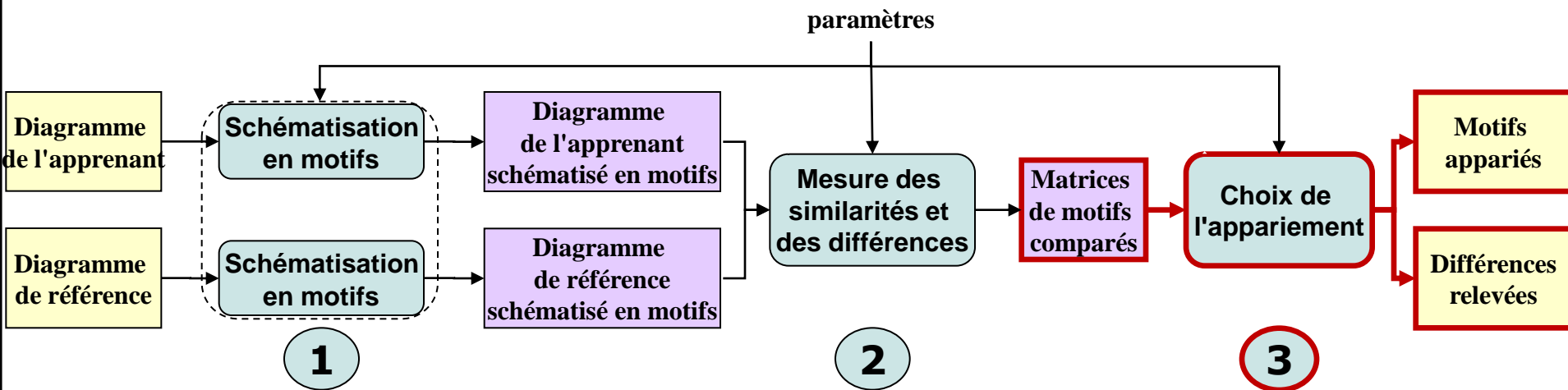
- Toutes les dimensions descriptives participent à la similarité
 - Principe de renforcement mutuel des motifs reliés
 - Propagation ascendante, descendante et transversale du contexte
- Problème de dépendance mutuelle et récursive dans la mesure (circularité)



Mesure des similarités et des différences

- Fonction de similarité combinant deux scores
 - Score simple : pondération des critères indépendants de tout autre motif
 - Score complexe : agrégation d'une partie des scores des motifs en contexte avec le couple de motifs comparés (motifs liés, contenus ou conteneurs)
- Fonction de similarité contextualisée par rapport aux critères propres au couple de motifs comparés
- Comparaison des motifs complexes dirigeant celle des motifs simples
- Comparaison spécifique des noms reposant sur la recherche de sous-chaînes communes
- Instanciation de la mesure sous forme d'une hiérarchie de comparateurs

Étapes de la méthode d'appariement ACDC

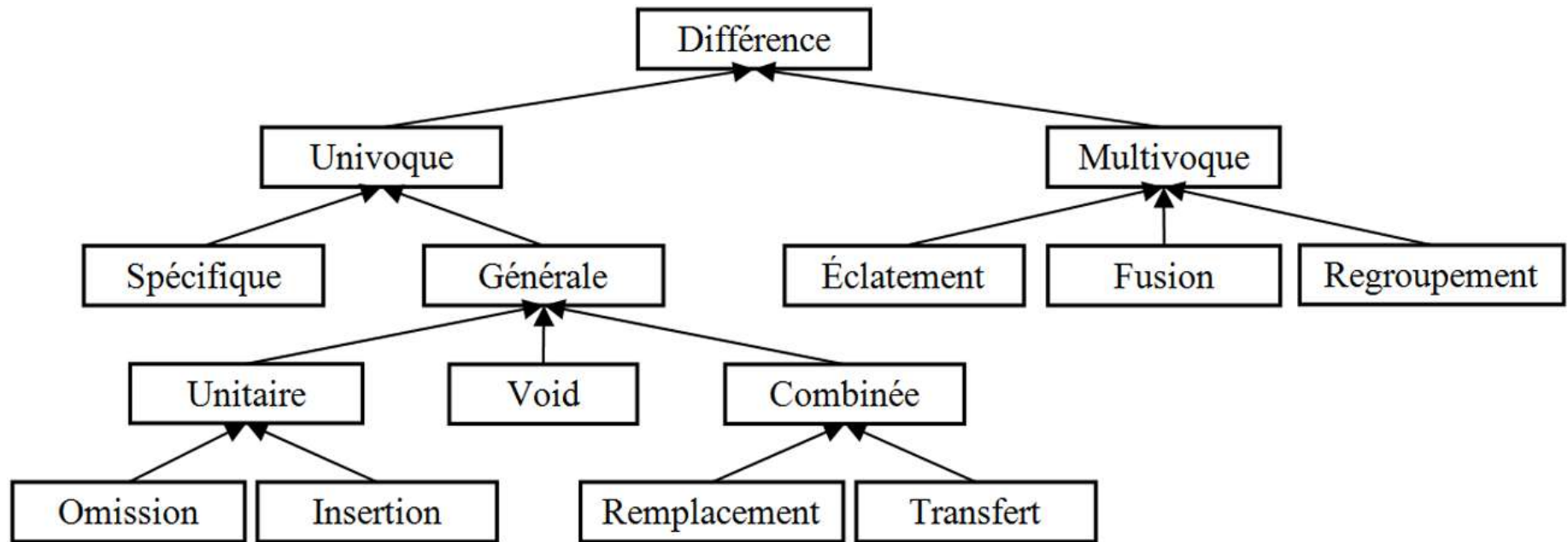




Choix de l'appariement des motifs

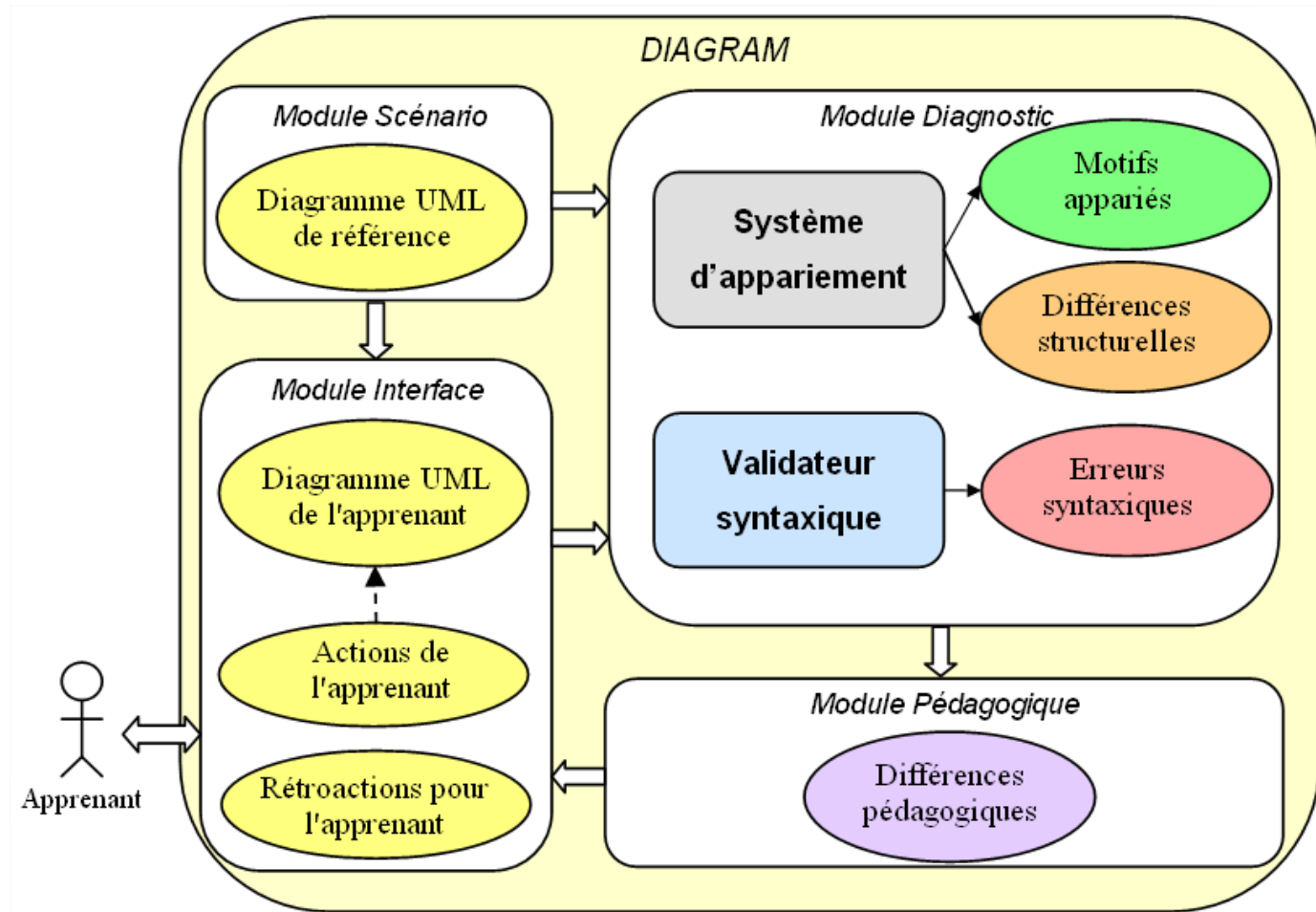
- Processus glouton sans retour en arrière
- Capacité à identifier des appariements univoques et multivoques
- Deux apparieurs appelés séquentiellement
- Comportement général de chaque apparieur
 - Définition d'une liste de couples de motifs candidats
 - Sélection des couples de motifs
 - Détermination des appariements multivoques
 - Finalisation du résultat
- Sorties de l'appariement : motifs appariés strictement et appariés moyennant des différences structurelles générales

Taxonomie des différences structurelles



- Différences factorisées des plus spécifiques aux plus générales

Application de la méthode au diagnostic



Exemple de diagnostic

Diagramme de l'apprenant

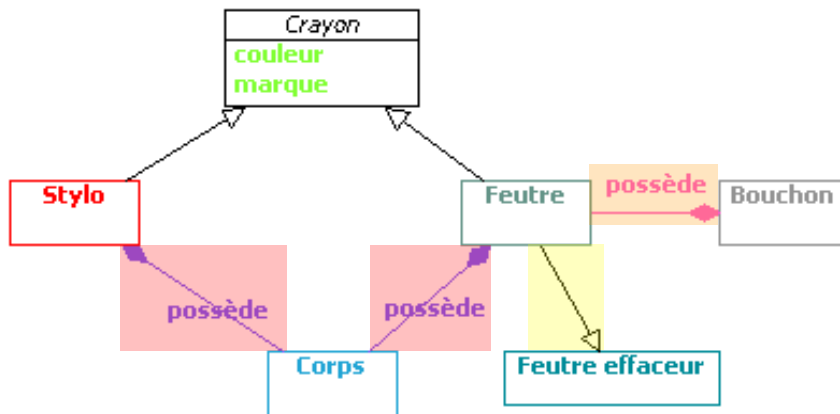
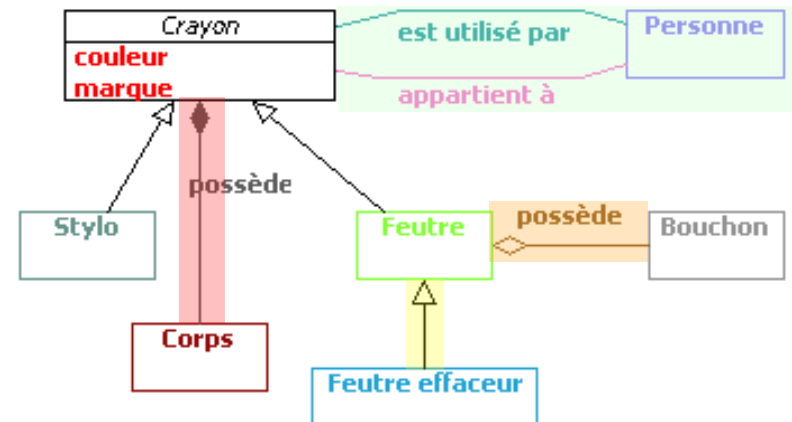


Diagramme de référence



- Douze appariements stricts (*Void*) et treize différences structurelles relevés par ACDC

Exemple de différence générale combinée

Diagramme de l'apprenant

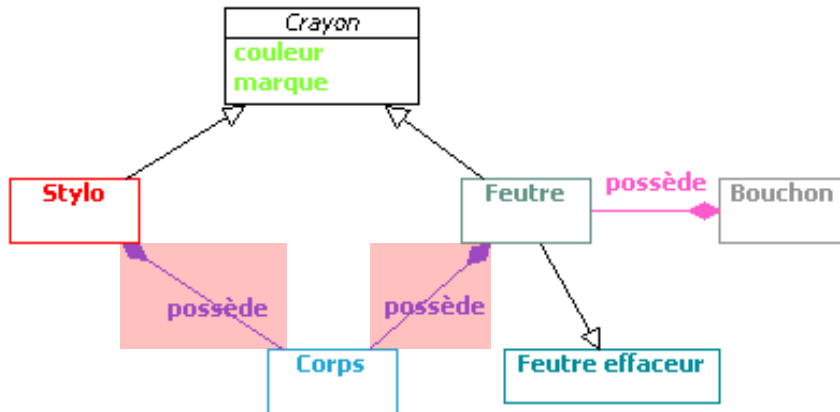
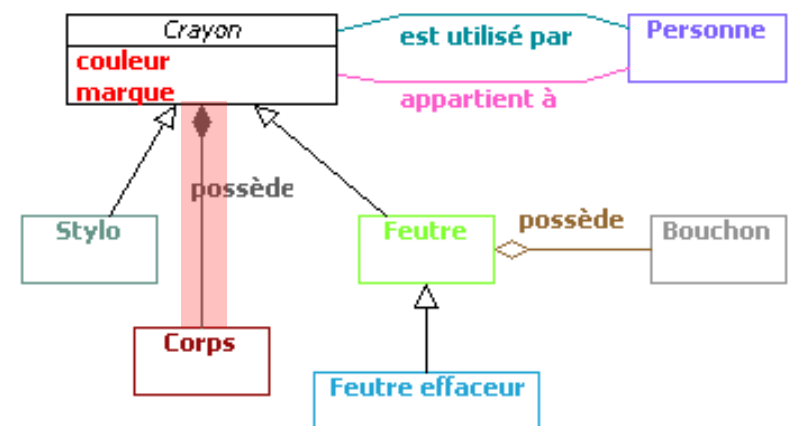


Diagramme de référence



Différences relevées par ACDC

{possède (Corps---Feutre) possède (Corps---Stylo)}
ECLATEMENT {possède (Corps---Crayon)}

{possède (Corps---Feutre)} TRANSFERT VERS_FILS
{possède (Corps---Crayon)}

{possède (Corps---Stylo)} TRANSFERT VERS_FILS
{possède (Corps---Crayon)}

Rétroactions pédagogiques

Dédoublage et transfert d'une relation

Exemple de différences générales unitaires

Diagramme de l'apprenant

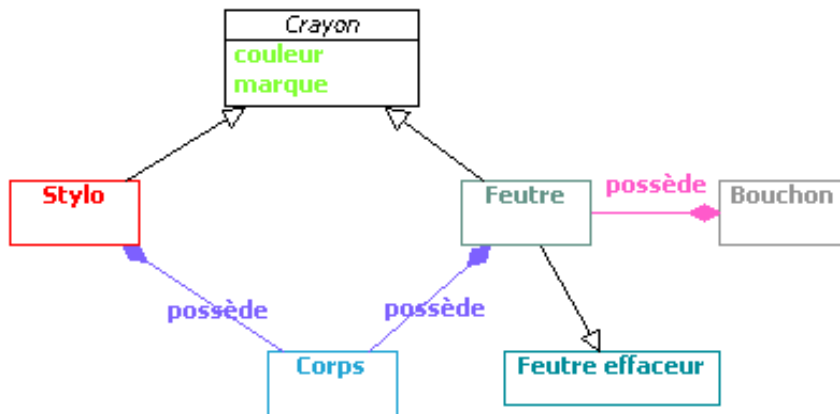
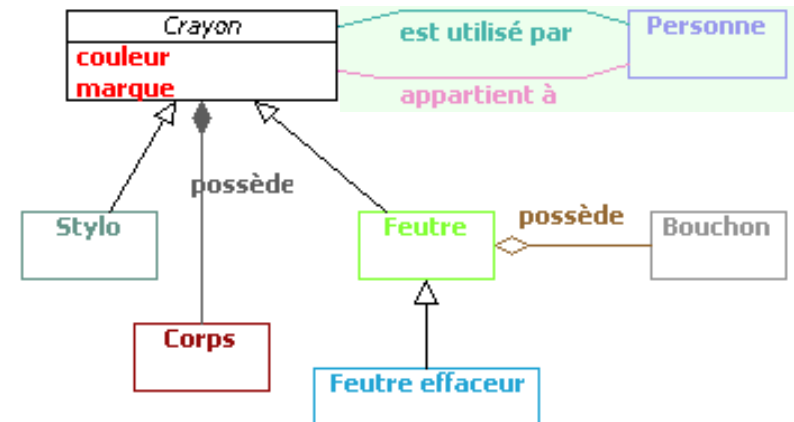


Diagramme de référence



Différences relevées par ACDC

OMISSION {Personne}

OMISSION {appartient à (Personne---Crayon)}

OMISSION {est utilisé par (Personne---Crayon)}

Rétroactions pédagogiques

Omission d'une classe et des éléments liés à cette classe



Évaluation hors-ligne d'ACDC [Auxepaules *et al.* 08]

○ Corpus

- 82 diagrammes de classes produits par des apprenants
- 3 exercices de difficulté croissante
- Diagrammes de référence fournis par l'enseignant de MOO

○ Méthode d'évaluation

- Calibrage empirique des fonctions et des algorithmes avec les diagrammes du premier exercice
- Tests et optimisations avec les diagrammes du deuxième exercice
- Validation avec les diagrammes du dernier exercice



Évaluation d'ACDC dans *Diagram* [Auxepaules & Py 10]

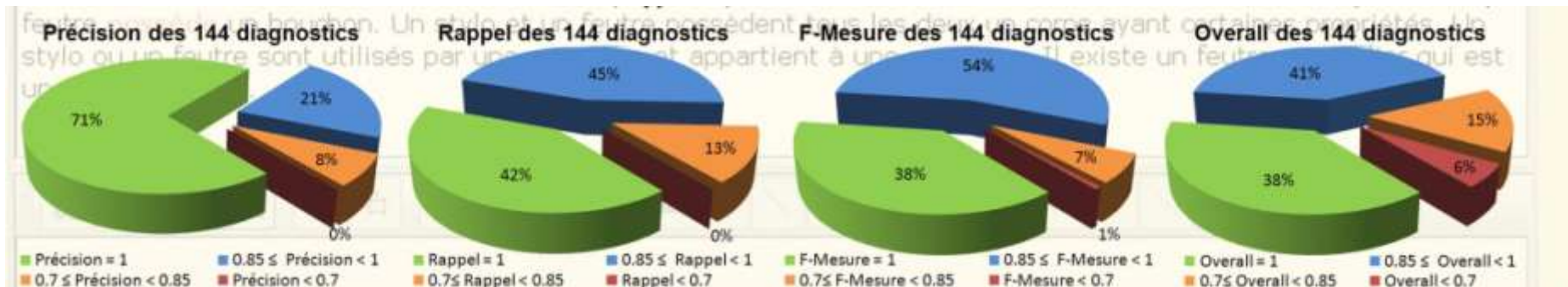
- Protocole d'évaluation
 - 18 étudiants en 2nde année de DEUST ISR
 - 4 sessions de travaux pratiques de 3 heures intégrées au cursus
 - Diagrammes de référence fournis par l'enseignant de MOO
 - Evaluation d'ACDC basée sur l'analyse des traces enregistrées pour 4 exercices
 - Soit 69 exercices d'étudiant (**144 appels différents du diagnostic**)
- Etude des rétroactions pédagogiques présentée dans [Alonso & Py 09]



Mesures de qualité utilisées

- Comparaison des appariements identifiés automatiquement aux appariements réels déterminés manuellement
 - *Faux négatifs A* : appariements non identifiés par l'apparieur
 - *Vrais positifs B* : appariements corrects identifiés automatiquement
 - *Faux positifs C* : appariements erronés ou typés de manière incorrecte
- Mesures de qualité [Do & Rahm 07] [Giunchiglia et al. 07] [Melnik et al. 02]
 - *Précision* = $|B| / (|B| + |C|)$ → une mesure de fidélité ou d'exactitude
 - *Rappel* = $|B| / (|B| + |A|)$ → une mesure de complétude
 - *F-Mesure* → une moyenne harmonique de la *Précision* et du *Rappel*
 - *Overall* = $(|B| - |C|) / (|B| + |A|)$ → une évaluation des efforts nécessaires pour ajouter les appariements omis et supprimer les appariements erronés

Résultats à l'échelle des 144 appels d'ACDC



- **38%** des diagnostics parfaits (*Overall* = 1)
- **92%** des diagnostics pertinents à 85% (*F-Mesure* ≥ 0.85)
- Au moins un oubli dans **58%** des cas (*Rappel* < 1)
- Au moins une erreur dans **28%** des cas (*Précision* < 1)
- Temps de calcul compris entre 0,2s et 6s (Temps moyen de 2s)

Résultats obtenus par type de différences

Différences structurelles	A : faux négatifs	B : vrais positifs	C : faux positifs	Précision	Rappel
Éclatement	43	71	23	0.76	0.63
Fusion	6	47	22	0.68	0.89
Void	31	3069	52	0.98	0.99
Remplacement	62	416	33	0.93	0.87
Transfert	141	133	18	0.88	0.49
Total	283	3736	148	0.96	0.93

- Une majorité des appariements univoques identifiés
- Plus de 70% des appariements multivoques identifiés
- Plus de 72% des appariements multivoques corrects



Bilan et apports des travaux

- Méthode de comparaison et d'appariement
 - Définition d'une nouvelle méthode d'appariement de diagrammes de classes
 - Introduction de motifs structurels pour diriger l'appariement
- Application au diagnostic dans un EIAH
 - Indépendance du module de diagnostic et du module pédagogique
 - Analyse en termes de différences
 - Production de résultats dans tous les cas de figure



Limites

- Appariement glouton sans *backtracking*
- Appariement de diagrammes distants moins performant
- Paramétrage *ad hoc* de la mesure de similarité et des apparieurs
- Prise en compte d'un seul diagramme de référence



Perspectives

- Utiliser plusieurs références pour analyser le diagramme construit par l'apprenant
 - Des diagrammes complets ou des parties de diagrammes
 - Des diagrammes corrects ou erronés
- Appliquer la méthode à d'autres types de modèles
 - La transposition est directe pour les modèles Entité-Relation
 - Pour les autres modèles, la transposition nécessite l'étude des métamodèles
- Mettre l'outil de diagnostic à disposition de l'enseignant
- Appliquer les résultats au suivi et à la modélisation de l'apprenant