

Cours 11 - Raisonnement à Partir de Cas

MARIE-HÉLÈNE ABEL

Table des matières

I - Introduction	5
A. Présentation d'un problème/cas.....	5
B. Analyse de la résolution du problème/cas.....	5
1. Etape 1.....	6
2. Etape 2.....	6
3. Etape 3.....	6
C. La réutilisation en résolution de problème.....	6
D. Définitions.....	6
E. Synthèse.....	7
II - Un peu d'histoire	9
A. Modèle de mémoire (Marvin Minsky).....	9
B. Modèle de mémoire dynamique (R. Shank).....	10
C. Avantages liés au RàPC.....	10
III - Le Raisonnement à Partir de cas	13
A. Les principes.....	13
B. Composantes du système à base de cas.....	15
1. Les processus.....	15
2. Les connaissances.....	16
C. Les modèles de RàPC.....	17
1. Le modèle structurel.....	17
2. Le modèle conversationnel.....	23
3. Le modèle textuel.....	24

Introduction



Présentation d'un problème/cas	5
Analyse de la résolution du problème/cas	5
La réutilisation en résolution de problème	6
Définitions	6
Synthèse	7

A. Présentation d'un problème/cas

Ma tondeuse auto portée refuse de démarrer.

Plus exactement, je peux la démarrer en tirant sur la ficelle, mais le démarreur, lui, ne fonctionne pas.

Pourtant, si je le branche directement sur la batterie, il fonctionne.

L'an dernier, c'était la machine à laver qui ne fonctionnait pas, alors que le moteur tournait si je le branchais directement sur le secteur.

Le réparateur m'a expliqué qu'il y a une sécurité qui empêche de fonctionner si la porte n'est pas fermée, et que c'est cette sécurité qui était grippée (un petit ressort s'était détendu).

Sur ma tondeuse aussi, il y a une sécurité, qui empêche de démarrer si le frein à main n'est pas mis.

Je shunte cette sécurité, et ça démarre.

B. Analyse de la résolution du problème/cas

Qu'ai-je fait?

- J'ai cherché en mémoire un cas de problème similaire.
- J'ai adapté la solution à mon problème d'aujourd'hui.
- Puis j'ai mémorisé ce nouveau cas, pour utilisation ultérieure.

1. Etape 1

La première étape est probablement la plus difficile à automatiser.

- Elle relève typiquement de l'analogie.
- Elle impose de savoir calculer une distance entre le problème actuel et les

problèmes en mémoire. Cela suppose que l'on sache discerner les paramètres pertinents (problème avec un moteur électrique, qui pourtant fonctionne, présence de sécurités) des autres (machine dans la maison contre machine en plein air, machine usagée contre machine quasi-neuve, machine utilisée toute l'année contre machine utilisée quand il fait beau...).

- Elle suppose également une représentation hiérarchique (démarreur = sorte de moteur électrique, etc.).
- Elle suppose, enfin, que l'on ait correctement énoncé le problème, éventuellement lors d'un dialogue.

2. Etape 2

La deuxième étape impose de posséder des opérateurs de transformation

- Il peut s'agir d'opérateurs de transformation de la solution : la sécurité de la machine à laver est dans la porte, celle de la tondeuse sous la pédale de frein.
- On peut également, plutôt que de transposer la solution (passif), transposer la résolution (actif).
- Enfin, il peut y avoir une étape "2bis", de réparation de la solution, si celle-ci n'est pas applicable.

3. Etape 3

La troisième étape peut ne pas consister simplement à stocker le nouveau cas à côté du précédent (mémoire plate), mais à les fusionner en un cas plus abstrait (mémoire hiérarchique).

C. La réutilisation en résolution de problème

La résolution du problème est d'abord la tâche de l'utilisateur, le système informatique **l'assiste** dans cette tâche.



Remarque

La réutilisation de **problèmes résolus** est souvent considérée comme plus efficace que la résolution « à partir de zéro ».

La **capitalisation** à partir de nouveaux cas de résolution de problème est un objectif du système.

D. Définitions



Définition : Le raisonnement à partir de cas (abrégé en RàPC ou CBR pour Case-Based Reasoning)

associe à la donnée d'un problème P une solution, qui est construite en réutilisant la solution mémorisée d'un problème P' jugé similaire à P.

- La remémoration : consiste à rechercher un problème similaire P' dans une base de cas.
- L'adaptation : consiste à adapter la solution du problème similaire P' .
- La mémorisation (éventuelle) du problème P et de sa solution, dans la perspective d'une réutilisation future (correspond à la phase d'apprentissage du cycle).



Un peu d'histoire



Fondements

Les fondements du RàPC proviennent de travaux en sciences cognitives menés par Shank et son équipe durant les années 80.



Remarque

C'est Shanck qui est l'auteur de l'expression "Case-Based Reasoning".

Ses travaux ont mené à la théorie de la mémoire dynamique :

Les processus cognitifs de compréhension, de mémorisation et d'apprentissage utilisent une même structure de mémoire.

A. Modèle de mémoire (Marvin Minsky)

Principe

"Quand on rencontre une nouvelle situation (décrite comme un changement substantiel à un problème en cours), on sélectionne de la mémoire une structure appelée "cadre" (frame). Il s'agit d'une structure remémorée qui doit être adaptée pour correspondre à la réalité en changeant les détails nécessaires."



Rappel : Les frames

Les frames sont des situations "idéales" regroupées en hiérarchie et sont reliés par les différences qui les séparent.



Méthode : Processus

- Sélectionner un frame,
- Tenter d'appliquer le frame (faire le bilan des buts non atteints),
- Appliquer une technique d'adaptation-correction,
- Synthétiser l'expérience pour l'ajouter à la bibliothèque de techniques de correction.



Remarque

L'utilisation des frames de Minsky s'applique à des problèmes statiques.

B. Modèle de mémoire dynamique (R. Shank)

- Notre mémoire contient non seulement de l'information, mais aussi la manière de traiter cette information.

- La mémoire est dynamique, car nos connaissances évoluent dynamiquement en fonction de l'expérience.
- Nos connaissances sont indexées, mais de manière imparfaite, d'où nos exaspérants trous de mémoire.



Définition : Shank distingue deux structures:

- Les scénarios sont des structures de mémoire qui décrivent comment les choses se passent dans telle ou telle situation (exemple: le fait de payer).
- Les MOP (Memory Organization Packets) lient les scénarios en plus grandes séquences.



Exemple : Au restaurant :

On demande une table
On s'assied,
On commande,
On mange,
On paye.



Remarque

On peut payer dans toute autre situation !

Si je fais une erreur dans l'utilisation de ma carte de crédit au restaurant, j'apprendrai de cette erreur et ne la répéterai pas dans d'autres contextes d'utilisation.

Principes

- "Comprendre c'est expliquer"
- Utilisation de scripts pour expliquer des situations.
- Utilisation de l'expérience concrète dans la construction des scripts.



Méthode : Processus de raisonnement

Dans une mémoire d'expériences, organisée en hiérarchie de généralisation :

- Chercher ce qui est le plus près du problème courant,
- Réutiliser le script trouvé en le spécialisant dans le contexte courant,
- Organiser la mémoire pour y intégrer le nouvel épisode.

C. Avantages liés au RàPC



Remarque

Pour certaines applications, la démarche RàPC est plus simple à mettre en œuvre que celle basée sur un modèle du domaine (e.g. base de règles).

Elle permet d'éviter les problèmes d'acquisition des connaissances qui rendent difficile la conception de bases de connaissances de grande taille.

Elle est particulièrement adaptée aux applications dont la tâche est accomplie par des humains expérimentés dans leur domaine :

les expériences sont disponibles dans une base de données, des documents ou chez un expert humain.

On l'utilise pour des domaines n'exigeant pas de solution optimale et dont les principes sont mal formalisés ou peu éprouvés.

Le Raisonnement à Partir de cas



Les principes	13
Composantes du système à base de cas	15
Les modèles de RàPC	17

A. Les principes

- Un cas est un couple (description de problème, description de solution).
- Un problème « bien » décrit permet de retrouver des solutions « réutilisables ».
- Les connaissances nécessaires sont essentiellement dans les cas, mais aussi dans tout ce qui « explique » les cas.
- Les connaissances duales de similarité et d'adaptation fondent le RàPC.

Objectif

Le RàPC a pour objectif de résoudre un problème cible en utilisant une base de cas constituée d'un ensemble fini de cas.



Définition : Un cas noté $(P, \text{Sol}(P))$ est défini par :

- Un problème P
 - les caractéristiques descriptives
 - la tâche à réaliser
- Sa solution $\text{Sol}(P)$
 - les caractéristiques descriptives
 - des éléments d'explication



Exemple : Je veux vendre un véhicule d'occasion

Description du problème

- Année de mise en service, nombre de kms, nombre de portes, ...
- A quel prix puis-je le vendre ?

Description de la solution

- Prix de vente proposé
- Explication du calcul du prix proposé : argus, etc.



Remarque

- Un cas de la base de cas est appelé cas source et noté $(\text{source}, \text{Sol}(\text{source}))$.

- Source est appelé le problème source.



Méthode

- La **remémoration** consiste à choisir dans la base de cas un cas (source, Sol(source)) similaire à cible.
- L'**adaptation** consiste à s'inspirer de la solution Sol(source) afin de résoudre cible.



Remarque : Autrement dit :

Etant donnés source, Sol(source), Cible et les relations entre ces objets, quelle est Sol(cible) ?



Définition : Le carré d'analogie

- Les relations α représentent les relations entre sources et cibles.
- Les relations β représentent les relations entre problèmes et solutions.

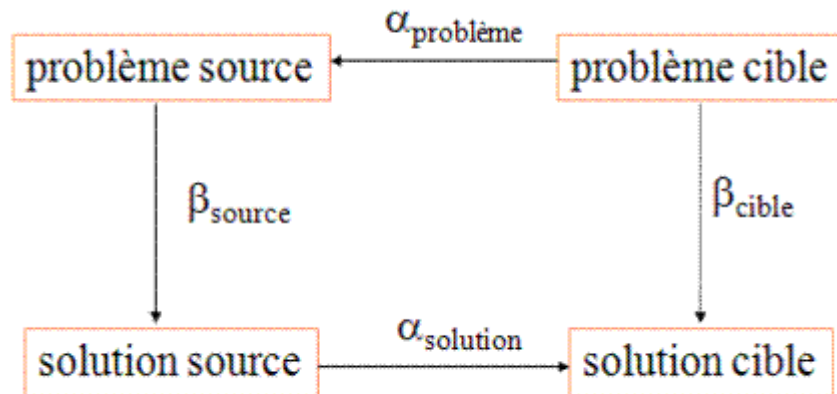


image1

- La relation α problème correspond au processus de remémoration
- La relation β solution correspond au processus d'adaptation



Remarque : La remémoration et l'adaptation peuvent être vus comme deux processus symétriques.

- L'objectif de la remémoration d'un cas similaire est de trouver un cas adaptable.
=> Il est donc important que la connaissance utilisée pour la remémoration soit directement reliée à la connaissance utilisée pour l'adaptation.
- Idéalement, un cas remémoré doit être adaptable afin que sa réutilisation soit garantie pour résoudre un nouveau problème.

B. Composantes du système à base de cas

On trouve des processus et des connaissances.

1. Les processus

- La recherche ou remémoration
- L'adaptation
- La maintenance ou mémorisation
- La construction ou élaboration

a) La recherche ou remémoration

Le processus de remémoration est habituellement implanté par une des deux solutions :

- Approche des plus proches voisins :
Utilise des métriques de similarité pour mesurer la correspondance entre chaque cas et le nouveau problème à résoudre.
- Approche par induction :
Génère un arbre qui répartit les cas selon différents attributs et qui permet de guider le processus de recherche.



Remarque : Approche des K plus proches voisins

Cette approche consiste à affecter au cas X la classe la plus représentée parmi les K plus proches voisins de X

b) L'adaptation

Suite à la sélection de cas lors de la phase de remémoration, le système RàPC aide l'utilisateur à modifier et à réutiliser les solutions de ces cas pour résoudre son problème courant.



Méthode : Deux approches peuvent être suivies

- Approche transformationnelle (ou structurelle) :
On obtient une nouvelle solution en modifiant des solutions antécédentes et en les réorientant afin de satisfaire le nouveau problème.
- Approche générative (ou dérivationnelle) :
On garde une trace des étapes qui ont permis de générer la solution. Pour un nouveau problème, une nouvelle solution est générée en appliquant l'une de ces suites d'étapes.



Remarque

Peu de systèmes RàPC font de l'adaptation complètement automatique.

=> Une intervention humaine est nécessaire pour générer partiellement ou complètement une solution à partir d'exemples.

c) La mémorisation (maintenance)

Durant le cycle de vie d'un système RàPC, les concepteurs doivent préconiser certaines stratégies pour :

- Intégrer de nouvelles solutions dans la base de cas.
- Modifier les structures de système RàPC pour en optimiser les performances.



Méthode

- Une stratégie simple est d'insérer tout nouveau cas dans la base.
- D'autres stratégies visent à apporter des modifications à la structuration de

la base de cas (e.g. indexation) pour en faciliter l'exploitation.

d) La construction (élaboration)



Rappel

On cherche une solution similaire à partir de l'énoncé d'un problème.

Processus en amont des activités de résolution de problèmes du système RàPC.

Il permet la structuration initiale de la base de cas et des autres connaissances du système à partir de différentes ressources (documents, BDD, interviews ...)



Méthode

- Compléter et/ou filtrer la description du problème en se fondant sur les connaissances disponibles sur l'adaptabilité.
- Commencer à résoudre le problème

=> Orienter la recherche d'une solution adaptable



Remarque

Il est souvent effectué manuellement car nécessite une connaissance du cadre applicatif pour guider, entre autre, la sélection du vocabulaire d'indexation et la définition des métriques de similarités.

2. Les connaissances

- Vocabulaire d'indexation
- Base de cas
- Mesures de similarité
- Connaissance d'adaptation



Définition : Vocabulaire d'indexation

Ensemble d'attributs qui caractérisent la description de problèmes et de solutions du domaine (rôle important dans la phase de remémoration).



Définition : Base de cas

Ensemble des expériences structurées, exploitées par les phases de remémoration, adaptation et mémorisation.



Définition : Mesure de similarité

Fonctions pour évaluer la similarité entre deux ou plusieurs cas.



Remarque

Ces mesures sont définies en fonction des attributs et sont utilisées pour la remémoration.



Définition : Connaissances d'adaptation

Heuristiques du domaine, habituellement sous forme de règles.



Remarque

Elles permettent de modifier les solutions et d'évaluer leur applicabilité à de

nouvelles situations.

C. Les modèles de RàPC

Il existe trois grandes familles de RàPC

- Structurelle.
- Conversationnelle.
- Textuelle.

1. Le modèle structurel

Il émerge avec les premières vagues applicatives du RàPC.

- Toutes les caractéristiques importantes pour décrire un cas sont déterminées à l'avance par le concepteur du système : élaboration d'un modèle de données du domaine applicatif.
- Les cas sont complètement structurés et sont représentés par des paires <attribut, valeur> (similaires à des frames ou des objets).

Calcul de la similarité

- La similarité entre deux cas est mesurée en fonction de la distance entre les valeurs de mêmes attributs. Cette distance est fréquemment estimée par les mesures euclidienne et de Hamming.
- La similarité globale entre deux cas est habituellement évaluée par une somme pondérée de la similarité de chacun des attributs.
 - Les attributs d'un cas n'ont pas toute la même importance.
 - L'importance peut varier d'une situation à l'autre
=> Un poids est attribué à chaque attribut.



Remarque

- Tous les travaux sur l'adaptation de cas sont menés dans le cadre du modèle structurel.
- L'adaptation peut varier d'une simple substitution de la valeur d'un attribut jusqu'à la restructuration complète d'une solution.

a) Exemple : configuration d'un ordinateur multimédia [BerWil98]

- L'utilisateur spécifie les applications qu'il souhaite exploiter (traitement de texte, musique, programmation, jeux).
- Chaque logiciel est noté selon l'importance accordée par l'utilisateur.
- L'objectif est d'élaborer la configuration idéale supportant les logiciels demandés en fonction de leur importance.
- La solution est représentée selon une structure "objet" ou frame d'un PC et de ses composants.

i - Adaptation transformationnelle



Méthode : Des éléments de la solution du cas retrouvé sont :

- Modifiés.
 - Supprimés
 - Ajoutés
- selon :



- Des écarts de contexte observés entre cas source et cas cible,
- Et grâce à un ensemble de règles d'adaptation.



Exemple

- Nouveau cas
- Jeux = 0
- Musique = 10
- Tdt = 5
- Prog = 5
- Puissance = 10

- Cas retrouvé
- Jeux = 10
- Musique = 0
- Tdt = 5
- Prog = 5
- Puissance = 10

- Solution
- Carte ASUS-3
- Processeur pentium 250
- Carte graphique Matrox G2
- Joystick JK600
- CD-Rom Sony 14X

image2



Exemple : Règles d'adaptation

- Si (source.jeu > 7) et (cible.jeu < 3) alors solution.supprimer(Joystick)
- Si (source.musique < 3) et (cible.musique > 7) alors solution.ajouter(carte_son_haut_de_gamme)
- Si (source.puissance < 3) et (cible.puissance > 7) alors solution.modifier(processeur.vitesse, delta_puissance, +)



Exemple

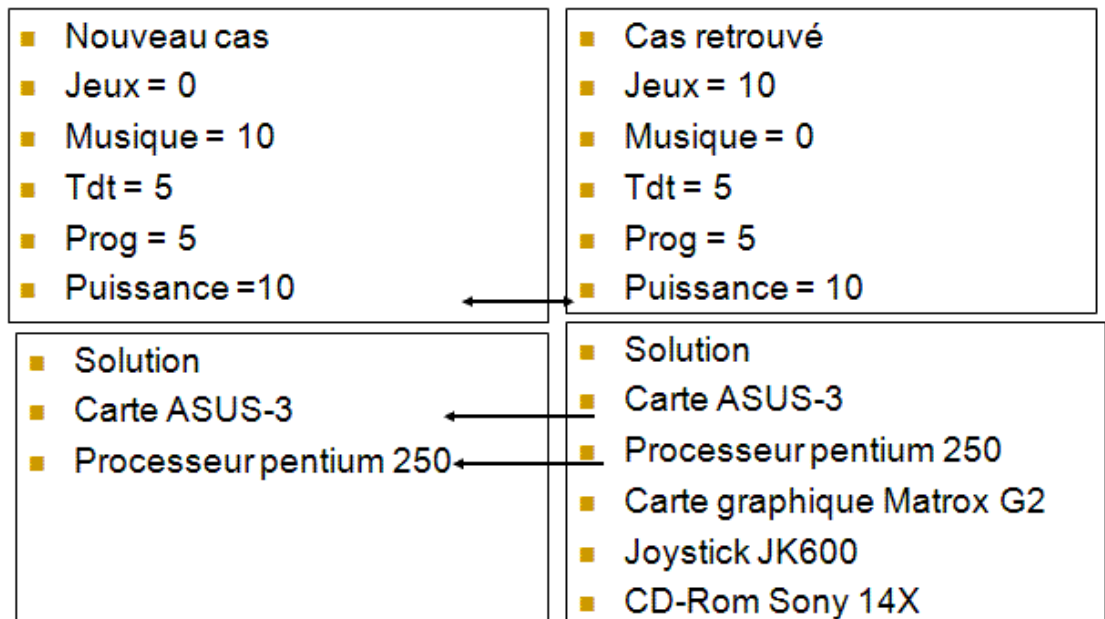


image3



Exemple

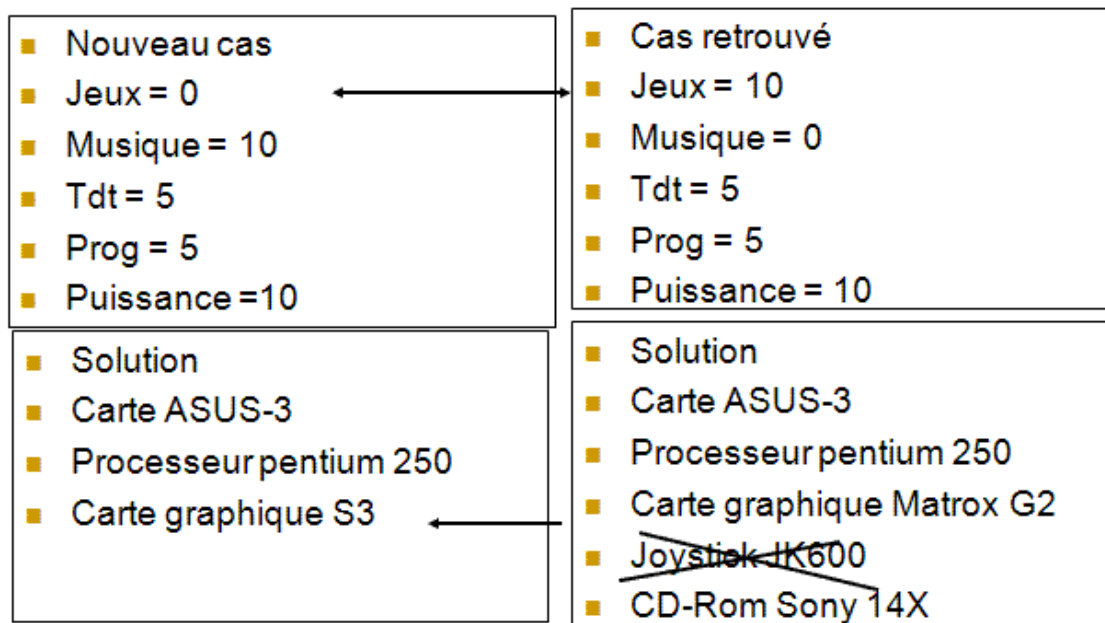


image4



Exemple

<ul style="list-style-type: none"> ■ Nouveau cas ■ Jeux = 0 ■ Musique = 10 ■ Tdt = 5 ■ Prog = 5 ■ Puissance = 10 	<ul style="list-style-type: none"> ■ Cas retrouvé ■ Jeux = 10 ■ Musique = 0 ■ Tdt = 5 ■ Prog = 5 ■ Puissance = 10
<ul style="list-style-type: none"> ■ Solution ■ Carte ASUS-3 ■ Processeur pentium 250 ■ Carte graphique S3 ■ Carte son midi 720 	<ul style="list-style-type: none"> ■ Solution ■ Carte ASUS-3 ■ Processeur pentium 250 ■ Carte graphique Matrox G2 ■ Joystick JK600 ■ CD-Rom Sony 14X

image5

<ul style="list-style-type: none"> ■ Nouveau cas ■ Jeux = 0 ■ Musique = 10 ■ Tdt = 5 ■ Prog = 5 ■ Puissance = 10 	<ul style="list-style-type: none"> ■ Cas retrouvé ■ Jeux = 10 ■ Musique = 0 ■ Tdt = 5 ■ Prog = 5 ■ Puissance = 10
<ul style="list-style-type: none"> ■ Solution ■ Carte ASUS-3 ■ Processeur pentium 720 ■ Carte graphique S3 ■ Carte son midi 720 ■ CD-Rom Sony 14x 	<ul style="list-style-type: none"> ■ Solution ■ Carte ASUS-3 ■ Processeur pentium 250 ■ Carte graphique Matrox G2 ■ Joystick JK600 ■ CD-Rom Sony 14X

image6

ii - Adaptation générative



Méthode

- Le cas retrouvé retrace le "raisonnement" ayant mené à la solution.
- On substitue les éléments de contexte du raisonnement retrouvé par les éléments différents du contexte du cas nouveau.
- On effectue de nouveau le "raisonnement" dans ce nouveau contexte.



Exemple

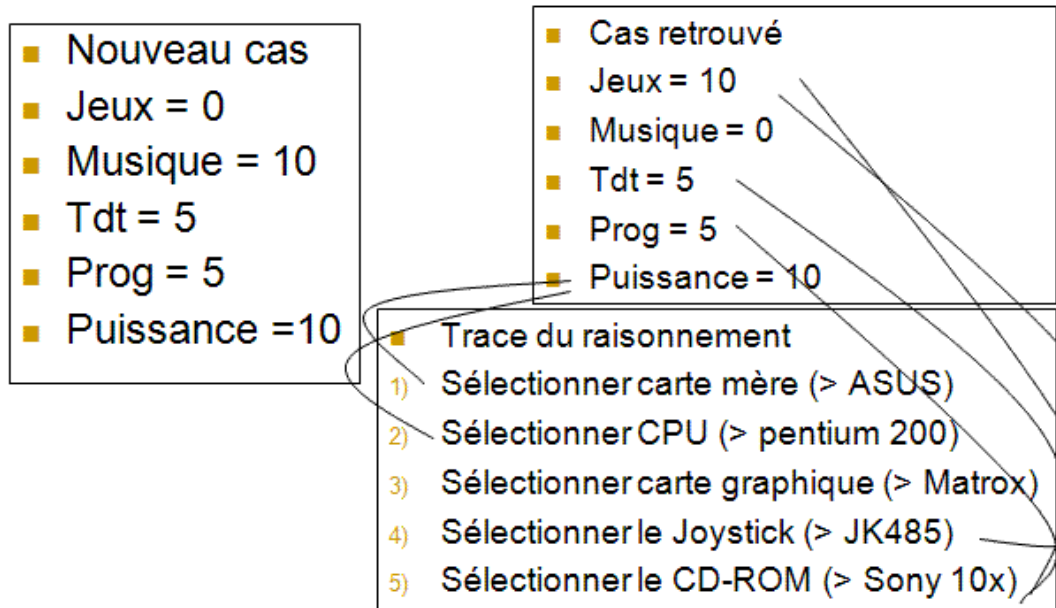


image7



Exemple

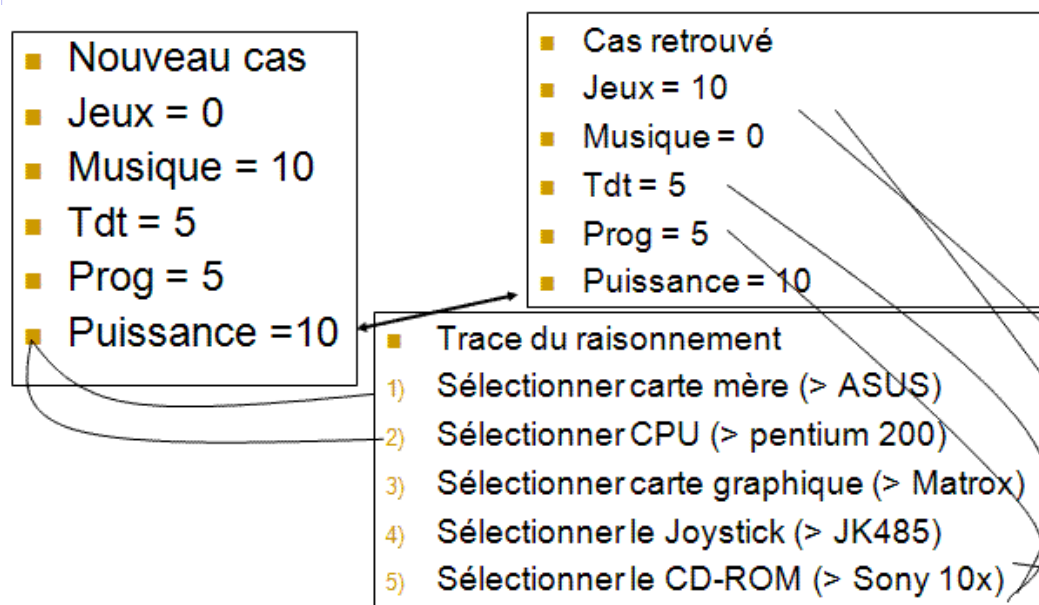
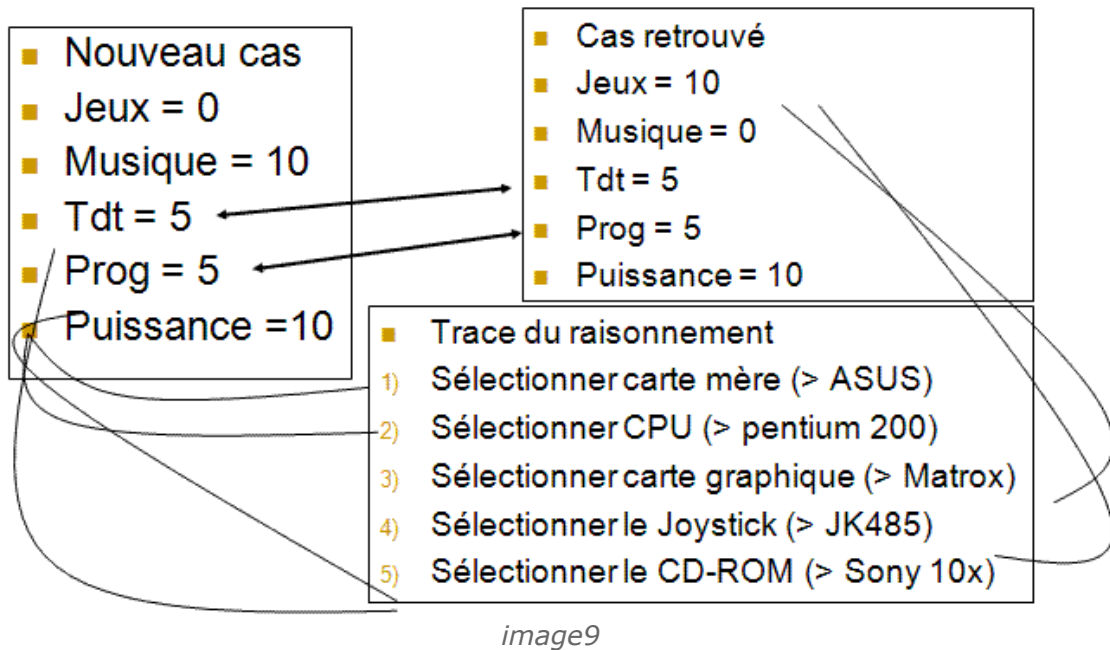


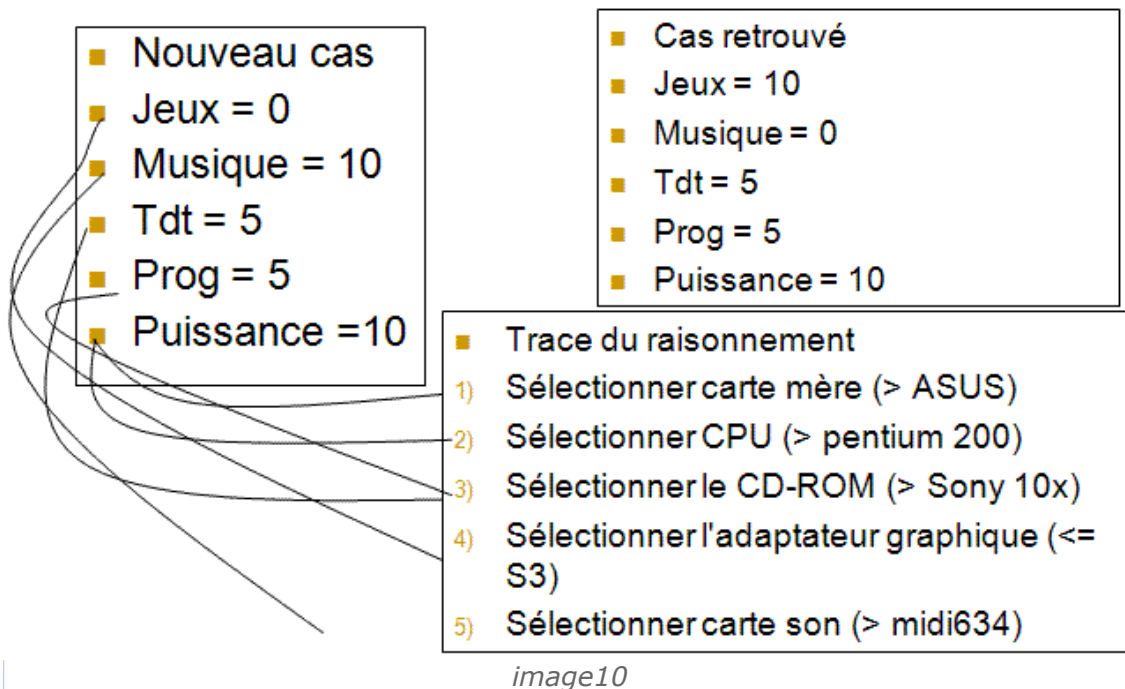
image8



Exemple



Exemple



Remarque

- Trace de raisonnement = plan de résolution + justifications (+ alternatives + tentatives ayant échoué ...)
- Moteur de résolution complet = système de résolution de contraintes, planificateur, recherche dans un espace d'états, ...

Réponse : oui Score : (-)

- Quels types de problèmes avez-vous ?

Réponse : trace noires Score : (default)

- Est-ce qu'un nettoyage de l'imprimante règle le problème ?

Réponse : non ...

Actions : vérifier la cartouche d'encre et la remplacer si le niveau d'encre est faible.



Remarque

La représentation de cas du modèle conversationnel est une extension de celle du modèle structuré avec des attributs de trois types :

- Description,
- Questions,
- Action.



Remarque

La notion de caractéristique (trait) est étendue à la notion de question afin de pouvoir interroger l'utilisateur.

L'interaction entre le système et l'utilisateur se fait comme suit :

- L'utilisateur fournit au système une brève description textuelle du problème à résoudre.
- Le système calcule la similarité entre cette description et les cas de la base de cas.
- Il propose une série de questions à l'utilisateur.
- Pour chaque réponse de l'utilisateur, le système réévalue la similarité de chacun des cas.
- Les questions n'ayant pas reçu de réponse sont représentées par ordre décroissant de priorité.
- Lorsqu'un des cas a atteint un degré suffisamment élevé, il est proposé comme solution.
- Si aucun cas n'a atteint un degré de similarité suffisant et qu'il n'y a plus de questions, le problème est stocké comme étant non résolu.



Remarque

Les systèmes conversationnels n'effectuent pas d'adaptation des solutions passées car la partie "solutions" des cas n'est pas structurée.

3. Le modèle textuel

Les travaux sur le raisonnement à base de cas textuels portent sur la résolution de problème à partir d'expériences dont la description est obtenue dans des documents textuels.

Les cas textuels sont de deux types :

- Non structurés : la description est complètement en "free-text".
- Semi-structurés : le texte est découpé en plusieurs portions étiquetées par des descripteurs tels que "problème", "solution".



Remarque

- Un cas textuel non structuré est un cas dont le seul attribut est textuel.
- Un cas textuel semi-structuré est un cas dont un sous-ensemble de ses attributs est textuel.

Le modèle textuel diffère de l'approche structurelle

Dans l'approche structurelle :

- Les textes sont des chaînes de caractères sans syntaxe ni sémantique,
- Une structuration complète des attributs d'un cas est imposée.

Le modèle textuel diffère de l'approche conversationnelle

- La phase préliminaire de l'approche conversationnelle consiste en une comparaison par mots-clé de courtes descriptions textuelles de problèmes.
- L'interaction avec l'utilisateur ne fait l'objet d'aucun traitement textuel. La langue est utilisée uniquement dans le but de rendre les questions plus intelligibles à l'utilisateur.
- L'approche textuelle date du début des années 90.
- Les efforts produits concernent la phase de recherche sur la base de cas.
- Il n'y a pas de proposition pour l'adaptation de solutions textuelles.