Couplage des approches par contraintes et par analogie : vers une application à la maintenance d'hélicoptères

A. Codet de Boisse[a,b], É. Vareilles[a], T. Coudert[b], P. Gaborit[a], M. Aldanondo[a], L. Geneste[b]

[a]Université de Toulouse - Mines d'Albi
Route de Teillet, Campus Jarlard
81013 Albi cedex 09

[b]Université de Toulouse - ENI de Tarbes
47, avenue d'Azereix, BP 1629
65016 Tarbes Cedex

[a]{nom}@mines-albi.fr et [b]{nom}@enit.fr

Abstract

Nous détaillons dans cette communication une des manières de prendre en compte simultanément de la connaissance générale et de la connaissance contextuelle pour aider à la décision. Nous nous rapprochons ainsi d'un processus cognitif fréquemment utilisé lors d'une prise de décision humaine. Dans nos travaux, la connaissance générale est modélisée comme un problème de satisfaction de contraintes (*CSP*) et la connaissance contextuelle par une adaptation d'un raisonnement par analogie (*RàPC*). Le couplage de ces deux approches est réalisé à l'aide de contraintes qualifiées de contextuelles qui permettent de calculer certaines contraintes du *CSP* grâce à des valeurs extraites des cas du *RàPC*. Dans un premier temps, nous posons le contexte et la problématique de nos travaux et présentons brièvement les différents couplages identifiés. Dans un deuxième temps, nous détaillons notre usage des *RàPC* et les contraintes contextuelles qui exploitent simultanément les deux types de connaissances. Cette proposition est illustrée sur un exemple tiré du projet FUI DGE Hélimaintenance R&D.

Introduction

La considération simultanée de connaissances générales et de connaissances contextualisées est un processus cognitif utilisé fréquemment lors d'une prise de décision humaine. En effet, la plupart des problèmes auxquels est confronté l'être humain peuvent être résolus en faisant appel à deux types de connaissances : des connaissances générales que la personne possède sur la situation, et des connaissances sur des situations similaires déjà vécues englobant tous les aspects qui ont été mémorisés. Des travaux tant en cognitique qu'en intelligence artificielle travaillent d'une part sur la représentation de ces deux types de connaissances et d'autre part sur leur utilisation conjointe en aide à la décision. Cette utilisation conjointe se décline généralement soit en extraction de connaissances abstraites d'un ensemble de situations vécues, soit en l'utilisation des connaissances générales pour adapter une situation vécue et mémorisée à une nouvelle situation à résoudre [4] et [14].

Dans cette communication, nous représentons les connaissances générales sous forme de problèmes de satisfaction de contraintes ou *CSP* et les connaissances contextualisées sous forme de cas utilisables par un raisonnement à partir de cas ou *RàPC*. Nous présentons ici nos réflexions sur la prise en compte de connaissances contextualisées issues d'un *RàPC* ajoutées de manière dynamique lors du filtrage d'un *CSP*.

Nous rappelons dans la section 1 les deux principes de formalisation de connaissances que sont les approches *CSP* et les approches *RàPC*. Puis dans la section 2, nous exposons comment il est possible d'associer celles-ci afin de prendre en compte les deux types de connaissances dans un processus d'aide à la décision. Dans la section 3, nous présentons notre extension des *RàPC* qui nous permet de conseiller l'utilisateur sur les décisions à prendre. Dans la section 4, nous exposons plus en détails les contraintes contextuelles

qui permettent d'évaluer certaines connaissances générales du *CSP* grâce à des connaissances contextualisées du *RàPC*.

Nous terminons par illustrer nos propositions sur une application industrielle d'aide à la décision en maintenance d'hélicoptères dans le cadre du projet Hélimaintenance R&D.

1 Capitalisation de connaissances et approches retenues

Dans cette communication, nous représentons les connaissances générales sous forme de problèmes de satisfaction de contraintes ou *CSP* et les connaissances contextualisées sous forme de cas utilisables par un raisonnement à partir de cas ou *RàPC*.

1.1 Connaissances générales et *CSP*

Selon [11], les *CSP* sont définis comme un triplet $(\mathbb{V}, \mathbb{D}, \mathbb{C})$ où $\mathbb{V} = \{v_1, v_2, ..., v_k\}$ est un ensemble fini de variables, $\mathbb{D} = \{d_1, d_2, ..., d_k\}$ est un ensemble fini de domaines de définition des variables, et $\mathbb{C} = \{c_1, c_2, ..., c_m\}$ est un ensemble fini de contraintes portant sur les variables où une contrainte décrit les combinaisons autorisées ou exclues des valeurs des variables. De nombreux auteurs tels que [5] [12] [6] ont démontré que ce type de modèles pouvait permettre l'aide à la décision dans le cas d'opérations routinières pour lesquelles des connaissances générales ou expertes peuvent être extraites, explicitées et formalisées.

Réaliser un outil d'aide à la décision à partir des concepts de *CSP* revient à traduire les connaissances soit sous forme d'élément unique de $\mathbb{C}$, soit sous forme de plusieurs éléments répartis sur le triplet $(\mathbb{V}, \mathbb{D}, \mathbb{C})$ [17]. Le modèle de connaissances est alors confondu avec le *CSP*. Trouver une solution d'un problème donné revient à instancier toutes les variables en respectant toutes les contraintes.

Il existe plusieurs manières de traiter un *CSP*. Nous pouvons vouloir :

- réduire les domaines des variables en supprimant les valeurs ne menant pas à une solution,
- trouver une ou toutes les solutions,
- trouver la solution optimale, ou du moins une bonne solution étant donné un objectif portant sur quelques variables ou toutes les variables.

Dans le cadre de nos travaux, nous souhaitons répercuter les choix utilisateur sur le problème courant en éliminant les valeurs devenues incohérentes de manière interactive : c'est la séquence de choix cohérents qui conduit à une ou plusieurs solutions. Nous utilisons donc des techniques de filtrage pour effectuer des déductions sur le problème en détectant les affectations partielles localement ou totalement incohérentes [1]. L'une des techniques les plus utilisées est la technique de renforcement de la cohérence locale ou cohérence d'arc [11] [10]. La cohérence d'arc vérifie que toute valeur du domaine d'une variable est compatible avec chaque contrainte prise séparément.

En pratique, un choix s'exprime comme une réduction du domaine d'une variable. Par application des contraintes liées à cette variable, cette réduction peut réduire le domaine des valeurs possibles des autres variables liées par ces contraintes. Et de proche en proche, tous les domaines peuvent se retrouver réduits. Lorsqu'on ne peut plus déduire aucune réduction de domaine supplémentaire, l'opération de filtrage est terminée. La réduction des domaines possibles pour chacune des variables guide ainsi l'utilisateur pour effectuer son choix suivant. Les itérations successives de ce processus l'amène à converger peu à peu vers une solution.

1.2 Connaissances contextuelles et *RàPC*

Le concept de raisonnement à base de cas ne nécessite pas d'extraction à proprement parler de connaissances. Celles-ci sont regroupées dans les instances du problème appelées « cas ». Chaque cas est caractérisé par un ensemble de valeurs de descripteurs qui va permettre de le différencier des autres cas présents dans la base et de le comparer ensuite à un nouveau problème. Cette comparaison est dans un premier temps faite localement descripteur par descripteur à l'aide d'une fonction de similarité locale. Puis, dans un deuxième temps, une mesure de distance ou similarité globale agrégeant la similarité locale de chaque descripteur, permet de classer les différents cas par rapport à leur similitude avec le problème soumis [9]. Le résultat fourni est un classement par ordre de similarité de l'ensemble des cas. Il faut alors sélectionner un ou plusieurs d'entre eux, les analyser et les adapter pour répondre au problème courant. La base est ensuite enrichie par le problème courant adapté et révisé, décrit au travers des descripteurs. Nous pouvons donc formaliser un *RàPC* comme un triplet $(\mathbb{D}, \mathbb{L}, \mathbb{F})$ où $\mathbb{D} = \{d_1, d_2, ..., d_k\}$ est un ensemble fini de descripteurs, $\mathbb{L} = \{l_1, l_2, ..., l_k\}$ est un ensemble fini de similarité locale, une par descripteur, et $\mathbb{F} = \{f_1, f_2, ..., f_m\}$ est un ensemble fini de fonctions de similarité globale.

Dans les systèmes à base de cas, les connaissances regroupées et noyées dans les cas n'ont pas besoin d'être explicitées pour être exploitées. Seule une description précise des cas est suffisante. Ces connaissances sont

1. Une affectation est partielle lorsque seul un sous-ensemble des variables est instancié.

contextuelles et plus riches que des connaissances générales. En conséquence, elles sont plus idiosyncrasiques et plus difficilement transposables dans un autre contexte ou une autre situation.

Dans le cadre de nos travaux, nous souhaitons conseiller l'utilisateur en nous basant sur des solutions déjà rencontrées et lui proposer un ensemble de valeurs possibles pour chaque descripteur d_i. Par conséquent, nous n'utiliserons que la partie recherche de cas, basée sur les fonctions de similarités locales et globale, des approches *RàPC*. Nous ajoutons à cette recherche de cas similaires, une fonction d'agrégation de valeurs, spécifique à chaque descripteur. Cette fonction permet, à partir de la liste des cas jugés les plus similaires, de proposer à l'utilisateur l'ensemble des valeurs renseignées pour ceux-ci, descripteur par descripteur et ainsi de le conseiller sur ses choix futurs. Nous nommons *EBA* pour *Experience Based Advice* cette utilisation spécifique du *RàPC*.

2 Présentation des types de couplage identifiés

Sept types de couplage possibles entre les approches par contraintes *CSP* et par analogie *RaPC* ont été identifiés [3] et parfois étudiés. Nous en faisons ici la synthèse :

1. Deux types de couplages visent à valider la connaissance stockée soit dans un *CSP*, soit dans un *RàPC*. Par exemple, les travaux de [7] utilisent des cas pour valider ou invalider les contraintes d'un *CSP*. Un *CSP* peut aussi être utilisé pour qualifier la cohérence de cas d'un *RàPC* par rapport à son modèle de contraintes.
2. Deux types de couplage permettent de créer ou modifier de la connaissance soit dans un *CSP*, soit dans un *RàPC*. En effet, l'analyse des cas d'un *RàPC*, peut conduire à ajouter des contraintes dans le *CSP*. De même, les approches par contrainte peuvent être utilisées pour compléter les cas d'un *RàPC* lorsque ceux-ci présentent des descripteurs non-renseignés.
3. Deux types de couplage utilisent les deux approches de manière séquentielle pour aider à la décision. Les travaux de [14] et [16] utilisent de manière séquentielle les *RàPC* pour identifier les cas les plus similaires au problème courant puis un *CSP* pour aider la phase d'adaptation des solutions retenues. De manière similaire, nous proposons d'utiliser un *CSP* pour limiter les valeurs possibles des descripteurs puis un *RàPC* pour détecter les cas les plus similaires au problème courant et en faire l'adaptation. Ce type de solution a déjà été étudiée dans les travaux de [8] et [2] qui proposent des techniques de solicitation de préférence pour le *RàPC* basées sur un pré-filtrage effectué par un *CSP*.
4. Le dernier type de couplage, objet de cette communication, utilise les deux approches de manière à entrelacer les connaissances générales et contextuelles. Nous définissons un type particulier de contraintes, nommées contraintes contextuelles, permettant d'injecter lors du filtrage d'un *CSP*, des valeurs de variables issues des cas passés. Cette approche se base sur notre extension du *RàPC* ou *EBA* pour fournir des conseils sur les valeurs des descripteurs, en considérant certains d'entre eux comme des contraintes pour le *CSP*. Cette considération est réalisée par une expertise lors de la création ou mise à jour des modèles de connaissance.

Dans la suite de nos travaux, comme l'utilisation de connaissances générales et contextuelles doit être transparente pour notre utilisateur, nous regroupons sous le terme **variable** les variables v_i du *CSP* et les descripteurs d_i de l'*EBA*. Nous définissons alors deux ensembles de variables dont l'intersection n'est pas forcément nulle : l'ensemble des variables apparaissant dans les cas ou CBV et l'ensemble des variables apparaissant dans les contraintes ou CSV.

3 Fonctionnement du moteur de recherche et d'agrégation de cas

3.1 Présentation générale de l'*EBA*

L'*EBA* doit permettre de réaliser deux opérations successives : la recherche et l'identification de cas stockés dans une base de cas puis l'agrégation des valeurs des variables sélectionnées dans les cas.

Pour la recherche et l'identification de cas, le principe est assez similaire à celui utilisé dans un système de *RàPC* sauf que le cas cible n'est pas représenté par un ensemble de couples (variable, valeur unique) mais par un ensemble de couples (variable, domaine cible). Le cas cible est confronté à chaque cas de la base de cas (appelés cas sources) afin d'évaluer leur similarité. Pour chaque variable V_{cible} pour laquelle on souhaite effectuer la recherche, une fonction de similarité permet de calculer la similarité locale entre le domaine des valeurs autorisées pour la variable V_{cible} (noté $\mathcal{D}_{V_{cible}}$) et la valeur de cette variable dans le cas source (il s'agit d'un singleton noté V_{source}). Ensuite, pour chaque cas source, l'ensemble des similarités locales est agrégé afin de calculer une similarité globale entre le cas cible et le cas source (cf. section 3.2). Enfin,

chaque cas source dont la similarité globale est supérieure à un seuil fixé (soit par l'utilisateur, soit par un expert) est sélectionné afin de réaliser la seconde opération de l'*EBA* : l'agrégation des valeurs des variables requises (cf. section 3.3).

In fine, l'*EBA* retourne la liste des variables ainsi que leurs domaines de valeurs agrégées.

3.2 Calcul de similarité

Le calcul de similarité s'effectue de manière globale entre un cas cible et un cas source de la base de cas à partir de mesures locales correspondant aux variables. La similarité locale correspondant à une variable V_{cible} est évaluée en considérant un singleton d'une part (la valeur de V_{cible} dans le cas source : V_{source}), et un domaine d'autre part (le domaine $\mathcal{D}_{V_{cible}}$ de la variable V_{cible}). Classiquement, il suffit de vérifier que $V_{source} \in \mathcal{D}_{V_{cible}}$ pour déterminer que la valeur source correspond exactement au domaine cible. Dans le cas contraire, elle ne correspond pas du tout et se voit donc pénalisée dans le calcul de similarité globale.

Cependant, si pour toutes les variables du cas source $V_{source} \notin \mathcal{D}_{V_{cible}}$, il est souhaitable de ne pas l'écarter trop rapidement en vérifiant d'abord la similarité des valeurs V_{source} avec les valeurs ou les bornes du domaine $\mathcal{D}_{V_{cible}}$ correspondant. En cas de valeurs proches, le calcul de similarité globale en tiendra compte. Nous considérons le cas des variables discrètes, celui des variables continues ainsi que le cas où une valeur V_{source} est absente d'un cas source.

1) Si la variable V_{cible} est discrète (symbolique ou numérique), le calcul de similarité locale correspondant consiste à retenir la similarité maximale entre chaque valeur discrète du domaine $\mathcal{D}_{V_{cible}}$ et le singleton V_{source}. Soit le domaine $\mathcal{D}_{V_{cible}}$ défini par l'ensemble des n valeurs autorisées v_i tel que $\mathcal{D}_{V_{cible}} = \{v_1, v_2, ..., v_n\}$. La similarité locale entre la valeur V_{source} et le domaine $\mathcal{D}_{V_{cible}}$ est donnée par l'équation :

$$sim_{loc}(\mathcal{V}_{source}, \mathcal{D}_{V\,cible}) = \max_i(sim(\mathcal{V}_{source}, v_i))$$

L'expression $sim(V_{source}, v_i)$ fait référence à la similarité entre deux valeurs d'une même variable et dont les valeurs sont comprises entre 0 et 1. Il s'agit d'une connaissance a priori fournie par un expert. Celle-ci peut être formalisée soit à l'aide d'une matrice de similarités, soit à l'aide d'une fonction de similarité comme par exemple une fonction seuil, linéaire, sigmoïde ou encore exponentielle (voir [15] pour un panorama).

2) Dans le cas d'une variable continue, trois cas de figures sont possibles. Soit $V_{source} \subset \mathcal{D}_{V_{cible}}$ alors la similarité est de 1 ; soit $V_{source} < \min(\mathcal{D}_{V_{cible}})$ alors la similarité est de $sim(V_{source}\ , min(\mathcal{D}_{V_{cible}}\ ))$; soit

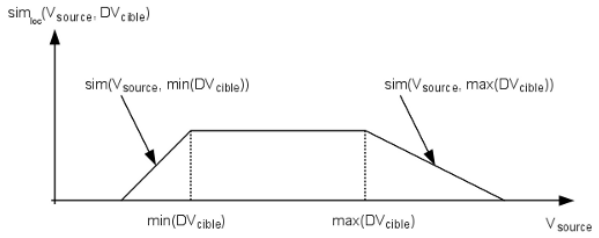

FIGURE 1 – Profil de la similarité entre un domaine $D_{V\,cible}$ et un singleton V_{source} avec variable continue

$V_{source} > max(\mathcal{D}_{V_{cible}}\ )$ alors la similarité est de $sim(V_{source}\ , max(\mathcal{D}_{V_{cible}}\ ))$. Ceci est représenté par l'équation 1.

Nous illustrons sur la figure 1 un exemple de similarité locale d'une valeur V_{source} par rapport à un domaine $\mathcal{D}_{V_{cible}}$ en utilisant une fonction de similarité linéaire. Cette fonction est utilisées pour évaluer la similarité entre une valeur V_{source} et la borne inférieure ou supérieure du domaine $\mathcal{D}_{V_{cible}}$.

Pour calculer la similarité globale (sim_g), nous utilisons une fonction de Minkowski [15] [1] [13] :

$$sim_g(\mathcal{V}_{source}, \mathcal{D}_{V\,cible}) = 1 - \sum_{i=O}^{Nbvar} [\omega_i * \left(1 - sim_{loc}(\mathcal{V}^i_{source}, \mathcal{DV}^i_{cible})\right)^p]^{1/p}$$

avec :
- $Nbvar$: le nombre de variables à considérer ;
- $\mathcal{V}^i_{source}$: la valeur de la variable i dans le cas source considéré ;
- $\mathcal{DV}^i_{cible}$: le domaine de la variable i dans la cible ;
- ω_i : le poids donné à la variable i.

Cette fonction permet, en faisant varier le paramètre p, de régler le mécanisme d'agrégation. Si $p = 1$, il s'agit d'une moyenne des similarités locales, pondérées par les poids ω_i. Si $p = 2$, l'agrégation fonctionne comme un calcul de distance Euclidienne. Si p devient grand, la fonction tend vers l'opérateur max.

3) Dans le cas où, pour une variable de la cible, la valeur V_{source} est absente d'un cas source, nous distinguons deux situations :
- la valeur n'a pas été renseignée par manque d'informations. Cela signifie donc que ce cas n'est pas complet et dans cette situation, nous considérons que la similarité locale est égale à 1. En effet, lorsque certaines informations sont manquantes, la valeur source peut être potentiellement égale à la valeur cible ;
- la valeur n'a pas été renseignée car elle n'avait pas de raison d'exister dans ce cas (exemple du carburateur qui n'est pas présent dans un moteur diesel). Dans cette situation, nous ne prenons pas

$$sim(\mathcal{V}_{source}, \mathcal{D}_{Vcible}) = \begin{cases} 1 & \text{si} \quad \mathcal{V}_{source} \subset \mathcal{D}_{Vcible} \\ sim(\mathcal{V}_{source}, min(\mathcal{D}_{Vcible})) & \text{si} \quad \mathcal{V}_{source} < min(\mathcal{D}_{Vcible}) \\ sim(\mathcal{V}_{source}, max(\mathcal{D}_{Vcible})) & \text{si} \quad \mathcal{V}_{source} > max(\mathcal{D}_{Vcible}) \end{cases} \tag{1}$$

en compte cette variable dans le calcul de similarité globale.

3.3 Agrégation de valeurs

La seconde opération réalisée par l'*EBA* permet d'agréger les valeurs de chaque variable à partir de l'ensemble des cas identifiés comme suffisamment similaires au cas cible. Deux situations sont donc prises en compte :

- celle de variable symbolique : l'ensemble des valeurs présentes dans les cas retenus est retourné ;
- celle de variable numérique : l'intervalle englobant l'ensemble des valeurs des cas sélectionnés est retourné.

À la fin de cette étape d'agrégation, il est possible de fournir à l'utilisateur des conseils basés sur les expériences passées.

4 Contraintes contextuelles

Les contraintes contextuelles permettent d'utiliser simultanément de la connaissance générale et de la connaissance contextualisée pour aider à la décision. Celles-ci permettent d'injecter lors du filtrage d'un *CSP*, des domaines de valeurs pour certaines variables, domaines issus des cas passés. Cette approche se base sur l'*EBA* pour proposer des conseils sur les domaines de valeurs des descripteurs. Dans le cas des contraintes contextuelles, la recherche des cas similaires est réalisée sur un sous-ensemble des variables de CBV. Certains domaines de valeurs sont ensuite considérés comme des contraintes unaires pour le *CSP*. Cette considération est réalisée à l'aide d'une expertise lors de la création ou mise à jour des modèles de connaissance.

4.1 Définition des contraintes contextuelles

Une contrainte contextuelle contient cinq paramètres :

$LVRe$ **:** La liste des variables de recherche $\subseteq CBV$, c'est-à-dire l'ensemble des variables sur lesquelles la recherche doit être effectuée dans l'*EBA*.

$LVRo$ **:** Les variables de retour de l'*EBA* $\subseteq CBV$, c'est-à-dire les variables dont on souhaite utiliser le domaine de valeurs dans une contrainte du *CSP*.

LVI **:** La liste des variables du *CSP* impactées par la contrainte contextuelle $\subseteq CSV$.

FC **:** La contrainte elle-même, composée des variables de LVI et de $LVRo$. Si plusieurs variables sont à la fois présentes dans le *CSP* et l'*EBA* ainsi que dans les ensembles $LVRe$ et LVI, nous indiçons leur domaines respectifs par le nom de l'ensemble considéré : V_{LVI} ou V_{LVRo}.

ms **:** La similarité minimale, c'est en fait un seuil endessous duquel la similarité entre la cible (les valeurs cherchées) et la source (les valeurs dans la base de cas) est trop faible pour que les cas soient considérés comme similaires.

Une contrainte contextuelle est donc représentée par un quintuplet $cc(LVRe, LVRo, LVI, FC, ms)$

4.2 Filtrage des contraintes contextuelles

Les contraintes contextuelles sont filtrées si l'un des deux évènements suivants apparaît :

- le contexte de recherche de l'*EBA* est modifié. Le contexte de recherche est défini par l'ensemble des variables de CBV réduit. Si l'une des variables de recherche de CBV est réduite par l'utilisateur ou par propagation de contraintes, le contexte de recherche de solution est modifié. Cette réduction va donc préciser le contexte et potentiellement réduire le nombre de cas à prendre en compte pour conseiller l'utilisateur. La réduction du nombre de cas à prendre en compte pourra avoir un impact sur l'évaluation des contraintes contextuelles du *CSP* et donc sur les domaines des variables de CSV.
- l'espace de solutions du *CSP* est modifié. Si l'une des variables de CSV est réduite, cette modification doit être propagée sur l'ensemble des variables liées à elle par une contrainte. Cette réduction de domaines va donc réduire l'espace de solutions défini par le *CSP*. Si cette réduction opère sur l'une des variables de LVI, il faut à nouveau évaluer la contrainte contextuelle par rapport au contexte défini par l'ensemble des variables de $LVRe$ valuées.

L'algorithme 1 représente le fonctionnement global du système étant donné les domaines et les ensembles précédemment définis.

Alg. 1 Processus d'utilisation du système

– ∴ – *RedV : liste des variables réduites non traitées*
– ∴ – *Vcour : variable courante*
Début
  $RedV \leftarrow$ Réduction utilisateur
  Tant Que ($RedV \neq \emptyset$) **Faire**
    Vcour ← SHIFT($RedV$)
    – ∴ – *Si la variable courante appartient à l'ensemble $LVRe$ d'une contrainte contextuelle et a été réduite, le contexte a été modifié. Il faut donc évaluer cette contrainte contextuelle par rapport à ce nouveau contexte.*
    Pour (chaque contrainte contextuelle cc dont $LVRe$ contient $Vcour$) **Faire**
      Recherche des cas similaires au contexte dans l'EBA selon les $LVRe$ réduites
      Sélection des cas les plus similaires
      Agrégation des valeurs des variables pour celles appartenant à $LVRo$
      Piltrage par la contrainte cc
      – ∴ – *La liste des variables réduites non traitées doit être mise à jour.*
      $RedV \leftarrow RedV \bigcup LVI$ nouvellement réduites
    Fin Pour
    – ∴ – *Si la variable courante appartient au* CSP*, il faut propager cette réduction.*
    Si ($Vcour \subset CSV$) **Alors**
      Pour (toutes les contraintes c portant sur $Vcour$) **Faire**
        Si (c est une contrainte contextuelle) **Alors**
          – ∴ – *Ici, le contexte de recherche est défini par l'ensemble des variables de $LVRe$.*
          Recherche des cas similaires au contexte dans l'EBA selon les $LVRe$
          Sélection des cas les plus similaires
          Agrégation des valeurs des variables pour celles appartenant à $LVRo$
        Fin Si
        Filtrage par la contrainte c
        – ∴ – *La liste des variables réduites non traitées doit être mise à jour.*
        $RedV \leftarrow RedV \bigcup CSV$ nouvellement réduites
      Fin Pour
    Fin Si
  Fin Tant Que
  – ∴ – *Les domaines des variables sont ici complétement filtrés par les contrainte du* CSP.
  – ∴ – *Réévaluation de l'*EBA *pour fournir de nouveaux conseils.*
  Recherche des cas similaires au contexte dans l'EBA selon les domaines réduits du CBV
  Sélection des cas les plus similaires
  Agrégation des valeurs des variables
  Présentation des domaines conseillés de l'EBA et déduits du CSP à l'utilisateur
Fin

5 Application à la maintenance d'hélicoptères

Cette section permet d'illustrer sur un cas concret nos propositions : l'estimation du temps et du coût pour de l'aide à la décision en maintenance d'hélicoptère.

5.1 Description du modèle de connaissances

Le modèle se compose de huit variables :

- Durée Théorique DT qui est la durée de maintenance théorique. Cette variable appartient à la fois au CSP et à l'EBA. Son domaine est défini dans le CSP, $D_{DT} = [50, 1000]$ heures et la valeur des cas de l'EBA sont cohérents avec cette plage de valeurs.
- Durée Réelle DR qui est la durée réellement passée à effectuer la maintenance. Cette variable n'appartient qu'à l'EBA et a pour domaine conseil $Dc_{DR} = [92, 1000]$.
- $Pays$ qui est le pays dans lequel l'hélicoptère a eu l'occasion d'évoluer. Cette variable appartient à la fois au CSP et à l'EBA. Son domaine est défini dans le CSP, $D_{Pays} = \{France, ArabieSaoudite, Norvège, Groenland\}$ et la valeur des cas de l'EBA sont cohérents avec cette liste de valeurs.
- $Ambiance$ qui est l'ambiance dans laquelle l'hélicoptère a principalement volé. Cette variable n'appartient qu'à l'EBA et a pour domaine conseil $Dc_{Ambiance} = \{Sable, Normal, Froid, Mer\}$.
- $Visite$ qui est le type de visite effectuée (exprimée en heures de vol). Cette variable n'appartient qu'au CSP et a pour domaine $D_{Visite} = \{25, 50, 100, 200\}$.
- $PrixMO$ qui est le prix de la main d'œuvre dans le pays dans lequel la maintenance s'effectue (on estime que la maintenance se déroule dans le pays où l'hélicoptère évolue). Cette variable n'appartient qu'au CSP et a pour domaine $D_{PrixMO} = \{10, 50, 70, 100\}$.
- Coût Prévisionnel CP qui est le coût théorique de la maintenance de l'hélicoptère. Cette variable n'appartient qu'au CSP et a pour domaine $D_{CP} = [0; +\infty[$.
- Durée Prévisionnelle DP qui est la durée prévisionnelle en fonction des aléas dûs à l'ambiance et au pays. Cette variable n'appartient qu'au CSP et a pour domaine $D_{DP} = [0; +\infty[$.

Visite	DT
25	[50 ; 150]
50	[150 ; 500]
100	[400 ; 750]
200	[700 ; 1000]

TABLE 1 – Contrainte C1

Pays	Prix MO
France	50
Arabie Saoudite	10
Norvège	100
Groenland	70

TABLE 2 – Contrainte C3

5.2 Modèles *CSP* et *EBA*

5.2.1 Modèle du *CSP*

Le *CSP* est composé de quatre contraintes dont une contrainte contextuelle :

C1 : la contrainte liant le type de visite $Visite$ à sa durée prévisionnelle DT. Cette contrainte est représentée par la table de compatibilité en table 1.

C2 : la contrainte liant le coût prévisionnel de maintenance CP à la durée prévisionnelle DP et au prix de la main d'œuvre $PrixMO$. Cette contrainte est représentée par l'équation 2.

$$CP = DP * PrixMO \tag{2}$$

C3 : la contrainte liant le pays de maintenance $Pays$ au prix de la main d'œuvre $PrixMO$. Cette contrainte est représentée par la table de compatibilité 2.

C4 : la contrainte contextuelle permettant de déduire la durée prévisionnelle DP en fonction de l'ambiance principale de vol $Ambiance$ et du pays $Pays$. Le calcul de la durée prévisionnelle DP se fait en multipliant la durée théorique de maintenance DT par un coefficient déduit du rapport entre les durées réelles déjà constatées DR et les durées théoriques initialement prévues DT enregistrées dans la base de cas. Le seuil de similarité des cas utilisable pour évaluer cette contrainte est fixé à 0.7. Cette contrainte est représentée par le quintuplet :

$$C4(LVRe = \{Ambiance, Pays\},$$
$$LVRo = \{DR, DT\}, LVI = \{DP\},$$
$$FC : DP_{LVI} = DT_{LVI} * \frac{DT_{LVRo}}{DR_{LVRo}}, ms = 0.95)$$

Cas	DT	Pays	Ambiance	DR
1	345	France	Sable	359
2	704	Arabie Saoudite	Normal	844
3	178	Norvège	Froid	182
4	371	France	Sable	484
5	250	France	Normal	250
6	932	Arabie Saoudite	Sable	1000
7	545	Groenland	Sable	575
8	366	France	Normal	366
9	246	Arabie Saoudite	Mer	301
10	671	France	Sable	883
11	631	Arabie Saoudite	Mer	768
12	314	France	Sable	343
13	614	Groenland	Normal	740
14	608	Arabie Saoudite	Sable	867
15	788	France	Sable	904
16	878	Arabie Saoudite	Sable	1000
17	485	Norvège	Normal	485
18	755	Norvège	Mer	875
19	73	Arabie Saoudite	Mer	92
20	726	Arabie Saoudite	Normal	867

TABLE 3 – Base de cas

Ambiance	Normal	Mer	Sable	Froid
Normal	1	0.8	0.7	0.5
Mer		1	0.8	0.2
Sable			1	0.3
Froid				1

TABLE 4 – Similarité pour l'ambiance

5.2.2 Modèle de l'*EBA*

La base de cas utilisée pour l'exemple illustratif est celle représentée dans le tableau 3.

Trois fonctions de similarités locales sont définies dans l'*EBA* :

- Une fonction de similarité locale $sim1$ identique pour les durées réelles DR et théoriques DT qui est représentée par l'équation 3. Cette fonction de similarité locale précise par exemple quelle est la similarité entre une durée source V_{source} de 222 heures et un domaine de recherche ciblé $\mathcal{D}_{V_{cible}}$ de $[260; 356]$: $\frac{222-260+100}{100} = 0.62$.
- La fonction de similarité locale pour l'ambiance $Ambiance$, représentée par la table 4. Cette fonction de similarité locale précise par exemple quelle est la similarité entre une valeur d'ambiance de type Mer et une valeur d'ambiance de type $Froid$: 0.2.
- La fonction de similarité locale pour le pays $Pays$, représentée dans la table 5. Cette fonction de similarité locale précise par exemple quelle est la similarité entre le $France$ et la $Norvège$: 0.5.

$$sim1 = \begin{cases} 1 & \text{si} \quad V_{source} \subset \mathcal{D}_{cible} \\ (V_{source} - min(\mathcal{D}_{cible}) + 100)/100 & \text{si} \quad V_{source} < min(\mathcal{D}_{cible}) \\ (-V_{source} - max(\mathcal{D}_{cible}) + 100)/100 & \text{si} \quad V_{source} > max(\mathcal{D}_{cible}) \end{cases} \tag{3}$$

Pays	France	Arabie Saoudite	Norvège	Groenland
France	1	0.6	0.5	0.2
Arabie Saoudite		1	0.2	0
Norvège			1	0.7
Groenland				1

TABLE 5 – Similarité pour l'ambiance

Notre exemple ne comprend qu'une seule fonction de similarité globale :

$$sim_g(\mathcal{V}_{source}, \mathcal{D}_{Vcible}) = 1 - \sum_{i=O}^{4} [1/4 * \left(1 - sim_i(\mathcal{V}^i_{source}, \mathcal{DV}^i_{cible})\right)$$

5.3 Exemples de fonctionnement

Nous illustrons nos propositions sur le modèle précédent. L'état initial du système d'aide à la décision en maintenance est le suivant :

- durée Théorique DT, $D_{DT} = [50, 1000]$,
- durée Réelle DR, $Dc_{DR} = [92, 1000]$,
- $Pays$, $D_{Pays} = \{France, ArabieSaoudite, Norvège, Groenland\}$,
- $Ambiance$, $Dc_{Amb} = \{Sable, Normal, Froid, Mer\}$,
- type de visite $Visite$, $D_{Visite} = \{25, 50, 100, 200\}$,
- prix de la main d'œuvre $PrixMO$, $D_{PrixMO} = \{10, 50, 70, 100\}$,
- Coût Prévisionnel CP, $D_{CP} = [0; +\infty[$,
- Durée Prévisionnelle DP, $D_{DP} = [0; +\infty[$.

5.3.1 Sélection de Visite

Dans un premier temps, l'utilisateur fixe la valeur de *Visite* à 200. Cette variable appartenant à CSV, il faut propager cette réduction. La contrainte **C1** est donc propagée et la valeur de DT est réduite à $[700; 1000]$.

La contrainte contextuelle **C4** est ensuite propagée car la variable DT appartient à cette contrainte. Une recherche dans l'EBA est donc effectuée sur le nouveau contexte. Étant donné que le contexte de recherche de la contrainte **C4** porte sur les variables *Ambiance* et *Pays* qui ne sont réduites, l'ensemble des cas de la base est remonté avec une similarité supérieure à 0.7.

Les ensemble englobant les valeurs de DR_{LVRo} et de DT_{LVRo} sont donc agrégés et remontés : $D_{DR_{LVRo}} = [92, 1000]$ et $D_{DT_{LVRo}} = [73, 932]$. La contrainte **C4** peut donc être propagée, entraînant ainsi $DP = (DR_{LVI} \otimes (DT_{LVRo} \oslash) DR_{LVRo}) \cap DP = ([700; 1000] \otimes ([73; 932] \oslash [92; 1000])) \cap [0; +\infty[= [51.1; 10130.4]$.

DP ayant été modifiée, la contrainte **C2** est alors propagée, entraînant $CP = DP \otimes MO = ([51.1; 10130.4] \otimes \{10; 50; 70; 100\}) \cap [0; +\infty[= [511; 1013043.47]$.

5.3.2 Sélection de Pays

$Visite$ reste fixé à 200, entraînant l'ensemble des déductions faites en section 5.3.1. L'utilisateur vient alors à fixer $Pays$ à *Norvège*.

La contrainte **C3** est tout d'abord propagée entraînant $PrixMO = 100$. La contrainte **C4** est propagée car $Pays \in LVRe$ a été valuée. Les cas $3, 17$ et 18 sont donc sélectionnés car le contexte de la recherche ne porte que sur la variable $Pays = Norvège$. Les ensembles englobant les valeurs de DR_{LVRo} et de DT_{LVRo} sont donc agrégés et remontés : $D_{DR_{LVRo}} = [182; 875]$ et $D_{DT_{LVRo}} = [178; 755]$. La contrainte **C4** peut donc être propagée, entraînant ainsi $DP = (DR_{LVI} \otimes (DT_{LVRo} \oslash) DR_{LVRo}) \cap DP = ([700; 1000] \otimes ([178; 755] \oslash [182; 875])) \cap [51.1; 10130.4] = [142.4; 4148.35]$.

DP ayant été modifiée, la contrainte **C2** est propagée, entraînant $CP = DP \otimes MO = ([142.4; 4148.35] \otimes \{100\}) \cap [511; 1013043.47] = [14240; 414835.16]$.

La variable $Pays$ appartenant à CBV, le contexte a été modifié. Une mise à jour des domaines conseils doit donc être réalisée. La recherche de cas similaires avec un seuil de 0.7 sélectionne les cas $1, 3, 4, 7, 8, 10, 12, 13, 15, 17$ et 18 de la base de cas, présentée par la table 3 et pour lesquels la mesure de similarité est supérieure au seuil fixé. La fonction d'agrégation de valeurs fournit donc les domaines conseils suivants : $DT = [178; 788], Pays = \{Norvège\}, Ambiance = \{Sable, Froid, Mer, Normal\}$ et $DR = [182; 904]$.

5.3.3 Sélection de Ambiance

$Visite$ et $Pays$ restent fixés et les valeurs restent les mêmes qu'à la fin de la section 5.3.2. L'utilisateur fixe alors $Ambiance$ à $Froid$.

La variable $Ambiance$ appartenant à $LVRe$ de la contrainte **C4**, celle-ci doit être propagée. Seul le cas 3 est sélectionné car le contexte de la recherche porte sur les variables $Pays = Norvège$ et $Ambiance = Froid$. L'ensemble englobant les valeurs de DR_{LVRo} et de DT_{LVRo} sont donc agrégés et remontés : $D_{DR_{LVRo}} = 182$ et $D_{DT_{LVRo}} = 178$. La contrainte **C4** peut donc être propagée, entraînant ainsi $DP = (DR_{LVI} \otimes (DT_{LVRo} \oslash) DR_{LVRo}) \cap DP = ([700; 1000] \otimes (178 \oslash 182)) \cap [142.4; 4148.35] = [84.61; 978.021]$.

DP ayant été modifiée, la contrainte **C2** est propagée, entraînant $CP = DP \otimes MO = ([84.61; 978.021] \otimes \{100\}) \cap [14240; 414835.16] = [8461.5384615385, 97802.197]$.

La variable $Ambiance$ appartenant à CBV, le contexte a donc changé. Une mise à jour des domaines conseils doit être réalisée. La recherche de cas similaires avec un seuil de 0.7 sélectionne les cas $3, 7$ et 17 de la base de cas présentée par la table 3. La fonction d'agrégation de valeurs fournit donc les domaines suivants : $DT = [178; 545], Pays = \{Norvège\}, Ambiance = \{Froid\}$ et $DR = [182; 575]$.

5.3.4 Estimation de la durée théorique DT

$Visite$, $Pays$ et $Ambiance$ restent fixés et les valeurs restent les mêmes qu'à la fin de la section 5.3.3. L'utilisateur estime alors la durée théorique de maintenance DT à 700. Nous aboutissons à l'état final suivant :

- durée Théorique DT, $D_{DT} = 700$,
- durée Réelle DR, $Dc_{DR} = 182$,
- $Pays$ $D_{Pays} = \{Norvège\}$,
- $Ambiance$, $Dc_{Amb} = \{Froid\}$,
- type de visite $Visite$, $D_{Visite} = \{200\}$,
- prix de la main d'œuvre $PrixMO$, $D_{PrixMO} = \{100\}$,
- Coût Prévisionnel CP, $D_{CP} = 68461.53$,
- Durée Prévisionnelle DP, $D_{DP} = 684.615$.

Conclusions et perspectives

Le but de cette communication était de présenter une manière de prendre en compte simultanément des connaissances générales et des connaissances contextuelles dans un unique système d'aide à la décision. Les connaissances générales sont modélisées comme un *CSP* tandis que les connaissances contextuelles dans une base de cas.

Dans un premier temps, nous avons rappelé les deux principes de capitalisation et de formalisation de la connaissance utilisés. Puis, nous avons fait la synthèse des différentes manières de coupler ces deux approches pour utiliser les deux types de connaissances dans un processus d'aide à la décision.

Dans un troisième temps, nous avons proposé une utilisation spécifique du *RàPC*, nommée *EBA*, qui nous permet de fournir des conseils à l'utilisateur basés sur les cas passés. Pour ce faire, nous utilisons le principe de recherche de cas similaires du *RàPC* étendue à des domaines cibles, basée sur des fonctions de similarités locales et globales. Nous ajoutons à ce principe de recherche la notion de seuil de similarité qui nous permet de ne retenir que les cas les plus similaires à notre recherche. Une fonction d'agrégation de valeurs est aussi proposée afin de fournir à l'utilisateur l'ensemble des valeurs renseignées pour les cas jugés les plus similaires comme domaines de conseils.

Dans un quatrième temps, nous définissons la notion de contrainte contextuelle. Ce type particulier de contrainte permet d'injecter les domaines de conseils de certaines variables comme domaines dans contraintes du *CSP*. Cette notion repose sur une hypothèse faite par un expert de cohérence entre les connaissances stockées dans le *CSP* et l'*EBA*. La définition des contraintes contextuelles est posée de manière formelle et le processus complet d'utilisation de notre système d'aide à la décision basée sur des connaissances générales et contextuelles est défini.

Dans un dernier temps, un exemple tiré de notre problématique industrielle illustre nos propositions.

Nous envisageons de recueillir des préférences émises par l'utilsateur sous forme de domaines pour améliorer les conseils fournis par l'*EBA*. Par ailleurs, pour aider l'expert à découvrir dans la base de cas des groupes de variables qui pourraient donner lieu à de nouvelles contraintes contextuelles et pour valider l'expression de cette contrainte, nous envisageons d'utiliser des techniques d'analyse de données et statistiques.

Remerciements

Les auteurs tiennent à remercier tous les partenaires du projet Hélimaintenance R&D, en particulier les experts d'*IXAIRCO* pour leur engagement dans la construction des modèles de *CSP* et *EBA*.

Références

[1] Y. Avramenko and A. Kraslawski. Similarity concept for case-based design in process engineering. *Elsevier - Computers and Chemical Engineering*, 30 :548–557, 2006.

[2] C. Boutilier, R. Brafman, C. Geib, and D. Poole. A constraint-based approach to preference elicitation and decision making. In *AAAI Spring Symposium on Qualitative Decision Theory*, page 14, 1997.

[3] Aurélien Codet de Boisse, Elise Vareilles, Michel Aldanondo, Paul Gaborit, Thierry Coudert, and Laurent Geneste. Couplage csp et cbr : premières identifications des modes de couplage. In *MOSIM 2010*, Mai 2010.

[4] A. Didierjean. Apprendre à partir d'exemples : extraction de règles et/ou mémoires d'exemplaires ? *Année psychologique*, 101 :325–348, 2001.

[5] S. Dutta, Wieranga B., and Dalebout A. Designing management support systems using an integrative perspective. *Communications of the ACM*, 40, Issue 6 :70–79, 1997. ISSN 0001-0782.

[6] H. Fargier, J. Lang, and T Schiex. Mixed constraint satisfaction : a framework for decision problems under incomplete knowledge. In *AAAI-96*, pages 175–180, 1996.

[7] Alexander Felfernig, Gerhard Friedrich, Klaus Isak, Kostyantyn Shchekotykhin, Erich Teppan, and Dietmar Jannach. Automated debugging of recommender user interface descriptions. *Applied Intelligence*, 31(1) :1–14, 2007.

[8] V. Ha and P. Haddawy. Toward case-based preference elicitation : Similarity measures on preference structures. In *Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence*, page 193–201, 1998.

[9] J. Kolodner. *Case-based reasoning*. Morgan Kaufman Publishers, 1993.

[10] O. Lhomme. Consistency techniques for numeric csp. *International Joint Conference on Artificial Intelligence*, 13 :232–238, 1993.

[11] U. Montanari. Networks of constraints : fundamental properties and application to picture processing. In *Information sciences*, volume 7, pages 95–132, 1974.

[12] H.R. Nemati, Steiger D. M., L. S. Iyer, and R. T et Herschel. An architectural integration of knowledge management. *Decision Support Systems*, 33 :143–161, 2002. ISSN 0167-9236, DOI : 10.1016/S0167-9236(01)00141-5.

[13] H. Nunez, M. Sanchez-Marre, U. Cortes, J. Comas, M. Martnez, I Rodrıguez-Roda, and M. Poch. A comparative study on the use of similarity measures in casebased reasoning to improve the classification of environmental system situations. *Elsevier - Environmental Modelling and Software*, 19 :809–819, 2004.

[14] Lisa Purvis and Pearl Pu. National science foundation grant iri. Technical report, Adaptation Using Constraint Satisfaction Techniques, 1995.

[15] Bergmann R. Experience management : Foundations, development methodology and, internet-based applications. *Springer-Verlag Berlin Heidelberg*, 2002.

[16] Magali Ruet and Laurent Geneste. Search and adaption in a fuzzy object oriented case base. In *Advances in Case-Based Reasoning*. ECCBR, September 2002.

[17] Y. Vernat. *Formalisation et qualification de modèles par contraintes en conception préliminaire*. Thèse de doctorat, École Nationale des Arts et Métier, Bordeaux, France, 2004.