

APPLICATIONS DES SYSTÈMES À BASE DE CONNAISSANCES¹**Par Gilbert Paquette****SOMMAIRE**

Introduction	1
1. Un outil interactif pour l'interprétation d'une loi sociale.....	2
Énoncé du problème.....	2
Représentation des connaissances et outils utilisés	3
Déroulement du projet.....	5
Évaluation du projet.....	6
2. Outil conseil pour l'attribution de prêts à la Caisse populaire	7
Énoncé du problème.....	7
Représentation des connaissances et outils utilisés	7
Déroulement du projet.....	10
Évaluation du projet.....	10
3. Système pour l'identification des mauvaises herbes.....	11
Énoncé du problème.....	11
Représentation des connaissances et outils utilisés	12
Déroulement du projet.....	13
Conclusion.....	14
Références	15

¹ Ce texte est extrait en majeure partie du volume G.Paquette et L. Roy, « Systèmes à base de connaissances, Télé-université et Beauchemin, pp.310-324

Introduction

Dans ce texte, nous examinons en détail trois exemples d'application réalisées au Québec au cours de la première phase d'implantation des systèmes à base de connaissances qui a précédé l'an 2000. Notre objectif ici est de répondre à un certain nombre de questions :

- Qu'est-ce qui caractérise un problème approprié, une implantation réussie?
- Qu'est-ce qui incitait les organisations à implanter des SBC?
- Quels sont les facteurs qui limitaient alors leur implantation?

Pour bien comprendre l'ampleur du phénomène au cours de cette période, mentionnons un article, publié en novembre 1986 dans la revue *La recherche*, présente un inventaire d'une soixantaine de systèmes experts établi par Bruce G. Buchanan, un des auteurs de DENDRAL, le premier système expert. Ces applications se répartissaient dans les secteurs suivants :

Informatique et électronique	24
Fabrication et ingénierie	14
Chimie	5
Géologie	4
Médecine	4
Finance	3
Agriculture	2
Formation	2
Divers	4

Au Québec, une recherche non exhaustive, réalisée au début de 1990 par la Télé-université a permis d'identifier vingt-sept projets de systèmes experts identifiés sur la base des publications dont ils ont fait l'objet ou par références personnelles. On constate qu'au Québec comme ailleurs, l'implantation des systèmes à base de connaissances était en plein essor.

Nous avons choisi de présenter trois cas d'implantation parmi les vingt-sept systèmes à base de connaissances québécois énumérés à la section précédente. Ces trois projets ont tous dépassé le stade d'une première validation. Ils ont été mis à l'essai auprès d'utilisateurs auxquels ils sont destinés, mais ils présentent des traits distinctifs importants quant au domaine d'application, au type de problème résolu et aux modes de représentation des connaissances utilisés.

Le premier cas est l'exemple typique d'un grand système à base de règles dans un domaine toutefois moins classique : l'interprétation de la loi de la protection de la jeunesse par des intervenants sociaux. Le second exemple est un outil de diagnostic à l'intention des conseillers en crédit des Caisses populaires Desjardins quant à l'attribution de prêts; il utilise une combinaison de règles et de schémas. Le troisième cas est un système de classification des plantules de mauvaises herbes conçu à l'intention des bureaux régionaux du ministère de l'Agriculture du Québec; il utilise une représentation des connaissances sous forme de faisceaux, adaptée à ce type de problèmes d'identification.

1. Un outil interactif pour l'interprétation d'une loi sociale

Le système à base de connaissances PRÉSENS (Programme d'Évaluation de Situations d'enfants Signalés) résulte d'une collaboration entreprise en avril 1988 entre des représentants du Centre de services sociaux Laurentides-Lanaudière et du Centre de recherche en gestion de l'UQAM. Il a également été développé avec la collaboration du Centre de services sociaux Richelieu.

La base de connaissances est constituée de plus de 2 000 règles. Celles-ci synthétisent les connaissances d'experts travail leurs sociaux qui ont à appliquer les dispositions de la Loi de la protection de la jeunesse du Québec, aux cas d'enfants en difficulté qui leur sont soumis. Il faut près de 200 règles en moyenne pour décrire chaque phrase de la loi; ceci révèle la marge laissée à l'interprétation et la très grande difficulté.

Énoncé du problème

La tâche d'un travailleur social consiste par essence en une activité professionnelle qui utilise des techniques basées sur les sciences sociales pour rendre plus acceptables les conditions de vie des individus, des groupes et des communautés. Au cœur du programme de protection de la jeunesse, on retrouve un processus d'intervention prescrit par la loi qui fait appel à des standards professionnels dans des cas aussi critiques que l'abandon d'un enfant, la délinquance ou l'abus sexuel.

Les travailleurs sociaux œuvrant dans le cadre de ce programme doivent appuyer leurs décisions sur des faits précis afin de respecter l'intention du législateur. Or, cette intention se résume en un court texte légal d'environ une page qui laisse une très large place à l'interprétation. La difficulté d'une telle intervention semble être la cause d'un taux élevé de roulement du personnel. Les débutants reçoivent leur formation sur le terrain et pour ce faire, ils disposent d'un guide d'interprétation établi par chacun des centres et qui vise à préciser le sens de la loi. Cependant, dans les cas un peu plus complexes, seuls quelques experts sont capables de s'acquitter correctement de cette tâche d'interprétation.

Les Centres de services sociaux ont donc besoin de méthodes pour accroître le rendement des **intervenants** en réduisant les erreurs d'interprétation et de diagnostic. La décision d'entreprendre la réalisation du système PRÉSENS était donc motivée, au départ, par les considérations suivantes :

- l'utilisation de ressources professionnelles rares dans un contexte de contraintes budgétaires peut être maximisée à l'aide d'outils plus adéquats;
- l'application de l'article 38 de la Loi de la protection de la jeunesse est une situation de diagnostic analogue à celles, notamment en médecine, où les systèmes experts ont fait leurs preuves;
- l'application de cette loi depuis 1979 a permis d'accumuler une expertise humaine significative et utilisable dans la conception d'un système expert;
- la rotation rapide des professionnels entraîne une dépense
- d'énergie récurrente pour la formation et la supervision du travail professionnel;
- l'existence d'une liste d'attente de cas d'enfants en difficulté entraîne la nécessité d'une évaluation à la fois plus rapide et plus fiable;
- la construction d'une base de connaissances permettrait de tracer un corridor plus cohérent et plus explicite pour l'Interprétation de la loi.

Représentation des connaissances et outils utilisés

L'outil utilisé pour le développement, la coquille DECIDEX/ARGUMENT a permis de modéliser naturellement les étapes du processus d'intervention prescrit par la loi. La coquille DECIDEX permet, en effet, de diviser les connaissances en bases locales constituant autant de petits systèmes experts, chacun contenant les faits et les règles sur un thème donné. Les conclusions de chaque module expert peuvent être utilisées par d'autres ou servir à construire des rapports à l'intention de l'utilisateur.

La figure 1 nous montre un agencement en chaînage avant de quatre systèmes experts correspondant aux différentes étapes du processus d'intervention. Ces différents « experts » permettent de :

- déterminer la matérialité des faits (étape d'évaluation);
- recommander ou non l'applicabilité de l'article 38 (étape d'évaluation);
- proposer un régime de protection, s'il y a lieu (étape d'orientation);
- proposer un type de placement, s'il y a lieu (étape de traitement).

Chacun des quatre systèmes experts fonctionne par contre en chaînage arrière; par exemple, le système expert permettant de déterminer la matérialité des faits dépend de sept critères :

- exercice des responsabilités parentales,
- développement physique,
- développement mental et affectif,
- mode de vie du gardien,
- troubles de comportement,
- abus physique,
- abus sexuel.

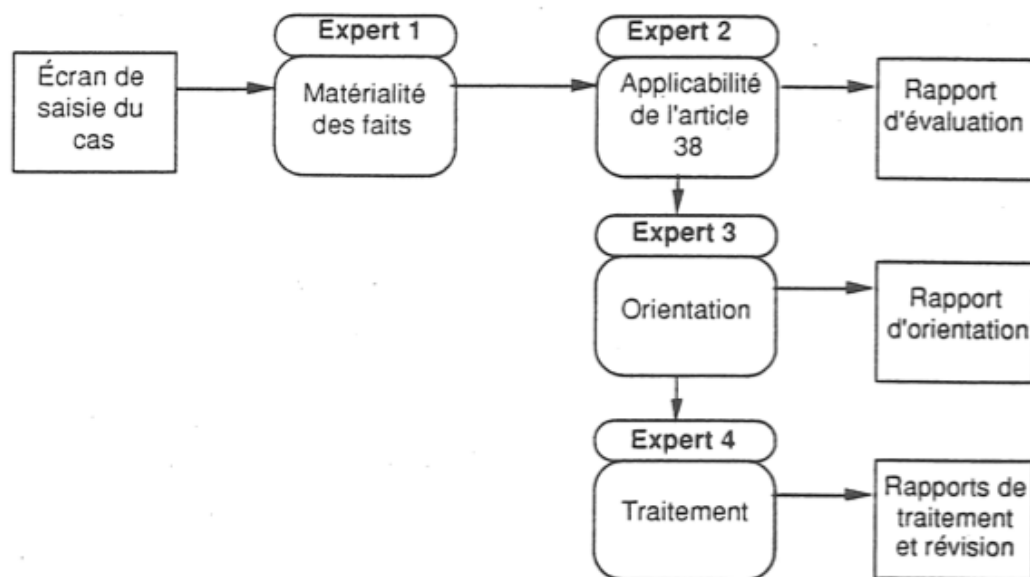


Figure 1 Chaînage avant des modules du système PRÉSENS.

Chacun de ces sept critères ou attributs est évalué par un ensemble de règles regroupées en modules. Par exemple, l'attribut « Abus sexuel » peut prendre une des valeurs suivantes : « ne s'applique pas », « situation à risque élevé » ou « s'applique ». Pour en déterminer la valeur, on utilise dix-huit questions de base et huit contraintes donnant lieu à autant de sous-groupes de règles. Par exemple, l'un de ces sous-groupes de règles consiste à établir s'il y a des faits à l'appui d'une situation d'abus sexuel selon la gravité, la fréquence, l'étalement des actes ou le caractère contemporain (récent) des actes. Voici, à titre d'exemple, deux règles de ce groupe :

RÈGLE 1178

Si ABUS-SEXUEL : GRAVITÉ DES ACTES (ENFANT) =
Activité élaborée

Si ABUS-SEXUEL : FRÉQUENCE DES ACTES (ENFANT) =
Geste isolé (1)

Si ABUS-SEXUEL : ÉTALEMENT DES ACTES (ENFANT) =
Indifférent

Si ABUS-SEXUEL : CONTEMPORANÉITÉ (ENFANT) =
Dernier acte : + de 5 ans

Alors ABUS-SEXUEL : FAITS (ENFANT) =
MATIÈRE INSUFFISANTE

RÈGLE 1172

Si ABUS-SEXUEL : GRAVITÉ DES ACTES (ENFANT) =
Sollicitation sexuelle ou Attouchement

Si ABUS-SEXUEL : FRÉQUENCE DES ACTES (ENFANT) =
Fréquents (11 +)

Si ABUS-SEXUEL : ÉTALEMENT DES ACTES (ENFANT) =
Plusieurs mois ou années

Si ABUS-SEXUEL : CONTEMPORANÉITÉ (ENFANT) =
Contemporain ou dernier acte : + de 6 mois

Alors ABUS-SEXUEL : FAITS (ENFANT) =
DE NATURE CRIMINELLE

Chacun des sept modules constituant le système expert sur la matérialité des faits comporte de 12 à 22 questions de base, de 5 à 11 contraintes ou groupes de règles et de 20 à 31 attributs. Le système expert permettant de compléter l'évaluation et d'établir l'applicabilité de l'article 38 comprend, en outre, 3 autres modules de même taille.

Évidemment, un tel système n'est pas destiné à n'importe quel usager, mais à des spécialistes du domaine. Une session type se conclut en 8 questions si la situation de l'enfant ne nécessite aucune protection. À l'opposé, pour le cas le plus grave, le travailleur social doit répondre à 150 questions. Cette éventualité est peu probable à moins que le signalement de l'enfant indique qu'il est confronté à toutes les situations prévues à l'article 38.

Déroulement du projet

L'article 38 de la loi contient huit alinéas décrivant autant de situations dans lesquelles l'enfant doit être protégé. Pour l'étude de faisabilité, l'alinéa h a été choisi parce qu'il semble présenter une complexité suffisante. Il se lit comme suit :

La sécurité ou le développement de l'enfant est considéré comme compromis si (h) il manifeste des troubles de comportement sérieux et que ses parents ne prennent pas les moyens nécessaires pour corriger la situation ou n'y parviennent pas.

À la fin de la première séance entre les experts et l'ingénieur de la connaissance, on a établi la contrainte principale pour cet article à la figure 2.

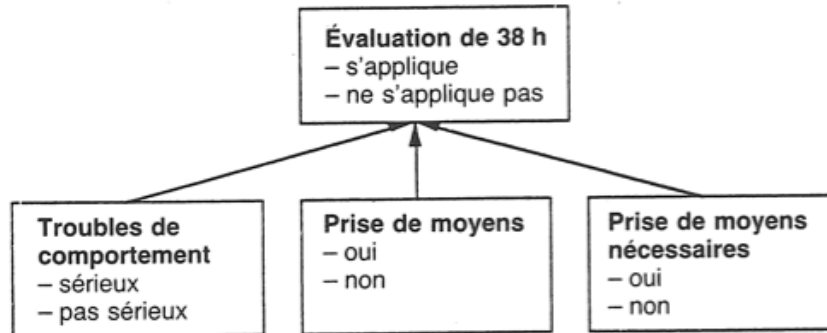


Figure 2 Une contrainte du système PRÉSENS.

Au cours des séances suivantes, on s'attaqua aux contraintes secondaires. Pour déterminer la valeur de l'attribut « troubles de comportement », on développa l'arbre des contraintes à la figure 3:

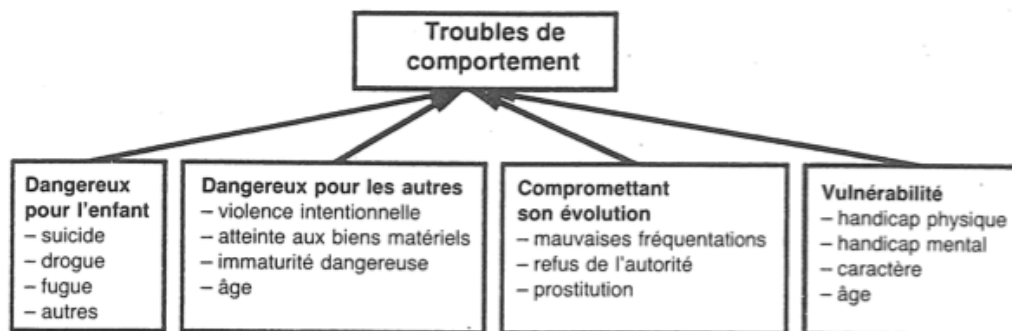


Figure 3 Développement d'une contrainte du système PRÉSENS.

Après la quatrième séance, les experts étaient en mesure de raffiner eux-mêmes le modèle et de le vérifier pour de nombreux cas réels (plus de 70) jusqu'à obtenir satisfaction.

Après quatre mois, l'étude de faisabilité avait réussi et on était en possession d'un premier prototype. Il fallait maintenant réaliser un système pleine grandeur. La démarche qui avait si bien réussi dans le cas de l'article h commença à ne plus être efficace. Il y avait des cas à la limite des deux alinéas et la liste des valeurs des différents attributs ne cessait de s'allonger. De plus, le moment allait bientôt arriver où il faudrait intégrer les différents modules en un tout cohérent.

Après de longues discussions, on a décidé d'abandonner la catégorisation par alinéas qui paraissait, à priori, plus naturelle aux experts. C'est ainsi qu'on en arriva à distinguer les informations concernant les parents (exercice des responsabilités, capacité parentale, implication

parentale, mode de vie des parents) de celles concernant l'enfant (développement physique ou psycho-affectif, comportement, vulnérabilité, enfant-victime).

Dès que le découpage de la connaissance en modules se fut imposé, les experts remarquèrent immédiatement que le système apparaissait comme le squelette conceptuel d'une théorie de la sécurité et du développement psycho-social de l'enfant. [...] Une fois réalisée, cette réaffectation des vieux matériaux, les cases vides furent très rapidement remplies. [...] Ce changement de rythme peut s'expliquer par la familiarisation croissante des experts avec l'écriture sous forme d'attributs, de contraintes et de règles. Avec l'expérience, ils pouvaient maintenant s'exprimer presque directement sous cette formulation de la connaissance. L'articulation par modules significatifs au regard d'une théorie de la sécurité et du développement psycho-social de l'enfant avait également grandement facilité la tâche des experts².

De la sorte, nous nous sommes retrouvés avec dix modules de base et chacun est devenu complet jusque dans la confection de ses règles. [...] La phase de construction finale a consisté à relier ces dix modules dans un système global en deux étapes : sept modules décrivant les situations vécues par l'enfant et permettant de conclure sur la « matérialité des faits » de chacun des alinéas. [...] Puis, étape suivante, les modules restant, confrontés à la matérialité des faits, permettent de se prononcer sur l'applicabilité de l'article 38³.

Par la suite, une étude de validation a été menée où on a comparé les conclusions du système PRÉSENS à celles d'experts locaux pour 109 dossiers réels d'enfants signalés. Les résultats ont permis de conclure que le système offrait une précision équivalente à celle d'un panel d'experts, ce qui constituait une garantie suffisante pour le mettre à l'essai auprès de non-experts.

Enfin, l'étape finale, une étude phénoménologique d'implantation auprès de cinq usagers non experts a permis de bien définir les conditions d'implantation pour un usage optimal du système.

Évaluation du projet

Voici comment le principal expert et animateur du projet en évalue les résultats :

« L'élaboration d'un système expert sur l'évaluation de la situation d'un enfant signalé s'est révélée une entreprise ardue et complexe, mais aussi extrêmement productive au plan professionnel. Une telle opération exige de bien définir l'ensemble et chacun des éléments qui sont utilisés dans de telles évaluations. Il s'agit en fait de monter un guide de pratique détaillant la moindre règle utilisée.

Toutefois, contrairement à un document écrit qui ne peut pas être consulté régulièrement dans le tumulte de la pratique professionnelle quotidienne, le mode conversationnel et la facilité d'adaptation de la connaissance qu'offre un système expert permettent d'y référer assez facilement et donnent une réponse très particularisée à chaque situation.

L'expérimentation sur son implantation permet de constater qu'en plus de constituer un instrument de formation prometteur, l'utilisation d'un système expert force la concertation sur des corridors officiels et partagés d'interprétation de la loi. Il devrait permettre une meilleure circulation de l'information et des remises en question plus systématiques des pratiques quotidiennes⁴. »

² Voir *Une ingénierie de la connaissance*, référence dans la bibliographie.

³ Tiré de : Marcel Frenette, *PRÉSENS, une ouverture sur l'informatique clinique via un système expert*, document du CSSLL, mars 1990.

⁴ Extrait du document venant d'être cité.

2. Outil conseil pour l'attribution de prêts à la Caisse populaire

Ce projet a été entrepris en avril 1988 grâce à une collaboration entre la firme de consultants LGS et les Caisses populaires Desjardins. La motivation principale des Caisses populaires était d'évaluer la technologie des systèmes à base de connaissances par rapport à ses besoins propres.

Il en a résulté un système de taille moyenne, contenant environ 500 règles et une trentaine de classes d'objets. Le système a été intégré à un système d'informations qui gère 150 écrans et comprend une base de données ORACLE. Il est présentement employé dans trois succursales.

Énoncé du problème

Les institutions financières, comme les caisses populaires, reçoivent chaque année des milliers de demandes d'emprunts. Dans chacune des institutions, des conseillers en crédit ont à évaluer le dossier de l'emprunteur et à faire une recommandation quant à la décision d'octroyer ou non le prêt.

Dans ce milieu hautement compétitif, il importe que la décision soit rendue le plus rapidement possible, tout en garantissant l'institution contre les risques découlant d'une mauvaise décision. Il faut donc également que la décision soit d'excellente qualité. Compte tenu des multiples facteurs influençant la décision, il est important de doter les conseillers en crédit, d'un instrument leur permettant de combiner correctement ces facteurs. Ceci est particulièrement vrai lorsque le conseiller est au début de sa carrière. L'expertise intégrée au système expert peut lui permettre d'accroître sa productivité tout en contribuant à sa formation.

Représentation des connaissances et outils utilisés

L'outil retenu par la firme de consultants est la coquille NEXPERT/Object de la firme américaine Neuron Data. Cette coquille, sur micro-ordinateur Macintosh, a l'avantage de combiner des représentations par règles et par schémas, tout en offrant au concepteur et aux usagers des interfaces conviviales.

La présence intégrée de ces deux représentations a fortement influencé, il va sans dire, le choix des modes de représentation des connaissances. Les données de base concernant l'emprunteur ont été organisées en une hiérarchie d'une trentaine d'objets ou schémas. Les règles utilisent les valeurs des attributs de ces objets dans leurs conditions. Lorsque celles-ci sont satisfaites, les règles ayant la plus haute priorité sont déclenchées et la conclusion est déduite. On procède ainsi en inférant d'autres valeurs d'attributs jusqu'à la décision finale d'accorder ou non le prêt.

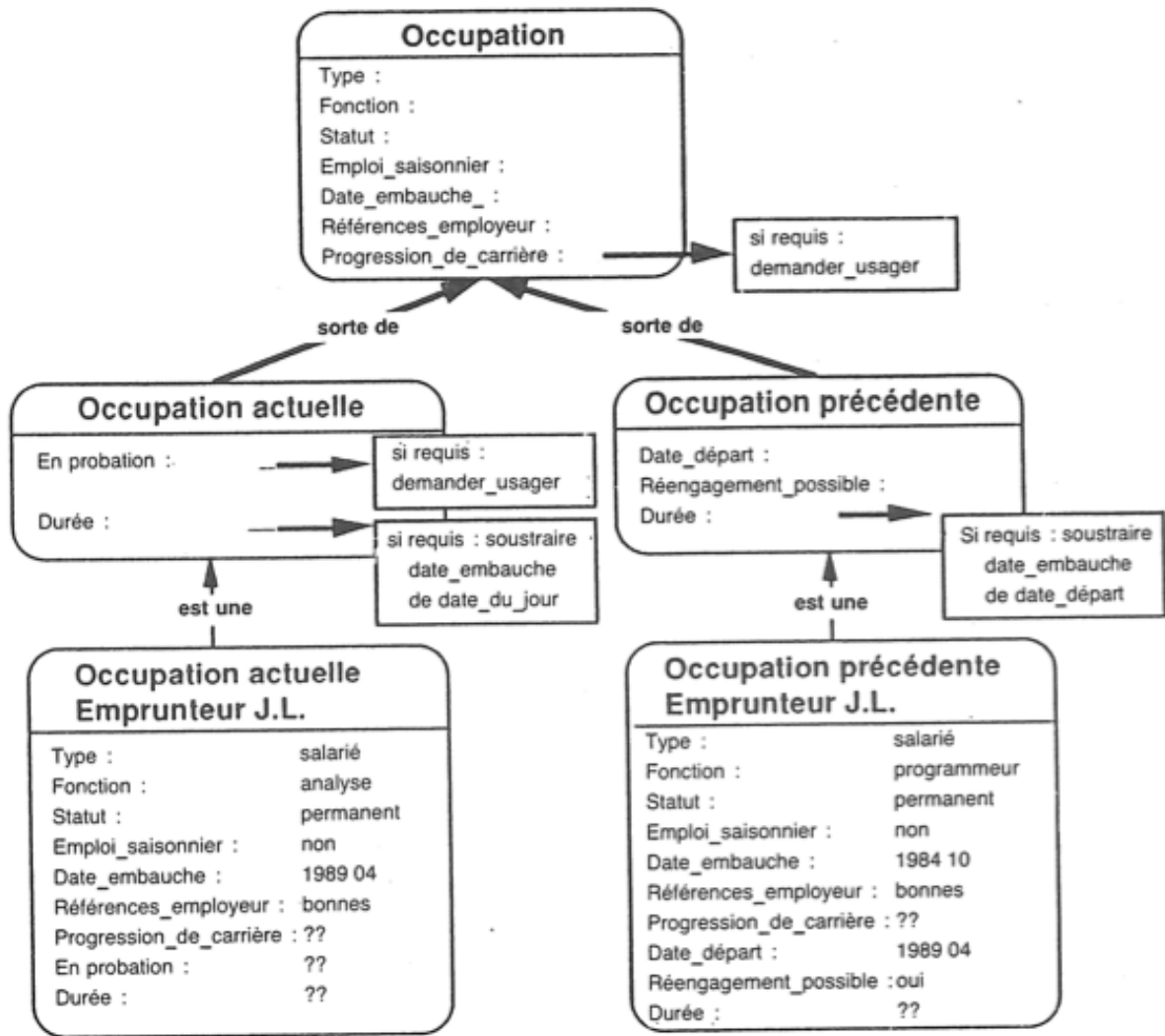


Figure 4 Quelques schémas du système expert en crédit.

La figure 4 montre trois schémas représentant autant de classes d'objets. Les schémas « occupation actuelle » et « occupation précédente » tirent leurs valeurs du schéma « occupation » et ont, en plus, des attributs qui leur sont propres. La plupart des attributs sont des valeurs à remplir à partir de la fiche de l'emprunteur. Par contre, certains schémas font l'objet d'un attachement procédural. Par exemple, si une règle requiert une valeur pour « progression de carrière » ou « à l'essai », la procédure « demander à l'utilisateur » est exécutée pour obtenir la valeur de l'attribut; si la durée de l'occupation est requise, une procédure de soustraction de dates est exécutée de façon à obtenir une valeur pour l'attribut « durée ».

Chaque emprunteur donne lieu à une instance de chacune des classes. La figure précédente nous montre les instances des schémas « occupation actuelle » et « occupation précédente » pour l'emprunteur J. L. Toutes les valeurs des attributs sont connues, sauf celles qui demandent l'exécution d'une procédure.

Examinons maintenant quelques règles concernant la détermination de la stabilité d'emploi de l'emprunteur, un critère important dans l'attribution d'un prêt. Nous utiliserons, dans les conditions comme dans les conclusions des règles, la notation :

objet : attribut = valeur

objet : attribut ≤ valeur

pour les triplets objets-attributs-valeurs.

Règle 1

Si occupation_actuelle_J.L. : type = salarié
et occupation_actuelle_J.L. : références_employeur = mauvaise
Alors évaluation_J.L. : stabilité d'emploi = faible
(Priorité 2)

Règle 2

Si occupation_actuelle_J.L. : type = salarié
et occupation_actuelle_J.L. : références_employeur = bonnes
et occupation_actuelle_J.L. : statut = permanent
et occupation_actuelle_J.L. : durée/5
Alors évaluation_J.L. : stabilité d'emploi = excellente
(Priorité 2)

Règle 3

Si occupation_actuelle_J.L. : type = salarié
et occupation_actuelle_J.L. : durée < 1
et occupation_actuelle_J.L. : à l'essai = oui
et occupation_actuelle_J.L. : durée/5
Alors évaluation_J.L. : stabilité d'emploi = faible
(Priorité 1)

Règle 4

Si occupation_actuelle_J.L. : type = salarié
et occupation_actuelle_J.L. : durée < 3
et occupation_actuelle_J.L. : type = salarié
et occupation_actuelle_J.L. : réengagement_possible = oui
et occupation_actuelle_J.L. : statut = permanent
et occupation_actuelle_J.L. : progression_de_carrière = oui
et occupation_actuelle_J.L. : durée/3
Alors évaluation_J.L. : stabilité d'emploi = bonne
(Priorité 2)

Analysons maintenant le fonctionnement du moteur d'inférence. Le système cherche à déterminer un certain nombre de critères telle la stabilité d'emploi et il place le résultat dans un objet « évaluation_J.L. » associé à chaque emprunteur. Ensuite, ces critères sont combinés pour obtenir la recommandation finale. Voici, de façon simplifiée, quelques étapes que parcourt le système pour déterminer la stabilité d'emploi à l'aide des règles énoncées ci-dessus :

- Faire une liste des règles permettant de trouver la stabilité d'emploi; on obtient (Règle 1, Règle 2, Règle 3, Règle 4).
- Éliminer de la liste les règles dont la condition est fausse. La règle 1 est éliminée, car l'employeur a donné de bonnes références (voir figure 6.7).

- Appliquer la règle prioritaire, la règle 3. Le fait « à l'essai » n'étant pas connu, exécuter la procédure « demander à l'utilisateur ». La réponse étant « non », la condition de la règle 3 n'est pas satisfaite et celle-ci est éliminée :
- Prendre la règle prioritaire dans celles qui restent. La règle 2 et la règle 4 ayant la même priorité, on prend la première. Le fait « durée » étant de même niveau on exécute la procédure « 1989-11 moins 1989-4 » et on obtient 7 mois. La condition de la règle 2 n'étant pas satisfaite, on l'élimine également.
- Reste la règle 4; le fait « progression_de_carrière » n'est pas connu, mais l'exécution de la procédure donne « oui ». Cette partie de la condition est vérifiée, mais la durée de l'occupation précédente est inconnue. On applique la procédure « 1989-04 moins 1984-10 » et on obtient 4 ans et 6 mois. Toutes les conditions étant satisfaites, la règle 4 permet de conclure que l'emprunteur
- J.L. a une bonne stabilité d'emploi.

Déroulement du projet

La première phase du projet a une durée de quatre mois. Pendant cette période, la cognitiennne de la firme LGS et l'experte provenant d'une des Caisses populaires se sont rencontrées chaque semaine. Il s'agissait pour la cognitiennne de faire exprimer à l'experte ses connaissances accumulées pendant ses 15 ans de travail. C'est une tâche difficile, car un savoir accumulé sur une longue période est toujours difficile à exprimer en raison de certaines connaissances devenues « évidentes » aux yeux de l'expert.

Au mois de juin 1988, un prototype comportant 150 règles et une dizaine de classes d'objets est réalisé. On passe à l'étape de validation auprès de trois experts en crédit n'ayant pas participé au développement. On constate alors qu'il y a une différence entre la décision de la succursale et la recommandation du système dans 20 % des cas. On a alors l'idée de confronter les normes de la succursale à celles du siège social.

Cette étape amène une révision de fond en comble du système. La plupart des objets et des attributs restent les mêmes, mais il faut aplanir certaines divergences entre les normes de l'institution et celles de la succursale. De plus, d'autres facteurs doivent être considérés. Selon la cognitiennne, l'expert de l'institution a ajouté de la profondeur aux connaissances : « Il parlait de données, une auto ou une maison par exemple, plutôt que d'un tout comme 50 000 \$ d'actif. »

Après avoir terminé la révision, le système contient 500 règles, trois fois plus qu'à l'origine et une trentaine de classes d'objets. Dans les trois succursales où il est installé, le système s'intègre à un système informatisé couvrant l'ensemble du processus d'attribution de prêts, de l'entrevue jusqu'à l'autorisation.

Évaluation du projet

Le projet a permis de démontrer qu'un système à base de connaissances peut être performant dans ce domaine et qu'il peut être intégré à un environnement informatique traditionnel.

Le projet n'a nécessité, pour ses deux phases, que 500 jours-personnes et il a coûté moins de 200 000 \$ à l'institution. Il a permis d'améliorer le service à la clientèle en réduisant le temps nécessaire à la prise de décision. De plus, l'analyse étant plus complète, l'évaluation est de meilleure qualité et le nombre d'erreurs et de pertes pour l'institution est moindre que par le passé.

3. Système pour l'identification des mauvaises herbes

Ce système a été amorcé à l'automne 1985, grâce à une subvention du ministère des Communications du Québec et un investissement en temps et en matériel du ministère de l'Agriculture, des Pêcheries et de l'Alimentation du Québec (MAPAQ). Ce financement a permis à des spécialistes du MAPAQ, en collaboration avec le centre ATO de l'UQAM, d'entreprendre la réalisation de deux systèmes experts : SADAP, un système d'aide au diagnostic des pathologies animales et SIP, un système d'identification des plantules de mauvaises herbes.

Le premier projet a permis au centre ATO de développer une coquille de système expert appelée Dérédec Expert avec laquelle d'autres projets ont pu être entrepris dans d'autres organisations.

Le second a favorisé le développement, par le centre ATO, d'un langage utilisant un mode de représentation appelé faisceau, lequel semble particulièrement bien adapté aux tâches d'identification. Ces traits distinctifs nous ont incités à décrire en détail ce dernier projet, même s'il n'a pas encore été implanté dans les bureaux régionaux du MAPAQ. Le système est cependant à l'étape de validation. Suite à la révision subséquente, il doit être terminé dans les mois qui viennent.

Énoncé du problème

La protection des cultures repose sur une identification hâtive et exacte des organismes indésirables. En particulier, les mauvaises herbes doivent être identifiées dès qu'elles ont atteint le stade de plantule, de façon que l'on puisse déterminer les traitements herbicides les mieux appropriés.

Or, l'identification des plantules n'est pas une tâche facile. Les systèmes de classification botanique ont été conçus pour les plantes adultes et reposent uniquement sur la similitude des caractères floraux d'espèces voisines. Avant la maturité, il peut être difficile de reconnaître à quelle famille appartient la plante parce que les caractères des espèces d'une même famille sont très variables et rarement exclusifs. Il arrive que des plantules d'espèces proches ont une apparence très différente. Il faut donc utiliser une autre classification que celle des plantes adultes.

L'information sur les plantules est aujourd'hui mieux documentée qu'elle ne l'était, il y a quelques années. Plusieurs publications récentes reposent sur la photographie ou l'image, comme support essentiel. Dans certaines, on trouve, au début de l'ouvrage, une clé ou une démarche d'identification des espèces, mais en pratique, on observe que l'utilisateur passe rapidement aux images, le menant ainsi à poser des diagnostics souvent erronés.

Le problème consiste donc, tout d'abord, à systématiser les connaissances sur les plantules en créant une classification qui leur est propre. Ensuite, cette connaissance, détenue actuellement par un seul expert du Ministère, doit être rendue accessible à des utilisateurs agronomes chargés de conseiller les agriculteurs dans quelque 90 bureaux régionaux du MAPAQ.

À cette fin, un système expert sur les plantules doit permettre aux agronomes de suivre, de façon interactive, la même démarche que l'expert. Comme un document sur papier, le système doit aussi présenter des images facilitant l'identification. Mais il les présente de façon ordonnée, au moment où l'utilisateur en a besoin pour progresser dans sa démarche d'identification.

Représentation des connaissances et outils utilisés

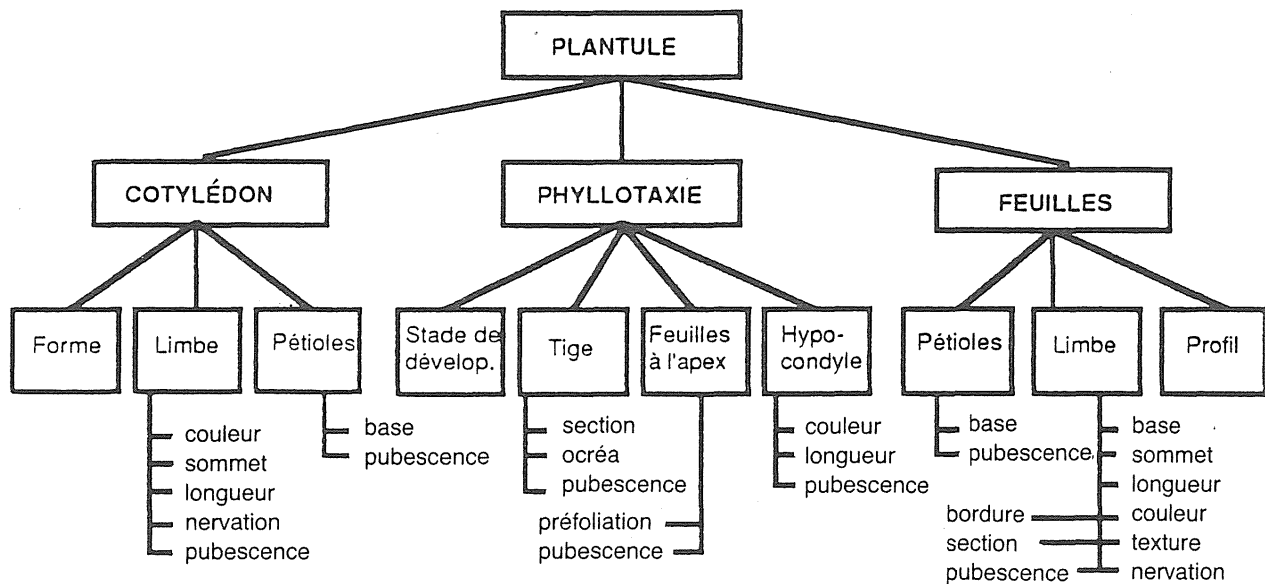


Figure 5 Attributs pour l'identification d'une plantule.

L'identification des plantules, telle que conçue par l'expert du MAPAQ et présentée à la figure 5, repose sur l'examen de leurs organes aériens (hypocotyle, cotylédons et feuilles) et l'observation de phénomènes reliés à leur développement (simultanéité ou non de l'apparition des feuilles à l'apex, évolution de certains caractères, etc.).

On a d'abord tenté de représenter ces connaissances sous la forme de règles déclaratives à l'aide de Dérédec Expert, approche qui avait donné de bons résultats pour le système de diagnostic en pathologie animale. Quelques mois après le début du projet, ce mode de représentation a été abandonné au profit d'un mode plus conforme aux schémas élaborés par l'expert. Certains éléments de l'atelier Dérédec, utilisés pour l'analyse grammaticale assistée, convenaient mieux à l'identification des plantules : les relations hiérarchiques entre les éléments et la modularité du processus d'identification.

Le moteur d'inférence classique des systèmes à base de règles fut donc remplacé par un moteur d'appariement. Celui-ci établit, en quelque sorte, le degré de parenté du spécimen à identifier avec les différentes espèces décrites par l'expert et constitue la base de faits du système. Cette base contient, pour le moment, une soixantaine d'espèces caractérisées chacune par des séquences d'attributs que fournit l'utilisateur au cours d'une identification. Les attributs décrivant les espèces ont été choisis en fonction de la facilité de leur observation et en tenant compte des interprétations différentes auxquelles ils peuvent se prêter. L'utilisateur choisit une valeur pour chaque attribut à l'aide de menus-images. Il pointe l'image qui ressemble le plus à l'attribut de la plantule qu'il tente d'identifier.

Au début ce nouveau moteur construisait, parallèlement à la promenade dans l'arbre de décisions, un autre arbre reflétant tous les choix effectués (une trace). Ce moteur a été remplacé par celui du nouveau langage FX. Dans ce langage, des faisceaux (nœuds de l'arbre) se déploient au fur et à mesure de la progression dans le système, intégrant en un seul concept de représentation la procédure de questionnement de l'utilisateur et sa trace (contenant les réponses). Celle-ci sera utilisée, plus tard, telle quelle pour tirer les conclusions de la consultation. Les espèces compatibles avec le spécimen à identifier sont reconnues par unification de caractères. Ce processus se répète de façon indépendante dans le module « cotylédon », le module « feuilles » et le module « phyllotaxie ». À la fin, l'ordinateur affiche la ou les espèces qui ont été

reconnues et donne la trace des questions et des réponses qui ont jalonné le parcours de l'utilisateur dans le système.

Déroulement du projet

Le projet a débuté au mois d'avril 1987 par une analyse du domaine de l'étude et une première ébauche de la représentation de l'expertise. À l'été, une étude sommaire de 25 espèces a été effectuée sur le terrain. Elle a permis la réalisation d'une première maquette sous forme de règles réalisée avec Dérédec Expert.

À l'hiver 1988, cette approche à base de règles a été abandonnée pour une représentation plus proche des schémas de l'expert, utilisant le système SAGA de l'atelier Dérédec. Puis, en février 1989, le système a été transféré, de nouveau, dans le langage FX, en utilisant une représentation sous forme de faisceaux.

La validation de ce système, réalisée au cours de l'été 1989 et à l'hiver 1990, a consisté à comparer les performances du système à celles des experts et à apprécier la convivialité de l'interface et son efficacité pour l'acquisition des faits de la part de l'utilisateur.

On travaille en ce moment à analyser les données de la validation, à décrire et à ajouter de nouvelles espèces au système et à mettre en œuvre les modifications et les corrections nécessaires. Actuellement, le processus d'unification ne reconnaît que les espèces dont tous les attributs ont été satisfaits. On prévoit assouplir cette procédure en affectant des poids aux descripteurs des espèces, ce qui a comme avantage de diminuer l'impact d'une mauvaise réponse de l'utilisateur.

On prévoit une implantation partielle du logiciel dans quelques bureaux de renseignements agricoles au cours de l'année 90; après quoi, on pourra procéder au remontage du système dans sa version définitive.

Conclusion

Tout Nous avons présenté trois exemples de systèmes experts élaborés et utilisés au Québec avant l'an 2000. Ces exemples, malgré la diversité des domaines d'application, ont en commun un certain nombre de caractéristiques dominantes qui ressortent de cette première phase d'implantation des SBC.

Un bon domaine d'application doit contenir des connaissances de plus haut niveau que des données factuelles, sinon, un système de gestion de bases de données suffirait amplement. Ensuite, le système doit être capable d'adapter ses questions à l'utilisateur en fonction des réponses antérieures qu'il a fournies, grâce à sa capacité d'inférence. Enfin, le domaine doit contenir des informations impliquant un certain degré d'incertitude. Le SBC doit pouvoir faire appel à des heuristiques pour ordonner ses recommandations selon un pourcentage de vraisemblance, sinon un algorithme classique pourrait faire l'affaire.

Les tâches ou problèmes de diagnostic, de classification ou de configuration conviennent bien à un SBC. Le problème ne doit pas être trop vaste. Au-delà de quelques milliers de règles, le système risque d'être trop lent. Mais le problème doit aussi être suffisamment complexe. Si un expert humain peut le résoudre en plus de quelques minutes et en moins de quelques heures, un SBC est sans doute un bon choix.

Les avantages des SBC sont nombreux, ce qui explique que tant d'organisations y ont consacré des efforts importants. Parmi ces avantages, on retrouve l'identification et la conservation des connaissances expertes, la diffusion de l'expertise chez le personnel, la formation du personnel novice, etc.

Par ailleurs, les difficultés d'acquisition des connaissances, de mise au point technique et de maintenance, ainsi que les coûts importants à investir ont ralenti l'implantation des systèmes experts.

Comme nous l'avons souligné dans un autre texte, depuis l'an 2000, la technologie des SBC s'est fondue dans la mouvance générale des technologies de l'information en se déployant sur trois plans : la création de modules à base de règles ou de schémas "embarqués" dans des systèmes plus larges; l'utilisation des systèmes de règles "métier" pour le conseil aux tâches et l'aide à la décision; l'utilisation des SBC et des moteurs d'inférence dans le développement de la troisième génération du Web fondée sur les ontologies, le Web sémantique.

Il n'en demeure pas moins que les concepts, les méthodes et les techniques des systèmes à base de connaissances présentés ici demeurent des éléments importants du bagage de tout professionnel des technologies de l'information.

Références

- Ajenstat, Jacques, Frenette, Marcel et St-Pierre, Alain, « Système expert d'aide à l'intervenant social », dans *Revue des Services sociaux*, Université Laval, février, 1990.
- Bobrow, D. G. et Hayes, P.J., (Éds), « Artificial intelligence, where are we? », dans *Artificial Intelligence Journal*, North Holland, Elsevier Science Publ., n° 25, 1985.
- Bouchard, Claude, Néron, Romain et Goyette, Normand, « Prototype d'identification par ordinateur des plantules de mauvaises herbes », dans *ICO*, Québec, septembre 1989.
- Caron, Carole, « Pour aider les conseillers en crédit », dans *Inférences*, automne 1989.
- Davis, Randall, « The Evolution of Expert Systems Ideas and Tools », dans *Software Tools Conference pour AI/Expert Systems*, novembre 1986.
- Lévy, Pierre, Ajenstat, Jacques et Frenette, Marcel, « Une ingénierie des connaissances », dans *ICO*, Québec, février 1990.
- Ministère des Communications du Québec, *L'intelligence artificielle au Québec*, répertoire, Direction générale des technologies de l'information, 1989.