

Gestion des Connaissances et Analyse Multicritère pour un Système Interactif d'Aide à la Décision en Organisation

Jacky Montmain, Michel Plantié, Abdellah Akharraz

Unité de Recherche sur la Complexité
Ecole des Mines d'Alès - Commissariat à l'Energie Atomique
Site EERIE - Parc scientifique Georges Besse, 30035 – Nîmes, France
{jacky.montmain, michel.plantie, abdellah.akharraz}@ema.fr

Résumé. Cet article propose un modèle conceptuel du processus d'apprentissage et de décision dans une organisation. La perspective d'analyse retenue relève d'une interprétation cybernétique du modèle cognitif de l'économiste H.A. Simon, qui montre que l'acquisition et le traitement de l'information apparaissent comme plus importants pour prendre une « bonne » décision dans une organisation que la recherche fine illusoire et trop simplificatrice d'une décision en apparence la meilleure. L'« homme administratif » n'est pas un « homo œconomicus » aux objectifs clairs et explicites, stables dans le temps, mutuellement indépendants. Une bonne partie de la littérature sur les sciences de l'organisation se borne à le constater. Cet article offre une alternative cybernétique au modèle du processus de décision vu par les sciences de l'organisation. Il propose une chaîne complète de traitement de l'information utile à la décision qui s'appuie sur des outils relevant du data-mining pour la phase d'apprentissage et sur des techniques multi critères pour la sélection.

1. Introduction

Face à un environnement incertain, la décision en situation complexe est affaire de bonne organisation collective. Elle devient l'émanation de savoirs multiples dans les domaines techniques, sociaux, culturels, organisationnels. Résultante complexe d'actions et de stratégies individuelles et de construits collectifs, elle se développe au travers de multiples structures de communication reliant les hommes. A ce titre, l'impact des Technologies de l'Information et de la Communication devient un facteur essentiel du développement de nos modes d'organisation, voire même de nos sociétés [Penalva *et al.*, 2002]. La complexité des systèmes, quelle qu'en soit la nature, dans lesquels l'homme est impliqué aujourd'hui conduit à l'émergence de systèmes dynamiques de traitement de l'information (STI) de plus en plus complexes et incontournables. Pour aider l'homme à faire face à la problématique de la décision en situation complexe, il faut instrumenter la relation que l'homme ou plutôt un collectif organisé, entretient avec son système d'information dans l'action [Simon, 1980 ; 1991]. Les fonctionnalités prioritaires à développer doivent donc être relatives aux processus de gestion des connaissances, d'apprentissage et de décision collectifs dans les grands systèmes sociotechniques [Montmain *et al.*, 2002].

L'abondance de l'information dans une société très consommatrice a changé le paradigme de la décision : on pouvait autrefois énoncer « Le pouvoir, c'est le savoir ». Seul un petit nombre d'élus qui « gouvernaient » savaient et pouvaient donc décider. Aujourd'hui, « tout le monde » est censé avoir accès à l'information, le paradigme tombe. Le véritable problème devient plutôt la capacité à gérer et traiter ce flux informationnel pour en tirer les connaissances pertinentes et utiles au bon moment.

Selon un premier modèle rationnel, la décision a longtemps été et est encore souvent considérée comme le résultat d'un choix comparatif exhaustif entre les diverses solutions possibles. Le décideur et ses conseillers doivent mesurer avec soin les risques et les issues probables de chaque formule, peser leurs avantages et inconvénients pour retenir finalement celle qui représente le meilleur rapport « coût/efficacité ». Cette analyse, en terme de calcul rationnel, postule l'existence d'un acteur unique qui agirait en vertu de préférences hiérarchisées en fonction de la meilleure utilité. Ce séduisant modèle théorique fait malheureusement abstraction des aspects organisationnels souvent implicites dans un processus décisionnel. Cet article propose un modèle conceptuel plus en adéquation avec les contraintes inhérentes à la décision en organisation. La perspective d'analyse retenue relève d'une interprétation cybernétique du modèle cognitif de l'économiste et prix Nobel, H.A. Simon, qui montre que l'acquisition et le traitement de l'information apparaissent comme plus importants pour prendre une « bonne » décision dans une organisation que la recherche fine illusoire et trop simplificatrice d'une décision en apparence la meilleure [Simon, 1983 ; 1997 ; Simon *et al.*, 1964].

L'organisation de l'article est la suivante. Dans un premier temps, nous précisons les spécificités de la décision en organisation. En particulier, nous insistons sur les aspects : acquisition et traitement de l'information, de la connaissance utile à la décision, signification de la notion d'optimalité dans un collectif, évaluation multidimensionnelle et multi-acteurs. Nous rappelons les caractéristiques du modèle de H.A. Simon pour la décision en organisation où il faut considérer que chacune des phases du processus décisionnel engendre des sous problèmes qui, à leur tour, appellent des phases d'intelligence, de conception, de sélection et de bilan : le processus de la décision y est perçu comme un processus dynamique non linéaire. C'est sur la base de cette dernière remarque que notre interprétation cybernétique du processus de la décision repose. La suite de l'article est alors consacrée à la description des fonctionnalités de notre alternative cybernétique au modèle du processus de décision vu par les sciences de l'organisation : chaque section de l'article est consacrée à un élément de la chaîne de traitement de l'information—assimilée à une boucle de commande—nécessaire aux processus d'apprentissage et de décision dans une organisation.

2. Décision en organisation

Les caractéristiques de la décision en organisation ont été introduites initialement par H.A. Simon (1983 ; 1991). Les fondements de sa théorie reposent sur les remarques suivantes :

1. La panoplie des solutions théoriquement possibles est plus large que celles réellement envisagées par une approche purement rationnelle. Les propositions

émanent de scénarios préétablis par des membres bien identifiés de l'organisation. De plus les informations disponibles sont généralement incomplètes, imprécises ou contradictoires ! Par conséquent il paraît bien difficile de pouvoir gérer simultanément dans le temps et partout dans l'organisation l'ensemble des flux d'informations nécessaires à la décision. Le décideur ne possède pas en fait une connaissance totale de la situation, d'où le terme de « rationalité limitée » cher à H.A. Simon [Simon, 1997] et que ses limitations dans la connaissance des faits et hypothèses proviennent principalement des contraintes de l'organisation qui sélectionne ou favorise tel ou tel scénario en fonction de ses intérêts.

2. Dès lors que des ensembles humains sont concernés, les décisions sont de l'ordre du «non-programmable». Dans sa théorie du surcode [Sfez, 1992], L. Sfez montre que derrière l'image trompeuse d'une décision consciente et unifiée, il y a en fait une multiplicité de rationalités différentes qui s'imbriquent, se superposent, se confrontent. La décision optimale apparaît dénuée de sens dans une évaluation multipoints de vue et multi-acteurs. L'incapacité (en tout cas la capacité limitée) de l'homme à traiter l'ensemble des flux d'informations imprécises, incertaines, incomplètes et contradictoires nécessaires à la décision semble montrer que la solution pour une aide à la décision efficace relève, dans ce cas, des systèmes de traitement de l'information. La rationalité «limitée» ou «procédurale» vient se substituer à la rationalité optimisante et «substantive». L'acquisition et le traitement de l'information apparaissant comme plus importants pour prendre une «bonne» décision que la recherche fine illusoire d'une décision en apparence «la meilleure».
3. Toujours selon le modèle de rationalité limitée de Simon, le décideur est naturellement tenté de s'orienter vers une approche monocritère, occultant la prise en compte de la complexité de la réalité et aboutissant au choix d'une solution satisfaisante mais non optimale. Lorsque l'évaluation globale d'un objectif est complexe, il est nécessaire de décomposer l'objectif à atteindre en structurant l'ensemble de critères d'évaluation. L'approche multicritère de l'aide à la décision permet de pallier à cette restriction en augmentant le niveau de réalisme et de lisibilité donné au décideur [Pomerol *et al.*, 1993]. Construire un modèle prenant explicitement appui sur plusieurs critères, traduit et formalise, un mode de raisonnement intuitif et naturel face à un problème de décision qui consiste à analyser séparément chaque conséquence [Roy, 1985].
4. Les différentes phases de la décision ne se présentent pas de façon linéaire, mais en boucles. De nombreuses itérations sont nécessaires, au vu de la faible capacité cognitive des hommes et de la complexité des problèmes de décision, avant qu'un terme ne puisse être apporté au processus de décision. Phases d'intelligence (*Intelligence*), de conception (*Design*), de sélection (*Choice*) et de bilan (*Review*) se succèdent sans logique chronologique préétablie possible (*modèle IDCR*). Les phases sont ainsi des «engrenages d'engrenages». Si le schéma linéaire du processus de décision (la vision *problem solving* de la Recherche Opérationnelle) est ici rejeté, il n'en reste pas moins que dans cette perspective, la planification conserve encore une place privilégiée dans la décision mais est gouvernée par la phase d'information.

Les phases d'acquisition et de traitement de l'information sont donc au cœur de la vision décisionnelle en organisation prônée par H.A.Simon et le terme de modèle STI de la décision désigne le contexte général du traitement de l'information et de la résolution de problème selon Newell et Simon [Newell et Simon, 1972]. Ce paradigme né en 1947 est qualifié de comportemental où la dimension d'expérience, d'apprentissage est essentielle—*The behavioral alternative* [Simon, 1983]. Le grand décideur mythique, *l'homo œconomicus*, y devient une aberration : aucun dirigeant ne peut tout calculer et tout prévoir dans une organisation un tant soit peu complexe aujourd'hui. Plus vraisemblablement, il examine les solutions séquentiellement et il s'arrête de réfléchir quand il trouve la solution lui procurant un niveau minimal de satisfaction et non un optimum. Ceci provient de ses limites physiques et intellectuelles, de ses capacités de conceptualisation, de mémoire, etc [Rojot, 2004].

Trois notions importantes se dégagent alors de ce constat pour notre modèle. Premièrement, l'idée d'apprentissage impose la caractéristique dynamique de notre modèle. Deuxièmement, il faut noter qu'une solution juste satisfaisante n'est pas a priori unique, sa justification devient donc l'enjeu prioritaire de la phase de sélection. Sous cette hypothèse d'équivocité : pour une situation décisionnelle donnée, plusieurs interprétations de la même situation sont possibles selon le point de vue, la fonction, les intérêts de l'acteur en jeu, etc., la délibération et l'argumentation se trouvent ainsi être les étapes majeures de la décision. Un raisonnement peut être défini comme un enchaînement d'énoncés ou de représentations symboliques conduit en fonction d'un but, ce but pouvant prendre des formes variées : démontrer, convaincre, élucider, interpréter, décider, justifier, expliquer, etc. Cet enchaînement est en général non linéaire et nécessite des « boucles » que l'on retrouve dans la quasi-totalité des processus cognitifs humains. L'idée de boucles de rétroaction pour rendre compte du modèle IDCR constitue la troisième grande caractéristique de notre représentation.

Néanmoins, le modèle que nous proposons ne se place pas dans l'optique descriptive et cognitive du modèle S.T.I. Il s'agit de montrer que ce dernier et ses causalités enchevêtrées peuvent trouver une interprétation cybernétique opérationnelle et pas seulement descriptive qui relève davantage de la commande de processus que de progrès techniques ou technologiques informatiques comme la vision informatique décisionnelle des SH pourrait le laisser entendre.

3. Gestion des connaissances utiles à la décision

3.1. Information, connaissance et décision

Toute décision qui se veut rationnelle repose d'abord sur un processus d'information plus ou moins élaboré selon la situation. L'information doit être interprétée en termes de finalité, d'objectifs et de missions. Les techniques de gestion de connaissances apportent alors de nouvelles perspectives d'évolution aux modèles S.T.I. La dualité action/décision et le recours aux connaissances se rejoignent dans le concept de connaissances utiles à l'action ou connaissances actionnables [Argyris *et al.*, 1983 ; Saad *et al.*, 2003]. Le couplage de systèmes de gestion des connaissances (SGC) et de systèmes d'aide à la décision ouvre la voie des *elucidative systems* ou systèmes automatiques d'argumentation où l'acceptabilité de la décision (on peut parler également de légitimité [Damart, 2003]) devient un enjeu essentiel, perpétuant ainsi la vocation explicative des systèmes intelligents, systèmes experts et autres systèmes à base de connaissances, à coloration Intelligence Artificielle (IA) des

années 80 [Dubois *et al.*, 1994 ; Giboin, 1995 ; 1996]. Cette notion de *elucidative systems* fait son chemin par ailleurs dans la communauté « Fusion d'Informations » [Dasarathy, 2000].

On a assisté ces dernières années, au passage de l'ère de l'information à celle de la connaissance. La connaissance à l'inverse de l'information repose sur un engagement, des systèmes de valeurs et de croyances, sur l'intention. La connaissance est bâtie à partir de l'information pour faire quelque chose, pour agir [Penalva *et al.*, 2002].

La gestion des connaissances a été définie comme la mise en place d'un système de gestion de flux cognitifs qui permet à tous les composants de l'organisation à la fois d'utiliser, et d'enrichir le patrimoine de connaissances de cette dernière. La gestion de connaissances se propose ainsi de repérer, formaliser, partager et valoriser les connaissances de l'organisation et en particulier celles qui revêtent un caractère stratégique et décisionnel [Ermine, 1996 ; Penalva *et al.*, 2002a].

Parmi toutes les démarches envisageables en gestion des connaissances, on constate deux tendances [Penalva *et al.*, 2002]:

1. La *capitalisation des savoirs et du savoir-faire*, dont l'objectif principal est de consigner les connaissances stratégiques, et qui porte donc l'effort sur la sélection et la structuration des connaissances;
2. Le *partage dynamique des connaissances* qui ne préjuge pas de leur utilisation future, et n'élimine pas des connaissances dont l'intérêt pourrait être révélé plus tard.

C'est l'IA qui a développé la première voie avec une ingénierie des connaissances fondée sur la constitution de structures de données exploitables et validables dans les domaines techniques où une expertise stable peut être dégagée. La mise en forme de cette expertise est conditionnée par la finalité : l'implémentation en machine à traiter l'information. L'utilisation massive des réseaux d'information et de communication a relancé récemment l'intérêt de constituer des corpus d'éléments de connaissances dynamiques partageables entre acteurs humains. Les systèmes de gestion dynamique des connaissances (SGDC) explorent en priorité cette seconde voie. Un corpus de connaissances y est vu comme recouvrant un domaine de connaissances qui est aussi un domaine d'action et de décision.

Les objectifs assignés à un SGDC doivent être les suivants [Penalva *et al.*, 2002a] :

1. Transformer les échanges d'informations en argumentation utile à l'action (aide à la décision) ;
2. Organiser le partage des connaissances (corpus dynamique d'éléments de connaissances) ;
3. Catalyser l'émergence d'une intelligence collective et l'amplifier.

Si l'on s'en réfère à la phase de perception du modèle S.T.I, il s'agit, en premier lieu, de doter le collectif d'une mémoire commune. Cette fonction doit être assurée par le (SGDC) qui permet de gérer et de contrôler l'évolution du corpus de connaissances produit par le collectif sur le problème à résoudre. Sur le plan technique, un outil intranet peut être utilisé : un serveur WEB s'interfaçant à des bases de données SQL et documentaires (figure 1). Sur le plan organisationnel, il est suggéré une administration centralisée, sachant que l'alimentation du système peut être distribuée sur les agents qui sont habilités à le faire [Penalva *et al.*, 2002]. L'apprentissage d'un savoir et d'une mémoire collectifs est le premier des processus cognitifs pour lequel le SGDC apporte des éléments de solution. Dans ce qui

suit, il est montré en quoi il constitue également un support naturel pour les processus cognitifs relevant d'une intelligence collective que sont l'évaluation et l'argumentation collectives.

L'utilisation des technologies d'information et de communication (TIC) donne une certaine transparence à la décision : en proposant un cadre et une instrumentation explicites pour les processus d'apprentissage puis de choix, d'argumentation, de sélection, elle permet (tant dans l'aspect gestion des connaissances que dans la décision elle-même) de favoriser l'échange hiérarchique et transversal de données, d'informations, de connaissances et de décisions à travers une modélisation systémique et mathématique des liens de subordination et de coordination qui lient les acteurs d'un processus de décision organisationnel [Montmain *et al.*, 2002]. Conformément au modèle S.T.I, instrumenter la délibération doit favoriser le consensus.

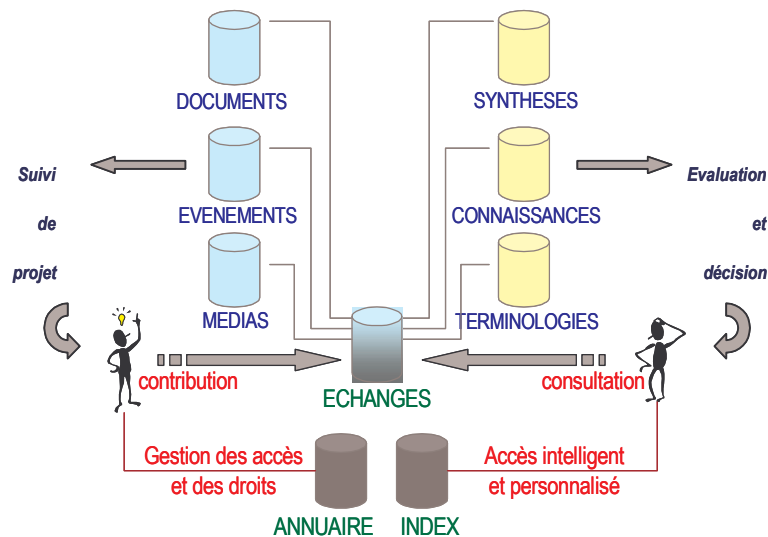


FIG. 1.- Bases et échanges du SGDC

3.2 Connaissances actionnables

Il s'agit d'un savoir élémentaire qui se distingue d'une simple information par le fait qu'il comporte une part d'interprétation liée à la personne qui l'énonce. Si cette interprétation tend à rendre l'information utile à l'action, on peut parler de connaissance actionnable élémentaire, *actionable knowledge*, noté CA [Schön *et al.*, 1983 ; Penalva *et al.*, 2002 ; Saad *et al.*, 2003].

Nous donnons maintenant notre interprétation de la CA dans le cadre de la phase de perception du processus de décision.

Une connaissance actionnable est définie tout à la fois comme :

1. Une donnée informative jugée utile et qui prend du sens dans un contexte;
2. Un savoir élémentaire interprété par la personne qui l'énonce ;
3. Une trace des raisonnements menés par les acteurs de processus décisionnel;
4. Une entité minimale intelligible, de sens partageable et réutilisable dans le contexte.

En résumé, un CA est l'interprétation en termes de finalité par son rédacteur d'une information dans le cadre de son projet : elle sera un élément d'*argumentation* et de *rhétorique* pour une justification de la décision. Le SGDC utilisé pour la phase de perception du processus de décision permet le partage des CAs dans l'organisation.

Pour notre modèle, précisons la forme opérationnelle de la connaissance actionnable dans laquelle la notion de connaissance interprétée permet de distinguer la valeur informative et l'utilité d'une connaissance dans le contexte du projet de l'organisation. Elle est caractérisée par :

- Un *objet* qui est la valeur informative de la CA. Il précise la source d'information, le sujet abordé, les références, hypothèses, etc. Il est en langage naturel.
- Le *commentaire* qui est l'expression d'un jugement de valeur. Il correspond à l'appréciation de la CA, par le rédacteur : il transcrit l'utilité de l'information délivrée dans l'objet de la CA pour le projet d'action. Il est en langage naturel et constitue un élément de rhétorique pour le projet.
- Une *date* qui permet de repérer dans le temps la perception de la valeur informative. L'interprétation de cette dernière peut évoluer dans le temps, un historique de sa perception dans le projet peut ainsi être construit.

3.3 Structure de la base de CAs

En amont de toute opération de gestion de connaissances, la cartographie des connaissances a pour but de mettre en valeur les connaissances critiques de l'organisation et de fournir une structuration des ressources cognitives de l'organisation. Une cartographie est une identification du patrimoine de connaissances [Peil *et al.*, 2001]. Les organisations désireuses de gérer ce patrimoine doivent donc en faire une analyse fine afin de déterminer, dans leur stratégie, quelles sont les connaissances qu'elles doivent développer, faire émerger, abandonner, etc. La cartographie devient elle-même un outil d'aide à la décision.

L'approche la plus courante de la cartographie consiste à une classification conceptuelle (ou par domaines) qui organise l'information autour de sujets, objets, concepts ou finalités. La représentation graphique de la cartographie des connaissances se base sur le principe de la visualisation qui permet de faciliter la navigation et donner une vue globale de la production des connaissances. Pour faciliter la lisibilité, il est toujours préférable de représenter la cartographie sur plusieurs supports (cartes, tableaux, ...).

Précisons maintenant ce que devient cette procédure de cartographie dans le cadre de la phase de perception du processus décisionnel.

Un projet complexe met généralement en jeu plusieurs problématiques (économie, durabilité, sûreté, acceptabilité d'une solution technologique, etc). Les décisions se prennent selon un certain nombre d'objectifs et de contraintes. En terme d'objectifs, il y a, en premier lieu, une stratégie de politique générale de l'organisation. Lorsque le problème est complexe ces objectifs sont hiérarchisés. Les objectifs les plus élémentaires (les « feuilles » de la hiérarchie) sont finalement déclinés en critères d'évaluation. Les critères doivent être suffisamment explicites pour donner lieu à une description exploitable des solutions ou actions envisagées dans le processus de décision.

La détermination de ces critères peut passer par une approche « terrain » qui nécessite dans un premier temps une fiche de description structurée de la solution proposée (ex : problématique de rattachement, objectifs, résultats escomptés, partenaires, état de l'art sur le

domaine, niveau d'innovation, découpage en tâches et jalons, risque technique et risque projet...). Sur la base de ces descripteurs, des critères peuvent être établis (ex : des macro-critères : les enjeux, le coût, le gain, le délai d'application, le risque projet, le risques technique... que l'on peut décliner en critères élémentaires).

L'autre possibilité pour l'identification des critères est de repartir d'une analyse systémique : définir dans un premier temps les problématiques générales relatives au pilotage du plan de l'organisation puis de les projeter dans une matrice multi-points de vue comme proposée dans la méthode d'analyse Sagace pour déterminer les problématiques spécifiques [Penalva, 1997]. De façon très résumée, il s'agit de considérer le système selon trois perspectives d'analyse, performances, stabilité et intégrité, croisées avec trois visions, fonctionnelles, structurelles et comportementales. Ensuite ce sont ces problématiques spécifiques qui sont déclinées en critères d'évaluation.

Parce que la CA est une donnée informative jugée utile et un savoir élémentaire interprété par son rédacteur, la CA doit correspondre à l'expression d'un jugement de valeur sur une solution envisagée vis-à-vis d'un critère dans le cadre du processus décisionnel. Ainsi, cartographier les CAs, consiste à intégrer les CAs dans un tableau ou grille à deux dimensions (figure 2) dont les lignes représentent les critères d'évaluation, les colonnes les solutions ou actions en concurrence dans le processus décisionnel. Cela permet de représenter graphiquement et globalement la dispersion de la production des CAs se rapportant à chaque case (critère j , solution k).

De cette façon, on peut associer à chacune des CA, sa compatibilité vis-à-vis d'un critère, i.e. une case de la matrice d'évaluation. Associer un score au commentaire d'une CA revient à quantifier l'utilité de la CA pour le projet, et permet d'introduire une intensité de préférence dans l'évaluation.

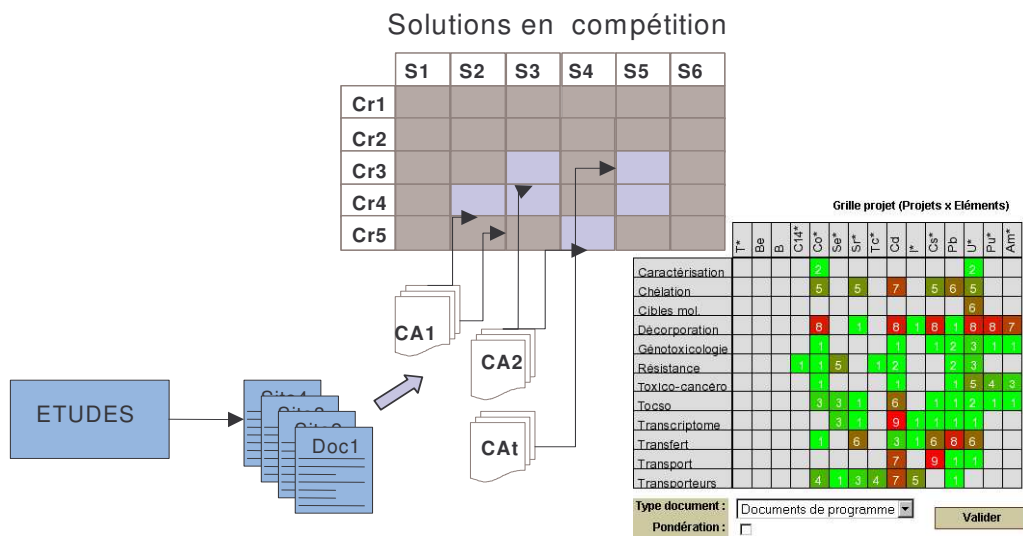


FIG. 2.- Cartographie des CAs

En ce qui concerne les caractéristiques de la CA, cela revient simplement à ajouter les *coordonnées* de la CA dans la grille d'évaluation.

3.4 Cartographie automatisée

La cartographie automatisée consiste à extraire automatiquement de la base documentaire (les données brutes) les connaissances relatives aux problématiques du projet, elles-mêmes déclinées en critères d'évaluation comme décrit précédemment. Il s'agit donc d'analyser de manière informatisée les documents pour en extraire des « fragments de texte » relatifs aux dimensions d'analyse de la grille d'évaluation du processus de décision. Ces fragments de texte constituent l'objet des CAs. Cette étape repose sur le concept de pertinence thématique : des documents sont extraits les passages les plus proches au sens d'une distance sémantique des différentes dimensions du problème de décision, les critères d'évaluation.

Dans cette vision automatisée de la cartographie, l'objet de la CA est donc en pratique un extrait ou fragment de texte contenant un élément informationnel utile à l'action puisqu'en rapport avec une des dimensions de l'analyse. On définit alors la pertinence du fragment de texte extrait de la base documentaire vis-à-vis du critère d'évaluation considéré. Les techniques décrites ci-après sont autant d'outils pour automatiser cette extraction à partir du corpus documentaire brut. Nous avons opté pour une combinaison de techniques relatives au TALN (Traitement Automatique de la Langue Naturelle) et au « Text-mining » pour cette étape de la chaîne de traitement de l'information (figure 3).

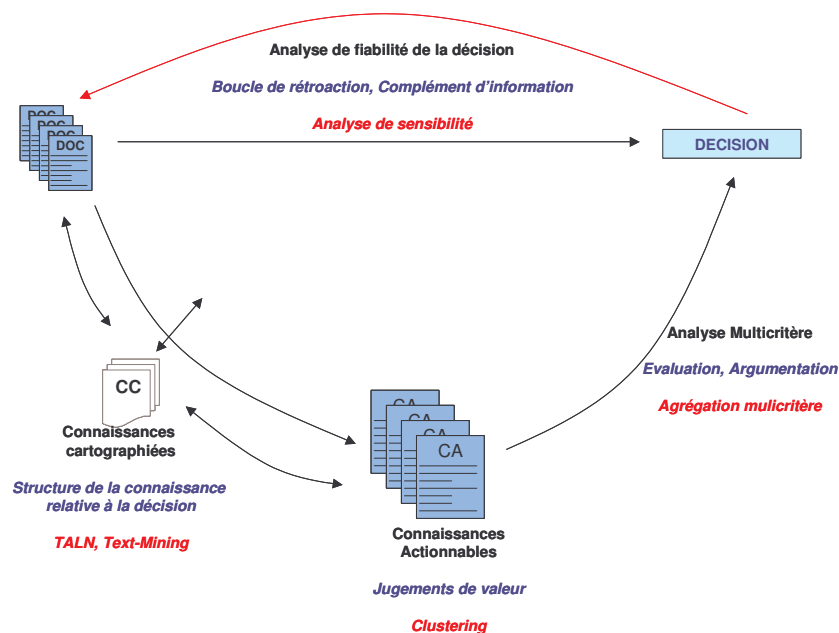


FIG. 3.- Vision générale du processus de traitement de l'information

La première étape consiste à effectuer une analyse morpho-syntaxique d'un texte pour localiser les mots de même racine, obtenir la fonction syntaxique et sémantique de chaque mot et enfin éliminer les ambiguïtés de sens.

La deuxième étape consiste à associer une valeur sémantique à chaque extrait de texte souhaité sous la forme d'un vecteur conceptuel [Chauché, 1990]. De manière très synthétique, à partir des concepts de la langue française, un espace vectoriel est construit dont la dimension est le nombre de concepts du thésaurus Larousse (873) [Larousse, 1992]. Chaque mot de la langue française est en correspondance avec un ou plusieurs concepts du thésaurus. Chaque mot est donc représenté par un vecteur normé dont les coordonnées sont non nulles pour les concepts avec lesquels il est en correspondance. Pour un extrait de texte, il faut ensuite effectuer une somme vectorielle de chaque mot lemmatisé de l'extrait s'il appartient à l'une des catégories grammaticales : nom, verbe, adjectif, adverbe. Les mots appartenant aux autres catégories grammaticales sont éliminés. La somme vectorielle respecte également la hiérarchie des termes, indiquée par le calcul de l'arbre morpho-syntaxique du texte étudié. Nous obtenons donc un vecteur unique pour un extrait de texte, ce vecteur est représentatif du sens attribué à l'extrait de texte [Chauché, 1990].

Par ailleurs, chaque critère d'évaluation du projet peut aussi être représenté par un vecteur conceptuel. Ainsi, nous avons une description homogène du contenu informationnel de la base documentaire et des critères de sélection à travers la notion de vecteur conceptuel. On peut alors calculer une mesure de proximité sémantique sous la forme de la distance angulaire entre le vecteur conceptuel associé à un extrait de texte et le vecteur de chacun des critères [Chauché, 1990 ; Salton, 2000]. Le calcul de distance angulaire peut être complété également par une distance ontologique ou sémantique, indiquant la proximité sémantique de deux concepts. Cette distance ontologique est obtenue par une combinaison de deux mesures : le nombre d'arcs nécessaires pour passer d'un concept à un autre dans un arbre ontologique [Chabert-Ranwez 2000] combiné à la mesure de profondeur relative entre deux concepts dans l'arbre, qui s'exprime par l'attribution d'un poids plus important pour un concept plus général.

Finalement, on obtient une mesure de proximité sémantique qui permet de discriminer les extraits de textes les plus pertinents à retenir dans un document relativement à un critère donné. Chaque document pourra donc être à l'origine d'une ou plusieurs CAs.

Cette technique est intéressante pour des documents entiers mais nécessite d'être améliorée dès qu'il s'agit d'analyser des fragments de textes. En effet, sur un texte trop court, le vecteur conceptuel associé est souvent trop « creux ». Dans une troisième étape, nous utilisons donc en complément des techniques dites de sémantique distributionnelle [Besançon *et al.*, 2000] permettant de définir un mot et son sens à partir de ses voisins. Le principe de cette technique est de dériver automatiquement le sens d'un mot à partir de l'ensemble de ses contextes dans un corpus. Les voisins sont ceux contenus dans la même unité syntaxique et ayant un même niveau de fonction grammaticale. Par exemple dans la même phrase, deux mots seront pris en considération pour le calcul de cooccurrence, s'ils sont tous les deux gouverneurs de la phrase. Le résultat de cette technique est l'obtention d'une matrice de cooccurrence. Elle est calculée sur un corpus de référence et donne les correspondances entre tous les mots (ou racines de mots) en relation dans le corpus. Cette matrice est ensuite utilisée pour calculer les proximités entre deux extraits de textes. Cette technique est très intéressante pour définir le sens d'éléments textuels dans un corpus spécialisé, ce qui est le cas quand on se place dans le cadre d'un projet spécifique.

Nous combinons les deux techniques : vecteurs conceptuels et matrice de cooccurrence pour affiner les résultats d'extraction. L'association des deux techniques présente un grand avantage : les vecteurs conceptuels donnent le sens d'un fragment de texte en référence à un ensemble de concepts généraux et permettent ainsi de positionner correctement son contexte. La sémantique distributionnelle raffine le sens de chaque mot et donc du fragment textuel, permettant d'obtenir une précision satisfaisante pour la suite des traitements.

La « taille » des fragments textuels à partir desquels sont calculés les vecteurs sémantiques ou autres grandeurs nécessaires à l'extraction des CAs n'est pas sans influence sur l'évaluation de la proximité sémantique entre la CA et une problématique d'évaluation. Avant tout autre chose, il semble donc pertinent de procéder à une étape préliminaire qui permettra un découpage du document avec une granularité rhétorique la plus pertinente possible. La question posée à cette étape est de savoir ce qui fait que l'on peut passer d'un fragment textuel modélisé par un vecteur conceptuel à un élément de connaissance « cartographiable ».

Une technique permettant d'apporter des éléments de solution à ce problème est actuellement explorée [Minel *et al.*, 2000]. Cette technique consiste à identifier dans le texte les éléments liant chaque phrase, comme les conjonctions, articles, etc., et à les considérer comme des indicateurs de la structure du texte. Ainsi pouvons-nous identifier, selon des règles précisées auparavant, par exemple, les « repères » dans le texte qui indiquent une introduction à un sujet ou à un développement ou bien à une conclusion. Ces éléments indiquent s'il est intéressant de regrouper deux ou plusieurs phrases ensemble pour en former un seul vecteur sémantique. Une phrase seule peut ne pas se rapporter significativement à un critère, mais associée à une autre phrase, elle peut donner une sémantique au groupement voisin de la sémantique de l'un des critères d'évaluation. Les vecteurs obtenus après ce traitement regroupent un ensemble de phrases, appelées « unité de sens » sur lesquelles s'effectuent les calculs précédents (vecteurs conceptuels, cooccurrences...).

Une autre technique [Belot, 2000], donne une méthode pour segmenter les documents du corpus en plusieurs parties qui chacune corresponde aux différentes thématiques abordées par le document. Cette méthode basée sur l'emploi d'arbres de décision et de modèles de Markov donne de bons résultats [Mittendorf *et al.*, 1994]. Dans notre contexte, cette méthode a l'avantage d'extraire directement les fragments textuels associés à un critère de jugement et permet de raffiner les éléments de connaissance obtenus précédemment.

4. Évaluation des actions potentielles par les CAs

Le SGDC doit être muni d'observateurs qui en permettent une exploitation à des fins d'évaluation et de contrôle dynamique de processus décisionnel. De façon pratique, pour passer d'une simple image à un observateur quantitatif du processus de décision, on définit un espace de mesures sur celui-ci : la *grille d'évaluation* des solutions selon un ensemble de critères en est l'élément de base.

4.1 Grille d'évaluation datée

L'idée de base de l'évaluation des solutions potentielles est de se référer à la théorie de la décision multicritère [Maystre *et al.*, 1994 ; Marichal, 1998 ; Schärli, 1998 ; Roy *et al.*,

1993 ; Dubois *et al.*, 1985]. Les jugements de valeurs ou scores attribués dans la grille d'évaluation via les CAs permettent de prendre en compte les intensités de préférence des critères : chaque score associé à la case (critère j , solution k) est une combinaison des jugements de valeurs portés par les CAs de la case, c'est l'évaluation partielle d'une solution k selon le critère j [Montmain *et al.*, 2002].

L'évaluation globale d'une action—son score global—correspond à l'agrégation des appréciations partielles qu'il a obtenues selon chaque critère et son obtention est donc assimilée à une procédure d'évaluation par les critères. Comme opérateur d'agrégation nous avons retenu l'intégrale de Choquet : elle permet de prendre en compte l'interaction entre les critères d'évaluation et reflète ainsi les nombreuses interactions, conflits et coalitions qu'il existe au sein de l'organisation. L'intégrale de Choquet dans le cas 2-additif (seules les interactions deux à deux entre les critères sont modélisées) prend la forme :

$$CI_g(x_s^1, x_s^2, \dots, x_s^n) = \sum_{\substack{I_{ij} > 0 \\ i > j}} \min(x_s^i, x_s^j) I_{ij} + \sum_{\substack{I_{ij} < 0 \\ i > j}} \max(x_s^i, x_s^j) |I_{ij}| + \sum_{i=1}^n x_s^i \left(\nu_i - \frac{1}{2} \sum_{j \neq i} |I_{ij}| \right)$$

avec n le nombre de critères, x_s^i le score partiel de la solutions S selon le

critère i , $\nu_i - \frac{1}{2} \sum_{j \neq i} |I_{ij}| \geq 0$ et où :

1. les ν_i sont les indices de Shapley et représentent l'importance globale de chaque objectif relativement à tous les autres ($\sum_{i=1}^n \nu_i = 1$).
2. les I_{ij} représentent les interactions entre les paires de critères (C_i, C_j) et prennent leur valeur dans l'intervalle $[-1; 1]$; une valeur de 1 signifie qu'il y a un effet synergique positif entre les deux critères, une valeur de -1 qu'il y a une synergie négative, et une valeur nulle que les critères sont indépendants.

Les propriétés mathématiques de l'intégrale de Choquet ont été largement traitées dans la littérature [Grabisch *et al.*, 1996 ; 1998 ; 2000 ; Marichal, 2000 ; Marichal *et al.*, 2000].

Parce qu'un système de notation numérique dans l'évaluation d'une solution i selon un critère j nous est apparu difficile à interpréter, nous avons opté pour une palette de couleurs plutôt qu'une notation purement quantitative (figure 4) : un rouge vif indique l'incompatibilité du *candidat* i avec le *critère* j alors qu'un vert prononcé exprime une adéquation parfaite. La palette de couleurs dépend bien sûr de la granularité de l'échelle d'appréciation que l'on souhaite ou que l'on est capable d'exprimer.

Une grille de ce type permet donc d'évaluer quantitativement à chaque *instant* chaque solution k selon un critère j . Chaque case de cette grille correspond à l'évaluation globale de chaque action suivant chaque critère ou sous-critère.

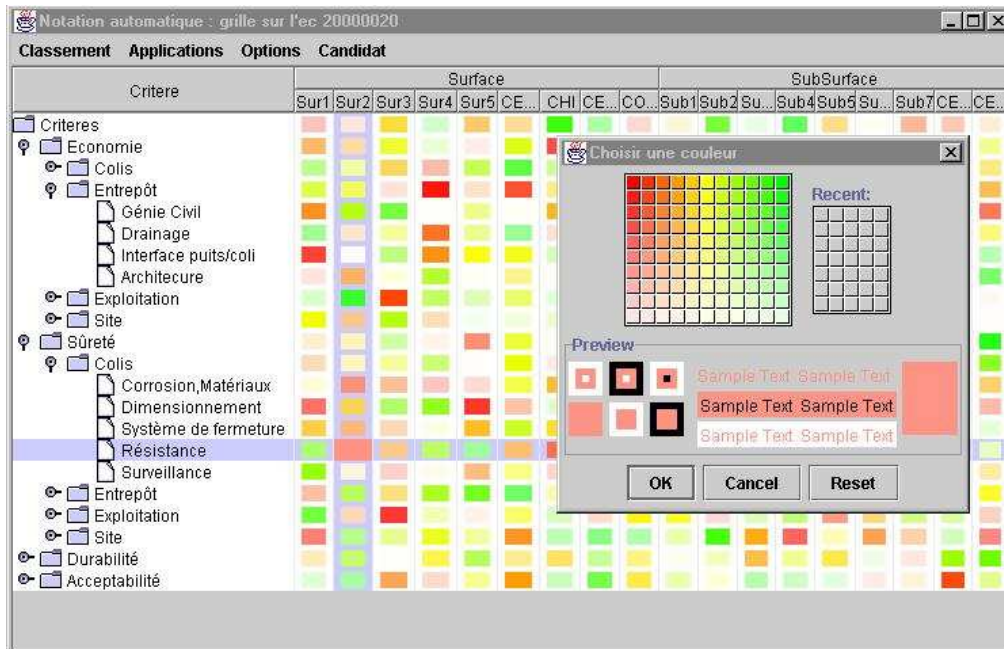


FIG. 4.- Grille d'évaluation et hiérarchie des critères

Chaque CA cartographiée porte un score qui est fonction de sa compatibilité vis-à-vis du couple (critère j ; candidat k) pointé. Cette compatibilité est mesurée par le degré de satisfaction du critère j par le candidat k . une échelle d'intervalle est utilisée pour exprimer ces degrés. La CA cartographiée, participe, par le jugement de valeur qui lui est associé, à l'évaluation des solutions candidates dès lors qu'elle est incorporée dans la base de connaissances du SGDC.

En pratique, cela revient à ajouter une dernière caractéristique à la CA : le *score partiel* qu'elle porte. Notons que ce score partiel doit en toute logique être cohérent avec le commentaire de la CA, lui-même étant une appréciation en langage naturel de l'objet de la CA.

En pratique, le rédacteur d'une CA semble être l'autorité légitime pour désigner quelles sont les cases (critère j ; candidat k) de la grille d'évaluation concernées par sa CA et pour estimer dans quelle mesure son CA est un élément favorable ou non à la solution k vis-à-vis du critère j . Ainsi, en pratique, le rédacteur doit apprécier la valeur de son EC pour les cases (j ; k) qu'il aura désignées au préalable.

4.2 Evaluation automatique par les CAs

L'objectif de cette fonctionnalité consiste à attribuer de manière automatique un jugement de valeur ou score à une CA pour informatiser la description précédente (figure 3).

Le principe est le suivant. Pour chaque dimension j du problème d'évaluation (un critère), on doit disposer d'une base de connaissances : B_j . Par conséquent, le corpus $\bigcup_j B_j$ correspond à

une base de CAs disponibles sur l'ensemble des problématiques du processus décisionnel. Ensuite, en fonction de la granularité fixée sur l'échelle d'intensité de préférence, on divise cette base B_j en classe d'appréciation $B_j(\omega_k)$: chacune des CAs de la base B_j est associée à une classe ω_k telle que le jugement de valeur porté par la CA selon le critère j est ω_k (ex : dans une base de critiques (B_j) portant sur le scénario de films (critère j), on considère les critiques positives (ω_1), neutres (ω_2) et négatives (ω_3)). Plusieurs techniques de text-mining sont utilisées pour générer les mots-clés représentatifs lemmatisés du corpus $\bigcup_j B_j$. Chacune

ayant ses avantages et inconvénients. En fonction du corpus de connaissance le choix sera effectué sur la technique la plus appropriée. [Salton, 1989], [Salton & Buckley, 1988,], [Besançon, R. and M. Rajman, 2000], [Mokrane A. *et al.* (2004)]. Chaque CA de $\bigcup_j B_j$ peut

être associée à un vecteur $V_{\bigcup_j B_j}(CA)$ dans l'espace $E_{\bigcup_j B_j}$ défini par les mots-clés. Chaque coordonnée i de $V_{\bigcup_j B_j}(CA)$ est la fréquence d'occurrence du mot-clé i dans la CA.

L'apprentissage consiste ensuite à définir un vecteur prototype $V_{B_j(\omega_k)}^*$ sur $E_{\bigcup_j B_j}$ pour chacune des classes ω_k des bases B_j . L'algorithme de clustering utilisé est décrit dans [Dray et al (2004)]. Cet apprentissage effectué, chaque nouvelle CA, se référant au critère j (cf section cartographie automatique), va être rattachée à une classe ω_k du critère j par mesure de proximité avec le prototype $V_{B_j(\omega_k)}^*$.

A chaque CA est ainsi associé automatiquement un score : la CA apporte un jugement de valeur sur une solution selon une dimension particulière du problème d'évaluation.

4.3 Tableau de bord

La grille d'évaluation doit être complétée par d'autres observateurs de l'évolution de la perception du projet. Le processus de décision étant un système organisationnel dynamique, son état à t_i s'explique en fonction de ses états précédents et des entrées de CAs entre t_{i-1} et t_i . Ainsi, peut-on construire un *tableau de bord* ou *grille de suivi* qui correspond à une table dont les lignes sont les solutions proposées, les colonnes symbolisent les dates d'observation, et dont les éléments rapportent l'évaluation globale d'une solution à une date t_i donnée, cette évaluation globale étant le résultat de l'agrégation des évaluations partielles de la solution dans la grille d'évaluation à t_i (figure 5).

Chaque colonne de cette grille, constitue donc une photographie de l'évaluation globale de toutes les solutions à t_i . A cette dernière on peut associer une décision instantanée dont on peut imaginer quantifier le risque associé par une fonction des évaluations agrégées à t_i (voir section 5).

On peut donc voir la grille d'évaluation à t_i comme *un niveau d'explication* de la colonne i du tableau de bord c'est-à-dire de la décision qui aurait été prise—quel candidat aurait été retenu— si le processus avait dû s'achever à cette date (figure 5).

Ce tableau de bord donne donc l'évolution de la perception de chaque solution au fil du processus décisionnel. L'interprétation graphique de cette grille de suivi est déjà intéressante en soi : elle permet d'observer l'évolution de l'évaluation d'une solution sur un intervalle de temps fixé [Montmain *et al.*, 2002].

Ce tableau de bord donne l'évolution de la perception de chaque solution au fil du projet. L'interprétation graphique de cette grille de suivi de projet est déjà intéressante en soi : elle permet d'observer, par exemple, que le premier candidat (Cp 1 de la figure 5), initialement mal perçu, finit par être largement accepté, l'évolution de cette évaluation se fait de façon continue et laisse penser qu'elle peut s'expliquer par une progression elle-même continue (exemple : les progrès dans le domaine du son numérique font que le CD a, petit à petit, remplacé le disque vinyle). Pour les candidats 2 ou 4, il n'en va pas de même : il y a rupture dans le suivi de leur perception, cette discontinuité laisse penser qu'un événement significatif a soudainement fait basculer l'image qu'on se faisait du concept (exemple : la catastrophe de Tchernobyl pour la production d'énergie nucléaire, le naufrage de l'Erika pour le transport d'hydrocarbures, la crise de la vache folle pour la consommation de viande rouge...).

4.4 Grille d'explication dans le temps

Un dernier type de grille—*grille d'explication*—sur un intervalle de temps permet de suivre les évaluations partielles d'une action particulière dans le temps. Si l'on prend l'exemple de l'action 4 (cp4) de la figure 5, pour expliquer la discontinuité qui a marqué sa perception entre t_{i+2} et t_{i+3} , on peut générer la grille d'explication de candidat 4 (Cp4, figure 5) sur l'intervalle $[t_i ; t_{i+5}]$. Cette dernière permet d'expliquer le brusque discrédit du Cp4 principalement par les critères 2 et 4, dimensions selon lesquelles apparaissent les plus significatives variations de perception entre t_{i+2} et t_{i+3} (figure 5).

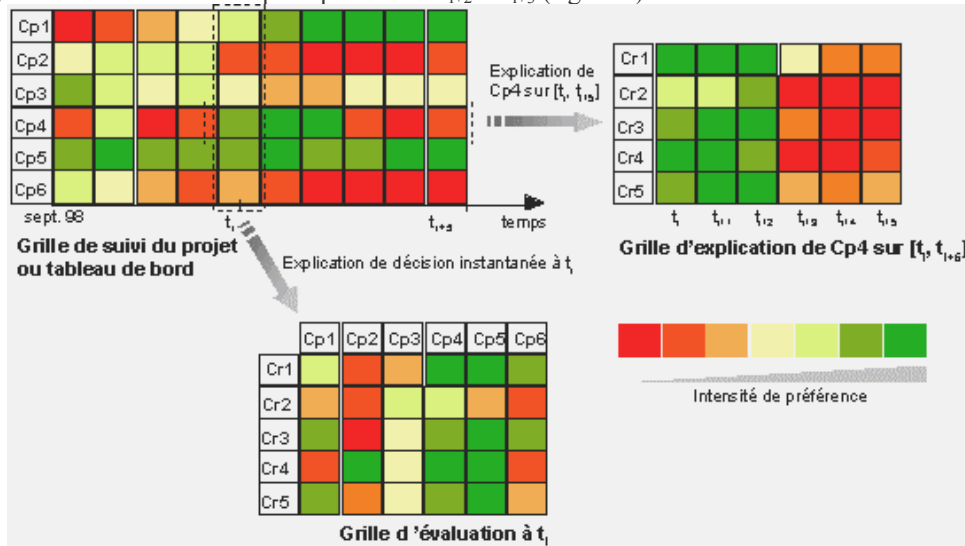


FIG. 5.- Grille d'évaluation, d'explication et tableau de bord et hiérarchie des critères

4.5 Logique décisionnelle et rhétorique argumentaire

Le processus d'évaluation repose sur la cartographie des CAs. Autrement dit, si l'on est capable d'identifier les critères qui ont joué un rôle particulièrement discriminant dans la décision, qui expliquent tel ou tel changement de perception d'une solution d'une date à l'autre, etc., alors cela revient à identifier des cases de la cartographie des CAs. Il est alors possible d'imaginer l'extraction automatique des CAs correspondant à ces critères ou plus précisément les commentaires de ces CAs, pour construire un argumentaire de la logique décisionnelle qui sous-tend le processus d'évaluation (dans l'hypothèse où commentaires et scores ont été déterminés de façon cohérente).

L'explication repose sur l'estimation de la contribution de chaque critère à l'évaluation globale : une analyse de sensibilité locale de l'opérateur d'agrégation (l'intégrale de Choquet dans notre cas) permet de hiérarchiser la contribution marginale de chaque critère et de déterminer ainsi les dimensions prépondérantes de l'évaluation au point d'analyse. Les critères les plus influents dans la décision sont ainsi remontés. Cette technique permet plusieurs niveaux d'argumentation : argumentation en absolu (pourquoi la Solution S_1 apparaît être la meilleure à une date donnée ?), argumentation en relatif (pourquoi la Solution S_1 apparaît meilleure que la solution S_2 à une date donnée ?) et argumentation dans le temps (pourquoi l'évaluation de la Solution S_1 a-t-elle évolué ainsi entre deux dates ?). Ces techniques ont été discutées dans [Montmain *et al.*, 2002 ; Akharraz *et al.*, 2002a, b].

Par exemple, si dans l'exemple de la catastrophe de Tchernobyl, le SGDC gère une base de presse électronique, le SIAD extraira de manière automatique entre mars et avril 1996, tous les articles qui traitent de l'accident et qui expliquent comment l'opinion publique a pu basculer du jour au lendemain concernant le nucléaire, et plus précisément la sûreté et la sécurité dans le nucléaire.

La légitimation d'une décision repose donc sur l'argumentation de celle-ci à partir de la base de connaissances du SGDC.

Une solution a un niveau de satisfaction minimale s'il existe dans la base de connaissances suffisamment d'éléments de rhétorique (des CAs) pour soutenir la logique décisionnelle qui a conduit à ce choix.

5. Une vision cybernétique de la problématique

Dans le déroulement du processus de décision, une dynamique interne se met en œuvre : le processus de décision est en évolution permanente au fur et à mesure que les CAs viennent alimenter la base de connaissances du SGDC. La perception des solutions candidates est susceptible de varier à chaque acquisition de CAs. Le classement des solutions candidates à un instant t dépend du classement antérieur à $(t-1)$ et du flux de CAs entre $t-1$ et t .

Le processus de décision est assimilé à un système dynamique dont les entrées sont les CAs et la sortie le classement des alternatives possibles.

5.1 Risque décisionnel

Notons \mathbf{x}_i le vecteur des scores partiels obtenus par le candidat X_i sur les n critères à une date donnée. \mathbf{u}_i est le vecteur des CAs relatives à X_i . Autrement dit u_i^k est l'information interprétée entrante concernant le candidat X_i selon le critère k .

La dynamique du vecteur d'évaluation \mathbf{x}_i , critère par critère, d'un candidat X_i peut s'écrire sous la forme :

$$\mathbf{x}_i(t+1) = \text{Agregation}_{case}(\mathbf{x}_i(t), \mathbf{u}_i.score(t+1))$$

où $\text{Agregation}_{case}(\dots)$ est un opérateur d'agrégation qui effectue une révision du score présent dans la base avant l'entrée des nouvelles CAs et $\mathbf{u}_i.score$ les scores partiels portés par \mathbf{u}_i .

Ainsi, plus la quantité d'informations nouvelles entre $t-1$ et t est grande (plus le flux de CAs entre $t-1$ et t est important), plus l'évaluation peut s'en trouver modifiée et le classement à t changé. Cette *quantité d'information* additionnelle fait varier l'entropie de la base de connaissances du SGDC or cette entropie détermine précisément l'état du processus de décision. Une modification significative de l'entropie de la base risque d'entraîner un changement dans le classement des alternatives : les variations d'entropie dues au flux entrant de CAs remettent en cause la stabilité du classement. La sensibilité du classement à toute nouvelle entrée dans la base du SGDC conduit à définir ce que l'on appellera le *risque décisionnel*.

Pour définir le risque décisionnel, i.e. la fiabilité du classement, nous proposons de nous baser sur une notion de distance entre le candidat X_1 , le meilleur au temps t , et le candidat X_i , basée sur la norme L_1 . Nous prenons une norme L_1 parce qu'elle correspond à l'idée que l'on somme les « progrès » nécessaires, critère par critère, pour estimer la distance qui sépare un candidat X_i de son vainqueur X_1 .

La notion de risque est donc ici assez éloignée de la définition probabiliste classique et relève plutôt de la sensibilité de l'ordre établi à toute perturbation externe. Sur la base de cette distance, on définit la notion de risque décisionnel relatif associé à la stabilité du classement des films par :

$$r_{D_R} = 1 - \min_{i=2\dots p} (d(X_i, X_1)) \text{ où } p \text{ est le nombre d'alternatives.}$$

Définissons encore le risque décisionnel absolu. La mesure précédente permet de s'assurer que le candidat retenu se distingue significativement des autres, mais n'assure pas pour autant que l'élu réponde à nos espérances. Le risque décisionnel absolu est lié à la distance qui sépare X_1 de l'objectif $Ob = (1 \dots 1)^T$:

$$r_{D_A} = d(Ob, X_1).$$

Le risque décisionnel peut être vu comme un observateur de l'entropie de la base de décision où les CAs sont les « quantités d'informations » élémentaires.

5.2 Contrôle de la dynamique du processus décisionnel

Plus la distance entre le candidat recommandé et ses principaux rivaux est faible, plus la sélection de X_1 est risquée : il suffirait d'« une faible quantité d'informations nouvelles et pertinentes » pour que le classement s'en trouve modifié. En effet, le classement peut être très sensible à toute nouvelle connaissance modifiant les scores partiels. On considère que l'on peut prendre une décision lorsque le risque décisionnel est au dessous d'un seuil fixé C_r . Ce seuil définit la notion d'*acceptabilité* de la décision ou de *situation décidable*.

Dans cette approche, la notion de contrôlabilité du risque décisionnel est reliée à la recherche du moindre effort à fournir (au sens de la norme L_1 pour les raisons sémantiques

invoquées précédemment) pour que l'évaluation de X_i soit au moins égale à celle de l'élue X_1 . Le calcul de cet effort minimal (basée sur une distance particulière [Akharraz *et al.*, 2004a]) permet d'identifier les dimensions sur lesquelles X_i doit concentrer son « effort » pour atteindre le meilleur score global. Autrement dit le calcul de l'effort fournit les critères selon lesquels X_i devra nécessairement progresser pour rejoindre X_1 avec un effort minimal.

Ainsi, à chaque instant, on connaît d'une part, les rivaux les plus dangereux, mais également les scores partiels les plus sensibles ou critiques (donc les critères associés) dans l'évaluation du candidat recommandé au vu des résultats obtenus par ces rivaux.

Si le candidat le plus proche de X_1 est X_i , alors on aura envie d'acquérir de nouvelles informations sur ces deux candidats et ce, selon les dimensions (critères) sensibles identifiées dans le calcul de l'effort minimal.

En résumé, cela signifie que sur la base du calcul du risque décisionnel, on va chercher à identifier à chaque pas de temps les cases de la grille d'évaluation (figure 5) où il faut faire rentrer de nouvelles CAs pour que la situation devienne décidable avec le niveau de risque fixé. On détermine ainsi l'information la plus pertinente pour l'évaluation suivante.

Le même raisonnement peut être appliqué pour le risque absolu : il s'agit d'identifier les dimensions les plus stratégiques selon lesquelles X_1 doit progresser pour devenir acceptable, c'est-à-dire raisonnablement « parfait ».

A travers les notions de risque et d'acceptabilité de la décision, on a défini une boucle de réaction sur le contenu de la base de connaissances utilisée pour l'évaluation des solutions. En indiquant simplement les candidats et les critères sur lesquels il est le plus pertinent de s'informer pour notre processus de décision, on modifie la dynamique de ce processus sans pour autant la contrôler (figure 3). En effet, l'ajout de nouvelles informations peut indifféremment faire augmenter ou diminuer la fiabilité de la recommandation. Par contre, si en plus des coordonnées des cases de la grille d'évaluation (critère k , candidat i , par exemple), on indique la nature de l'évaluation portée par l'information à acquérir (*positive* si elle fait augmenter le score ou *négative* si elle fait diminuer le score), alors on induit un contrôle qualitatif de la dynamique du processus de décision (à condition que l'on fasse l'hypothèse que l'information existe nécessairement et que le problème est seulement de la trouver).

En tout état de cause, cette boucle de rétroaction est un problème de commande *multivariable*. En fait, il s'agit davantage de fournir une recommandation qualitative au manager de l'organisation plutôt que d'établir une loi de commande globale, aussi ramènerons-nous le problème localement à un système monovariable dans le chapitre 2, qui fournit la mathématique de ces propos, pour ne pas perdre de vue l'idée que la sortie de notre système de décision est destinée à un être humain et non pas à un système numérique de contrôle commande !

Choisir des informations correspondantes à ce contrôle qualitatif accentuera au temps suivant la différence d'évaluation entre les candidats X_i et X_1 , on "creusera" ainsi volontairement l'écart entre les scores globaux des deux candidats : les chances que X_i revienne sur X_1 diminueront encore, la distance entre les deux candidats augmentera et le risque décisionnel relatif encouru à recommander X_i diminuera. Cette attitude correspond au comportement d'un décideur qui cherche à se conforter dans son choix ou à légitimer sa position (mais ce mode de fonctionnement n'est pas obligatoire).

Dans une optique plus *objective*, la recherche de toutes nouvelles informations, positives ou négatives, peut être intéressante pour augmenter la fiabilité des scores partiels, même si celles-ci ne font pas évoluer la distance entre les scores globaux. La dynamique du processus de décision reste alors libre, on s'intéresse simplement à la fiabilité de l'évaluation.

Les éléments mathématiques et algorithmiques nécessaires au calcul du risque décisionnel sont fournis dans [Akharraz et al., 2004a, b, c].

5.3 Les éléments de la boucle de contrôle

Nous avons ainsi défini une boucle de contrôle (figure 6).

Le processus à contrôler est le processus de décision assimilé à un système dynamique dont les entrées sont les CAs et la sortie le classement des alternatives possibles.

Dans cette interprétation, le risque décisionnel peut être vu comme une variable régulée. Compte tenu des remarques précédentes, remarquons que contrôler le risque décisionnel, n'est autre que contrôler l'entropie de la base de connaissance utilisée pour la décision.

Le SGDC est l'actionneur de ce schéma de contrôle : le SGDC est l'organe d'action qui « ouvre ou ferme la vanne d'informations utiles à la décision ».

Le calcul du risque et des critères sensibles du classement constitue le régulateur du système.

L'acceptabilité de la décision prend une interprétation bien particulière dans ce cadre : c'est la consigne fixée sur le risque décisionnel. Si l'on est en dessous de ce seuil, la situation est décidable, aucune étude supplémentaire n'est requise pour proposer un choix qui ne pourra être « stratégiquement » désavoué (arrêt ou ralentissement des études devenues inutiles compte tenu de l'état des connaissances suffisant). Par contre, si l'on ne parvient pas à ramener le risque décisionnel en dessous de la consigne fixée, indépendamment de nouveaux flux d'informations, la stratégie d'agrégation doit être révisée (les paramètres la déterminant peuvent être réajustés, les critères d'évaluation réexaminés) : aucun consensus ne pourra être trouvé avec cette stratégie.

Toute information que l'on rentre dans le SGDC sans que la dimension à laquelle elle se réfère soit une sortie du régulateur est une perturbation : ainsi, l'inertie psychologique ou la controverse sont typiquement deux comportements organisationnels identifiables à des perturbations de la dynamique d'évolution du processus de décision (ce sont des perturbations sur l'actionneur ; l'incertitude sur la stratégie appliquée pourrait constituer une perturbation en sortie).

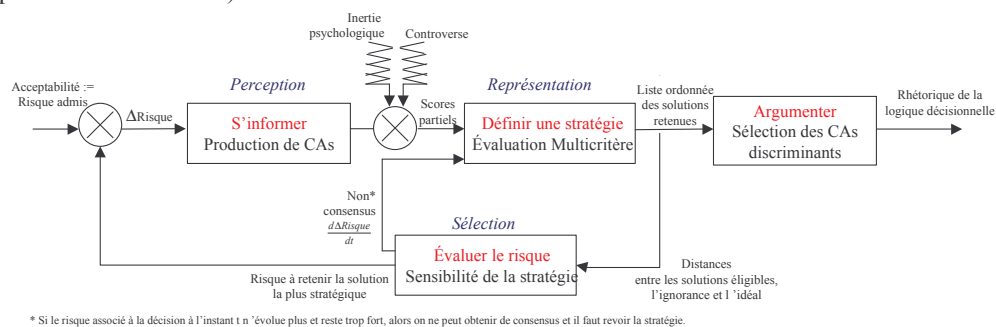


FIG. 6.- Contrôle de l'information décisionnelle ou pilotage par le risque

Revenons au modèle S.T.I de H.A.Simon. Les causalités enchevêtrées mises en exergue dans le modèle de l'économiste trouvent ici une interprétation en terme de boucle de contrôle. La phase d'information ou de perception est à associer à l'actionneur de notre schéma, la phase de représentation (le processus d'évaluation multicritère ici) est le système à contrôler, la phase de sélection (définition de la logique de choix) est le régulateur.

Notons enfin que la sortie du processus de décision est le classement des alternatives sur la base des scores qui leur ont été attribués au cours de l'évaluation. Ce classement n'est qu'un artifice mathématique pour rendre compte de la combinaison des intensités de préférence des alternatives en fonction des critères à un instant donné. Les scores sont assimilables à des *variables d'état* dont le décideur « n'a que faire » pour justifier ses actions. Comme nous l'avons expliqué dans la section précédente, une dimension essentielle de notre système d'aide à la décision est sa capacité à justifier ses évaluations et par conséquent un classement résultant des CAs du SGDC. La fonctionnalité de légitimation joue le rôle d'une *matrice d'observation* : elle transforme le « vecteur d'état », les scores, le classement, en un « observable », la rhétorique de la logique décisionnelle appliquée pour obtenir ce classement. Cette rhétorique de logique décisionnelle, qui prendra la forme de rapports d'argumentation remis par le décideur à ses administrés, aux pouvoirs publics, constitue l'observable du processus d'évaluation par les CAs.

6. Conclusion

L'idée générale de notre approche est d'automatiser au mieux la chaîne complète de traitement de l'information qui permet de passer de la base documentaire associée au processus de décision (les documents d'ingénierie pour un projet par exemple) à la sélection d'une solution candidate (l'action) et à l'argumentation automatisée de la logique décisionnelle (la légitimation) ayant conduit dans le processus de décision à retenir cette solution plutôt qu'une autre. Autrement dit la question est de construire de façon automatisée un choix raisonné sur une masse d'informations non formalisées.

La décision d'un collectif organisé nécessite une rationalité cognitive fondée sur l'adéquation entre les informations possédées et les représentations adoptées. C'est l'objectif de notre approche qui repose sur le couplage d'un SGDC, qui favorise les processus cognitifs d'apprentissage du savoir et des connaissances, avec un système interactif d'aide à la décision de groupe (SIADG) : tout choix proposé par le système trouve sa justification dans le corpus des connaissances développé par l'organisation dans le cadre du processus décisionnel en jeu. Les techniques d'extraction et de traitement informatisées du corpus géré par le SGDC permettent la transformation automatique de l'information en connaissances actionnables, c'est-à-dire en connaissances utiles à la décision.

Le SIADG est utilisé afin d'argumenter, justifier les décisions jalonnant le processus décisionnel : il est capable de fournir à tout instant la connaissance pertinente gérée par le SGDC pour alimenter la rhétorique d'une logique décisionnelle à faire valoir. Si l'approche de l'aide à la décision consiste, à proposer des solutions au décideur en vue d'effectuer un choix, elle a aussi pour objectif d'aider le décideur à justifier et argumenter ses choix. Il est souvent aussi important, si ce n'est plus, de pouvoir justifier une décision que de pouvoir l'obtenir.

La reconnaissance de la complexité des processus de décision, des rationalités multiples qui s'y côtoient, des conflits qui s'y déroulent et des transformations qui s'y opèrent ne peut que faire douter de la possibilité de toujours prouver qu'une décision est ou non, la

meilleure : meilleure pour qui (dans un processus multi-acteurs) ? Selon quels critères ? A quel moment ? Aucune démarche objective fondée sur la seule raison ne peut démontrer l'optimalité ni même le bien-fondé d'un système de valeurs ou d'un mode d'anticipation de l'avenir [Roy *et al.*, 93].

Autrement dit, en admettant qu'une décision optimale puisse exister, ce qui est déjà une hypothèse forte, le caractère optimal de cette décision n'est pas forcément démontrable. De plus, l'optimalité par rapport à un modèle et à une méthode d'une décision ne correspond pas nécessairement à l'optimalité de cette décision par rapport à la réalité. C'est pourquoi, on s'intéresse davantage ici à la capacité du SIADG à justifier d'une stratégie simplement satisfaisante dans le sens qu'elle s'appuie sur des connaissances partagées et délibérées par le collectif.

Le risque décisionnel nous a paru être une grandeur à sémantique simple pour observer l'entropie de la base de connaissance du processus décisionnel. Il nous a permis par ailleurs de réinterpréter le modèle S.T.I, non causal, sous la forme d'un schéma de contrôle ce qui donne une dimension opérationnelle à la représentation cognitive de H.A. Simon.

Notre modèle conceptuel de SIADG sous la forme d'une boucle de contrôle, dont le SGDC est l'organe d'action, semble bien rendre compte des différentes exigences d'une « bonne aide à la décision » au sens de H.A. Simon : évaluation multicritère, processus dynamique non séquentiel, légitimation de la décision et apprentissage du décideur.

Enfin, nous concluons que pour qu'un SIADG soit accepté par ses utilisateurs, il faut :

1. qu'il soit capable de gérer toutes les informations disponibles utiles à la décision ;
2. qu'il propose la trace de sa logique décisionnelle à des fins argumentatives ;
3. qu'il évalue la fiabilité de ses décisions et identifie les points précis sur lesquels le décideur doit acquérir davantage d'informations pour garantir cette fiabilité ultérieurement.

L'acceptabilité d'une décision repose sur la légitimation du risque associé à celle-ci.

Références

- [Akharraz *et al.*, 2002a] Akharraz, A., Montmain, J., Mauris, G (2002a). Fonctionnalités explicatives d'un système de fusion de connaissances collectives par intégrale de Choquet 2-additive. *LFA'2002*, Montpellier, France
- [Akharraz *et al.*, 2002b] Akharraz, A. , Montmain, J., Mauris, G (2002b). A project decision support system based on an elucidative fusion system. *Fusion 2002, 5th International Conference on Information Fusion*, Annapolis, Maryland, USA.
- [Akharraz *et al.*, 2004a] Akharraz, A. , Montmain, J., Mauris, G. (2004). A cybernetic modeling of a multi-criteria decision process in organization. *EURO XX , 20th European Conference on Operational Research*, Rhodes, Greece.
- [Akharraz *et al.*, 2004b] Akharraz, A. , Montmain, J., Mauris, G (2004). Elucidation and risk expressions of a movie recommendation based on a multi-criteria aggregation with a Choquet integral. *IPMU 2004, 10th International Conference on Information processing and Management of uncertainty in Knowledge-Based Systems*, Genova, Italy.

- [Akharraz *et al*, 2004] Akharraz, A. , Montmain, J., Denguir, A., Mauris, G. (2004). Information System and Decisional Risk Control for a Cybernetic Modeling of Project Management. *Modelling, Computation and Optimization in Information Systems and Management Sciences*, Metz, France.
- [Belot, 2000] Belot P., (2000). Méthodes de classification et de segmentation locales non supervisées pour la recherche documentaire. Thèse de doctorat informatique de l'Université d'Avignon et des pays de vaucluse.
- [Besançon et Rajman, 2000] Besançon, R. and M. Rajman (2000). Le modèle DSIR : une approche à base de sémantique distributionnelle pour la recherche documentaire. *TAL'2000, vol. 41/2, 1-27*
- [Chabert-Ranwez, 2000] Chabert-Ranwez (2000) Sylvie Chabert-Ranwez. Composition Automatique de Documents Hypermédia Adaptatifs à partir d'Ontologies et de Requêtes Intentionnelles de l'Utilisateur. *Thèse de doctorat en informatique de l'université Montpellier II, 21 décembre 2000.*
- [Chauché, 1990] Chauché, J. (1990). Détermination sémantique en analyse structurale : une expérience basée sur une définition de distance. *TA Information, vol 31/1, p 17-24.*
- [Damart, 2003] Damart, S. (2003). Aide à la décision et légitimation des décisions. *ROADEF'03*, Avignon, France.
- [Dasarathy, 2000] Dasarathy, B. (2000). Elucidative fusion systems – an exposition. *Information Fusion, 1*, 5-15.
- [Dray *et al*, 2004] Dray et al (2004). "Comparing concept-based and statistical representations for textual categorization". Simon Jaillet, Maguelonne Teisseire, Gérard Dray, Michel Plantié. *IPMU*, july, 4-9 , 2004, Perugia, Italy
- [Dubois et Gentil, 1994] Dubois, D. and S. Gentil (1994). Intelligence Artificielle pour l'Automatique. *Revue d'Intelligence Artificielle, 8* , (1), 1-27.
- [Dubois et Prade, 1985] Dubois, D. and H. Prade (1985). A review of fuzzy set aggregation connectives. *Information Sciences, 36*, 85-121.
- [Ermine, 1996] Ermine, J-L. (1996) Les systèmes de connaissances, *Editions Hermès*, Paris.
- [Giboin, 1995] Giboin, A. (1995). Les explications destinées aux utilisateurs de systèmes à base de connaissances. *Groupe explication In Bulletin de l'AFIA, 20*, 21-47.
- [Giboin, 1996] Giboin, A. (1996). L'explication homme - machine : explication ou correspondance? *In Actes des 3èmes Journées Explications*, Sophia-Antipolis, 77-90.
- [Grabisch et al, 1995] Grabisch, M., Nguyen, H.T., Walker, E.A.(1995). Fundamentals of Uncertainty Calculi, with Applications to Fuzzy Inference (Chap. 8). *Kluwer Academic*.
- [Grabisch et Roubens, 1996] Grabisch, M., Roubens, M. (1996). The application of fuzzy integrals in multicriteria decision making. *European Journal of Operational Research, 89*, 445-456.
- [Grabisch et al, 1998] Grabisch, M., Orlowski, S.A., Yager, R.R.(1998). Fuzzy aggregation of numerical preferences, in R. Slowinski, *Ed., Fuzzy Sets in Decision Analysis, Operations Research and Statistics, Kluwer Academic*.
- [Grabisch et Roubens, 2000] Grabisch, M. and M. Roubens (2000). Application of the Choquet Integral in Multicriteria Decision Making, in *Fuzzy Measures and Integrals : Theory and Applications* (Grabisch, Murofushi and Sugeno), Physica-Verlag.
- [Larousse, 1992]. Thésaurus Larousse - des idées aux mots - des mots aux idées. *Larousse, ISBN 2-03-320-148-1.*
- [Marichal, 1998] Marichal J.-L., Aggregation Operations for Multicriteria Decision Aid. (1998). *Ph.D. Thesis, University of Liège.*

- [Marichal, 2000] Marichal, J-L. (2000). An axiomatic approach of the discrete Choquet integral as a tool to aggregate interacting criteria, *IEEE Transactions on Fuzzy Systems* **8** (6), 800-807
- [Marichal et Roubens, 2000] Marichal, J-L, Roubens, M. (2000). Theory and Methodology : Determination of weights of interacting criteria from a reference set, *EJOR*, **124**, 641-650.
- [Maystre *et al*, 1994] Maystre L.Y., Pictet J., Simos J. (1994). Méthodes multicritères Electre. Description, conseils pratiques et cas d'application à la gestion environnementale. *Presses polytechniques et universitaires romandes*, Lausanne, Suisse.
- [Mittendorf et Schäuble, 1994] Mittendorf E., Schäuble P. (1994), Document and Passage Retrieval Based on Hidden Markov Models. *ACM/SIGIR'94 Conference on Research and Development in Information Retrieval, Dublin, Irlande*, 318-325.
- [Minel *et al*, 2000] Minel, J.L. *et al.* (2000). Résumé par filtrage sémantique d'information dans les textes. *Techniques et science informatique*, **20**, (3), 369-395.
- [Mokrane *et al*, 2004] Mokrane A. *et al.* (2004). Mokrane A., Arezki R., Dray G. et Poncelet P.. Cartographie automatique du contenu d'un corpus de documents textuels. JADT'04, Le poids des mots, vol 2 pp816-823, Presses Universitaires de Louvain, Mars 2004, Belgique. Argyris, C., Schön, D. (1978). Organizational learning : a theory of action perspective. *Addison Wesley*
- [Montmain *et al*, 2002] Montmain, J., Akharraz, A. and Mauris, G. (2002). Knowledge management as a support for collective decision-making and argumentation processes. *IPMU'2002*, Annecy, France.
- [Newel et Simon, 1972] Newell, A., Simon, H.A. (1972). Human problem solving. *Englewood Cliffs, NJ: Prentice-Hall*.
- [Peil *et al*, 2001] Peil O, Aubertin G, Ermine J-L, Matta N. (2001). La cartographie des connaissances critiques, un outil de gestion stratégique des connaissances. *CITE'2001*, Troyes, 287-308.
- [Penalva, 1997] Penalva, J-M. (1997). La modélisation par les systèmes en situations complexes. *Thèse de l'Université de Paris XI-Orsay*.
- [Penalva et Montmain, 2002] Penalva, J-M. and Montmain, J. (2002). Travail collectif et intelligence collective : les référentiels de connaissances. *IPMU'2002*, Annecy, France.
- [Penalva et Montmain, 2002a] Penalva, J-M, Montmain, J. (2002a). Les référentiels de connaissances : travail collaboratif et intelligence collective. *La Lettre d'Adeli, la maîtrise des systèmes complexes*, 48, 40-46
- [Pomerol et Barba-Romero, 1993] Pomerol, J-C. and S.Barba-Romero (1993). Choix multicritère dans l'entreprise. *Ed. Hermès*.
- [Rojot, 2003] Rojot, J. (2003) Théorie des organisations. *Edition Eska*.
- [Roy, 1985] Roy, B. (1985). Méthodologies multicritères d'aide à la décision. *Economica*.
- [Roy et Bouysou, 1993] Roy, B., Bouysou, D. (1993). Aide Multicritères à la décision : Méthodes et Cas. *Ed. Economica*.
- [Saad et Rosenthal-Sabroux, 2003] Saad, I., Rosenthal-Sabroux, C. (2003). Méthode de classification multicritère des connaissances pour le choix d'investissement dans des opérations de capitalisation des connaissances. *ROADEF'03, Avignon, France*.
- [Sfez, 1992] Sfez, L. (1992). Critique de la décision. *Presses de la Fondation des sciences politiques*, 4^e ed. (1^{re} ed. 1973)

- [Salton, 1989] Salton, (1989). G.Salton, « Automatic Text Processing : the transformation, analysis and retrieval of information by computer », Reading, Mass. Addison-Wesley, 1989.
- [Salton & Buckley, 1988] G. Salton, C. Buckley, « *Term weighting approaches in automatic text retrieval* », Information Processing and Management, vol. 24, n° 5, 1988, p. 513 à 523.
- [Simon, 1980] Simon, H. A. (1980). Le nouveau management. La décision par les ordinateurs. *Économica, Paris*, 159 p.
- [Simon, 1983] Simon, H. A. (1983). Administration et processus de décision. (*trad. de Administrative Behavior, 1947*), *Économica, Paris*.
- [Simon, 1991] Simon, H.A. (1991). Sciences des systèmes, sciences de l'artificiel. *Ed. Dunod, 1969 (réédition)*.
- [Simon, 1997] Simon, H.A. (1997). Models of bounded rationality. *MIT Press, Cambridge Massachusetts*. 3.
- [Simon, 1964] Simon, H. A., March, J. J. (1964). Les organisations. *Dunod, Paris*.

Summary

The acceptability of a decision-making support (DMS) system relies on its capacities to process all the knowledge pieces necessary and useful to the decision-making process, but also to propose traceability functions for strategic choices and logical argumentation. Indeed, when a collective is concerned with a decision-making, a great number of interpretations and satisfying decisional strategies can be associated to a same situation. The notion of optimality is generally meaningless. Group decision-making should lead to consensual decisions. This remark legitimates the need of arguing decision strategies in the DMS. The DMS must also evaluate the reliability of its own decisions and identifies the crucial information points for which a deeper investigation would be of great interest to control this reliability in the future. This paper proposes a conceptual the decision phase is modeled by a multi-criteria evaluation with a Choquet integral. Then a method is proposed to argue and estimate the risk associated to the decision.

This paper proposes a conceptual model of the learning and decision-making processes in an organization. The analysis relies on a cybernetic interpretation of the cognitive model of decision proposed by the economist H.A.Simon that claims that it might be more effective to develop a more satisfactory information processing and reasoning procedure rather than attempting to designate an illusory “optimum” decision.

The *administrative decision-maker* isn't a *homo œconomicus* with clear and explicit objectives, reliable in time, mutually independent. A great part of the Sciences of Management literature is merely dedicated to this statement. This paper proposes an operational cybernetic alternative to the purely cognitive model issued from the Sciences of Management: it provides a complete information processing system to decision-making in organizations.