

Méta Modèle adaptatif de la pertinence d'un modèle de connaissance

Adaptive Meta Model of Pertinence in a Model of Knowledge

Michel Crampes

LGI2P-Ecole des Mines d'Ales
6 avenue de Clavières, 30319 Ales Cedex
Michel.Crampes@ema.fr

Résumé

La modélisation de la décision a fait l'objet de multiples approches. Alors que beaucoup de modèles mettent en avant la dimension rationnelle, l'approche bayésienne à l'inverse en souligne le caractère incertain. Nous enrichissons ce type de modèle en considérant la possibilité de son adaptation par rapport à la situation et au profil du décideur. Nous proposons un méta modèle d'adaptabilité lui-même bayésien qui permet à la fois d'envisager des applications opérationnelles, et une formalisation de défauts de raisonnement associés à des décisions inopportunes.

Mots Clef

Décision, adaptabilité, réseaux bayésiens, interfaces adaptatives.

Abstract

Multiple approaches have been used to model decision. When many models rely on rationality in decision formalisation, a bayesian approach underlines the uncertain dimension of decision. We extend this type of bayesian models considering the possibility to adapt them according to the situation and the decision maker. We propose a meta model of adaptivity based on another bayesian model which permits both to expect operational applications, and to formalize pitfalls in reasoning during decision making leading to unexpected results.

Keywords

Decision, adaptivity, bayesian networks, adaptative interfaces.

1 Introduction et problématique

De nombreuses approches ont été proposées pour modéliser un processus de décision, qu'il soit individuel ou collectif. On peut citer les méthodes classiques d'optimisation, de recherche opérationnelle (études combinatoires, choix aléatoires, concurrence et théorie des jeux), les méthodes multicritères, les méthodes informatiques à base de connaissances ou à base de cas, la systémique, la simulation, l'approche agents rationnels, etc. Cette multiplicité est le reflet de la difficulté à observer et à appréhender un processus complexe à différents titres. Il doit prendre en compte des données de nature très diverses, incomplètes, imprécises, changeantes, aléatoires, etc. Dans ce contexte, il apparaît intéressant de voir la décision comme une démarche individuelle dans un cadre collectif essentiellement basée sur la psychologie du décideur, son retour d'expérience, et l'observation objective ou subjective qu'il fait d'une situation avec une grande part d'inconnu.

Ainsi posé, le processus de décision trouve des modes d'expression particulièrement pertinents avec les réseaux bayésiens (RB) et les diagrammes d'influence (DI) [2] [17]. Nous engloberons ces diverses approches sous le terme de Modèles Graphiques Probabilistes (MGP).

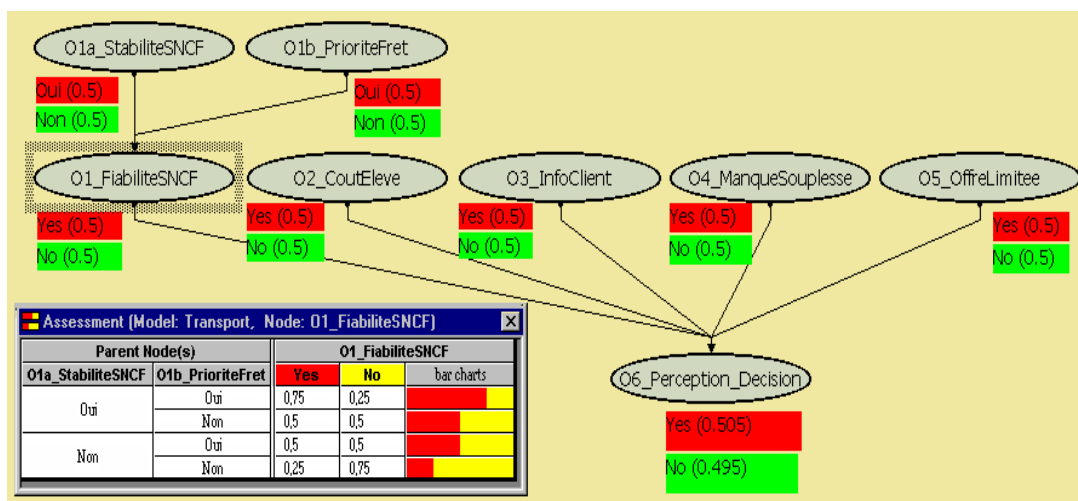


Figure 1 : exemple d'un modèle graphique de décision probabilistique dans le domaine du ferroutage

La figure 1 présente l'exemple concret d'un RB qui modélise de manière synthétique le processus de décision d'un chef d'entreprise vis à vis du ferroutage lorsqu'il doit faire transporter des marchandises¹. La structure du modèle que nous avons construit reflète les informations qualitatives contenues dans un document qui nous a été fourni par le Ministère des Transports avec mission d'élaborer des modèles de décision nouveaux là où les modèles traditionnels ne semblent pas pertinents. Ces informations sont issues d'enquêtes auprès d'un échantillon de chefs d'entreprise.

La question posée est : comment représenter les variables sur lesquelles il est possible de jouer pour amener le chef d'entreprise à penser positivement au ferroutage ? Le mode de raisonnement (implicite) du décideur est ici représenté par un diagramme de causalité (RB). La probabilité de l'état (choix ferroutage OUI) dans le nœud final ('Perception-Décision') est déterminée, selon un principe bayésien, à partir d'un tableau de probabilités conditionnelles et des valeurs des fonctions de probabilité des nœuds parents connues 'a priori', ou qu'il est possible d'inférer à partir d'autres nœuds parents. Nous avons nous-même établi les valeurs des probabilités conditionnelles en fonction de l'ordre d'influence des variables tel qu'il apparaît dans le document du Ministère. Par ailleurs, nous avons mis les 'a priori' à 50% sur chaque nœud source pour représenter une incertitude maximale du décideur.

Dans la figure 1, nous présentons uniquement le tableau des probabilités conditionnelles pour le nœud 01 (Fiabilité de la SNCF) ainsi que les probabilités (a priori ou calculées) pour chacun des nœuds sources.

Parmi les qualités souvent citées des MGP figure la lisibilité. En effet, un réseau bayésien ou un diagramme d'influence est un **modèle graphique** qui offre une **représentation visuelle** intuitive des différents paramètres entrant dans une décision. Or, on voit sur la figure 1 qu'avec assez peu de variables le graphe devient vite compliqué. De plus, la taille d'un tableau de probabilités conditionnelles associé à un nœud est une fonction exponentielle du nombre de ses nœuds parents et des différents états de ces nœuds. Ainsi, la qualité première d'un MGP est-elle discutable si l'on considère qu'un graphe trop complexe n'est pas un graphe facile à expliquer et à justifier. Plus généralement, nous posons la question du niveau de lisibilité d'un graphe par rapport à différentes catégories de lecteur. En effet, le caractère subjectif d'un MGP suppose une interaction entre les modélisateurs et les sujets de la modélisation détenteurs de l'expertise. Si le graphe est trop complexe, il ne peut pas servir de support à cette interaction.

¹ Tous les diagrammes présentés dans cet article ont été réalisés avec l'outil MSBNx de Microsoft [11]

Le présent article pose le problème de l'adaptation d'un modèle de connaissance, ici un modèle bayésien de la décision, selon le lecteur et l'intention de lecture. Nous définissons un méta modèle de la pertinence qui permet de réaliser des adaptations du modèle de connaissance, et nous discutons sa légitimité. Notre approche utilise des modèles de diagrammes d'influence pour décider de la pertinence de chaque variable, et donc de l'opportunité de la présence du nœud correspondant. Ainsi, dans l'exemple que nous utilisons ici, chaque nœud du RB est conditionné par un DI. Nous proposons un schéma général de ces DI afin d'obtenir un modèle d'adaptation générique.

2. Travaux antérieurs

Les Réseaux Bayésiens [17][12] et leurs divers dérivés (Réseaux Bayésiens Dynamiques, [13], arbres de décision, graphes de dépendance [7], Diagrammes d'Influence) connaissent un intérêt croissant étant donné les récentes avancées tant théoriques (modélisation, apprentissage, ...) que pratiques (outils en ligne) en la matière. Ils sont largement utilisés dans des domaines où prévaut la décision automatique ou la décision assistée dans un environnement incertain et éventuellement évolutif : interprétation du langage naturel [5], analyse de transcriptomes [20][6], diagnostic médical, maintenance [21][18], modélisation d'un utilisateur [14][19], coordination d'agents autonomes [25], etc.

Les MGP sont particulièrement adaptés à la modélisation de la prise de décision dans un environnement incertain étant donné le caractère à la fois subjectif et objectif de la démarche, appuyée par une représentation graphique. Deux types de travaux nous interpellent ici. Le processus de modélisation porte sur la décision stratégique : affectation de ressources, déclenchement d'opérations, etc. [16] [3]. Dans ces travaux, le modèle ne dépend pas de l'utilisateur. Il est donné a priori, et peut éventuellement être modifié par le modélisateur dans sa topologie ou dans ses valeurs de répartition des probabilités. Ceci correspond à notre démarche pour établir un modèle initial tel que celui exposé dans la figure 1. Plus intéressantes sont les applications des MGP pour la modélisation d'interfaces ou d'applications adaptatives qui relèvent de l'initiative mixte [8] [9] [10]. Selon cette approche, un modèle de l'utilisateur est construit a priori, à partir d'observations lors de manipulations en laboratoire. L'interface ou l'application s'adapte ensuite au contexte d'utilisation à partir du modèle et des observations des actions de l'utilisateur. Cependant, il n'est pas donné de modèle particulier de la pertinence des informations à afficher en fonction du profil de l'utilisateur, de ses intentions, et du niveau de détail qu'il souhaite dans le modèle. [1] [15] présentent des travaux de recherche qui placent l'intérêt et la préférence de l'utilisateur au centre d'un processus de recherche d'information adaptatif basé sur une topologie évolutive

de réseaux bayésien. Cependant, aucun de ces modèles ne considère un méta modèle bayésien capable de modifier la topologie d'un autre graphe de décision selon la pertinence des nœuds en fonction d'un modèle probabiliste de la situation.

En effet, chez tous ces auteurs, le modèle de décision est soit expertisé a priori, soit appris en situation, mais n'est pas conditionné par un méta modèle qui prend en compte le décideur. Nous voyons dans notre approche d'une part un moyen d'assister un utilisateur dans sa décision par rapport à son contexte puisqu'il est possible de présenter un modèle de décision simplifié et personnalisé sans perte d'information. D'autre part, nous rendons en grande partie explicite ici la part d'irrationnel qui semble souvent accompagner un processus de décision. Notre modèle dévoile la personnalité du décideur et la rend partie prenante du processus de décision, dimension qui selon nous échappe aux approches des auteurs cités.

3. Modèle bayésien de la pertinence

Quand on observe le graphe de la figure 1, un processus de décision paraît complexe, et la communication à son sujet encore plus difficile. Nous postulons qu'un décideur ne peut prendre en compte toutes ses composantes, et que nécessairement il opère, le plus souvent inconsciemment, une réduction du problème en s'attachant à certaines variables plutôt qu'à d'autres selon son contexte et sa personnalité. Cela revient à opérer une simplification du graphe en supprimant certains nœuds. Quel statut donner aux variables qui sont ainsi ignorées ?

Une première méthode consiste à donner *a priori* une valeur à une variable dans son ensemble de définition (vrai ou faux par exemple), ce qui, selon le principe de calcul dans les Réseaux Bayésiens, revient à transformer une probabilité en une évidence. Celle-ci se propage, écartant par là même la nécessité de prendre en compte la variable dans la suite des calculs. Par exemple, si l'on se pose la question de savoir s'il est bon de prendre un encombrant parapluie pour la journée, on peut se contenter de regarder le temps qu'il fait sur le moment parce qu'il est trop compliqué et coûteux de rechercher l'information à la radio ou dans les journaux (principe d'économie cognitive). Parce qu'aussi le soleil du moment est un bon alibi pour ne pas s'encombrer du parapluie (principe de préférence personnelle). Pourtant, la météo qu'on ignore peut prévoir la pluie pour plus tard dans la journée. Par économie cognitive et par goût personnel, on transforme la probabilité de pluie en une évidence, qui peut s'avérer par la suite trompeuse. Nous verrons que c'est cette démarche réductrice qui est mise en œuvre dans l'outil MSBNx de Microsoft pour simplifier un RB. La simplification s'accompagne d'une perte d'information en matière de probabilités, ce qui interdit par la suite une éventuelle révision si d'autres évidences s'imposent.

Nous proposons une autre approche à la fois explicative et opératoire de la simplification du processus de décision en

nous basant sur l'analyse de la 'pertinence' d'une variable dans un contexte particulier. Appliqué à la complexité du graphe de la figure 1, la question qui est posée est la suivante : est-il possible de le simplifier en masquant certains nœuds faiblement pertinents à visualiser afin de le rendre plus compréhensible, tout en ne perdant pas l'information contenue dans ces nœuds masqués qui peuvent s'avérer pertinents à prendre en compte sur le moment ou plus tard ? Ainsi il est important de distinguer la pertinence d'un nœud pour la prise de décision, de la pertinence de visualiser ce nœud dans un contexte particulier.

Nos travaux sur les interfaces adaptatives dans le domaine de la supervision industrielle ou dans celui de la télévision interactive proposent un modèle de la 'pertinence' d'une information en préalable à la décision de sa présentation, [22][23][24]. Nous reprenons ici ce modèle de la pertinence de la visualisation pour lui donner une forme probabilistique en fonction de diverses variables.

La figure 2 présente ce procédé qui à chaque nœud du réseau initial associe un autre graphe de décision, un méta-modèle, qui calcule la pertinence de la visualisation du nœud correspondant. Il convient donc d'affecter à chaque nœud du graphe de décision initial, une fonction de probabilité de la pertinence de visualisation de ce nœud. Il correspond à cette fonction un méta nœud que nous appellerons VisualisationDuNoeud.

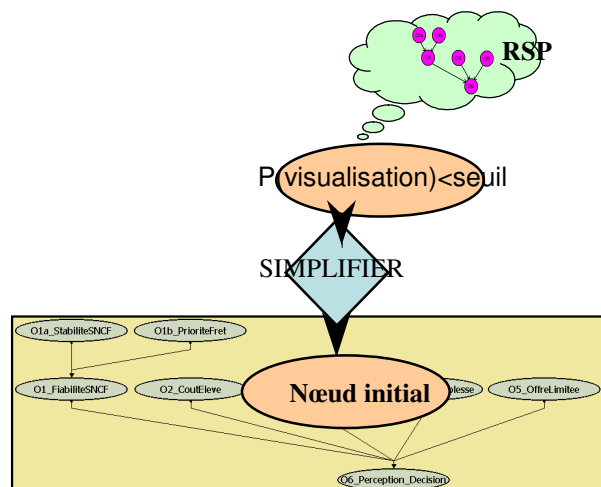


Figure 2 : meta-modèle associé à un nœud initial

La figure 3 propose un méta-modèle où ce nœud figure comme enfant unique d'un réseau bayésien qui comporte les différentes variables susceptibles d'influer sur la pertinence de la visualisation. Nous appelons ce réseau le Réseau Secondaire de Pertinence (RSP). Chacun des nœuds du réseau d'origine, que nous appellerons le Réseau Principal (RP), est ainsi doté (en arrière plan) de son propre RSP. Nous appellerons un nœud du RP ainsi doté d'un RSP un 'nœud conditionné' (NC). Nous n'avons retenu dans le RSP que les variables qui nous paraissent les plus ... pertinentes, car il n'échappera pas au lecteur

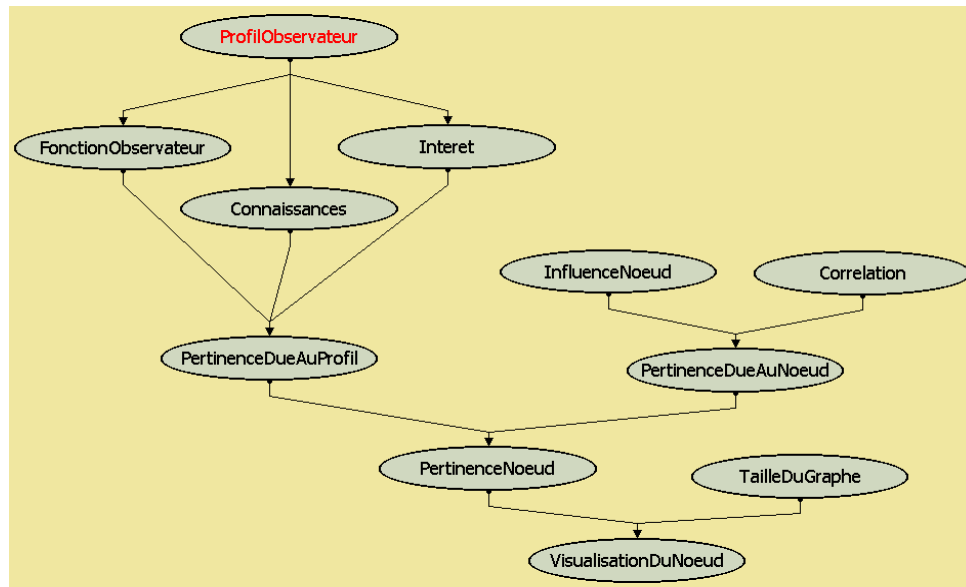


Figure 3 : meta-modèle d'adaptation du modèle de décision

que le même principe de calcul de pertinence peut aussi s'appliquer au RSP qui devient alors un RP à un niveau méta.

3.1 Structure du méta-modèle

Nous posons que la valeur de probabilité de la pertinence pour la visualisation (nœud *VisualisationDuNoeud* du RSP) d'un nœud NC est conditionnée par des classes de variables selon les principes suivants. Deux classes influencent la pertinence du NC et donc convergent vers un nœud *PertinenceNoeud*. Il s'agit d'une classe de variables qui relèvent de la pertinence intrinsèque du NC d'une part, et d'une pertinence perçue par le décideur d'autre part. Une autre classe de variables concerne uniquement l'opportunité de visualiser le nœud si l'on considère l'encombrement du graphe. Il s'agit d'une fonction coût de la visualisation du NC.

3.2 Pertinence intrinsèque

Les nœuds du sous graphe du RSP qui concernent la pertinence intrinsèque du NC convergent sur le nœud *PertinenceDueAuNoeud*. Plus le NC a une influence propre sur la décision finale dans le RP, plus il convient de le conserver visible dans ce réseau. Cette influence propre se décline selon deux aspects. D'une part la part d'influence du nœud sur la décision finale (nœud *InfluenceNoeud* du RSP). Celle-ci peut être observée au travers de l'influence de la variable dans le tableau des probabilités conditionnelles du nœud enfant du NC. On peut aussi envisager la valeur a priori de la variable du NC, sachant qu'une variable qui se trouve presque

toujours dans un certain état pourrait être supprimée puisqu'elle relève d'une évidence².

Le second aspect du nœud *PertinenceDueAuNoeud* porte sur la corrélation du NC avec un éventuel autre nœud du RP (nœud *Corrélation*). Dans un Réseau Bayésien, les nœuds parents d'un même nœud sont sensés représenter des variables indépendantes. Mais ceci suppose un certain degré de certitude du modélisateur dans l'indépendance des variables. Or, dans un univers très incertain, des corrélations inattendues et non directement observables peuvent se produire. Par exemple, un décideur peut recevoir une information identique de deux sources différentes auxquelles il accorde un certain crédit. Il peut alors augmenter sa perception de la véracité de l'information étant donné l'existence de deux sources distinctes. En fait il se peut que l'une des sources (la télévision par exemple annonce une grève à la SNCF) ait transmis antérieurement l'information à l'autre source (le directeur commercial)³. Le réseau présente donc un cycle qui n'est pas connu du décideur. Par contre, le décideur peut soupçonner une certaine corrélation entre les deux sources, et considérer que l'information est redondante entre plusieurs NC (par exemple, sur la figure 1, entre les variables 'manque de souplesse' et 'fiabilité').

3.3 La pertinence perçue

Les nœuds du sous graphe du RSP qui concernent la pertinence perçue du NC convergent sur le nœud

² Nous utilisons le terme 'pourrait' car il peut y avoir d'autres nœuds dans le RSP qui justifient d'un point de vue subjectif pour le décideur la présence du nœud, même si celui ci est intrinsèquement peu pertinent.

³ Notre modèle peut ainsi avoir valeur explicative du phénomène de rumeur.

PertinenceDueAuProfil. Dans notre modèle, nous considérons que le profil du décideur intervient sur la pertinence du nœud NC de trois manières différentes. Nous détaillons ces trois variables.

Le nœud *Connaissance* du RSP. Un nœud sur lequel le décideur a peu d'informations est perçu comme moins pertinent. Autrement dit, par méconnaissance du domaine, soit le décideur l'ignorera, soit il projettera la situation la pire (peur du risque). Cette vision est très réductrice, mais elle représente des comportements classiques de décideurs. Une étude poussée pourrait permettre de classer les décideurs selon leur manière de réagir en situation de méconnaissance d'un domaine. On pourrait tout aussi bien envisager une situation inverse où toute méconnaissance d'une variable qui par ailleurs apparaît pertinente en raison d'autres considérations, peut être vue comme un danger et sa pertinence surestimée. On voit bien que la modélisation du perçu de la pertinence du nœud NC doit particulièrement tenir compte du type de décision et de la psychologie du décideur (prudent, fonceur, etc.) face à l'inconnue. Notre approche montre ici qu'un modèle de décision en milieu incertain ne peut se contenter de données purement statistiques et techniques et que la psychologie du décideur doit en être l'une des variables.

Le nœud *Intérêt*. Un décideur peut aussi marquer son intérêt pour une variable NC indépendamment de la connaissance qu'il en a. Cette variable renvoie aux goûts du décideur, ses penchants, ses attentes.

Il est évident qu'il est difficile d'éviter une corrélation entre la variable *Connaissance* et la variable *Intérêt*. On peut être intéressé par ce que l'on connaît bien (en matière de musique par exemple), ou à l'inverse être intéressé par ce que l'on ne connaît pas (un film déjà vu est en général moins intéressant).

Pourtant, il semble judicieux de séparer la connaissance du domaine de l'intérêt qu'il présente. En cela, nous donnons au décideur une certaine forme de rationalité qui lui fait distinguer ses goûts personnels des nécessités professionnelles. Notre méta modèle isole et met en lumière un aspect de la décision ignoré de tout autre modèle.

Le nœud *FonctionObservateur*. La troisième composante de la perception de la pertinence du NC concerne le rôle social du décideur. Nous pensons qu'un acteur décide différemment selon le poste qu'il occupe. Ceci revient à considérer qu'il ne prend pas en compte de la même manière les paramètres d'une décision, qu'il n'observe pas le monde de manière unique. On peut voir ici aussi une corrélation forte entre cette variable et les variables *Connaissances* et *Intérêts*. Selon le décideur, l'indépendance sera plus ou moins réelle. Eventuellement, un méta méta modèle qui travaillerait sur la pertinence de ces trois nœuds du RSP pourrait, via le nœud *Corrélation*, considérer qu'il ne font qu'un.

3.4 La coût de visualisation

La troisième classe de nœuds parents du nœud *VisualisationDuNoeud* concerne le coût cognitif de la présence du NC. Cette charge cognitive peut être analysée de différentes manières selon que l'on considère un coût objectif (qui fait faiblement appel à un modèle psychologique du décideur) et un coût subjectif lié au décideur. Nous nous limiterons ici au coût objectif, le coût subjectif étant renvoyé au modèle du décideur que nous avons analysé. Nous pouvons résumer le coût objectif de la manière suivante : plus il y a d'informations (de nœuds) dans le RP, plus pénible est son analyse, et donc plus faible est la probabilité qu'un nœud quelconque soit pertinent, indépendamment de sa pertinence intrinsèque et de sa pertinence perçue. Le coût d'un nœud est donc fonction du nombre de nœuds, ou, autrement dit, la pertinence de la visualisation d'un nœud est une fonction inverse du nombre de nœuds. Comme ceci s'applique quelque soit le nœud, ce principe n'est pas discriminant entre les nœuds NC. Cependant, il est possible de l'appliquer de deux manières. Soit on donne a priori un seuil de pertinence, et les nœuds les moins pertinents ne sont pas visualisés jusqu'à dépasser ce seuil. La probabilité de pertinence de chacun des nœuds augmente alors avec la réduction du graphe (que l'on pourrait modéliser par une fonction d'entropie). On arrête de réduire le graphe quand la perte marginale de pertinence due à la suppression d'un nœud équilibre le gain de la simplification.

Soit on fixe a priori le nombre de nœuds, et on ne tient pas compte de leur pertinence. Le risque est de voir disparaître des nœuds à la limite de la pertinence de visualisation. La première solution semble la plus intéressante pour réduire le graphe tout en perdant le minimum de pertinence intrinsèque et perçue des nœuds. Cependant, à des fins de calcul, la solution où il est intéressant de fixer a priori le nombre de nœuds à visualiser est la plus simple et c'est celle que nous avons retenue pour l'instant. En pratique, nous ne retenons que six nœuds dans le RP, ce chiffre correspondant au chiffre traditionnel conseillé pour la complexité d'un graphe en matière de communication à des tiers.

Se limiter à considérer le coût de lecture du graphe comme une fonction du nombre de nœuds est particulièrement réducteur. Il serait intéressant de considérer plus attentivement la lisibilité d'un graphe à partir des autres considérations suivantes : (i) sa topologie, (ii) sa disposition, (iii) la connaissance préalable du graphe par le lecteur. Nous renvoyons le troisième point au modèle du décideur. Le second point a fait l'objet de nombreux travaux dans les communautés IHM et Graph Drawing [4]. Quand à la relation entre la topologie du graphe et sa lisibilité, elle nous apparaît très complexe, en particulier quand le graphe est un réseau bayésien. C'est pourquoi ces différents aspects, quoique importants, restent pour l'instant en dehors de notre modèle.

4 Application et mode de réduction du réseau

4.1 Application

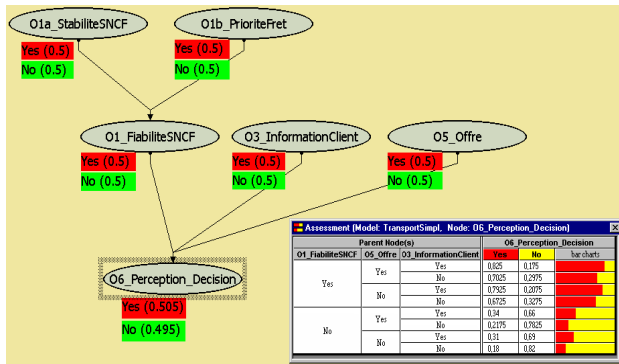


Figure 4 : modèle de décision simplifié de la figure 1

La figure 4 montre comment le graphe de décision de la figure 1 est réduit après application du méta modèle sur tous les nœuds. Les évidences suivantes ont été intégrées par ajout de la valeur de probabilité 1 correspondante dans les a priori : la 'FonctionObservateur' du RSP est 'chef d'entreprise' et les nœuds NC ont été actualisés pour ce profil ; les nœuds 2 et 5 du RP (à savoir 'coût du fret élevé' et 'offre limitée') ont été considérés comme redondants avec l'informationClient ce qui entraîne une probabilité de corrélation égale à 1 dans les RSP de ces deux nœuds et donc une perte de leur pertinence. Finalement, il ne reste que six nœuds du réseau initial dans le réseau adapté.

4.2 Mode de réduction du réseau initial

Lorsque le réseau initial est réduit, il se pose la question de savoir ce que l'on fait de l'influence des nœuds qui disparaissent sur les autres nœuds qui restent. Une première manière de faire est celle à laquelle force l'outil MSBNx. Quand on souhaite effacer un nœud d'un réseau, l'outil demande l'évidence qu'il convient de mettre sur la variable. Autrement dit, la topologie du réseau est transformée parce qu'une variable devient connue et donc elle n'intervient plus, après propagation de l'évidence, sur les autres variables.

Or, ce procédé n'est pas applicable dans notre cas, car on considère qu'il n'y a pas d'évidence, mais qu'une variable masquée est toujours présente sans en connaître la valeur. Ceci revient à bien distinguer entre la suppression d'une variable parce qu'elle est faiblement pertinente, ou bien sa dissimulation pour simplifier le réseau. En supprimant purement et simplement une variable, on interdit qu'elle puisse par la suite revenir au devant de la scène (ceci doit être pris au sens propre). A l'inverse, selon notre approche, en la dissimulant provisoirement, elle reste latente et active. La manière de la dissimuler provisoirement tout en conservant son influence est d'opérer une opération de marginalisation [Pearl 88]

dessus, ce qui donne un résultat sur les probabilités des nœuds enfants en général différent du résultat qui découle de la simple introduction d'une évidence.

Nous voyons là la puissance explicative de notre modèle sur certains choix décisionnels erronés et surtout entêtés. Un décideur peut ignorer certaines variables à un moment donné par manque d'information pour disposer d'un modèle de décision calculable. Mais dans la mesure où le modèle est non adaptatif, même si des évidences viennent remettre en question les conclusions, il ne peut revenir sur le fait d'avoir écarté certaines variables, car elles l'ont été parce que considérées comme non pertinentes. En réalité elles n'étaient pas suffisamment exprimées à un moment donné mais restaient actives.

Un modèle adaptatif permet de revoir le modèle lui-même en fonction des évidences qui peuvent remettre en question les préjugés de départ. Nous voyons là dans notre approche un outil puissant non seulement de raisonnement adaptatif, mais encore un modèle de certaines formes de défauts de raisonnement, en particulier les raisonnements entachés de présupposés (les centres d'intérêts du décideur influent sur la décision), et les entêtements (des variables initiales ont été ignorées par manque d'information, et restent ignorées même si l'information devient disponible). Des faits de société récents sont des exemples concrets de ce type de raisonnement erroné.

On peut trouver une illustration de ce phénomène d'occultation définitive d'une variable dans notre exemple du ferroutage. Un chef d'entreprise considérera comme une évidence la non fiabilité de la SNCF suite à une grève, alors que statistiquement cette variable n'est pas déterminante.

5 Acquisition des valeurs des variables

Nous avons ainsi défini la topologie et la sémantique du méta modèle de pertinence de la visualisation d'un nœud NC. Nous n'avons rien dit sur le mode d'acquisition des valeurs des probabilités conditionnelles et des 'a priori' des RSP. Selon le type de variable et selon le contexte, la méthode d'acquisition peut être très différente.

La méthode classique est celle de valeurs subjectives données par le modélisateur. Cette approche est tout à fait justifiée par les théoriciens des réseaux bayésiens [17]. Mais ceci pose le problème de la difficulté de la tâche de modélisation et de sa visibilité, surtout si le modélisateur doit tenir compte de la psychologie du décideur, ce qu'impose notre modèle. C'est ce que nous avons fait dans l'exemple que nous avons proposé pour certains nœuds en nous laissant guider par notre intuition à défaut de disposer de méthodes d'apprentissage ou de méthodes de calcul.

Pour ce qui concerne les nœuds *Intérêt* et *Connaissance*, nous proposons un mode de calcul générique. Un vecteur de mots clefs correspondant à un profil de décideur peut

être associé à chacun de ces nœuds⁴. Par ailleurs, un autre vecteur peut être associé à chaque NC correspondant à son domaine de connaissance. Ainsi, il est possible de calculer la proximité (ou l'éloignement selon le point de vue) sémantique du décideur par rapport aux différentes variables qui influencent la décision. Cette valeur peut ensuite être exploitée comme 'a priori' de pertinence pour ces nœuds. L'intérêt de cette approche que nous avons utilisée dans notre exemple est qu'elle permet d'envisager un phénomène d'adaptation dynamique (apprentissage), des informations nouvelles pouvant modifier le profil du décideur, et donc sa manière de penser la décision. Elle est de plus générique, un même modèle pouvant servir à modéliser plusieurs acteurs d'une décision, voire même à définir quel est le collectif d'acteurs qu'il est nécessaire de mettre en place pour peser une décision. Pour les autres nœuds, la détermination des valeurs est plus sujette à discussion. Nous voyons alors deux approches : la détermination de stéréotypes de décideurs d'une part, et éventuellement l'utilisation de fonctions d'apprentissage pour des processus de décision récurrents.

6. Conclusion

L'approche que nous avons proposée pour modéliser la décision repose sur un double principe. L'utilisation, classique, des réseaux bayésiens permet de rendre compte du caractère incertain et subjectif d'un processus de décision. Mais ce processus est prédéterminé dans sa topologie par le modélisateur. Ceci ne permet pas de rendre compte de nombreux effets qui paraissent non rationnels, et personnels d'un processus de décision. Certaines variables qui influent une décision peuvent apparaître non pertinentes à un moment du fait que la charge cognitive du décideur ne peut les englober toutes. Nous considérons une modélisation adaptative qui réduit la complexité de l'espace de décision tout en gardant toutes ses dimensions. Il s'agit de réduire cet espace pour le rendre accessible au décideur et à la décision. Considérant que réduire un diagramme de décision consiste à opérer des micro-décisions sur ce qu'il est important ou pas de considérer à moment donné, nous affectons à chaque nœud du graphe de décision un réseau secondaire qui en mesure la pertinence. Celle-ci doit s'entendre comme une opportunité de prise en compte de chaque nœud, et non sa suppression. Ce modèle nous apporte un outillage précieux de modulation du problème de la décision selon l'interlocuteur, la situation et son

⁴ Nous avons utilisé pour cela dans MSBNx la possibilité de rattacher des propriétés aux nœuds d'un graphe. Nous avons ensuite sauvé ces valeurs au format XML, les avons récupérés sous EXCEL pour faire les calculs, puis avons rapporté les résultats dans le modèle MSBNx à la main. La démarche pourrait être entièrement automatisée avec ces environnements à l'aide de feuilles de style XSLT, en attendant un outil qui intègre l'ensemble.

évolution. Ce modèle nous paraît un outil particulièrement pertinent tant pour assister la décision, que pour l'expliquer. Nous avons montré en particulier qu'il pouvait servir à expliquer des processus de décision fallacieux ou tendancieux. Ainsi ce méta modèle adaptatif d'un modèle de décision peut-il jeter une lumière formelle sur la dimension psychologique de cette activité cognitive, chose rarement explicitée dans les modèles antérieurs.

Annexe

L'auteur remercie les élèves de l'Ecole des Mines d'Ales qui ont participé avec application à l'ensemble des travaux dont témoigne cet article:

Bibliographie

- [1] S. Brown and E. Santos, *Active User Interfaces for Building Decision-Theoretic Systems*. In Proceedings of the 1st Asia-Pacific Conference on Intelligent Agent Technology, Hong Kong."1999
- [2] E. Charniak *Bayesian Networks Without Tears*. AI Magazine, vol. 12, 1991.
- [3] B. Das, *Representing Uncertainties Using Bayesian Networks*, Information Technology Division Electronics and Surveillance Research Laboratory DSTO-TR-0918 www.dsto.defence.gov.au/corporate/reports/DSTO-TR-0918.pdf, 1999.
- [4] Di Battista, G., Eades, P., Tamassia, R., Tollis, I., Graph drawing. Algorithms for the visualisation of graphs. Prentice Hall, 1999.
- [5] M. R. Eizirik, Valmir C. Barbosa, Sueli B. T. Mendes: *A Bayesian-Network Approach to Lexical Disambiguation*. Cognitive Science 17(2): 257-283 (1993)
- [6] N. Friedman, M. Linial, I. Nachman, D. Pe'er, Using Bayesian Networks to Analyse Expression Data, RECOMB 2000.
- [7] Hekerman, D.M. Chickering, C. Meek, R. Rounthwaite, C. Kadie, Dependency Networks for Inference, Collaborative Filtering, and Data Visualization. Journal of Machine Learning Research 1 (2000), p. 49-75.
- [8] Horvitz, E. Principles of Mixed-Initiative User Interfaces. Proceedings of CHI '99, ACM SIGCHI Conference on Human Factors in Computing Systems, Pittsburgh, PA, May 1999.
- [9] Horvitz, E., Breese, J., Heckerman, D., Hovel, D., Rommelse, K.. The Lumiere Project: Bayesian User Modeling for Inferring the Goals and Needs of Software Users. Proceedings of the Fourteenth

- Conference on Uncertainty in Artificial Intelligence, July 1998.
- [10] E. Horvitz and A. Seiver. Time-Critical Action: Representations and Application. Proceedings of the Thirteenth Conference on Uncertainty in Artificial Intelligence, August 1997.
- [11] C.M. Kadie, D. Hovel, and E. Horvitz. MSBNx: A Component-Centric Toolkit for Modeling and Inference with Bayesian Networks. Microsoft Research Technical Report MSR-TR-2001-67, July 2001.
- [12] K. Murphy, A Brief Introduction to Graphical Models and Bayesian Networks
<http://www.cs.berkeley.edu/~murphyk/Bayes/bayes.html>
- [13] Murphy K. et S. Mian, Modelling Gene Expression Data using Dynamic Bayesian Networks. Technical report, Computer Science Division, University of California, Berkeley, CA.
<http://citeseer.nj.nec.com/murphy99modelling.html>
- [14] Murray W.R. An easily implemented, linear-time algorithm for bayesian student modeling in multi-level trees. In actes de la conférence Artificial Intelligence in Education, Le Mans 1999.
- [15] Nguyen H., G. Mitchell Saba, Eugene Santos, Jr., Scott M. Brown Active user interface in a knowledge discovery and retrieval system. University of Connecticut 2000°
- [16] S. Noh and Piotr J. Gmytrasiewicz, Bayesian Belief Update in Anti-air Defense, In the Workshop on Machine Learning for User Modeling of the Sixth International Conference on User Modeling, Italy, June 1997.
- [17] J. Pearl, Probabilistic Reasoning in Intelligent Systems, San Mateo: Morgan Kaufmann, 1988.
- [18] J. Pearl, "Causal Inference in the Health Sciences: A Conceptual Introduction" Contributed to Health Services and Outcomes Research Methodology, 2001.
- [19] M. Sahami, S. Dumais, D. Heckerman, and E. Horvitz. A Bayesian approach to filtering junk e-mail. AAAI Workshop on Learning for Text Categorization, July 1998, Madison, Wisconsin.
- [20] Segal E, Taskar B, Gasch A, Friedman N, Koller D. "Rich probabilistic models for gene expression." Bioinformatics. 2001;17 Suppl 1:S243-52
- [21] Stangl DK, Berry DA. Meta-analysis: past and present challenges. In: Stangl DK, Berry DA, editors. Meta-Analysis in Medicine and Health Policy. New York: Marcel Dekker; 2000. p. 1-28.
- [22] Vaudry, C., Crampes, M., AMEBICA : une interface adaptative pour la supervision industrielle. Actes de la Conférence JIM 2001, Interaction Homme Machine et Assistance, Metz, 2001.
- [23] Vaudry, C., Crampes, M., Modèle de la composition multimédia adaptative dans l'interaction homme-machine. Actes de MFI 2001, Modèles Formels de l'Interaction, Volume III, B. Chaaib-Draa et P. Enjalbert editors, Toulouse, 2001.
- [24] Vaudry, C., Crampes, M., Composition adaptative dans la supervision industrielle. RIHM (octobre 2002)
- [25] A. Zunino, A. Amandi, *Representing Coordination Relationships With Influence Diagrams*. In Proceedings of the Argentine Symposium on Artificial Intelligence (ASAI2001), JAIIO pp. 130-144.