Article paru dans la Revue de l'Electricité et de l'Electronique (REE) de février 2000

Formalisation de l'argumentaire de sûreté d'un système programmé critique à l'aide d'un réseau bayésien, selon la méthode SERENE

Marc Bouissou • EDF, Div. R&D/ESF • Clamart

Mots-clés : fiabilité, logiciel critique, SERENE, réseau bayésien, expertise

L'accroche :
Le projet européen SERENE a donné naissance à une méthode totalement nouvelle, qui rend plus lisibles les raisonnements des experts en évaluation de la sûreté des systèmes programmés critiques, en les modélisant grâce aux réseaux bayésiens.

L'essentiel :
- L'évaluation de la sûreté d'un système programmé critique (tel que le système de protection d'une centrale nucléaire, le contrôle-commande d'un avion...) ne peut être faite uniquement à l'aide des méthodes classiques de la sûreté de fonctionnement, fondées sur une analyse de retour d'expérience et des modèles de fiabilité.
- La phase de synthèse de cette évaluation est essentiellement qualitative, et est réalisée entièrement par des experts, qui ont à pondérer des informations de sources et natures très disparates, portant à la fois sur le processus de fabrication du système et sur les résultats des analyses, tests et audits dont il a fait l'objet.
- Les réseaux bayésiens (modèle mathématique probabiliste) permettent de modéliser les raisonnements faits par ces experts en prenant en compte le caractère incertain des déductions faites, de façon à les rendre plus lisibles, plus systématiques, plus reproductibles...
- L'intérêt de la méthode SERENE dépasse largement le domaine dans lequel elle a été développée initialement. Son caractère très générique en fait un support de choix dans tout domaine où l'on est amené à faire une synthèse experte d'un ensemble d'évaluations qualitatives.

Synopsis :
- The safety assessment of a safety critical software intensive system (like the protection system of a nuclear power plant, an aircraft command and control system...) cannot rely only on conventional reliability analyses, based on reliability models and feedback of experience data.
- The synthesis part of this evaluation is essentially qualitative, and is entirely done by experts, who must take into account information from quite diverse sources and natures, relative to both the system building process and the results of analyses, tests and audits the system had to pass.
- Belief Networks (a mathematical probabilistic model) make it possible to model the expert's reasoning, to make it more readable, systematic, repeatable, while taking into account the uncertainty in the deduction process.
- The SERENE method scope goes far beyond the domain in which it was initially devised. It is quite generic, making it a first choice tool in any domain where one has to perform an expert synthesis of a set of qualitative evaluations.

1. INTRODUCTION

Ces dernières années ont été marquées par une tendance à remplacer de plus en plus de systèmes électromécaniques de contrôle commande des installations industrielles par des systèmes programmés. Bien que cette nouvelle technologie ait clairement de nombreux avantages, elle n'est pas encore utilisée largement dans certaines applications critiques, telles que les systèmes de protection des centrales nucléaires.
L'explication de cette réticence est la complexité de l'évaluation de la sûreté et de la certification de ces systèmes, due en particulier aux caractéristiques du logiciel.

En effet, on ne peut prétendre que l'effet des erreurs de conception est négligeable par rapport aux défaillances matérielles (qui sont bien plus faciles à estimer) : en fait, ce type d'erreurs joue souvent un rôle important.
Ceci est malheureux, car les erreurs de conception sont de loin les plus difficiles à prédire. Ceci est dû en particulier au fait qu'un système numérique a un comportement discontinu : le moindre changement dans ses entrées, ou dans sa conception peut conduire à des changements énormes (et potentiellement catastrophiques) de son comportement.

Les modèles de croissance de fiabilité (qu'ils soient dédiés au logiciel, ou plus généraux) ne sont pas utilisables pour les systèmes critiques pour la sûreté, car le nombre de défaillances observées est (fort heureusement), beaucoup trop faible pour autoriser une quelconque exploitation statistique de ces données.

Donc, le seul moyen de certifier un tel système est de construire un "argumentaire de sûreté", qui est une collection de tous les types de preuves de bon fonctionnement, relatives aussi bien au processus de développement qu'au produit final.

Le type d'information à réunir, et la façon de le faire sont maintenant relativement bien définis, au moyen par exemple de documents types et de check-lists telles que celles que l'on donne dans les projets ESPRIT SHIP [1] et CASCADE [2].

Mais jusqu'au projet SERENE, aucune méthode n'avait été proposée pour combiner les divers éléments de preuve en un argument global. Ce travail crucial reposait entièrement sur l'expertise de l'évaluateur.

C'est pour tenter de résoudre ce problème que le projet de recherche SERENE (juin 96-juin 99) a été monté, associant EDF aux partenaires suivants : ERA Technology (chef de projet, UK), Centre for Software Reliability (UK), Objectif Technologie (France), TÜV Nord (Allemagne), et HUGIN (Danemark). Il s'agissait de construire une méthode et un outil contribuant à rendre la partie "synthèse" d'un argumentaire de sûreté plus compréhensible, et plus reproductible.

SERENE s'appuie sur l'utilisation des réseaux bayésiens, un formalisme mathématique bien connu pour son aptitude à modéliser des raisonnements portant sur des faits incertains, pour modéliser et formaliser l'expertise des évaluateurs.

Ce projet a été l'occasion de confronter des approches pratiquées dans différents pays (les partenaires apportant l'expertise sur le domaine étaient Anglais, Français, et Allemand) ; ces

approches étant très différentes, le projet SERENE a produit une méthode et un outil possédant un grand degré de généralité, car adaptable à toutes sortes de contextes.

L'objet de cet article est la présentation de la méthode SERENE à travers un exemple d'application : le modèle construit par EDF, et les résultats de l'expérimentation de ce modèle sur onze projets informatiques réels.

L'article est organisé comme suit :

- La partie 2 donne une définition des réseaux bayésiens, et explique sur un cas d'école le type d'utilisation qu'on peut en faire,
- La partie 3 présente brièvement la méthode SERENE,
- La partie 4 décrit le réseau bayésien que nous avons construit,
- La partie 5 donne un aperçu des expérimentations faites et des résultats obtenus grâce à ce modèle.

2. QU'EST CE QU'UN RESEAU BAYESIEN ?

Le développement théorique des réseaux bayésiens date des années 70, mais à cette époque, le manque d'algorithmes efficaces et d'outils pratiques a empêché le développement d'applications.
Aujourd'hui, grâce à des outils puissants et conviviaux, les réseaux bayésiens sont une technique en pleine expansion, dans beaucoup de domaines où l'on a besoin d'aides à la décision dans un contexte de connaissance incertaine du "monde réel", par exemple dans les domaines médical, militaire, financier, robotique, météorologique, etc. [3].

Un réseau bayésien est un objet mathématique relativement simple ; cependant, nous allons le présenter ici d'une manière plutôt informelle, afin de ne pas dérouter les lecteurs qui ignorent tout de ce concept.

Les lecteurs intéressés par une présentation complète et mathématique des réseaux bayésiens, ainsi que des algorithmes nécessaires à leur traitement, pourront se référer à [4], tandis que ceux plus intéressés par les applications, et par la façon de construire un réseau bayésien en partant d'une expertise humaine pourront plutôt lire [5], et, bien sûr, le manuel de la méthode SERENE [8].

2.1 Un problème simple faisant intervenir des faits incertains

Le petit exemple que nous allons maintenant utiliser pour introduire les concepts des réseaux bayésiens est une adaptation d'un extrait du manuel de la méthode SERENE.
Imaginons que nous devions modéliser la **connaissance** suivante : « Fantasio et Gaston vont à leur travail en utilisant des moyens de transport différents. Gaston utilise sa voiture, alors que Fantasio voyage en train. Fantasio manque rarement son train, et le train est presque toujours à l'heure, *sauf les jours de grève*. Toutefois, une grève de train n'implique pas forcément que Fantasio soit en retard (il peut partir tôt en voiture). Une grève de train peut aussi retarder Gaston car elle provoque des embouteillages. Mais Gaston est de toute façon souvent en retard parce qu'il n'entend pas la sonnerie de son réveil, et de ce fait, une grève n'augmente la probabilité de son retard que d'une faible quantité. En cas de grève, Gaston a moins de chances d'être en retard que Fantasio. »

Maintenant, étant donné cette connaissance, comment pourrions nous modéliser les **inférences** suivantes, issues d'un raisonnement intuitif ?

1- Si nous savons que Fantasio est en retard, nous pensons qu'il y a une grève des trains, et donc que Gaston risque (un peu) plus que d'habitude d'être en retard,
2- Supposons que nous sachions que Gaston est en retard. Cette constatation augmente notre croyance en les deux causes possibles de ce retard (grève, réveil non entendu). Mais si nous apprenons que Fantasio est également en retard, nous serons tentés d'en déduire qu'une grève de train est en cours, et a été la cause du retard de Gaston, ce qui fait retomber quelque peu notre croyance en le fait qu'il n'a pas entendu son réveil.

2.2 Un réseau bayésien pour représenter la connaissance

Un réseau bayésien (RB) est un graphe (constitué de noeuds et d'arcs), associé à un ensemble de tables de probabilités de noeuds (TPN), ainsi nommées car il y en a une et une seule par noeud du graphe.

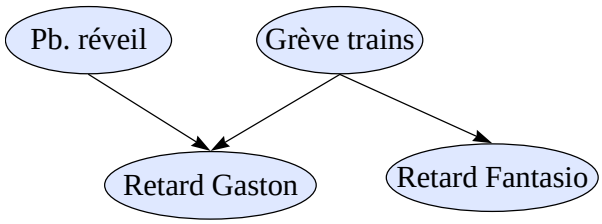

Figure 1 : Un RB représentant la connaissance décrite dans le texte du § 2.1

Les noeuds représentent des variables aléatoires discrètes (il y a quelques extensions des réseaux bayésiens vers le domaine des variables continues, mais elles sont soumises à de fortes limitations sur les types de distributions utilisables : c'est pourquoi nous nous limiterons au cas des variables discrètes). Dans notre exemple, (cf. Figure 1 ci-dessus) les quatre variables ont seulement deux états : 'Vrai' et 'Faux'.

Les arcs représentent des relations de cause à effet entre variables. Comme une grève peut provoquer le retard de Fantasio, nous modélisons cette relation par un arc allant du noeud 'Grève trains' au noeud 'Retard Fantasio'.

Le grand avantage des réseaux bayésiens est de permettre de modéliser des relations non déterministes. Dans notre exemple, voici quelle pourrait être la table de probabilités modélisant la dépendance entre le retard de Fantasio et la grève des trains :

	Grève trains = Vrai	Grève trains = Faux
Pr(Retard Fantasio = Vrai)	0.6	0.1

C'est en fait la distribution de probabilité de la variable 'Retard Fantasio', conditionnelle à la variable 'Grève trains'. La table donne seulement la probabilité de l'événement 'Retard Fantasio' = Vrai, car Pr(Retard Fantasio = Faux) = 1- Pr(Retard Fantasio = Vrai) [1]. Cette table exprime d'une manière formelle et précise le fait que Fantasio a très peu de chances d'être en retard en temps ordinaire, mais que s'il y a une grève des trains, au contraire, il risque fort d'être en retard (la probabilité est de 0.6).

[1] A partir de ce point, nous ne donnerons que les probabilités pour la valeur 'Vrai', car celles pour la valeur 'Faux' en seront les complémentaires.

Afin de formaliser notre connaissance décrite plus haut, nous modélisons la relation entre le retard de Gaston et ses deux causes possibles par la table de probabilités suivante :

Pb. réveil	Vrai		Faux	
Grève trains	Vrai	Faux	Vrai	Faux
Pr(Retard Gaston = Vrai)	0.7	0.5	0.4	0.1

Les tables de probabilités associées aux noeuds 'Grève trains' et 'Pb. réveil' ont une nature quelque peu différente. Ces noeuds n'ont pas de noeud parent dans ce modèle (ce sont des noeuds racines), et nous n'avons donc qu'à leur assigner des probabilités pour leurs deux valeurs 'Vrai' et 'Faux'. En fait, nous supposerons que Pr(Grève trains = Vrai) = 0.1, et que Pr(Pb. réveil = Vrai) = 0.4.

Les noeuds racines modélisent des variables indépendantes entre elles. C'est bien le cas ici.

Il peut y avoir diverses manières de déterminer les probabilités des TPNs. Par exemple, nous pourrions tirer Pr(Grève trains = Vrai) d'un historique sur les jours de grève. Mais en l'absence de telles données de retour d'expérience, il est toujours possible de faire appel à des valeurs de probabilités subjectives, évaluées par des experts.

L'avantage des réseaux bayésiens est de pouvoir mêler dans un cadre théorique unique (la théorie des probabilités) les probabilités issues d'un traitement statistique de retour d'expérience, et les probabilités subjectives.

2.3 Exploitation du réseau bayésien pour faire des déductions

Les **inférences** données au § 2.1 (et beaucoup d'autres !) peuvent être obtenues par des **calculs** (faits selon la théorie des probabilités) réalisés sur le RB que nous venons de décrire.
Le RB est en fait une **représentation concise de la distribution de probabilité conjointe des variables du RB, qui est l'information la plus complète que l'on puisse donner au sujet de cet ensemble de variables aléatoires.**

Cette distribution pourrait, dans cet exemple simple, être donnée par un grand tableau, unique, donnant les probabilités des 2^4 combinaisons possibles de valeurs de l'ensemble des variables.

Construire cette table serait un travail infaisable pour un gros RB, mais il n'est pas nécessaire, car les calculs peuvent être faits directement sur le RB. Par exemple, nous pourrions vouloir calculer la probabilité de l'événement 'Retard Fantasio' :

Pr(Retard Fantasio) =
Pr(Retard Fantasio | Grève) * Pr(Grève) + Pr(Retard Fantasio | pas de Grève) * (1 - Pr(Grève))
= (0.6 * 0.1) + (0.1 * 0.9) = 0.15

Cette probabilité est un exemple de ce qu'on pourrait appeler une probabilité "a priori" : elle reflète notre connaissance sur l'événement 'Retard Fantasio' en l'absence d'information sur l'état réel d'un quelconque des noeuds du RB.

Le tableau suivant récapitule les probabilités a priori pour les quatre variables du RB :

Pb. réveil	Grève trains	Retard Fantasio	Retard Gaston
0.4	0.1	0.15	0.286

En fait, l'utilisation la plus importante des réseaux bayésiens est la révision des probabilités à la lumière de l'observation d'événements.

- Supposons, par exemple que nous sachions qu'il y a une grève. Nous pouvons "saisir cette observation" dans le modèle en mettant Pr(Grève trains = Vrai) à 1. Alors les tables de probabilités conditionnelles nous donnent directement les probabilités du retard de Fantasio (0.6) et un calcul indique que la probabilité du retard de Gaston devient 0.52, ce qui est cohérent avec la phrase "En cas de grève, Gaston a moins de chances d'être en retard que Fantasio" du § 2.1. Si ce n'était pas le cas, cela signifierait qu'il faudrait réviser les TPNs.

Le fait de saisir des observations en mettant certaines probabilités (correspondant aux événements observés) à 1 et de réactualiser toutes les autres probabilités dans le RB s'appelle **propagation**.

- Supposons, maintenant, que nous ne savons pas s'il y a une grève, mais nous savons que Fantasio est en retard. Alors, nous pouvons saisir l'observation 'Retard Fantasio' = Vrai et nous pouvons utiliser cette observation pour calculer la probabilité (révisée) qu'il y ait une grève, et la probabilité (révisée) que Gaston soit en retard. En fait, l'observation saisie augmente significativement la probabilité qu'il y ait une grève (qui passe de 0.1 à 0.4), et augmente légèrement la probabilité que Gaston soit en retard (de 0.286 à 0.364). Ceci est la version "formelle" de l'inférence (intuitive) n°1 du § 2.1.

- Voici un dernier exemple de propagation, correspondant à l'inférence n°2 du § 2.1. Supposons que nous ayons vu que Gaston est en retard. Cette constatation augmente notre croyance en les deux causes possibles de ce retard (grève, réveil non entendu). En fait, l'application du théorème de Bayes permet de calculer que la probabilité révisée de grève est de 0.182 (rappelons que sa probabilité a priori était de 0.1) et la probabilité révisée de non réveil de Gaston est de 0.727 (probabilité a priori : 0.4). Donc, si nous avions à parier, nous miserions bien plus sur la deuxième cause. Supposons maintenant que nous découvrions que Fantasio aussi est en retard. En saisissant cette observation et appliquant le théorème de Bayes, nous trouvons des probabilités révisées (pour la deuxième fois) de 0.571 pour une grève des trains, et de 0.637 pour le non réveil de Gaston : cette fois, les deux causes sont presque à égalité.

Ce petit exemple montre que le fait d'utiliser un modèle simple en RB nous permet de réaliser des déductions *bien plus étayées* que ce qu'il aurait été possible de faire en utilisant juste des phrases contenant des expressions telles que : "très probable", "peu probable", "augmente légèrement", etc.

La même remarque est vraie au *sujet des raisonnements que l'on peut faire sur la sûreté d'un système programmé critique*.

En fait, la puissance des réseaux bayésiens apparaît pleinement lorsqu'on réalise que la théorie des probabilités permet de propager de manière **cohérente** l'impact des observations faites sur l'ensemble des issues incertaines. C'est ainsi que les réseaux bayésiens peuvent déjouer les erreurs communes faites dans des raisonnements intuitifs, erreurs dues à une mauvaise connaissance des probabilités.

3. LA METHODE SERENE

3.1 Construction du modèle

Lorsqu'on cherche à construire un RB à partir de connaissances expertes, on procède habituellement en trois temps :

- Identification des noeuds : il s'agit de repérer quelles sont les variables importantes par rapport aux questions que l'on va être amené à résoudre grâce au modèle. On ne peut se contenter des noeuds correspondants aux variables observables : en effet, en général, c'est une structure sous-jacente faisant intervenir des variables non observables qui permet de relier entre elles les observables. Dans le cas de l'évaluation de la sûreté des systèmes programmés critiques, on s'intéresse par exemple à la pertinence du cahier des charges, la compétence de l'équipe de développement, l'effort dépensé en validation...
- Identification de la structure du RB : la structure correspond aux arcs du graphe. Elle s'appuie en général fortement sur des relations de cause à effet connues des experts, relations non déterministes en général. Par exemple, le fait que les modules logiciels soient de petite taille augmente en principe leur testabilité, mais on peut très bien avoir un module de grande taille et néanmoins très facile à tester.
- Choix du nombre d'états des noeuds et détermination du contenu des tables de probabilités.

L'expérience des partenaires EDF, ERA et TÜV du projet SERENE, qui ont dû, après formation aux concepts des réseaux bayésiens, en construire pour modéliser leurs approches d'évaluation des systèmes programmés critiques, a montré que sans le support d'une méthode, ce processus de construction est extrêmement laborieux et consommateur de temps, et ceci même si l'on connaît bien le domaine à modéliser.
Il peut, de plus, conduire les néophytes à faire des erreurs ; en voici deux courantes :

- choix d'une mauvaise orientation pour les arcs,
- production d'une structure de RB inexploitable, dans laquelle on trouve des TPNs de trop grande taille, impossibles à renseigner par des experts car faisant intervenir trop de facteurs simultanément.

Les modèles construits par les différents partenaires étaient si éloignés les uns des autres, qu'il est rapidement apparu qu'il serait impossible de trouver un modèle unique, sur lequel l'ensemble des partenaires pourraient se mettre d'accord. Mais l'examen de ces modèles a permis d'identifier un très petit nombre de structures typiques, correspondant à des schémas de pensée "universels", que l'on retrouvait maintes fois.

Cette constatation a permis de forger le **premier concept clé de la méthode SERENE : celui d'idiome**.
Les idiomes identifiés ont été répertoriés et décrits dans le manuel de la méthode. Ils se présentent sous une forme très générique, abstraite : dans chaque situation particulière, ils seront instanciés sous une forme un peu différente, par le nombre de noeuds du sous réseau correspondant, ou bien par les tables de probabilités associées.

Voici deux exemples d'idiomes (ce sont les deux les plus utilisés) :

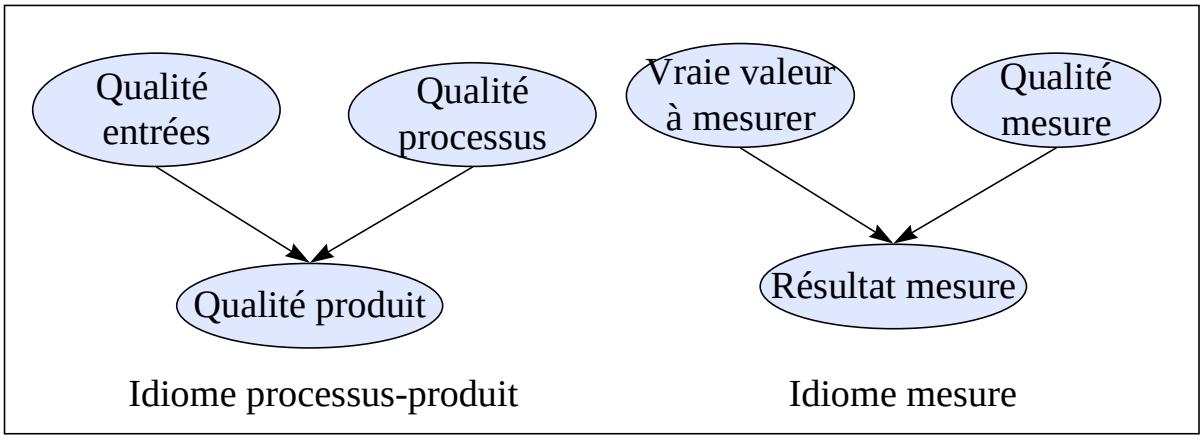

Figure 2 : Deux idiomes de la méthode SERENE

L'idiome processus-produit exprime le fait que la qualité d'un produit (par exemple la spécification d'un système) dépend à la fois de la qualité des données d'entrée qui ont servi à le réaliser (dans le même exemple, ce serait le cahier des charges), et du processus de construction. On entend par processus de construction l'ensemble des moyens humains et matériels, ainsi que les méthodes, qui ont servi à la réalisation.

L'idiome mesure, quant à lui, exprime la relation entre la vraie valeur d'une propriété mesurée (par exemple la présence/absence d'erreurs dans un programme), la qualité de la mesure (dans le même exemple, ce serait l'exhaustivité des tests), et le résultat obtenu pour la mesure. La prise en compte de la qualité de la mesure est fondamentale.
En effet, il est clair que des tests très superficiels peuvent donner à tort l'impression qu'un logiciel est sans erreur : on a alors un résultat de mesure positif, mais on ne peut pratiquement rien en déduire quant à la qualité réelle du programme.

Le **deuxième concept clé est l'assemblage de sous-réseaux**. Cette opération permet de définir des représentations à différents niveaux de détail d'un réseau complexe. Celui-ci est alors représenté comme une arborescence de sous-réseaux, dont les feuilles sont des fragments de réseaux bayésiens au sens classique : ce sont habituellement des instanciations d'idiomes.

L'assemblage de deux sous-réseaux A et B est une opération très simple, consistant à fusionner un (ou plusieurs) noeud(s) (dit(s) de "sortie") du réseau A avec autant de noeuds (dit(s) d' "entrée") du réseau B. L'opération n'est possible que si les noeuds d'entrée n'ont pas de parent avant la fusion. Il n'y a aucune condition relative aux noeuds de sortie.

Par exemple, on peut assembler les deux idiomes cités ci-dessus, en fusionnant le noeud 'Qualité produit' de l'idiome processus-produit avec le noeud 'Vraie valeur à mesurer' de l'idiome mesure.
Lors de cette fusion, la TPN du noeud 'Vraie valeur à mesurer' au sein de l'idiome mesure est **remplacée** par la TPN du noeud 'Qualité produit' de l'idiome processus-produit. Voici le modèle obtenu après assemblage (nous l'avons instancié dans le cas où le produit considéré est la spécification d'un système) :

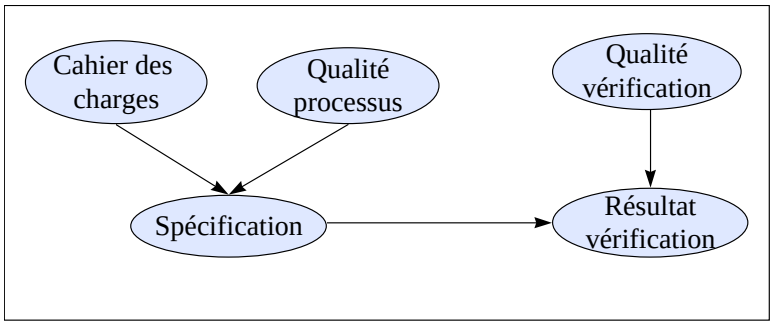

Figure 3 : Construction et vérification de la spécification d'un système

Ce type d'assemblage permet de profiter du caractère local des informations saisies dans un RB. Il permet ainsi la réutilisation de sous-réseaux, appelés "templates" dans l'outil SERENE.

La méthode SERENE apporte une aide efficace dans les trois phases de la construction du modèle.

- Pour l'identification des noeuds : le novice trouvera dans les modèles d'arguments complets décrits dans les annexes du manuel (et fournis avec l'outil) une bonne part des entités que les partenaires du projet considèrent comme importantes dans la prédiction de la sûreté de fonctionnement des systèmes programmés.

- Pour l'identification de la structure : là encore, les modèles complets donnent des exemples de structures (avec tous leurs niveaux de détail, ce qui permet de les découvrir progressivement) ; en outre, pour le niveau le plus fin, la méthode comporte un organigramme aidant à trouver un idiome correspondant au type de relation que l'on veut exprimer entre des variables.

- Pour l'élicitation des tables de probabilités : la méthode contient un répertoire des biais et erreurs fréquemment rencontrés lorsqu'on interroge des experts sur des probabilités, et elle propose des solutions pour les atténuer. Par exemple, on obtient de bien meilleurs résultats en demandant aux personnes interrogées de s'exprimer en termes de fréquences (pourcentages...) qu'en termes de probabilités.

3.2 Exploitation du modèle

Une fois le modèle construit à l'aide de l'éditeur de l'outil SERENE, il ne reste plus qu'à le compiler (une opération qui peut prendre d'une seconde à un temps... déraisonnable suivant la complexité du modèle) pour pouvoir l'exploiter.
Cette exploitation consiste en la saisie d'observations, et le calcul des probabilités révisées (propagation), cette opération pouvant être renouvelée de nombreuses fois.

En effet, pour utiliser un modèle pour évaluer globalement la sûreté de fonctionnement d'un système, il faut faire d'abord des évaluations (au moyen de tests, audits, mesures) des différentes propriétés observables, afin de saisir ces observations dans le RB, et d'obtenir par propagation les probabilités révisées des propriétés inobservables (telles que la sûreté du système).

Plus on peut donner d'informations sur les propriétés observables, mieux c'est. Mais le fait de ne rien pouvoir dire sur certaines n'est pas bloquant : le RB donne de toute façon un résultat : simplement, ce résultat est d'autant plus proche de la valeur a priori qu'on dispose de peu d'information.

Cette capacité des réseaux bayésiens à traiter des informations incomplètes est très précieuse, car elle permet de garder un même modèle dans des situations diversifiées, ainsi que cela a été fait lors de l'expérience décrite au § 5.2.
Ces situations peuvent d'ailleurs être tout simplement les différents stades d'avancement d'un projet. Ainsi, il est possible de voir se dessiner une tendance dès les premières phases : la tendance espérée est une amélioration d'une estimation au départ assez pessimiste. Tout écart par rapport à cette tendance doit inciter un chef de projet à se poser des questions...

La fin de cet article est consacrée à la description d'un modèle construit à EDF selon la méthode SERENE, et à son exploitation au moyen de l'outil SERENE, afin d'illustrer les généralités qui viennent d'être données.

4. MODELE D'EVALUATION D'UN SYSTEME DEVELOPPE POUR EDF PAR UN FOURNISSEUR

4.1 Objet du réseau bayésien

Le RB décrit ci-après est relatif à un système réalisé par un fournisseur extérieur d'après un cahier des charges écrit par EDF. Il présente une vue synthétique de la façon dont EDF évalue les produits et les processus appliqués par elle-même et par ce fournisseur afin d'assurer la sûreté fonctionnelle d'un système de criticité moyenne. Il est axé seulement sur la sûreté fonctionnelle. En particulier, il laisse de côté :

- la fiabilité matérielle (défaillances aléatoires, ou dues à un environnement agressif : haute température, vibrations...),
- les problèmes de facteur humain,
- des problèmes d'ordre systémique global, faisant intervenir d'autres systèmes et la conception d'ensemble de l'installation contenant le système,
- la modification des systèmes existants,
- la façon dont sont évaluées dans le détail les caractéristiques techniques du système (le modèle ne s'intéresse qu'à la **synthèse** des résultats de ces évaluations détaillées).

Ces limitations sont justifiées par la nécessité de limiter la taille et la complexité du RB.

4.2 Méthodologie

Le réseau a été construit grâce à un travail en groupe, faisant intervenir cinq personnes d'EDF impliquées dans des contextes divers dans l'évaluation de systèmes programmés. Nous avons toujours travaillé en réunion, en suivant les étapes données au § 3, mais sans l'aide de la méthode SERENE, car elle n'existait pas encore !
En fait, la méthode a servi, une fois qu'elle a été disponible, à restructurer le modèle et à l'organiser en trois niveaux de détail, le niveau le plus fin étant exclusivement constitué d'idiomes. Cela a permis de rendre le modèle lisible, alors qu'il n'était initialement qu'un "plat de spaghetti", incompréhensible par toute personne étrangère à sa construction.

En diverses occasions, le groupe a pu confronter ses idées à celles des autres partenaires du projet, ce qui a provoqué des débats très intéressants, et quelquefois assez vifs.
Une des premières options qui ont été discutées a été le nombre d'états qu'il conviendrait d'affecter aux variables du RB.

Nous avons décidé que ce nombre devrait être pair, pour forcer les évaluateurs à donner des jugements plutôt positifs ou plutôt négatifs. En choisissant 4 états pour chaque noeud, on aurait eu à remplir des TPNs de grande taille (par exemple, il aurait fallu déterminer 192 valeurs pour un noeud à 3 parents !). Il nous apparaissait impossible de justifier un tel nombre de valeurs, et de toute façon, leur élaboration serait excessivement longue et ennuyeuse.

Nous avons donc choisi de prendre seulement deux états, signifiant bon ou mauvais, acceptable ou inacceptable, oui ou non... De plus, afin de simplifier la description du RB, la plupart des noeuds correspondent à des phrases du type *l'objet x est y*, avec les deux valeurs possibles : Vrai ou Faux.
Exemple : "Le cahier des charges représente correctement le besoin"

Nous avons également décidé, après quelques tentatives pour évaluer les valeurs des probabilités, de les choisir dans l'ensemble : {0, 0.1, 0.25, 0.5, 0.75, 0.9, 1}.
La raison de ce choix est qu'il nous paraissait impossible de justifier un plus grand nombre de niveaux dans cette échelle, qui était juste ce qu'il fallait pour exprimer les concepts suivants : impossible, très improbable, improbable, moyennement probable, probable, très probable, certain.

Un autre avantage de ce choix est que le complément à 1 de toute probabilité prise dans cette échelle est aussi dans l'échelle, ce qui assure une parfaite "symétrie" pour des variables booléennes : s'il est très improbable que A soit 'Vrai', c'est qu'il est très probable qu'il soit 'Faux', etc.

Il est important de remarquer que les valeurs 0 et 1 (qui modélisent une connaissance déterministe) jouent un rôle tout à fait particulier.

Pour illustrer ce point, imaginons deux variables booléennes A et B, telles que A $\Leftrightarrow$ Non B.
Cela peut se modéliser en RB par (par exemple), la structure suivante : A $\rightarrow$ B,
et la TPN suivante pour B (la TPN associée à A importe peu) :

A	Vrai	Faux
P(B=Vrai)	p	1-p

(avec p = 0)

Dans un tel RB, il est impossible (car c'est incohérent) d'imposer la même valeur à A et B. De telles incohérences sont détectées et signalées à l'utilisateur par les outils de calcul.
En revanche, dès que p prend une valeur un tant soit peu différente de 0, on peut imposer la même valeur à A et B ; simplement, une telle situation aura une très faible probabilité.

Il apparaît donc une sorte de discontinuité du comportement du RB pour les valeurs 0 et 1 de probabilités. On peut donc craindre des problèmes de précision numérique, ou simplement de grande sensibilité des résultats aux données lorsqu'on utilise ces valeurs.
C'est pourquoi nous avons réalisé des études de sensibilité autour de ces valeurs quand nous avons testé notre modèle (cf. § 5.1).

Grâce au principe général donné plus haut (chaque noeud correspond à une proposition, qui peut être vraie ou fausse), la description du RB peut se limiter à sa structure, et, pour chaque noeud, la liste des éléments suivants :

- une description détaillée de la signification de la proposition,
- la table de probabilités du noeud,
- une explication des valeurs données dans cette table.

Le RB est organisé en trois niveaux de détail. Le § 4.3, ci-après, donne la description du niveau le plus haut, et quelques indications sur les structures des niveaux inférieurs, qu'il est impossible, faute de place, de détailler ici.

4.3 Modèle du plus haut niveau

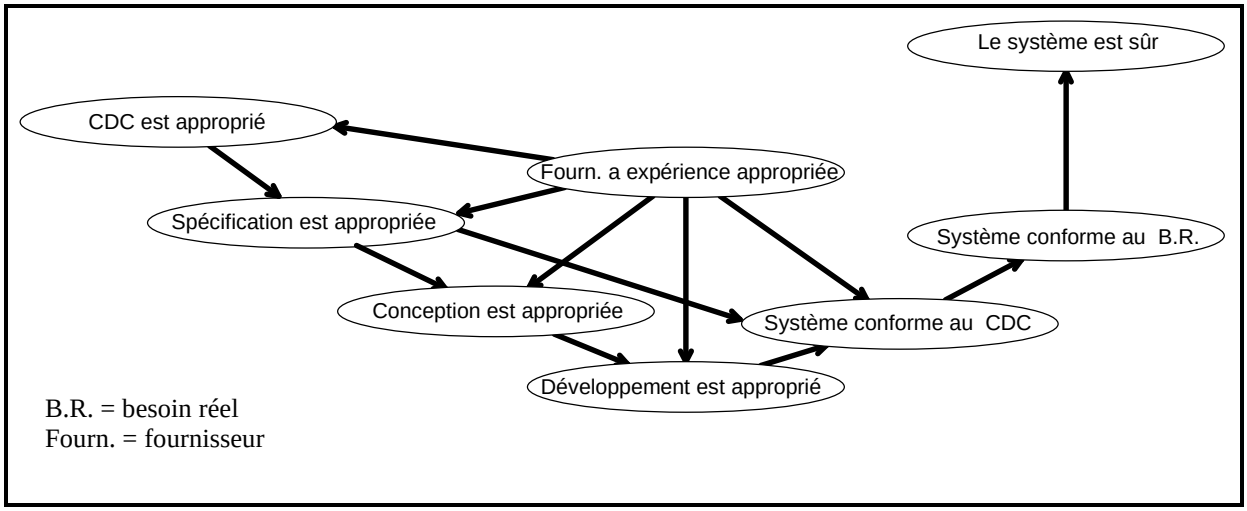

Figure 4: structure globale du RB

Dans la Figure 4, les noeuds correspondent à des sous-réseaux, et les arcs à l'existence de noeuds partagés par deux sous-réseaux, suivant le principe décrit au § 3.1.

La structure globale du modèle est une chaîne qui suit les étapes du processus de développement. Chaque étape sert de point d'entrée pour la suivante, et a une grande influence sur elle.
Chaque étape se décompose en un processus de construction et un processus de vérification (en général, il y a deux vérifications, faites respectivement par EDF et le fournisseur), et les produits qui résultent de cette étape sont influencés par le caractère plus ou moins approprié de ces deux processus.
Rappelons que pour une vérification, "être approprié" signifie que le résultat de la vérification est pertinent : la vérification a été suffisamment exhaustive, et a couvert tous les aspects envisageables.
On voit tout de suite que l'expérience du fournisseur joue un rôle très important, et introduit un certain degré de dépendance entre toutes les étapes du processus de développement. Elle intervient même dans la qualité du processus de vérification du cahier des charges. En effet, un fournisseur ayant une bonne expérience du domaine auquel appartient le système qu'il aura à construire est plus à même de poser des questions pertinentes et de soulever les points délicats dès le cahier des charges, ce qui peut amener à réviser celui-ci.

Le "sous-réseau" final, intitulé 'Le système est sûr', est en fait réduit à un seul noeud : c'est le noeud de "sortie" du modèle, celui dont la probabilité d'être à Vrai donne le jugement global sur le système, prenant en compte toutes les observations saisies dans le modèle.

Les différents sous-réseaux ont des structures et des TPNs assez semblables. Ce sont essentiellement des assemblages d'idiomes processus-produit et d'idiomes mesure. Au niveau de détail le plus fin (le troisième), on trouve les instances des idiomes.

5. UTILISATION DU MODELE

5.1 Scénarios hypothétiques

Nous avons commencé par tester le modèle sur un certain nombre de scénarios hypothétiques, allant du meilleur (toutes les observations saisies sont positives) au pire des cas (elles sont toutes négatives). Les résultats pour les cas extrêmes sont sans surprise : on obtient pour Pr(le système est sûr) des valeurs proches respectivement de 1 et de 0.1. L'essai sur des situations plus nuancées, détaillé dans [1] et [8], donne des résultats cohérents et en accord avec l'intuition.
Nous avons également fait des études de sensibilité au voisinage des valeurs 0 et 1, pour les raisons exposées au § 4.2. Les valeurs 0 et 1 (strictement) donnent des résultats très proches de ceux que l'on a avec des valeurs qui s'en écartent un peu (respectivement 0.01 et 0.99), **sauf dans un cas** : celui où les observations saisies sont **fortement contradictoires** (par exemple, on a un processus de construction optimal, et une vérification qui donne un mauvais résultat). Cela signifie que dans une telle situation, les prévisions faites par le modèle ne sont pas fiables ; **mais il est probable que de telles données feraient aussi hésiter un expert**.
Nous avons constaté le même phénomène en faisant ce type d'études de sensibilité sur les projets réels. Heureusement, seul un cas sur 11 a produit des résultats "instables".

5.2 Comparaison entre une évaluation experte et à l'aide du modèle, sur 11 projets réels

Pour cette expérience, un expert en évaluation des projets informatiques a sélectionné 11 projets des six dernières années, selon les critères suivants :

- le projet a été développé spécialement pour EDF, par un fournisseur extérieur,
- il était soumis à des exigences de sûreté,
- il a été évalué avec une précision suffisante, aussi bien en termes de processus de développement que de produits,
- il existe un minimum de retour d'expérience relatif au projet, donnant des éléments factuels pour estimer sa sûreté fonctionnelle.

Puis, pour chaque projet, il a (dans cet ordre) :

- collecté l'information relative à l'évaluation du projet,
- rempli un questionnaire de 59 questions (avec trois réponses possibles : Vrai, Faux, non réponse), qui correspondent aux noeuds observables du RB.
- rempli un questionnaire d'évaluation experte de 9 questions assez globales relatives aux différentes phases du projet, avec pour chaque question, une évaluation subjective de la probabilité que le noeud correspondant du RB prenne sa valeur positive.

Une fois toutes les données disponibles, l'outil SERENE a été utilisé pour calculer les probabilités des noeuds correspondant aux 9 questions globales, à partir des réponses au questionnaire détaillé. Il a donc été possible de comparer les évaluations faites directement par l'expert à celles qui ont été calculées.

L'accord entre les deux séries de nombres est assez, voire très satisfaisant. En outre, les quelques disparités importantes qui ont été mises en évidence ont pu être expliquées facilement par des insuffisances du modèle qu'il devrait être possible de corriger.

A titre d'exemple, la Figure 5 représente graphiquement la corrélation entre l'évaluation experte (en abscisses) et celle calculée grâce au RB (en ordonnées), pour le noeud 'Le système est sûr'.

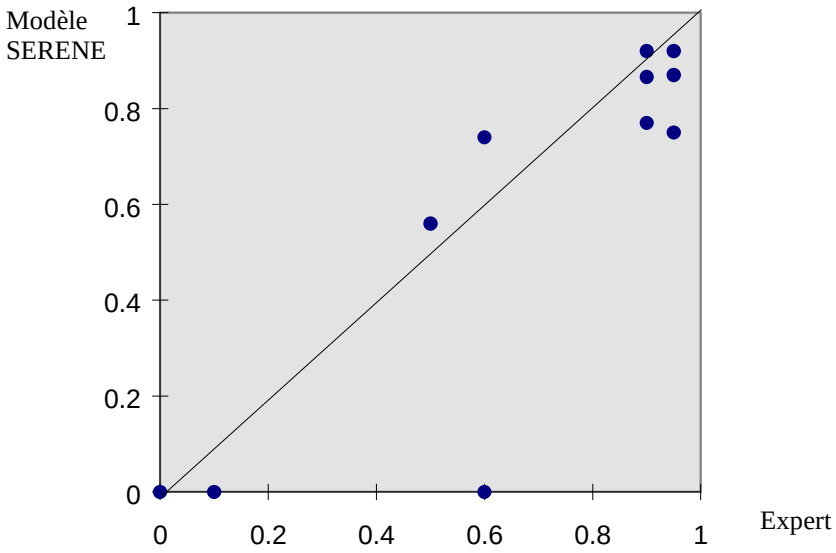

Figure 5 : Evaluation experte/à l'aide du modèle pour 11 projets réels

Les deux points pour lesquels l'expert a donné une évaluation non nulle alors que le modèle a prédit un cas désespéré correspondent à des projets très mal partis, mais qui ont finalement pu être "récupérés" grâce à des actions correctives. Cette dimension temporelle manque dans le modèle, qui est adapté seulement à un développement qui progresse normalement d'étape en étape, sans grande remise en question ni retour en arrière.

6. CONCLUSION

Notre expérience est qu'un réseau bayésien est facile à comprendre, et peut améliorer considérablement la qualité de la communication entre experts : un tel langage a un pouvoir expressif bien plus grand que, par exemple, un ensemble de check-lists, même si elles sont sophistiquées, usant de pondérations et de seuils. Ainsi, les réseaux bayésiens pourraient faciliter grandement les discussions entre les experts qui évaluent les systèmes programmés critiques.

Dans cet article, nous avons montré un exemple de réseau bayésien qui peut être utilisé pour modéliser un argumentaire de sûreté, avec quelques éléments concourant à un début de validation de ce modèle.
Ce réseau bayésien est un prototype ; sa construction a été une bonne occasion de rendre explicite et de capitaliser la connaissance d'un groupe d'experts en évaluation des systèmes programmés critiques.

Ce savoir-faire, au même titre que le savoir-faire versé par les autres partenaires dans le projet, fait maintenant partie intégrante de la méthode SERENE et de l'outil qui la supporte.

Grâce à leur aide, il est désormais possible de construire un modèle en réseau bayésien pour chaque nouveau type d'argumentaire de sûreté, de façon à le rendre plus compréhensible, et ceci en un temps bien inférieur à celui qui nous a été nécessaire.

L'utilisation de la méthode SERENE, si elle se répand, devrait aider à déterminer quelles sont les variables qui ont le plus d'influence dans le processus de développement d'un système, ce qui est un pré-requis indispensable pour mettre en place un quelconque système de retour d'expérience.
Un objectif à plus long terme pour SERENE pourrait être de réaliser des prédictions exactes, mais atteindre ce but prendra, à n'en pas douter, de nombreuses années de calibrage, à faire en comparant le niveau de sûreté prédit par la méthode, et celui observé sur le terrain.

7. REMERCIEMENTS

L'auteur de cet article exprime ici toute sa reconnaissance envers les différentes personnes d'EDF qui ont participé à l'élaboration, puis à l'expérimentation du modèle décrit plus haut. Il s'agit de (par ordre alphabétique) :

F. Ficheux-Vapné, P. Fourcade, F. Martin, Nguyen N. Q. Thuy, A. Ourghanlian.

REFERENCES

[1] P.G. BISHOP and R.E. BLOOMFIELD, "The SHIP Safety Case", *in SafeComp 95 Proceedings*, 1995

[2] CASCADE, ESPRIT project n° 9032, "Generalized Assessment Method (GAM)"
part 1 (rules, doc. ID: CAS/LR/WP2.T3/SM/D2.3.1) and
part 2 (guidelines, doc. ID: CAS/IC/MK/D2.3.2/V3), Jan. 1997.

[3] CACM, "Real world applications of Bayesian Networks", *Communication of the ACM, special issue*, 38 (3), pp. 24-57, 1995.

[4] PEARL J., "Probabilistic reasoning in intelligent systems: Networks of plausible inference", Morgan-Kaufmann, San Mateo, California, 1988

[5] JENSEN F.V., "An introduction to Bayesian Networks", Springer Verlag, New York, 1996

[6] FENTON N., LITTLEWOOD B., NEIL M., STRIGINI L., SUTCLIFFE A. and WRIGHT D., "Assessing dependability of safety critical systems using diverse evidence", *IEEE Proceedings on Software Engineering, Vol. 145, No.1*, Fév. 1998.

[7] M. BOUISSOU, F. MARTIN, A. OURGHANLIAN, "Assessment of a Safety Critical System Including Software : a Bayesian Belief Network for Evidence Sources", *Proceedings of the RAMS'99 Conference, Washington*, Jan. 1999.

[8] adresse internet www.hugin.com, où l'on peut télécharger la méthode SERENE, les modèles élaborés par les partenaires, et une version de démonstration de l'outil.