



République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche
Scientifique
Université d'Oran
Institut de Maintenance et de Sécurité Industrielle

Thèse de Magister en Sciences des Risques et Matériaux

Approche Bayésienne pour la Sûreté de
Fonctionnement de la chaudière N°2 du Complexe
GNL4Z

Présentée par : ZOUAIRI Saim

Soutenue le 17/03/2009 devant le jury composé de :

SEBBANI Mohamed	<i>Président</i>	<i>Professeur</i>	<i>Université d'Oran</i>
BELAIB Lekhmissi	<i>Examineur</i>	<i>Maître de conférence</i>	<i>Université d'Oran</i>
LOUNIS Zoubida	<i>Examineur</i>	<i>Maître de conférence</i>	<i>IMSI/Université d'Oran</i>
HASSINI Noureddine	<i>Directeur de thèse</i>	<i>Maître de conférence</i>	<i>Université d'Oran</i>

Année universitaire 2008/2009

*A ma mère,
A mon père.*

Remerciements

Je remercie mon directeur de thèse Monsieur HASSINI Noureddine pour la confiance et le soutien qu'il m'a accordé ainsi que pour la justesse des conseils qu'il m'a prodigués. Je lui suis reconnaissant de m'avoir fait partager son enthousiasme et son goût pour la recherche.

Je remercie Monsieur SEBBANI Mohamed Professeur à l'Université d'Oran, qui a accepté de présider ma soutenance, je lui dois toute ma gratitude.

Mes remerciements s'adressent aussi à Madame LOUNIS Zoubida, Maître de conférence à l'Institut de Maintenance et de Sécurité Industrielle de l'université d'Oran et Monsieur BELAIB Lekhmissi, Maître de conférence à l'université d'Oran, et Monsieur BOIURAHLA Ahmed, Maître de conférence à l'Université de Mostaganem, d'avoir accepté d'être des examinateurs de ce travail.

Je tiens à remercier aussi, Mademoiselle Khadidja GUENACHI, de nous avoir offert la possibilité de faire partie de cette post graduation et de nous avoir initié avec son équipe pédagogique à la Science du Danger.

Un remerciement particulier à Monsieur BAHAR Ali, pour ses conseils et de m'avoir consacré du temps pour les explications du mode de fonctionnement de la chaudière.

Je remercie aussi, Monsieur FERHAT, Monsieur ADDA BOUDJELLAL, Monsieur KHATIR, Monsieur SAADI, Monsieur BOUAMRANE, Monsieur BELKHATIR, Monsieur HADJEL pour tous leurs efforts qu'ils ont mis et leurs sérieux pour nous former.

Mes vifs remerciements s'adressent aussi à Monsieur TAHRAOUI Mohamed, Chef de département de Sécurité industrielle et Environnement pour son aide durant mon travail

Je remercie la Direction AVAL de Sonatrach pour son parrainage de cette formation et pour les stages effectués par toute l'équipe de la post graduation durant notre formation et dans les différentes unités d'AVAL.

Ce travail a été réalisé avec l'aide de la Direction de Sonatrach, particulièrement le complexe GL4Z à ARZEW, où j'ai collecté les informations nécessaires à la réalisation de cette thèse. J'exprime ma gratitude à Monsieur TEBBAL Abdelkrim pour son soutien et son aide précieux, aussi je remercie Monsieur ZORKANI Mourad pour son accompagnement et ses explications du processus de fonctionnement de la chaudière, Monsieur KHEMLICHE

pour son sens rigoureux et ses explications des pannes fréquentes de la chaudière, Madame la DRH qui nous a permis et faciliter la tâche pour l'accès au complexe. Sans oublier tout le personnel en général du complexe qui nous a orienté durant toute la période du stage.

Je remercie mes collègues thésards : N. NOUREDDINE, S. ASSAF, S. AMEUR, F. BOUKEZZI , M. BOUANANI pour leur soutien moral, et pour l'ambiance agréable qu'ils ont créé.

Merci pour tous ce qui m'ont aidé de proches ou de loin !

Enfin, je remercie tous ceux qui ont contribué à ma formation.

Merci à Dieu qui m'a donné la force d'aller jusqu'au bout de cette thèse

Table des matières

Remerciements	i
Table des figures	v
Liste des tableaux	viii
Liste des abréviations	ix
Introduction	1
Problématique	2
1 Sûreté de Fonctionnement	5
1.1 Evolution de la discipline	5
1.2 Les principales composantes de la sûreté de fonctionnement	6
1.2.1 La fiabilité	6
1.2.2 La maintenabilité	7
1.2.3 La disponibilité	7
1.2.4 La sécurité	7
1.3 Le but de la sûreté de fonctionnement	8
1.3.1 Quelques indicateurs	8
1.4 Les études de sûreté de fonctionnement	9
1.4.1 Les étape pour une analyse SdF	10
1.5 Définitions et concepts de base	10
1.5.1 Défaillances, missions et fonctions d'un système et de ses composants	10
1.5.2 Missions et fonctions	11
1.5.3 Description des procédés industriels	13
1.6 Lois de probabilité rencontrées dans les études de fiabilité	15
1.6.1 Loi binomiale de paramètres p et n	15
1.6.2 Loi de Poisson de paramètre λ	16
1.6.3 Loi exponentielle de paramètre λ	16
1.6.4 Loi log -normale ou loi de Galton	16
1.6.5 Loi de Weibull	17
1.6.6 Loi normale	17
1.6.7 Loi uniforme	17
1.7 Concepts de base et sûreté de fonctionnement	18

1.7.1	Fiabilité (Reliability)	18
1.7.2	Disponibilité (Availability)	20
1.7.3	Maintenabilité (Maintainability)	21
1.7.4	Sécurité	25
1.8	Démarches et méthodes fondamentales d'une approche SdF	27
1.8.1	Analyse préliminaire de risques (APR)	28
1.8.2	Analyse des modes de défaillance, de leurs effets et de leurs criticités (AMDEC)	29
1.8.3	Arbres de causes, d'événement, de défaillances	30
1.8.4	La méthode MADS-MOSAR	32
1.8.5	Les Réseaux Bayésiens	33
2	Les Réseaux Bayésiens	35
2.1	Rappel de Probabilités	35
2.1.1	Les Axiomes de probabilités	36
2.1.2	La probabilité jointe et les événements indépendants	38
2.1.3	Distinction entre événements indépendants et événements mutuellement exclusifs	39
2.1.4	Probabilité marginale	39
2.1.5	La probabilité conditionnelle	39
2.1.6	Probabilité conditionnelle pour les événements indépendants	40
2.1.7	La règle de Multiplication	41
2.1.8	Le théorème de Bayes	41
2.2	Les Modèles graphiques	41
2.2.1	Les graphes	42
2.2.2	Notion de parents et d'enfants	42
2.2.3	Graphe Acyclique Dirigé	42
2.2.4	Graphe Moral	43
2.2.5	Graphe Triangulé	43
2.3	Réseaux Bayésiens	44
2.3.1	Indépendance Conditionnelle	44
2.3.2	Notations et définitions	48
2.4	Le critère de d- séparation	51
3	Inférence dans les Réseaux Bayésiens	55
3.1	Spécification d'un réseau	55
3.1.1	Exemple	55
3.1.2	Étape qualitative	56
3.1.3	Étape probabiliste	57
3.1.4	Étape quantitative	57
3.2	Inférence	57
3.2.1	Les principaux algorithmes	58
3.2.2	Approche générale de l'inférence	59
3.3	Algorithme de l'arbre de jonction dit JLO	62
3.3.1	Moralisation	62

3.3.2	Triangulation	63
4	Application pour la chaudière	65
4.1	Identification de la chaudière	65
4.1.1	Chaudière à tube d'eau	66
4.2	Retour d'expérience	68
4.3	Application de la méthode MADS-MOSAR à la chaudière	69
4.3.1	Module A	70
4.3.2	Décomposition du système en sous systèmes	70
4.4	Etude Statistique	75
4.4.1	L'analyse de la variance	75
4.4.2	Analyse en Composante principale	78
4.4.3	Construction du réseau Bayésien	79
4.4.4	Discussions	84
4.4.5	Conclusion	86
	Conclusions et perspectives	86
	A Réglementation	89
	Bibliographie	89

Table des figures

1.1	Représentation des MTTF, MDT et MUT.	9
1.2	Les différentes étapes pour l'étude de sûreté de fonctionnement.	9
1.3	Cas de figure conduisant tous à la défaillance.	11
1.4	Allure de la fonction $R(t)$ pour la loi exponentielle.	19
1.5	Allure d'un taux de défaillance " en baignoire ".	20
1.6	Disponibilité en fonction du temps t	21
1.7	Représentation des MTBF, MDT et MUT.	21
1.8	Chronologie des temps des activités de maintenance.	22
1.9	Allure de la courbe de maintenabilité.	23
1.10	Relations entre les liens temporels en fiabilité, disponibilité et maintenabilité.	24
1.11	Chânage temporel des activités de détection et de remise en service.	25
1.12	Relations entre fiabilité, maintenabilité, disponibilité et sécurité.	27
1.13	Extrait d'un tableau d'AMDEC.	30
1.14	Un arbre de causes pour le scénario " panne de batterie ".	31
1.15	Un arbre d'événement pour le scénario " évitement d'un obstacle sur l'autoroute ".	32
2.1	l'événement F est un sous ensemble de l'événement E	36
2.2	E et F sont des événements en intersection.	37
2.3	les événements E et F sont mutuellement exclusifs ou disjoints.	37
2.4	Deux événements A et B dans l'univers U	38
2.5	l'univers réduit, sachant que l'événement A est arrivé.	40
2.6	Exemple d'un graphe dirigé illustrant les relations parents - enfants.	43
2.7	Exemple de graphe triangulé. La corde $\{a, b\}$ change le cycle $\{c, a, d, b\}$ en deux cycles de longueur 3.	44
2.8	un exemple d'un réseau bayésien.	45
2.9	Représentation graphique de l'équation 2.8 avec des Réseaux Bayésiens.	45
2.10	Trois RBs élémentaires.	47
2.11	Exemple de chaîne divergente en un sommet X	48
2.12	Exemple de chaîne en série en un sommet X	49
2.13	Exemple de chaîne convergente en un sommet X	49
2.14	Exemple de chaîne bloquée par un ensemble.	50
2.15	Autre exemple de chaîne bloquée par un ensemble.	50
2.16	Exemple d'ensembles d - séparés.	51

2.17	Un réseau bayésien représentant les dépendances entre cinq variables. . . .	52
3.1	Représentation du problème d'asphyxie du nouveau-né.	56
3.2	Graphe de la figure 2.17 transformé en arbre en instanciant X_1	59
3.3	Un réseau bayésien simple.	61
3.4	Graphe moralisé. Les arcs en pointillés ont été rajoutés au cours de la moralisation.	63
4.1	Schéma général retenu par le complexe GL4Z.	66
4.2	Vue d'ensemble intérieur chaudière centrale.	67
4.3	Condition de fonctionnement de la chaudière.	67
4.4	Statistique des pannes des chaudières dans les années antérieures [10] . . .	68
4.5	Statistique des accidents, blessures et morts, entre les années 1996 et 2002 [10]	68
4.6	Schéma de processus de danger [19]	70
4.7	Niveau macroscopique de MADS-MOSAR, décrit par le module A.	70
4.8	découpage du système en sous système.	71
4.9	Valeurs propres en rapport avec les composantes principales.	79
4.10	Correspondance entre appréciations qualitatives et probabilités.	80
4.11	RB correspondant à l'état de fonctionnement de la chaudière.	81
4.12	simplification du RB par changement de variable de la figure 4.11.	83
4.13	Principe de l'évaluation des risques adopté par GL4Z.	85

Liste des tableaux

1.1	Principales démarches et méthodes de base de SdF.	28
4.1	Module A pour le SS1.	72
4.2	Module A pour le SS2.	72
4.3	Module A pour le SS3.	73
4.4	Module A pour le SS4.	73
4.5	Module A pour le SS5.	74
4.6	Module A pour le SS6.	74
4.7	Module A pour le SS7.	74
4.8	Legendre des pannes par composant.	76
4.9	données statistiques des pannes de la chaudière.	77
4.10	représentation du tableau 4.9 en procédant à un changement de variables.	77
4.11	Resultat de l'ANOVA.	78
4.12	Probabilités marginales des variables Vd,Tf,Id.	81
4.13	Probabilité de L sachant Vd.	82
4.14	Probabilité de Pcc sachant Tf.	82
4.15	Probabilité de Dysf sachant L et Pcc.	82
4.16	Probabilité de Expl sachant Id et Dysf.	83

Liste des abréviations

Pour des raisons de lisibilité, la signification d'une abréviation ou d'un acronyme n'est souvent rappelée qu'à sa première apparition dans le texte d'un chapitre. Par ailleurs, puisque nous utilisons toujours l'abréviation la plus usuelle, il est fréquent que ce soit le terme anglais qui soit employé, auquel cas nous présentons une traduction.

RB	<i>Réseau Bayésien.</i>
SdF	<i>Sûreté de Fonctionnement.</i>
MTTF	<i>Mean Time To first Failure.</i>
MDT	<i>Mean Down Time.</i>
MUT	<i>Mean Up Time.</i>
TMD	<i>temps moyen de disponibilité.</i>
TMI	<i>temps moyen d'indisponibilité.</i>
MTBF	<i>Mean Time Between Failure.</i>
MTTR	<i>Mean Time To Repair.</i>
APD	<i>Analyse Préliminaire de Danger.</i>
AMDE	<i>Analyse des modes de défaillance, de leurs effets.</i>
AMDEC	<i>Analyse des modes de défaillance, de leurs effets et de leurs criticités.</i>
GAD	<i>Graphes Acycliques Dirigés.</i>
DPJ	<i>Distribution de Probabilité Jointe.</i>
DP	<i>Distribution de Probabilité.</i>
...	

Introduction

Les techniques de modélisation et d'évaluation des systèmes industriels ont suscité une attention croissante durant les vingt dernières années.

De nombreuses activités dépendent du bon fonctionnement de ces systèmes et génèrent le besoin de concevoir des machines de plus en plus performantes et fiables. En particulier cela concerne les systèmes dits critiques, à hauts risques humains ou économiques (surveillance de centrales nucléaires, systèmes de commande de navigation spatiale ou aérienne, réseaux de communication, etc.).

Divers concepts ont été développés pour analyser ces systèmes et sont largement utilisés. En particulier *la sûreté de fonctionnement d'un système est la propriété qui permet à ses utilisateurs de placer une confiance justifiée dans le service qu'il leur délivre*. Les principales dimensions de la sûreté de fonctionnement sont la fiabilité, la disponibilité, la maintenabilité, et la sécurité. Pour quantifier ces dimensions, les systèmes sont en général approchés par des modèles probabilistes.

Les Réseaux bayésiens permettent d'analyser de grandes quantités de données pour en extraire des connaissances utiles à la prise de décision, contrôler ou prévoir le comportement d'un système, diagnostiquer les causes d'un phénomène, etc.

Les réseaux bayésiens sont utilisés dans de nombreux domaines : santé (diagnostic, localisation de gènes)[22], industriels (contrôle d'automates ou de robots)[25, 7, 8], marketing (data mining, gestion de la relation client)[2], banque et finances (analyse financière), management (aide à la décision, gestion de risques)[27, 11, 18, 23], etc., .

Dans le domaine industriel de la sûreté de fonctionnement, les incidents et accidents sont des événements rares mais qu'il faut savoir gérer. Le management des risques nécessite le développement d'analyses statistiques et probabilistes qui utilisent toute l'information pertinente disponible pour pallier les limites d'un retour d'expérience par nature peu important, et donc insuffisant pour une analyse statistique fréquentielle classique.

En effet, la perspective bayésienne offre des avantages nombreux par rapport à la perspective fréquentiste ; l'"objectivité" dans les statistiques fréquentistes a été obtenue en ignorant toute connaissance à priori du processus à mesurer, mais dans le domaine de la science, il y a toujours quelques connaissances à priori du processus à mesurer. Le rejet de ces données à priori à part est un gaspillage de connaissances (qui peut se convertir en coût financier).

Les statistiques bayésiennes utilisent deux sources d'informations, l'information à priori qu'on a à propos du système et l'information contenue dans les données. Elles sont combinées en utilisant le théorème de Bayes.

L'approche bayésienne est aujourd'hui perçue comme un cadre très adapté pour les applications devant gérer un certain niveau d'incertitude. En effet, les probabilités des réseaux bayésiens traduisent des degrés de confiance dans la vérité des propositions.

Notre travail se veut une contribution basée sur une modélisation par les réseaux bayésiens pour apporter une aide à la décision dans le domaine du diagnostic, de la simulation et de l'optimisation de maintenance.

Le document issu de nos travaux sera axé autour de quatre parties :

Partie 1 : ou sera évoqué la notion de Sûreté de Fonctionnement et les différents concepts qui lui sont liés,

Partie 2 : ou sera abordé les graphes, les Réseaux bayésiens et les notions de probabilité,

Partie 3 : sera focalisée sur l'inférences et l'apprentissage des réseaux bayésiens,

Partie 4 : elle sera consacré à l'identification des risques liés à la chaudière et l'application de l'inférence pour prévoir et minimiser à un seuil acceptable les risques liés à son fonctionnement.

L'élément objet de notre étude sera une chaudière.

La chaudière est un système composé de plusieurs entités, la défaillance de l'une d'entre elles peut engendrer des risques potentiels à gravité importante. Par conséquent il est nécessaire de bien définir une politique de maintenance, afin de garantir le maximum de sécurité lors de son fonctionnement. Ceci peut être obtenu par la modélisation de toute cause de dysfonctionnement tout au long de son service.

Problématique

Les incidents, accidents et événements catastrophiques liés aux activités industriels font presque quotidiennement les grands titres des médias.

En effet, le zéro défaut ou le risque zéro n'existe pas pour les activités industrielles à cause de l'occurrence de défaillances humaines ou matérielles.

La chaudière est un composant important dans l'industrie pétrochimique. Elle joue un rôle important dans les procédés nécessitant de l'eau chaude ou de la vapeur surchauffée. Par conséquent comme tout autre matériel, elle est assujetti à des dysfonctionnement de degré de gravité variable, générant ainsi des risques pour l'opérateur (dommage physiques et psychique), et sur l'installation elle-même (surtout pour les moyens de stockage de gaz), pour la population avoisinante et l'environnement. Les accidents, liés à ce genre d'installation sont, à travers le monde, très significatifs.

La prise en compte de l'incertitude est un enjeu important dans le domaine de l'intelligence artificielle puisqu'il faut être capable de déterminer quelles sont les causes ou les conséquences les plus vraisemblables à partir d'un état particulier du système pour en déduire les meilleures actions possibles. Or, pour cela, l'usage de règles du type "*si ... alors ...*" est trop rigide et doit être modulé. De nombreux formalismes ont alors été créés dans cet objectif. Citons par exemple la théorie de l'évidence de Dempster- Shafer [20]. Pour parer à ces problèmes, un ensemble de méthodes à caractère scientifique ont émergés l'aube du troisième millénaire, englobés dans la discipline de la "*Sûreté de Fonctionnement*", qui consiste à évaluer, prévoir, mesurer et maîtriser les défaillances des systèmes technologiques.

Probabilités et modèles graphiques : Nos travaux ne se concentrent pas exclusivement sur le formalisme des probabilités et sur les méthodes statistiques. En effet, pour pouvoir interpréter simplement les résultats formels et pour les expliquer aux non spécialistes, il faut pouvoir présenter ces résultats sous forme synthétique.

Les techniques pour définir ce type de résultats sont souvent graphiques (schémas, diagrammes en camembert, etc.). Les modèles que nous allons introduire ont cet avantage d'être à la fois probabilistes et graphiques. La partie graphique de ces modèles permettra d'identifier simplement quels ensembles de variables sont dépendants ou conditionnellement indépendants. Ces modèles graphiques possèdent par ailleurs un grand pouvoir expressif et permettent d'effectuer des raisonnements probabilistes efficaces.

Les systèmes experts peuvent être utilisés pour de telles applications, puisque qu'ils sont

capables de raisonner sur des faits et des règles incertains. Dans le cadre des systèmes experts, les réseaux bayésiens constituent une approche possible pour intégrer l'incertitude dans le raisonnement. D'autres méthodes existent, mais les réseaux bayésiens présentent l'avantage d'être une approche quantitative.

Notre travail consiste à utiliser les réseaux bayésiens pour une prédiction des événements non souhaités, en se basant sur l'observation, jugement d'experts et le retour d'expérience. Un état des lieux devra nous permettre de mettre en évidence les avantages liés à l'utilisation de l'approche bayésienne dans l'analyse de la sûreté de fonctionnement de la chaudière, en tant qu'objet macroscopique et en tant qu'ensemble d'éléments microscopique, par rapport aux autres méthodes d'évaluation.

Une étude expérimentale basée sur des statistiques réelles, nous permettra d'aborder et de mettre en place notre modèle bayésien.

Dans un premier temps nous utiliserons des modèles existants, donc très généraux. Dans une seconde étape, nous essayerons de proposer un modèle de prévision d'occurrence des risques, basé sur l'inférence statistique et le retour d'expérience liée à l'utilisation de la chaudière.

Justification du choix : Le complexe GL4Z dispose de cinq (05) chaudières dont quatre (04) ont été installées lors de la construction du complexe, la cinquième chaudière a été installée en 1983 pour parer à un besoin de vapeur, à cause de l'apparition des premiers signes de fatigue (perçement des tubes) des anciennes chaudières.

C'est ainsi que durant les années 1991-1992 des opérations de retubage ont été lancées sur les quatre premières chaudières.

La chaudière numéro 2, présentait beaucoup de pannes aux niveau des tubes, ainsi on dispose de statistiques suffisantes pour son arrêt, nous permettant de bien envisager le réseau bayésien correspondant.

Chapitre 1

Sûreté de Fonctionnement

Introduction

Le zéro défaut ou le risque zéro n'existe malheureusement pas pour les activités industrielles à cause de l'occurrence de défaillances humaines ou matérielles.

Toutefois, pour tenter de réduire les risques à un niveau le plus faible possible et acceptable, des méthodes, des techniques et des outils scientifiques ont été développés dès le début du 20ème siècle pour évaluer les risques potentiels, prévoir l'occurrence des défaillances et tenter de minimiser les conséquences des situations catastrophiques lorsqu'elles se produisent.

La discipline de la *Sûreté de fonctionnement* consiste à connaître, évaluer, prévoir, mesurer et maîtriser les défaillances des systèmes technologiques et les défaillances humaines.

1.1 Evolution de la discipline

La sûreté de fonctionnement est appelée la science des " défaillances " [24]. D'autres désignations existent suivant les domaines d'applications : analyse de risque (milieu pétrolier), FMDS (Fiabilité, Maintenabilité, Disponibilité, Sécurité), en anglais RAMS (Reliability, Availability, Maintainability and Safety)[28]. Elle se caractérise à la fois par les études structurelles statiques et dynamiques des systèmes, du point de vue prévisionnel mais aussi opérationnel et expérimental (essais, accidents), en tenant compte des aspects probabilistiques et des conséquences induites par les défaillances techniques et humaines. Cette discipline intervient non seulement au niveau de systèmes déjà construits mais aussi au niveau conceptuel pour la réalisation des systèmes [28].

Introduite en 1962 pour traduire le terme anglais reliability, la fiabilité est la probabilité de non -défaillance d'un équipement sur un intervalle de temps donné (du latin *fidare* :

faire confiance, *fidus* : fidèle et du latin médiéval *fiabete* ce qui est digne de confiance). La disponibilité se définit par la probabilité d'être en état d'accomplir sa fonction à un instant donné. Anglicisme introduit vers 1965, la maintenabilité est l'aptitude d'un système à être maintenu en état. Elle correspond à la probabilité que la remise en état d'une entité en panne soit effectuée dans un intervalle de temps qui soit le plus réduit possible. Les mots sûreté et sécurité ont en fait la même racine étymologique (latin *securus* : sûr). La sécurité, implique actuellement les aspects réglementaires de la sécurité des personnes [28]. C'est ainsi que les " comités hygiène et sécurité (CHS) " ont vu leur apparition dans les établissements commerciaux et industriels. Le terme sûreté est plutôt utilisé par les techniciens pour la conception ou l'exploitation de biens et de services pour qualifier la fiabilité et la disponibilité du fonctionnement des installations.

La sûreté de fonctionnement s'est développée principalement au cours du 20ème siècle pour être actuellement un domaine incontournable pour les industries à risques mais aussi, de plus en plus, pour toute l'industrie, en raison de sa corrélation avec la notion de qualité, les problèmes ergonomiques (relation homme -machine) et l'impact sur l'environnement. Jusqu'avant la Seconde Guerre mondiale, la sûreté de fonctionnement est demeurée plus un art empirique qu'une science exacte [24].

1.2 Les principales composantes de la sûreté de fonctionnement

Le terme "sûreté de fonctionnement", inventé voici trente ans pour englober plusieurs concepts, n'a pas d'équivalent exact en langue anglaise. En France, la sûreté de fonctionnement regroupe quatre composantes.

1.2.1 La fiabilité

Aptitude d'un système à rester constamment opérationnel pendant une durée donnée, ou aptitude d'une entité à accomplir une fonction requise, dans des conditions données, pendant une durée donnée [24]. Elle est mesurée par la probabilité qu'une entité E accomplisse une fonction requise, dans les conditions données, pendant l'intervalle de temps $[0, t]$, elle est notée :

$$R(t) = P[E \text{ non défaillante sur } [0, t]]$$

L'aptitude contraire sera dénommée " défiabilité " : sa mesure est notée $\bar{R}(t)$.

$$\bar{R}(t) = 1 - R(t)$$

1.2.2 La maintenabilité

C'est l'aptitude d'une entité à être maintenue ou rétablie dans un état dans lequel elle peut accomplir une fonction requise, lorsque la maintenance est accomplie dans des conditions données avec des procédures et des moyens prescrits [24]. C'est aussi l'aptitude d'un système à être remis rapidement dans un état opérationnel. Ainsi les systèmes dont les composants sont très facilement démontables peuvent bénéficier d'une meilleure maintenabilité que les autres.

Elle est mesurée par la probabilité que la maintenance d'une entité E accomplie dans des conditions données, avec des procédures et des moyens prescrits, soit achevée au temps t, sachant que l'entité est défaillante au temps t=0 :

$M(t) = P$ [la maintenance de E est achevée au temps t] ;

On écrit, pour simplifier : $M(t) = P$ [E est réparée sur $[0, t]$] ;

L'aptitude contraire sera dénommée " immaintenabilité " : sa mesure est notée $\overline{M}(t)$.

$$\overline{M}(t) = 1 - M(t)$$

1.2.3 La disponibilité

C'est l'aptitude d'une entité à être en état d'accomplir une fonction requise dans des conditions données et à un instant donné [24], c'est aussi l'aptitude d'un système à être opérationnel au moment où il est sollicité. C'est une notion importante pour un appareil de sécurité tel qu'un disjoncteur par exemple. Une disponibilité importante est compatible avec une fiabilité faible, pour peu que l'appareil puisse être réparé très rapidement.

Elle est mesurée par la probabilité qu'une entité E soit en état d'accomplir une fonction requise dans des conditions données et à un instant donnée, elle est notée $A(t)$ (A pour Availability) ;

$A(t) = P$ [E non défaillante à l'instant t] ;

L'aptitude contraire sera dénommée " indisponibilité " : sa mesure est notée $\overline{A}(t)$;

$$\overline{A}(t) = 1 - A(t)$$

1.2.4 La sécurité

C'est l'aptitude d'un système à ne pas connaître de pannes considérées comme catastrophiques pendant une durée donnée, ou l'aptitude d'une entité à éviter de faire apparaître, dans des conditions données, des événements critiques ou catastrophiques [24]. Elle est mesurée par la probabilité qu'une entité E évite de faire apparaître, dans des conditions données, des événements critiques ou catastrophiques. L'aptitude contraire sera dénommée " insécurité ".

On trouvera aussi l'acronyme FMDS pour désigner la sûreté de fonctionnement (comme

fiabilité, maintenabilité, disponibilité et sécurité).

Les Anglo-Saxons utilisent le terme *dependability*, qui recouvre la fiabilité (*reliability*), la disponibilité (*availability*) et la maintenabilité (*maintainability*). La sécurité est traitée à part.

Abusivement, on assimile le mot "*dependability*" à "sûreté de fonctionnement". On préférera le terme anglais de RAMS (pour *Reliability, Availability, Maintainability and Safety*) pour votre part.

1.3 Le but de la sûreté de fonctionnement

La sûreté de fonctionnement est une notion générique qui mesure la qualité de service délivré par un système, de manière à ce que l'utilisateur ait en lui une confiance justifiée [14].

Cette confiance justifiée s'obtient à travers une analyse qualitative et quantitative des différentes propriétés du service délivré par le système, mesurée par les grandeurs probabilistes associées : fiabilité, maintenabilité, disponibilité, sécurité.

1.3.1 Quelques indicateurs

Certains indicateurs vont caractériser le fonctionnement prévu du système, tels que le MTTF, le MDT et le MUT (voir figure 1.1).

- Le MTTF (Mean Time To first Failure) est l'estimation de la durée moyenne s'écoulant entre la mise en service du système et la survenance de la première panne
- Le MDT (Mean Down Time) est le temps moyen séparant la survenance d'une panne et la remise en état opérationnel du système. Il se décompose en plusieurs phases :
 - durée de détection de la panne ;
 - durée de diagnostic de la panne ;
 - durée d'intervention jusqu'au début de la réparation ;
 - durée de la réparation ;
 - durée de remise en service du système.
- Le MUT (Mean Up Time) est le temps moyen qui sépare une remise en service opérationnelle du système de la survenance de la panne suivante.

Ces deux derniers indicateurs ne sont pertinents que dans le cas de systèmes réparables. Leur somme $MUT+MDT$ représente le temps moyen qui sépare deux pannes consécutives du système. On le note MTBF (Mean Time Between Failures).

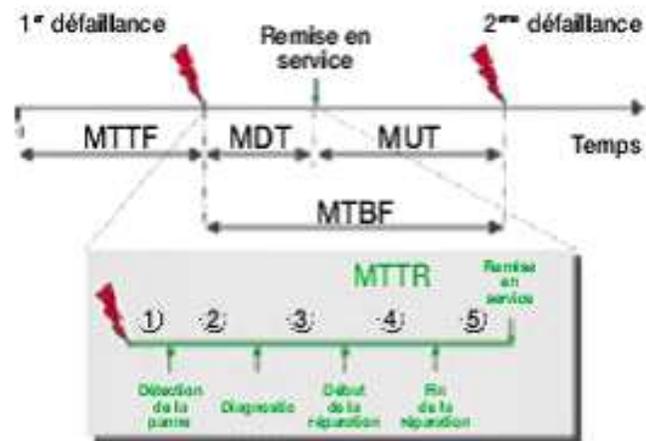


FIG. 1.1 – Représentation des MTTF, MDT et MUT.

1.4 Les études de sûreté de fonctionnement

Elles constituent un préalable indispensable à la conception d'un système voulu sûr, et permet une aide à la décision en :

- comprenant et identifiant les risques ;
- optimisant l'architecture et comparant des solutions différentes ;
- optimisant les moyens de soutien en comparant des solutions ;
- justifiant les choix de façon rationnelle et démontrée ;
- vérifiant la bonne atteinte des objectifs de sûreté de fonctionnement.

Elle peuvent aussi aider à l'optimisation en :

- diminuant le nombre de pannes qui seront observées durant la vie du système ;
- optimisant économiquement la conception par le dimensionnement des équipements et des architectures au "juste nécessaire" ;
- rendant la maintenance plus ciblée et plus efficace ;
- dimensionnant au plus juste les moyens de soutien nécessaires (stocks de pièces de rechange).

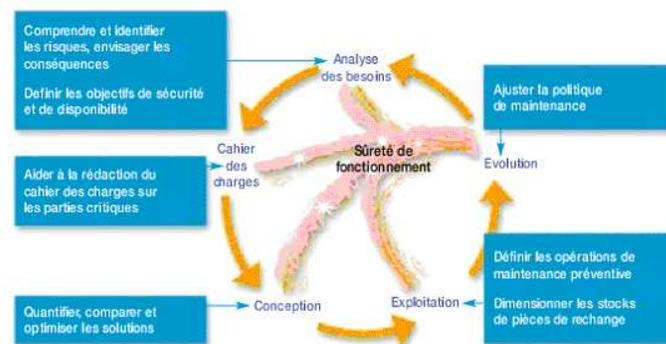


FIG. 1.2 – Les différentes étapes pour l'étude de sûreté de fonctionnement.

1.4.1 Les étapes pour une analyse SdF

La première étape consiste à analyser rigoureusement le besoin pour comprendre et identifier l'ensemble des risques, et envisager leurs conséquences. Ensuite, des niveaux d'acceptabilité sont attribués pour ces risques (on parle d'objectifs de F, M, D et/ou S selon les systèmes).

L'identification précise de ces risques va aider à la rédaction du cahier des charges du système, précisément sur ses parties critiques. Il faudra alors imaginer des solutions techniques, des architectures adaptées qui, toutes, seront quantifiées d'un point de vue sûreté de fonctionnement, comparées entre elles et, si nécessaire, optimisées. Une fois la solution retenue, il sera nécessaire de préciser les conditions d'une exploitation la plus efficace possible en [5] :

- définissant les opérations de maintenance préventive nécessaires pour maintenir les caractéristiques de sûreté de fonctionnement au niveau voulu, sans dégradation des équipements préjudiciable à l'une des quatre composantes ;
- dimensionnant les stocks de pièces de rechange au plus juste, sans dégrader la disponibilité du système (figure 1.2).

1.5 Définitions et concepts de base

1.5.1 Défaillances, missions et fonctions d'un système et de ses composants

Définition de la défaillance fonctionnelle

C'est la cessation de l'aptitude d'un dispositif à accomplir une fonction requise. Une défaillance est "*l'altération ou la cessation de l'aptitude d'un ensemble à accomplir sa ou ses fonction(s) requise(s) avec les performances définies dans les spécifications techniques*". L'ensemble est indisponible suite à la défaillance. La cessation de l'aptitude conduit l'entité à être dans un état appelé "panne" [28].

Un ensemble est défaillant si ses capacités fonctionnelles sont interrompues (panne ou arrêt volontaire par action d'un système interne de protection ou une procédure manuelle équivalente).

Dans le cas d'une dégradation sans perte totale de la fonction, on considère qu'il s'agit d'une défaillance si sa performance tombe au-dessous d'un seuil défini, lorsqu'un tel seuil minimal est contenu dans les spécifications fonctionnelles du matériel.

Il s'ensuit qu'un ensemble est défaillant s'il est considéré ou déclaré incapable d'assurer les fonctions requises par l'exploitant utilisant des critères fonctionnels simples [28].

Toute étude de fiabilité implique l'acceptation de deux états totalement exclusifs : le fonctionnement normal et le fonctionnement défaillant. Les passages d'un état de fonc-

tionnement normal à un état défaillant pouvant se manifester en fonction du temps de manière progressive, soudaine ou de façon aléatoire, la fiabilité ne connaît pas la notion de défaillance partielle ou progressive. La figure 1.3 représente trois cas conduisant tous à une défaillance.

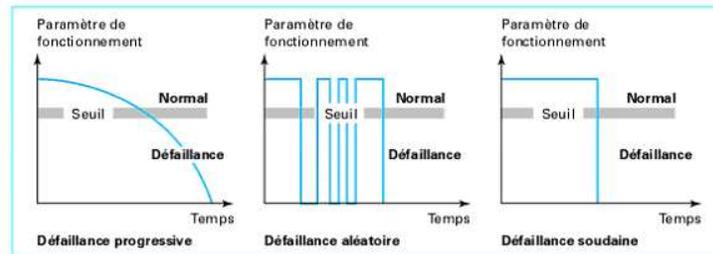


FIG. 1.3 – Cas de figure conduisant tous à la défaillance.

1.5.2 Missions et fonctions

Une étude de Sûreté de fonctionnement passe nécessairement par une analyse exhaustive des différentes phases de fonctionnement faisant appel à des termes précis tels que missions et fonctions assurées par un bien.

Il est particulièrement indispensable de définir les notions de fonction et de mission. Ces définitions sont, en effet, très variables d'un secteur d'activité à un autre ou d'un pays à un autre. Les missions et fonctions d'une entité découlent directement de la définition et de la spécification de ses exigences opérationnelles [24] :

- quelles sont les exigences opérationnelles et comment l'entité sera-t-elle utilisée ?
- où l'entité sera-t-elle utilisée ?
- combien de temps l'entité sera-t-elle utilisée ?

Missions

Le profil de la mission d'une entité se décompose en plusieurs phases distinctes dans lesquelles il est indispensable pour chacune d'entre elles de disposer d'un ensemble de fonctions bien définies.

Fonctions

La compréhension de la notion de fonction et de ses paramètres est l'élément clé sur lequel reposent les analyses de Sûreté de fonctionnement car on recherche les éléments matériels qui vont entraîner la perte ou la dégradation des fonctions.

L'AFNOR définit une fonction comme " *l'action d'une entité ou de l'un de ses constituants exprimée en termes de finalité* ".

Cette définition de nature qualitative est indispensable pour décrire de façon globale une fonction. La fonction fait appel à des notions qualitatives et quantitatives. Pour un système tel qu'un propulseur à poudre, la description de sa fonction est simple : fournir une poussée nominale de 15 t pendant 30 s. Cette définition donne ses caractéristiques en termes quantitatifs.

Un moteur électrique possède une fonction de base consistant à convertir de l'énergie mécanique en énergie électrique.

Les différences entre ces fonctions sont quelquefois subtiles et l'acceptation de leur terminologie doit toujours faire l'objet d'une acceptation au sens d'une même entreprise et de ses prestataires. Pour des systèmes plus complexes, il est indispensable de classer et de hiérarchiser la nature des fonctions [28] :

- principales ;
- secondaires ;
- de protection ;
- redondantes.

Fonctions principales Une fonction principale peut se définir comme étant la raison d'être d'un bien ou d'un système défini souvent avec ses caractéristiques associées (durée, caractéristiques physiques, chimiques...).

Par exemple, une première définition générale de la fonction principale d'une chaudière est de fournir de la vapeur.

Une seconde définition de la fonction de la chaudière à vapeur est de fournir de la vapeur à 110 °C avec un débit de 40 t/heure pendant 24 h.

Ces deux exemples de définitions montrent la difficulté de décrire une fonction quand on s'intéresse à ses modes de défaillance. En effet, dans la première définition, la perte de la fonction correspondra à une non -fourniture de vapeur. Il n'existera dans ce cas qu'un seul mode de défaillance. Par contre, en utilisant la deuxième définition, trois modes de défaillance de la fonction apparaissent [28] :

- non fourniture de vapeur à 110 °C (température supérieure ou inférieure) ;
- non fourniture du débit de 40 t/h (débit supérieur ou inférieur) ;
- non fourniture de la vapeur à une température de 110 °C, avec un débit de 40 t/heure, pendant 24 h.

Sous l'aspect sémantique une fonction est définie par un verbe suivi d'un complément d'objet.

Fonctions secondaires Dans de nombreux cas, un système assure d'autres fonctions que la fonction principale. Ces fonctions sont appelées fonctions secondaires et leur perte peut également avoir des conséquences catastrophiques. En reprenant l'exemple de la chaudière, une fonction secondaire est de maintenir l'intégrité du confinement de la vapeur. L'existence d'une fuite ou le risque d'une explosion entraînerait une défaillance de la

fonction principale. Le calorifugeage de la chaudière est une autre fonction secondaire de la chaudière ayant pour but de minimiser les pertes thermiques.

Fonctions de protection Les fonctions de protection ont pour but de garantir, par des moyens de signalisation ou la mise en route de systèmes redondants, la sécurité des biens, des personnes et de l'environnement. Ces fonctions de protection sont assurées par des systèmes de signalisation, d'alarme ou de protection automatique.

Dans le cas d'une chaudière à vapeur alimentée par un brûleur à gaz, les soupapes de sécurité assurent une protection passive contre le risque de surpression, le système de mesure de pression et de température permet aux opérateurs de contrôler les anomalies de fonctionnement et le système de détection de gaz à l'intérieur de la chaudière a pour rôle d'éviter son explosion [28].

Fonctions redondantes Dans les industries telles que celles des secteurs aéronautiques, nucléaires et spatiaux, des systèmes ou des matériels redondants (doublés, triplés ou quadruplés) sont couramment mis en oeuvre pour assurer le niveau requis de sécurité ou de sûreté.

Ces systèmes redondants peuvent fonctionner en permanence (redondance active) ou être en attente (redondance passive). Dans l'industrie automobile, c'est ainsi que l'on équipe certains véhicules avec un double circuit de freinage.

Pour un avion bimoteur, les deux moteurs illustrent la redondance active. En effet, en cas de panne d'un des moteurs, le propulseur restant a été conçu pour pouvoir ramener l'aéronef sur un terrain d'atterrissage avec toutes les conditions de sécurité.

1.5.3 Description des procédés industriels

La notion de processus industriel recouvre des secteurs industriels très variés. Au sens très large, un processus assure la fabrication d'un produit ou fournit un service. Ainsi une raffinerie a pour objectif de fournir des produits pétroliers élaborés; une centrale électrique a pour vocation de produire du courant électrique; un avion assure un service de transport de biens ou de personnes.

Un ordinateur s'assimile également à un processus puisqu'il fournit un service de calculs scientifiques ou de gestion. Ces quelques exemples illustrent la notion très large de processus industriels [28].

Conceptuellement, on appellera processus industriel, une installation complexe assumant un objectif fonctionnel de haut niveau (production de biens ou de services).

Pour assurer ces objectifs fonctionnels de haut niveau, le processus fait appel à un ensemble de systèmes interconnectés ou en interaction.

Chaque système assure une ou plusieurs fonctions bien définie(s). Les systèmes peuvent être décomposés en sous-systèmes assumant à leur tour généralement un seul objectif fonctionnel. L'étape suivante concerne la décomposition des systèmes ou sous-systèmes

en composants ou matériels bien déterminés.

A titre d'illustration, des moteurs électriques, des moteurs Diesel, des vannes, des pompes ou des unités centrales d'ordinateurs sont des composants ou des matériels bien identifiés. Selon les spécificités des secteurs industriels concernés, les méthodes de description des éléments constitutifs d'un procédé industriel font appel à des découpages matériels ou fonctionnels.

Les méthodes fonctionnelles utilisent des formalismes ou des arborescences basées sur des fonctions principales, secondaires, externes et redondantes permettant de comprendre le fonctionnement sans se préoccuper des réalisations matérielles. Ce sont ces méthodes qui sont à l'heure actuelle appliquées par les concepteurs et les équipes chargées des études de Sûreté de fonctionnement.

Description générale

Les termes suivants s'emploient pour décrire les entités constitutives d'un processus ou d'une installation industrielle sans vouloir leur affecter un sens matériel ou fonctionnel.

Bien durable Tout élément, composant, équipement, sous-système, système, matériel de processus, etc., que l'on peut considérer individuellement et qui a pour objectif d'assurer une fonction donnée pendant un temps relativement long, compte tenu de la qualité des opérations de maintenance. Un bien durable peut être relativement simple (machine à laver) ou complexe (avion, centrale nucléaire, ouvrage d'art, etc.).

Élément Partie constitutive d'un ensemble ou sous-ensemble quelles qu'en soient la nature ou la dimension.

Exemple : tuyère d'un propulseur.

Sous-ensemble Groupement d'éléments associés en fonctionnement entrant dans la constitution d'un ensemble.

Exemple : propulseur d'une fusée. Le sous-ensemble peut avoir une signification matérielle et/ou fonctionnelle.

Ensemble Groupement de sous-ensembles assurant une ou plusieurs fonctions techniques qui le rendent apte à remplir une fonction opérationnelle.

Exemple : les propulseurs d'une fusée permettent le lancement en orbite d'un satellite (fonction opérationnelle); les fonctions techniques consistent à réaliser la poussée nécessaire. L'ensemble peut avoir également une signification matérielle et/ou fonctionnelle.

Description fonctionnelle

Trois termes principaux sont largement utilisés pour décrire fonctionnellement une installation industrielle complexe : les systèmes, les sous-systèmes et les composants.

Système C'est l'association de sous-systèmes constituant un tout organique complexe, destiné à remplir une fonction générale d'un bien durable complexe.

Exemple : le système de propulsion d'un avion quadriréacteur comporte les 4 réacteurs et sa fonction est de propulser l'avion.

Sous-système Le sous-système représente une association de composants destinée à remplir une ou plusieurs fonctions opérationnelles.

Exemple : un réacteur d'un quadriréacteur remplit une partie de la fonction de propulsion durant le décollage et pendant le vol. Il assure par inversion de poussée la fonction de freinage à l'atterrissage.

Composants Le composant représente un élément matériel ou un ensemble matériel remplissant une fonction particulière dans un système ou sous-système.

Exemple : le compresseur d'un réacteur d'avion est un composant qui comprime l'air avant son injection dans les chambres de combustion.

Si l'installation industrielle est complexe il est possible de compléter la description par des niveaux inférieurs tels que sous -sous-systèmes, etc.

1.6 Loïs de probabilité rencontrées dans les études de fiabilité

1.6.1 Loi binomiale de paramètres p et n

La variable aléatoire discrète prend des valeurs entières entre 0 et n avec une probabilité :

$$P(X = i) = C_n^i p^i (1 - p)^{n-i} \quad (1.1)$$

avec $p + q = 1$.

Elle correspond à la probabilité de réalisation d'un événement de probabilité p au cours de n expériences [24]. La valeur moyenne et la variance sont données par :

valeur moyenne $\mu = np$.

et $Var(X) = npq$.

1.6.2 Loi de Poisson de paramètre λ

La variable aléatoire discrète prend des valeurs entières entre 0 et $+\infty$, la valeur moyenne et la variance sont données par :

valeur moyenne $\lambda = \mu$.

et $Var(X) = \mu$.

La loi de probabilité est définie par :

$$P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda} \quad (1.2)$$

Elle correspond au nombre d'occurrences sur une période donnée d'un événement dont la probabilité par unité de temps est constante [24].

On démontre que la loi de Poisson est la limite d'une loi binomiale de paramètre $\mu = np$ quand $n \rightarrow +\infty$.

1.6.3 Loi exponentielle de paramètre λ

La variable aléatoire est dans ce cas une variable continue $t \in [0, +\infty[$ dont la loi de densité de probabilité est donnée par [24] :

$$f(t) = \lambda e^{-\lambda t} \quad (1.3)$$

avec λ paramètre de la loi exponentielle.

La valeur moyenne et la variance sont données par :

$\mu = \frac{1}{\lambda}$ et $Var(x) = \frac{1}{\lambda^2}$

1.6.4 Loi log -normale ou loi de Galton

La variable aléatoire est dans ce cas une variable continue $t \in [0, +\infty[$ dont la loi de densité de probabilité est donnée par :

$$f(t) = \frac{1}{\sigma\sqrt{2\pi}} \frac{1}{t} \exp \frac{(-\ln(t) - \mu)^2}{2\sigma^2} \quad (1.4)$$

avec μ moyenne des $\ln(t)$,

σ écart-type des $\ln(t)$. La valeur moyenne et la variance sont données par :

$m = \exp(\mu + \frac{\sigma^2}{2})$

et $Var(t) = \frac{1}{\lambda^2}$

1.6.5 Loi de Weibull

Cette loi est très utilisée pour représenter le comportement des matériels pendant toute leur période de vie avec une loi de densité de probabilité définie par :

$$f(t) = \frac{\beta(t - \gamma)^{\beta-1}}{\eta^\beta} \exp\left(-\frac{t - \gamma}{\eta}\right)^\beta \quad (1.5)$$

avec β paramètre de forme (sans unité),

η paramètre d'échelle (en unités de temps),

γ paramètre de position (en unité de temps).

La valeur moyenne et la variance sont données par :

$$m = \gamma + \eta \Gamma\left(\frac{1+\beta}{\beta}\right)$$

$$\text{et } Var(t) = \eta^2 \left[\Gamma\left(\frac{2}{\beta} + 1\right) - \Gamma^2\left(\frac{1+\beta}{\beta}\right) \right]$$

$$\text{avec } \Gamma(b) = \int_0^{+\infty} x^{b-1} \exp(-x) dx$$

1.6.6 Loi normale

La variable aléatoire est dans ce cas une variable continue $t \in [0, +\infty[$ dont la loi de densité de probabilité est donnée par :

$$f(t) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(t - \mu)^2}{2\sigma^2}\right) \quad (1.6)$$

La valeur moyenne et la variance sont données par :

$$m = \mu \text{ et } Var(t) = \sigma^2$$

1.6.7 Loi uniforme

La variable aléatoire t est, dans ce cas, une variable continue dans l'intervalle $[t_1, t_2]$ dont la loi de densité de probabilité est donnée par :

$$f(t) = \frac{1}{t_2 - t_1} \quad (1.7)$$

La valeur moyenne et la variance sont données par :

$$m = \frac{t_1+t_2}{2} \text{ et } Var(t) = \frac{(t_2-t_1)^2}{12}$$

1.7 Concepts de base et sûreté de fonctionnement

1.7.1 Fiabilité (Reliability)

La norme NF X60-500 définit la fiabilité comme " *l'aptitude d'une entité à accomplir une fonction requise, dans des conditions données, pendant un intervalle de temps donné*". L'entité (E) désigne au sens large un composant, sous-système ou système, et la fonction requise est la ou les fonctions que doit accomplir le dispositif pour pleinement remplir la tâche qui lui est assignée.

Considérons l'instant T d'occurrence de la défaillance ; cette variable aléatoire permet de définir la notion de fiabilité qui s'interprète comme la probabilité que l'entité considérée ne tombe pas en panne avant un instant t donné ou bien comme la probabilité qu'elle tombe en panne après l'instant t.

Par extension, on appelle également fiabilité la probabilité associée $R(t)$ à cette notion alors qu'elle n'en est qu'une mesure. Elle est définie par :

$R(t) = P(E \text{ non défaillante sur la durée } [0, t], \text{ en supposant qu'elle n'est pas défaillante à l'instant } t = 0)$.

Ce qui peut s'exprimer par : $R(t) = P(T > t)$

L'aptitude contraire est appelée défiabilité, et est définie par :

$$\bar{R}(t) = 1 - R(t) = P(t < T) = F(t)$$

On distingue plusieurs types de fiabilité (termes spécifiques) :

- la fiabilité opérationnelle (observée ou estimée) déduite de l'analyse d'entités identiques dans les mêmes conditions opérationnelles à partir de l'exploitation d'un retour d'expérience ;
- la fiabilité prévisionnelle (prédite) correspondant à la fiabilité future d'un système et établie par son analyse, connaissant les fiabilités de ses composants ;
- a fiabilité extrapolée déduite de la fiabilité opérationnelle par extrapolation ou interpolation pour des conditions ou des durées différentes ;
- la fiabilité intrinsèque ou inhérente qui découle directement des paramètres de conception. Sans modification de conception des entités, il n'est pas possible d'obtenir un niveau de fiabilité au plus égal à la fiabilité intrinsèque [28].

La figure 1.4 représente une allure de la fiabilité $R(t)$ en fonction du temps pour une loi exponentielle définie par :

$$R(t) = \exp(-\lambda t)$$

avec $t \geq 0$ et $\lambda > 0$

La connaissance des lois de fiabilité des différentes entités constituant les systèmes industriels est indispensable pour mettre en oeuvre une démarche de Sûreté de fonctionnement.

Densité de défaillance : La densité de probabilité de l'instant de la défaillance T s'obtient en dérivant la fonction de répartition F (t) :

$$f(t) = \frac{df(t)}{dt} = -\frac{dR(t)}{dt}$$

Ce terme est appelé la densité de défaillance.

Taux de défaillance : A partir de la connaissance des termes $R(t)$, $f(t)$ et $F(t)$, on peut définir la notion de taux de défaillance au temps t qui est noté universellement

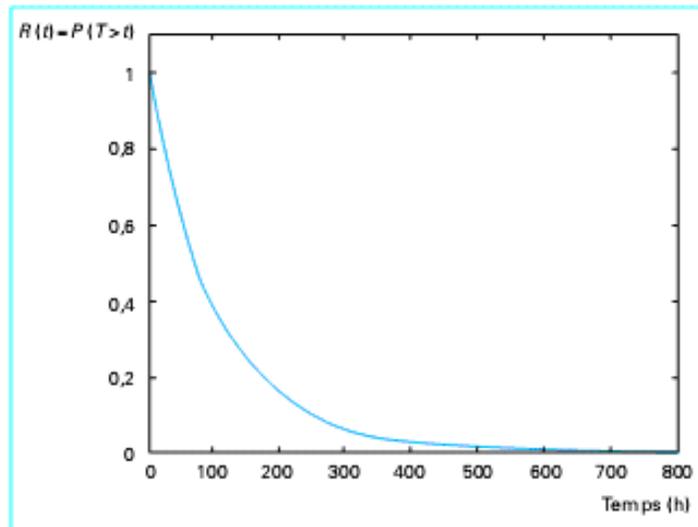


FIG. 1.4 – Allure de la fonction $R(t)$ pour la loi exponentielle.

par $\lambda(t)$. Formellement $\lambda(t)dt$ représente la probabilité d'avoir une défaillance entre $(t, t + dt)$, sachant qu'il n'y a pas eu de défaillance entre sur $[0, t]$. En appliquant le théorème des probabilités conditionnelles, il vient, si dt est petit [24] :

$$\lambda(t) = -\frac{1}{R(t)} \frac{dR(t)}{dt}$$

Taux de défaillance pendant les périodes de jeunesse, d'exploitation et

de vieillissement : Pendant la durée de vie d'une entité, on constate que son taux de défaillance évolue au cours du temps. Après la première mise en service, sur une période appelée période de jeunesse, le taux de défaillance est élevé puis à tendance à décroître.

Cela correspond à la période de rodage des systèmes mécaniques ou au déverminage de cartes électroniques. Ensuite le taux de défaillance de la majorité des entités est caractérisé par une valeur constante, ce qui signifie que la probabilité d'une défaillance est identique pour chaque instant considéré. C'est la période de défaillance à taux constant.

Ensuite, en raison du vieillissement des matériaux des composants ou de leur usure, on observe généralement un taux de défaillance qui se met à croître de façon significative. Cette période de la vie de l'entité est appelée la période de vieillissement ou d'usure [28].

Cette évolution est connue sous le nom de " courbe en baignoire " comme le montre la figure 1.5.

Moyenne de temps de vie avant la première défaillance (MTTF) Une grandeur moyenne associée à la fiabilité souvent utilisée est le temps moyen de fonctionnement d'une entité ou moyenne de temps de vie avant la première défaillance (Mean operating Time To Failure) :

$$MTTF = \int_0^{+\infty} R(t)dt$$

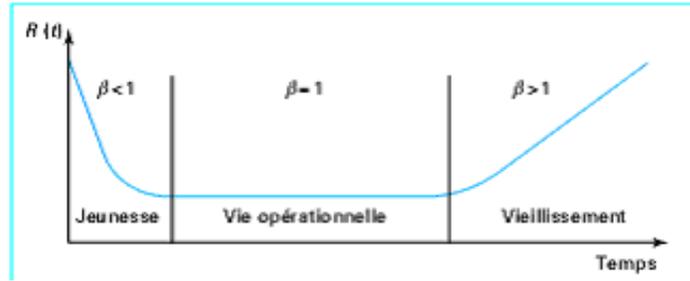


FIG. 1.5 – Allure d'un taux de défaillance " en baignoire ".

1.7.2 Disponibilité (Availability)

La norme AFNOR X 60-500 définit la disponibilité comme " l'aptitude d'une entité à être en état d'accomplir une fonction requise dans des conditions données, à un instant donné ou pendant un intervalle de temps donné, en supposant que la fourniture des moyens extérieurs nécessaires de maintenance soit assurée ".

La probabilité associée $A(t)$ à l'instant t est aussi appelée disponibilité et s'exprime par : $A(t) = P(E \text{ non défaillante à l'instant } t)$.

L'aptitude contraire est appelée indisponibilité et est définie par : $\bar{A}(t) = 1 - A(t)$

La disponibilité $A(t)$ est une grandeur instantanée. L'entité peut donc avoir subi une panne puis une réparation avant l'instant t , contrairement à la fiabilité $R(t)$ qui est une grandeur mesurée sur une durée (intervalle $[0, t]$). La confusion entre disponibilité et fiabilité est due au fait que ces deux concepts sont équivalents quand le système est non réparable [28].

L'interprétation physique est illustrée sur la figure 1.6 où sont représentés les cas d'entités disponibles et non disponibles à l'instant $t = 0$.

Dans l'industrie cela revient au constat que, sous réserve d'une politique de maintenance efficace, au bout d'un temps donné les entités ont atteint leur limite asymptotique de disponibilité. Comme la fiabilité, plusieurs types de disponibilités peuvent être utilisés :

- la disponibilité instantanée prévisionnelle (définie précédemment) ;
- la disponibilité moyenne : moyenne sur un intervalle de temps donné $[t_1, t_2]$ de la disponibilité instantanée prévisionnelle, ou mesurée en phase opérationnelle par la durée de fonctionnement effectif divisée par la durée donnée.

Les grandeurs moyennes associées à la disponibilité les plus courantes sont :

- le temps moyen de disponibilité (TMD) ou durée de bon fonctionnement après réparation, ou mean up time (MUT) : durée moyenne de fonctionnement après la réparation et la défaillance suivante ;
- le temps moyen d'indisponibilité (TMI) ou durée moyenne d'indisponibilité, ou mean down time (MDT) : durée moyenne entre une défaillance et la remise en état suivante ;
- la durée moyenne entre défaillance notée MTBF (mean time between failure) : durée moyenne entre deux défaillances consécutives de l'entité. En général, on a la relation : $MTBF = MUT + MDT$

La figure 1.7 fournit les représentations graphiques de ces définitions en fonction du temps. La disponibilité asymptotique se déduit du MUT et du MTBF par la relation :

$$A_{\infty} = \frac{MUT}{MTBF}$$

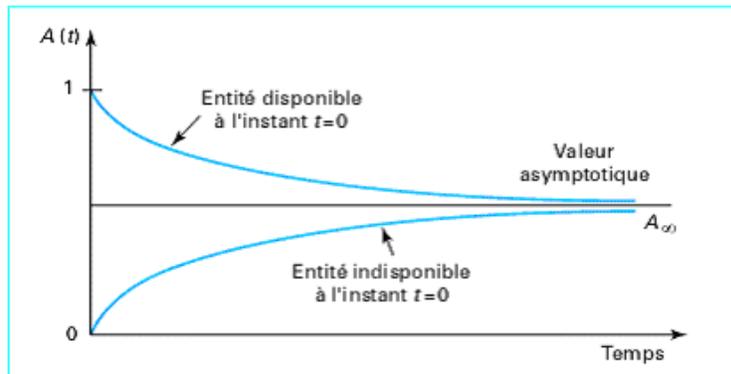


FIG. 1.6 – Disponibilité en fonction du temps t .

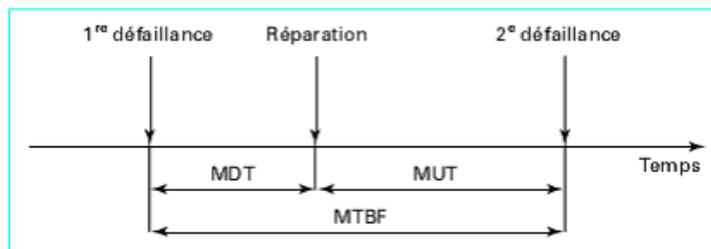


FIG. 1.7 – Représentation des MTBF, MDT et MUT.

1.7.3 Maintenabilité (Maintainability)

Les défaillances étant par définition subies sans que l'on puisse prévoir leur instant d'apparition, il importe à tout responsable d'une installation industrielle de faire face rapidement aux conséquences d'une défaillance. Dans de nombreuses situations, sauf en cas de matériels redondants, il sera indispensable de réparer les défaillances dans un temps le plus court possible pour réduire les temps d'indisponibilité. Ce paragraphe décrit les termes utilisés en maintenabilité pour évaluer les temps techniques et administratifs entre l'instant d'apparition de la défaillance et la remise en service de l'installation industrielle. Compte tenu des enjeux liés à la maintenabilité, on dissociera le cas des systèmes déjà opérationnels de ceux en cours de conception. La compréhension des termes utilisés en maintenabilité rend nécessaire l'établissement d'un diagramme chronologique des temps entre l'instant de l'apparition de la défaillance et l'instant de la remise en service de l'installation. Le diagramme de la figure 1.8 résume tous les instants importants de cette chronologie [28]. Dans cette séquence, l'instant t_0 correspond à l'instant de l'apparition

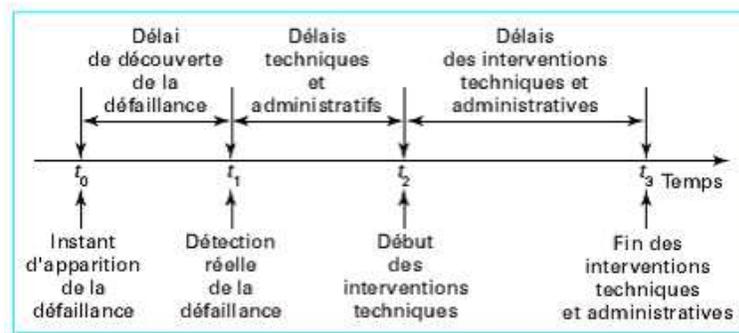


FIG. 1.8 – Chronologie des temps des activités de maintenance.

réelle de la défaillance. En fonction des moyens mis à la disposition des opérateurs (systèmes d'alarme ou bien informations venant de rondes de surveillance), il s'écoulera un délai $t_1 - t_0$ allant de quelques secondes à quelques heures pour réaliser le diagnostic de la présence d'une défaillance. La confirmation de la défaillance ayant été réalisée, il s'écoule des délais techniques et administratifs pour réunir les personnels, les pièces détachées et les autorisations administratives (par exemple, consignation d'autres matériels) pour débiter les opérations de réparation. A partir du temps t_2 , les opérations de maintenance peuvent se dérouler et incluent également les procédures d'assurance qualité et l'obtention des autorisations administratives éventuelles (par exemple, pour les appareils à pression soumis à réglementation). Ce n'est qu'à partir du temps t_3 que l'on peut considérer que l'installation est devenue à nouveau opérationnelle [28].

Définitions de la maintenabilité

Définition 1.1. *la maintenabilité* suivant la norme AFNOR X-06-010 c'est : " Aptitude d'un dispositif à être maintenu ou rétabli dans un état dans lequel il puisse accomplir une fonction requise lorsque la maintenance est accomplie dans des conditions d'utilisation données avec des moyens et des procédures prescrits ".

On notera que la norme américaine MIL-STD-721C est presque identique dans sa formulation mais inclut le niveau requis de qualification des personnels :

" La maintenabilité est la mesure de l'aptitude d'un dispositif (" item ") à être maintenu ou remis dans des conditions spécifiées lorsque la maintenance de celui-ci est réalisée par des agents ayant les niveaux spécifiés de compétence, utilisant les procédures et les ressources prescrites, à tous les niveaux prescrits de maintenance et de réparation ". La maintenabilité (maintainability) d'une entité réparable est caractérisée par une probabilité $M(t)$ que la maintenance d'une entité E accomplie dans des conditions données, avec des procédures et des moyens prescrits, soit achevée au temps t , sachant que E est défaillante au temps $t = 0$:

$M(t) = P(\text{la maintenance de } E \text{ est achevée au temps } t);$

$M(t) = 1 - P(E \text{ non réparée sur la durée } [0, t]);$

Il s'agit donc d'un équivalent à la fiabilité mais appliqué à la réparation au lieu de la défaillance. L'immaintenabilité correspond à la probabilité contraire, soit :

$$\overline{M}(t) = 1 - M(t);$$

La figure 1.9 représente l'allure de la maintenabilité $M(t)$ en fonction du temps.

On constate qu'à l'origine des temps $M(0) = 0$, ce qui est évident car l'entité est défaillante. Elle possède une asymptote égale à 1 car l'on peut supposer qu'elle sera réparée au bout d'un temps donné sinon cette entité ne serait d'aucune utilité. La maintenabilité d'une entité dépend étroitement des moyens et compétences mis en oeuvre. Comme pour la fiabilité, la notion de maintenabilité dépend de l'état initial de l'entité étudiée. Au plan pratique, la comparaison des performances d'entités identiques n'a de sens que si les méthodes et outils de maintenance sont appliqués dans des conditions strictement identiques.

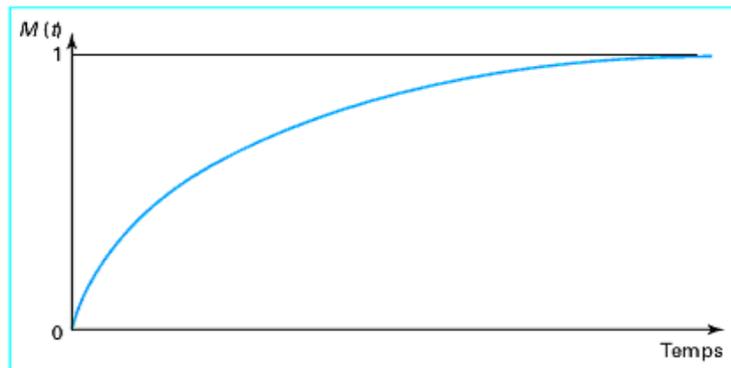


FIG. 1.9 – Allure de la courbe de maintenabilité.

Taux de réparation $\mu(t)$

On appelle taux de réparation $\mu(t)$ d'un système réparable au temps t la probabilité que l'entité soit réparée entre t et $t+dt$ sachant qu'elle n'était pas réparée sur l'intervalle $[0, t]$. Elle se note : $\mu(t) = P(\text{entité réparée sur } [t, t+dt] \text{ sachant qu'elle n'était pas réparée sur } [0, t])$. Le taux de réparation peut s'assimiler à la proportion d'entités réparées sur $[t, t+dt]$ rapportées aux entités non réparées à l'instant t . En prenant un raisonnement similaire à celui utilisé pour le calcul de l'intensité de défaillance $\lambda(t)$ et à l'aide du théorème des probabilités conditionnelles, on déduit, après des calculs simples [24] :

$$\mu(t) = \frac{1}{1 - M(t)} \frac{dM(t)}{dt}$$

Pour obtenir un taux de réparation constant, il suffit de remplacer instantanément toute entité réparée par une nouvelle entité défaillante. Le taux de défaillance $\mu(t)$ est donc proportionnel au nombre de réparations relevées sur un intervalle de temps très court dt autour de t .

Intensité de réparation $g(t)$

A partir de la définition de la maintenabilité $M(t)$ on peut définir aussi l'intensité de réparation $g(t)$ représentant la densité de probabilité de la variable aléatoire correspondant au temps de réparation.

$$g(t) = \frac{dM(t)}{dt}$$

MTTR

Le terme MTTR (mean time to repair) est la durée moyenne jusqu'à la réparation d'une entité réparable. Pour cette variable aléatoire, le MTTR se calcule par la formule :

$$MTTR = \int_0^{+\infty} (1 - M(t)) dt$$

Le MTTR s'assimile ainsi à la durée moyenne jusqu'à la première réparation et requiert la connaissance de l'état initial de l'entité.

Comme il y a en général plusieurs modes de défaillances, il faut définir plusieurs MTTR d'une entité : à chaque mode de défaillance correspondra un MTTR spécifique. Cela implique de définir clairement le ou les états de l'entité pour lesquels celle-ci est considérée comme réparée. Dans certaines industries, on définit le mean time to restore (MTR) qui est le temps nécessaire à retirer physiquement une entité en panne (avion, train) qui bloque un processus séquentiel (file d'attente pour une piste d'avion ou rame de métro en panne bloquant le trafic).

Dans des secteurs de l'aéronautique et des transports terrestres, on utilise le terme MTBR (mean time between replacement) qui correspond au temps moyen entre deux remplacements d'équipements modulaires par échange standard [28]. Les figures 1.10 et 1.11

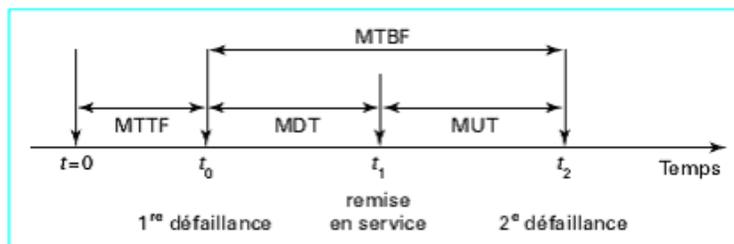


FIG. 1.10 – Relations entre les liens temporels en fiabilité, disponibilité et maintenabilité.

récapitulent les liens temporels entre les différents termes définis en fiabilité, disponibilité et maintenabilité. Pour les systèmes existants, les performances en maintenabilité sont principalement dépendantes de la configuration des matériels et de leurs dispositions, des

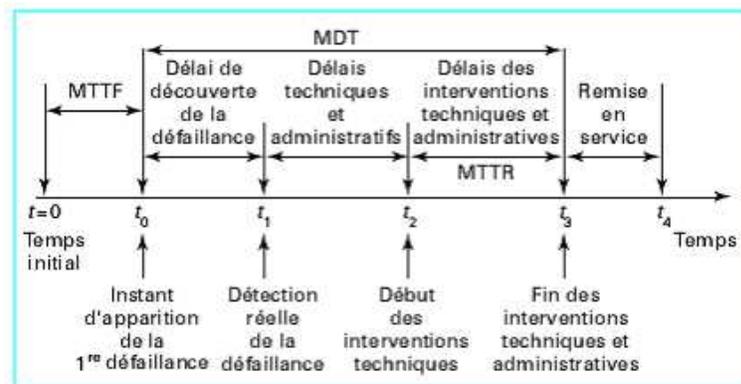


FIG. 1.11 – Chaînage temporel des activités de détection et de remise en service.

compétences et des organisations des équipes de maintenance (internes ou prestataires de service).

L'expérience montre que, très souvent, les accès physiques aux pièces défaillantes sont difficiles. Il est utile de savoir que des normes définissent les espaces nécessaires pour réaliser les opérations de maintenance avec un opérateur humain [28].

1.7.4 Sécurité

La sécurité restant un terme très général, il n'existe pas actuellement de consensus pour une normalisation. La terminologie en usage en France ne fait pas la différence entre les termes anglais " *security* " et " *safety* ".

Le terme " *security* " concerne les aspects réglementaires de la sécurité (respects des normes, contrôle des accès à des locaux ou à des systèmes informatiques) tandis que le terme " *safety* " enseigné aux États-Unis sous le nom d' " *industrial safety* " recouvre les aspects techniques de la sécurité [28].

La Sûreté de fonctionnement, est " *l'aptitude d'une entité à éviter de faire apparaître, dans des conditions données, des événements critiques ou catastrophiques* " [24].

Les circonstances et les conséquences des catastrophes et accidents sont variables. Elles montrent que le risque présente deux aspects : probabilité et conséquences. Au niveau des conséquences, celles-ci se caractérisent par la sécurité : protection des personnes, de l'environnement mais aussi protection de l'outil de production (aspect économique et, par extension, social) [28].

Deux voies peuvent être pratiquées pour réduire les risques :

- diminution de la probabilité d'occurrence de " l'événement indésirable " ;
- atténuation des conséquences de " l'événement indésirable " ;

L'évaluation de la sécurité est actuellement encore limitée et est effectuée pour les installations chimiques, les centrales nucléaires, les plates-formes pétrolières et l'aéronautique. Elle est basée sur des études statistiques des impacts des accidents (réels, expérimentés ou simulés) sur l'homme et l'environnement (notion de gravité). La démarche de la construction de la sécurité implique au départ la maîtrise des risques à un niveau acceptable, le

risque zéro étant un concept qui ne peut se réaliser dans les systèmes industriels, contrairement à ce que de nombreux discours ou écrits ont tenté de le laisser croire, il y a seulement quelques années. Le niveau de risque acceptable prend en compte des paramètres techniques, économiques, médiatiques, sociaux voire politiques. Ces niveaux acceptables sont, pour des industries à risques, définis par les autorités administratives sous la direction des autorités ministérielles de tutelle. Ainsi, dans le domaine nucléaire, le risque acceptable pour la probabilité de fusion du coeur d'une centrale nucléaire est fixé dans la fourchette $10^5 - 10^6$ par réacteur et par an pour la plupart des exploitants américains, européens et asiatiques.

Dans le domaine aéronautique, le risque de catastrophe aérienne est d'un accident par 10^7 vols [28].

Le rôle d'un spécialiste en Sûreté de fonctionnement est de ramener le risque industriel à son niveau acceptable en définissant :

- les critères d'acceptabilité des risques ;
- des méthodes de conception en sécurité ;
- des méthodes d'évaluation des risques résiduels et de vérification de leur niveau d'accessibilité.

Les études de sécurité visent essentiellement à évaluer la probabilité de l'occurrence d'un événement indésirable en prenant en compte dès la conception tous les facteurs initiateurs :

- facteurs techniques : matériels et produits manipulés (incluant les problèmes de conception, de fabrication, d'assurance qualité, de conduite et de maintenance) ;
- facteurs humains : qualité de la formation, ergonomie, procédures ;
- facteurs environnementaux : risques naturels, milieux ambiants (poussières, gaz, électricité statique...).

Les études de sécurité, où la maintenance joue un rôle non négligeable dans la mesure où de nombreux accidents sont liés à des défaillances techniques ou humaines, couvrent un spectre technique étendu. Les méthodes utilisées en sécurité doivent en particulier :

- identifier les modes de fonctionnement anormaux pouvant conduire à une situation dangereuse ;
- analyser la combinaison et l'enchaînement d'événements peu probables, pris isolément, qui conduisent à des accidents. L'expérience montre, en effet, que de nombreuses catastrophes ont été le résultat de séquences de défaillances mineures (techniques ou humaines) ;
- évaluer la probabilité d'occurrence d'un accident et lui assigner une gravité sur une échelle appropriée pour juger si le risque est acceptable économiquement ou écologiquement compte tenu des enjeux de la mission ;
- maintenir le risque à son niveau acceptable grâce, par exemple, à la maîtrise de la fiabilité des matériels obtenue par des politiques efficaces de maintenance.

La figure 1.12 résume les liens entre fiabilité, maintenabilité, disponibilité et sécurité.

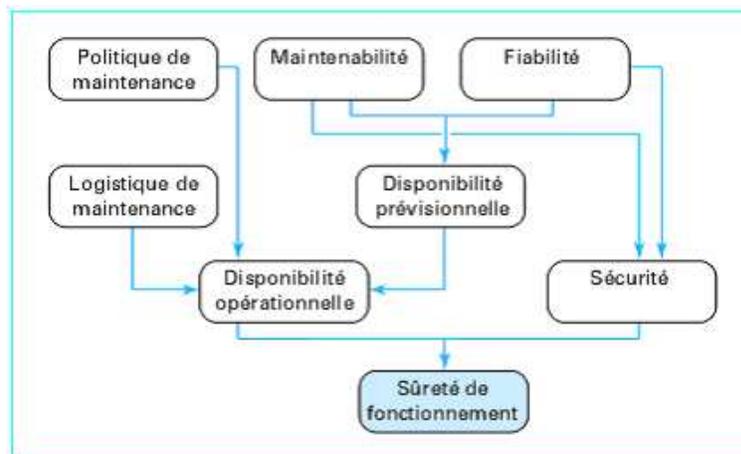


FIG. 1.12 – Relations entre fiabilité, maintenabilité, disponibilité et sécurité.

1.8 Démarches et méthodes fondamentales d'une approche SdF

Les méthodes qui décrivent des liens causes/conséquences partent des causes pour en déduire les conséquences, on les dit inductives, ou partent des conséquences pour remonter aux causes, on les dit déductives. L'un est exclusif de l'autre, mais une approche par une méthode inductive est judicieusement complétée par une approche déductive et réciproquement.

Il n'y a pas de recherche de quantification sans analyse qualitative. Par contre, il peut y avoir analyse qualitative sans quantification. Nous avons donc qualifié de quantitatives les méthodes qui offraient une possibilité importante de quantification (de fréquence) et de qualitatives les méthodes qui l'excluaient ou dans lesquelles cet aspect est marginal. L'objectif de la présentation des méthodes et démarches est de les caractériser rapidement et de les situer les unes par rapport aux autres. Leur mise en oeuvre relève d'articles spécialisés, que le lecteur trouvera dans la littérature. Dans le tableau 1.1, nous avons essayé de caractériser la démarche ou méthode citée : inductive ou déductive, quantitative ou qualitative et sa fonction [24].

Démarches/Méthodes	In/dé ductive	Quanti/quali tative	Phrase clef
Retour d'expérience	Déductive	Quantitative	Alimenter sa connaissance du système
Analyse préliminaire de risques (APR)	Inductive	Qualitative	Repérer a priori les risques à étudier les risques à étudier
Méthode d'Analyse de Dysfonctionnement des Systèmes (MADS)	Inductive	Qualitative	Recenser les différents types de scénarios possibles
Analyse des modes de défaillance et de leurs effets (AMDE)	Inductive	Qualitative	Recenser les conséquences des défaillances
Analyse des modes de défaillance, de leurs effets et de leurs criticités (AMDEC)	Inductive	Quantitative	Évaluer les conséquences des défaillances
Arbre de causes	Déductive	Qualitative	Organiser les éléments ayant contribué à un accident
Arbre d'événement	Inductive	Qualitative	Évaluer les conséquences possibles d'un événement
Arbre de défaillances	Déductive	Quantitative	Évaluer les scénarios d'un accident potentiel
Graphes d'état	Inductive	Quantitative	Évaluer les états possibles d'un système réparable

TAB. 1.1 – Principales démarches et méthodes de base de SdF.

1.8.1 Analyse préliminaire de risques (APR)

Cette démarche a pour but de produire une liste d'événements redoutés qui doivent être étudiés. Elle cherche, d'une part, à faire un tour aussi complet que possible des événements redoutés (accidents ou défaillances) et, d'autre part, à écarter rapidement ceux qui ne pourraient conduire à des conséquences assez importantes pour justifier les études.

Dans l'habitude, elle est qualifiée de " méthode de SdF ". Dans le sens où méthode évoque des règles, un ordre, une procédure normée reproductible, l'APR en est très éloignée. Elle prend des formes très différentes selon les pratiques : cela va d'un court tableau très

synthétique, bâti en une heure ou deux à la lumière de listes types, à une étude très complète et assez détaillée des risques qui concernent un système, en passant par tous les intermédiaires imaginables.

Elle est un peu à elle seule un résumé d'une démarche de maîtrise des risques par la SdF [14] :

- identification des événements redoutés, mais sans entrer dans les détails ;
- évaluation et acceptation des risques : il s'agit sommairement de retirer de la liste les événements qui sont, a priori, sans doute clairement en dessous du seuil à partir duquel ils méritent attention, évaluation aussi précise que possible et arbitrage ;
- maîtrise : mesures propres à réduire et contenir les risques aux niveaux acceptables.

Toute APR a pour objectifs :

- l'identification des événements redoutés (souvent limités à ce qui concerne la sécurité, généralement dans toutes les phases de vie du système) ;
- une évaluation permettant une hiérarchisation de ces risques ;
- l'identification des dispositions de réduction de risques qu'il faudra prévoir (donc spécifier, évaluer en termes de SdF en particulier).

Elle est particulièrement importante dans un projet de système complexe réunissant plusieurs partenaires ayant un rôle significatif dans la sécurité. Elle doit permettre, grâce à la vision globale des risques qu'elle offre, une vision commune à tous ces partenaires des risques, des besoins de réduction de risques, des rôles et responsabilités de chacun.

Cette démarche APR, ou APD (analyse préliminaire de danger) dans certains domaines avec quelques différences, est généralement une première étape indispensable quand des questions de sécurité sont posées. Elle l'est beaucoup moins s'il n'est question que de Fiabilité, Maintenabilité, Disponibilité. Quand elle est réalisée dès le début, elle sert aussi de référence tout au long du projet.

Une APR se fonde sur l'identification des situations que peut connaître le système pour isoler les situations dangereuses. La connaissance des éléments qui constituent le système est la base pour repérer les composants potentiellement dangereux, les sous-systèmes de sécurité et leurs fragilités [24].

1.8.2 Analyse des modes de défaillance, de leurs effets et de leurs criticités (AMDEC)

C'est la méthode la plus citée au point d'être parfois confondue avec la sûreté de fonctionnement. Son principe est le suivant :

- décomposer le système en éléments bien connus ; le niveau de détail de la décomposition se détermine sur ce critère : le niveau auquel on sait associer des modes de défaillance et des fréquences si possible ;
 - associer à chaque élément ses modes de défaillance ; chaque mode de défaillance est une façon dont il se comporte à part son comportement prescrit ;
 - identifier les effets sur le système de chaque mode de défaillance de chaque élément ;
 - éventuellement (c'est cette étape qui distingue l'AMDEC de l'AMDE, Analyse des
-

modes de défaillance et de leurs effets), associer à chacun des modes de défaillance de chaque élément sa criticité, en fonction des effets qu'il produit.

La figure 1.13 donne un extrait d'un tableau d'AMDEC [14]

ANALYSE DES MODES DE DÉFAILLANCE, DE LEURS EFFETS ET DE LEUR CRITICITÉ (AMDEC)										
Origine : Pièce :		AMDEC : Produit/Projet NUMÉRO : AGRAFE			Analyste :		Date : 17/06/95	Folio		
FONCTIONS (ET/OU) PROCESSUS	DÉFAILLANCE				GRAVITÉ	FRÉQUENCE	DÉTECTI- ON	ACTION	Responsable	Délai
	MODE	CAUSE	EFFET	DÉTECTION						
Ôter une agrafe d'une feuille de papier	Impossible d'accrocher l'agrafe avec les mors	Forme de l'extrémité des mors trop arrondie P1 : 10 % P21 : 100 %	Impossible d'arracher l'agrafe	Essais de formes	9	9	2			
	Impossible d'exercer la traction sur l'agrafe	Frottement entre les deux méchôires P1 : 1 % P21 : 0,1 %			6	3	10	Utilisation de deux méchôires chromées	D. Dupont	20/06/95
		Coincement de l'agrafe entre les deux méchôires P1 : 1,0 P21 : 1				5	10	10		

FIG. 1.13 – Extrait d'un tableau d'AMDEC.

1.8.3 Arbres de causes, d'événement, de défaillances

Les représentations de liens logiques causes/conséquences partant d'un événement ont une forme arborescente. Nous évoquons ici trois formes très utilisées.

- Un arbre de causes explicite la façon dont des combinaisons d'événements et de conditions ont conduit à la réalisation d'un accident. C'est une représentation logique de l'analyse des causes d'un accident (il n'utilise que le connecteur logique ET). À une extrémité, on trouve l'accident ou l'incident analysé et aux extrémités des ramifications, les événements, conditions ou circonstances qui se sont rencontrées pour qu'il se produise.

L'appellation " *arbre de causes* ", qui aurait aussi bien pu être retenue pour désigner les arbres de défaillances a été retenue pour désigner la méthode de représentation de l'analyse a posteriori d'un accident. Elle a été très bien promue dans le domaine de la sécurité au travail et largement diffusée par l'INRS (Institut National de Recherche et de Sécurité) dans ce cadre.

La construction d'un arbre de causes (figure 1.14) consiste à partir de l'événement final et à remonter, étape par étape, en se demandant à chaque fois, quelle est la

combinaison d'événements, d'actions, de décisions, de conditions qui a provoqué l'événement immédiatement supérieur.

- Un arbre d'événement part d'un événement (origine), et on s'interroge sur les conséquences qu'il peut avoir. Quand on se pose cette question, on suit un chemin jalonné de " si telle condition est remplie, alors..., dans le cas contraire..." [14].

On représente ce raisonnement sous forme d'un schéma qui se lit usuellement de gauche (une entrée unique : l'événement dont on cherche les conséquences) vers la droite (les événements ultimes dont il peut être une cause) en progressant de bifurcation en bifurcation, le nombre de branches croissant au fur et à mesure que l'on progresse vers les conséquences en explorant à chaque fois les deux branches des alternatives qui s'ouvrent (figure 1.15).

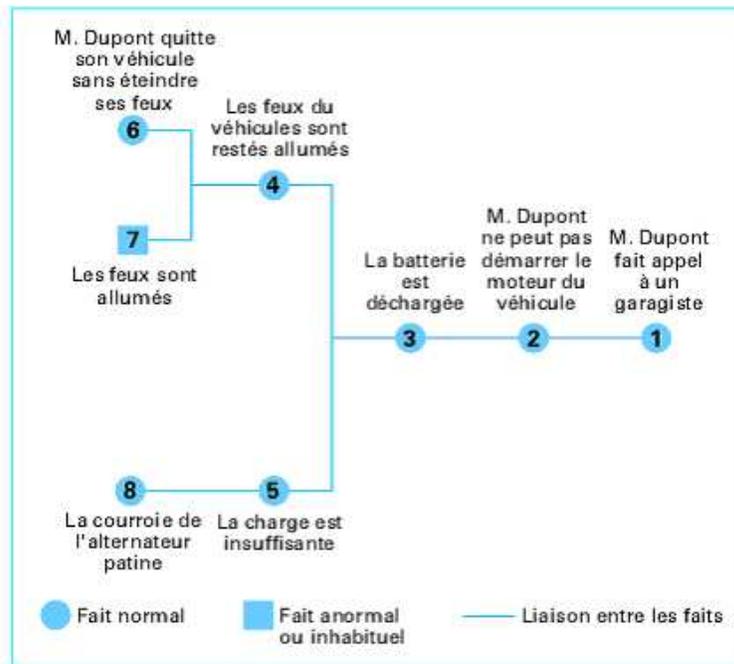


FIG. 1.14 – Un arbre de causes pour le scénario " panne de batterie ".

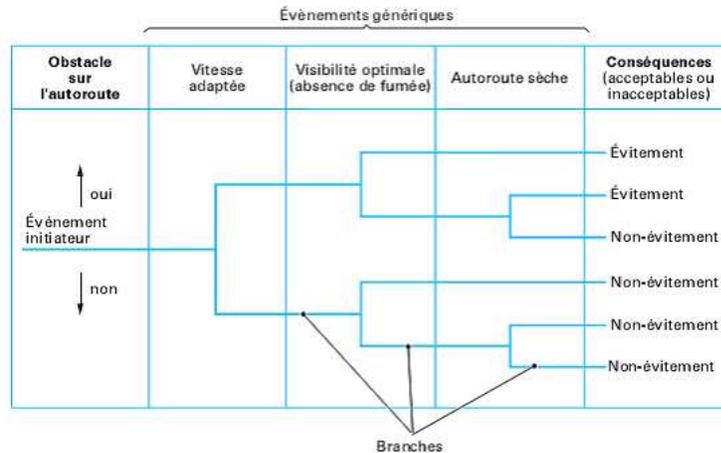


FIG. 1.15 – Un arbre d'événement pour le scénario " évitement d'un obstacle sur l'autoroute ".

L'événement initiateur EI initie le scénario.

L'action correctrice de l'événement générique EG peut atténuer ou aggraver l'EI.

Branche : notion qui particularise les EG en fonction du passé, en fonction des conditions initiales de la séquence.

Séquence : chemin qui conduit aux conséquences acceptables ou inacceptables

1.8.4 La méthode MADS-MOSAR

MOSAR est une méthode générique qui permet d'analyser les risques techniques d'une installation humaine et d'identifier les moyens de prévention nécessaires pour les neutraliser. Elle s'applique aussi bien dès la conception d'une installation nouvelle qu'au diagnostic d'une installation existante [19].

Elle constitue aussi un outil d'aide à la décision par les choix qu'elle met en évidence. Elle comprend deux modules.

Le premier module ou module A permet de réaliser une analyse des risques principaux. A partir d'une décomposition de l'installation en sous-systèmes, on commence par identifier de manière systématique en quoi chaque sous-système peut être source de dangers. Pour cela, on fait référence à une grille de typologie des systèmes sources de dangers et on utilise le modèle MADS qui relie source de dangers et cibles. La négociation d'objectifs entre les acteurs concernés, par construction de graphes probabilités -gravité permet de hiérarchiser les scénarios identifiés.

La recherche des moyens de prévention (barrières techniques et barrières opératoires) nécessaires pour neutraliser les scénarios assure la prévention des risques. Ce premier module se termine par la qualification dans le temps des barrières identifiées.

Le deuxième module ou module B permet de réaliser une analyse détaillée de l'installation et notamment il met en oeuvre les outils de la sûreté de fonctionnement pour la recherche des dysfonctionnements techniques des machines et appareils.

1.8.5 Les Réseaux Bayésiens

Les incertitudes et des risques étaient, et resteront toujours présents; tant dans la vie de chacun que dans les stratégies de travail. Une préoccupation concerne une variété d'incertitudes pratiquement irréductibles, qui défient nos modèles traditionnels et exigent de nouveaux concepts et des outils analytiques [13].

Le raisonnement probabiliste avec des modèles graphiques, aussi connus comme réseaux Bayésiens ou réseaux de croyance, est devenu un domaine actif de la recherche et de la pratique dans l'intelligence artificielle [26], on désigne par l'activité de deviner l'état du système de la connaissance antérieure et des observations comme le *raisonnement ou l'inférence*.

La tâche du raisonneur n'est pas une de déduction déterminée, mais plutôt du raisonnement incertain. C'est-à-dire le raisonneur doit déduire l'état du système basé sur la connaissance incomplète et incertaine du domaine et des observations.

Beaucoup de théories existent sur la façon d'exécuter un tel raisonnement incertain. Dans ce qui suit, on se concentre sur des méthodologies fondées sur la théorie de probabilité Bayésienne. Autrement dit, on se concentre sur le *raisonnement probabiliste*.

Conclusion

La Sûreté de Fonctionnement représente pour des systèmes opérationnels ou en cours de conception, un des leviers importants pour réduire les coûts de possession tout en contribuant de façon significative à des gains de productivité des biens ou des services.

La démarche rationnelle et structurée de cette méthode exige de la part des utilisateurs potentiels une vision globale, donc systémique, de l'entreprise pour inclure tous les facteurs contribuant à une implantation couronnée de succès du programme de la Sûreté de Fonctionnement.

La richesse et la variété des techniques et méthodes nécessaires pour le développement d'une étude de Sûreté de Fonctionnement rendent obligatoire une approche transverse qui fédère toutes les compétences de l'entreprise. En effet, avoir une vision globale du coût de possession fait appel à des disciplines aussi variées que l'analyse de la valeur, l'analyse fonctionnelle, les méthodes de conception intégrant la Sûreté de Fonctionnement (fiabilité, disponibilité, sûreté, maintenabilité).

Chapitre 2

Les Réseaux Bayésiens

Les Réseaux Bayésiens (RBs) sont des modèles graphiques interprétables et flexibles pour représenter des relations probabilistes entre de multiples entités interagissantes. À un niveau *qualitatif*, la structure d'un réseau Bayésien décrit les relations entre ces entités sous forme de relations d'indépendance conditionnelles. À un niveau *quantitatif*, des relations (locales) entre les entités interagissantes sont représentées par des distributions de probabilités (conditionnelles). Formellement, un RB est défini par une structure graphique, G , une famille de distributions de probabilité (conditionnelles), P et leurs paramètres, spécifiant ensemble une distribution jointe sur un ensemble de variables aléatoires. Dans ce chapitre on présentera les réseaux bayésiens, leur propriétés, et la notion de séparation qui nous permettra, plus tard de faire l'inférence ou par la disponibilité de l'observation (appelée aussi évidence) [1], de faire la propagation de l'information. Mais avant il est utile de faire un rappel de probabilité.

2.1 Rappel de Probabilités

Nous commençons cette section avec l'idée d'une expérience aléatoire. Quand nous répétons l'expérience dans des conditions identiques, nous pouvons obtenir un résultat différent. Nous commençons par les définitions suivantes :

- *Expérience aléatoire*. Une expérience qui a un résultat qui n'est pas complètement prévisible. Nous pouvons répéter l'expérience dans les mêmes conditions et sans obtenir le même résultat. Le jet d'une pièce de monnaie est un exemple d'une expérience aléatoire [3].
 - *Résultat*. Le résultat d'un essai simple d'une expérience aléatoire.
 - *Espace d'échantillon*. L'ensemble de tous les résultats possibles d'un essai simple d'une expérience aléatoire. Nous le dénotons Ω . L'espace d'échantillon contient tout ce que nous considérons dans cette expérience, donc nous pouvons aussi l'appeler l'Univers, noté par U [3].
-

- *Événement*. N'importe quel ensemble de résultats possibles d'une expérience aléatoire [3].
- *Union de deux événements*. L'union de deux événements E et F est l'ensemble de résultats dans E ou F (ou inclusif). noté $E \cup F$.
- *Intersection de deux événements*. L'intersection de deux événements E et F est l'ensemble de résultats tant dans E que F simultanément. Dénomé $E \cap F$.
- *Complément d'un événement*. Le complément d'un événement E est l'ensemble de résultats qui ne sont pas dans E . Dénomé \tilde{E}

On va utiliser le diagramme de Venn pour illustrer la relation entre événements. Les événements sont représentés par des régions dans l'Univers. La relation entre deux événements dépend du résultat qu'ils partagent. Si tous les résultats d'un événement F sont aussi dans l'autre événement E , le premier événement est un sous ensemble de l'autre événement (Figure 2.1).

Si les événements ont quelques résultats en commun, mais chacun d'eux a des résultats n'appartenant aux autres, ils sont des événements en intersection (Figure 2.2).

Si deux événements ne partagent aucun résultat, ils sont des événements mutuellement exclusifs, dans ce cas l'occurrence d'événement exclut l'autre et vice versa. Ils sont dits des événements disjoints (Figure 2.3).

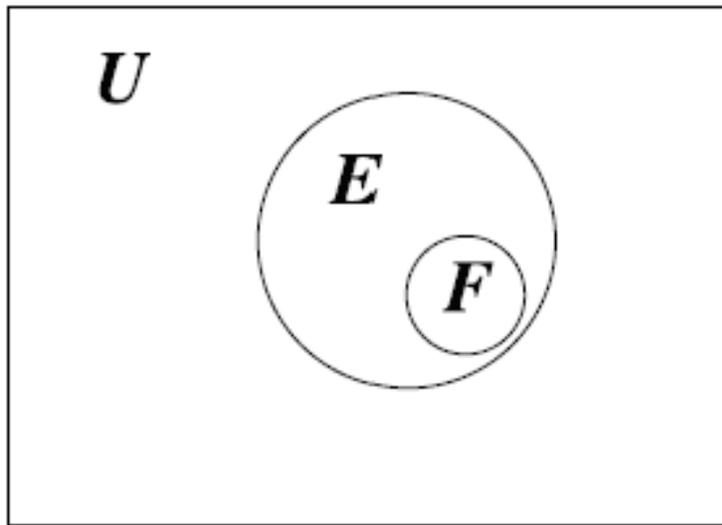
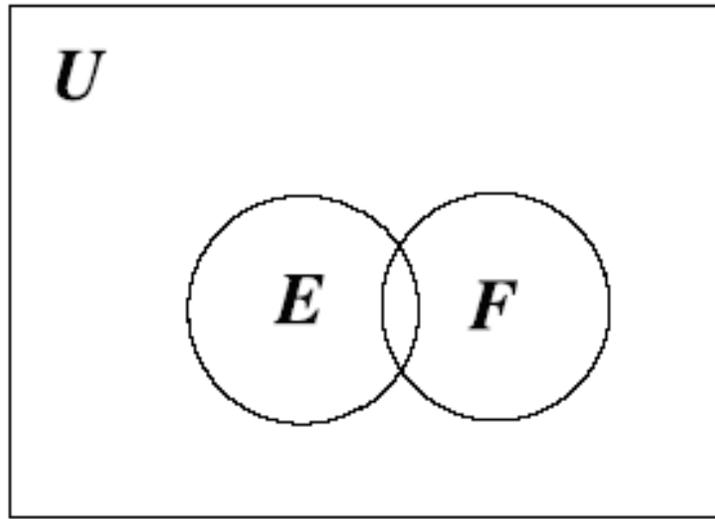
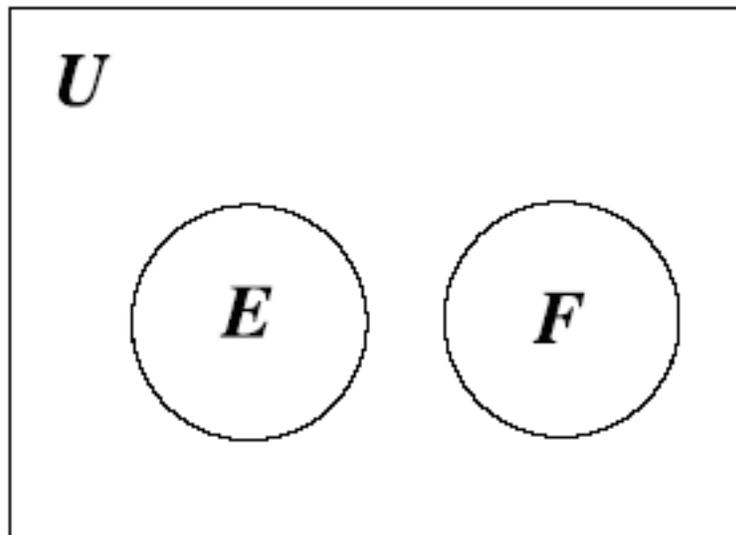


FIG. 2.1 – l'événement F est un sous ensemble de l'événement E .

2.1.1 Les Axiomes de probabilités

La désignation de probabilité pour une expérience aléatoire est une désignation de probabilités à tous les événements possibles générés par l'expérience. Ces probabilités sont des nombres réels entre 0 et 1. Plus la probabilité d'un événement est proche de 1, plus probablement l'événement doit arriver [3]. Une probabilité égale 1 signifie que

FIG. 2.2 – E et F sont des événements en intersection.FIG. 2.3 – les événements E et F sont mutuellement exclusifs ou disjoints.

l'événement est certain d'arriver et une probabilité de 0 signifie que l'événement n'a pas de possibilité d'arriver. Pour être cohérent, la désignation de probabilités aux événements doit satisfaire les axiomes suivants.

1. $P(A) \geq 0$ pour tout événement A. (Les Probabilités sont non négative.)
2. $P(U) = 1$. (Probabilité de l'univers = 1.)
3. Si A et B sont des événements mutuellement exclusifs, alors $P(A \cap B) = P(A) + P(B)$. (La probabilité est additive sur les événements disjoints.)

2.1.2 La probabilité jointe et les événements indépendants

La figure 2.4 montre le diagramme de Venn pour deux événements A et B dans l'univers U. La probabilité jointe des deux événements A et B est la probabilité que les deux événements arrivent simultanément. Ce serait la probabilité de l'ensemble des résultats qui sont tant dans l'événement A que dans l'événement B, l'intersection $A \cap B$. Autrement dit la probabilité commune d'événements A et B est $P(A \cap B)$, c'est la probabilité de leur intersection. Si l'événement A et l'événement B est indépendant, donc $P(A \cap B) = P(A) \times P(B)$, la probabilité commune est le produit des probabilités individuelles. Le cas contraire ils sont appelés des événements *dépendants*.

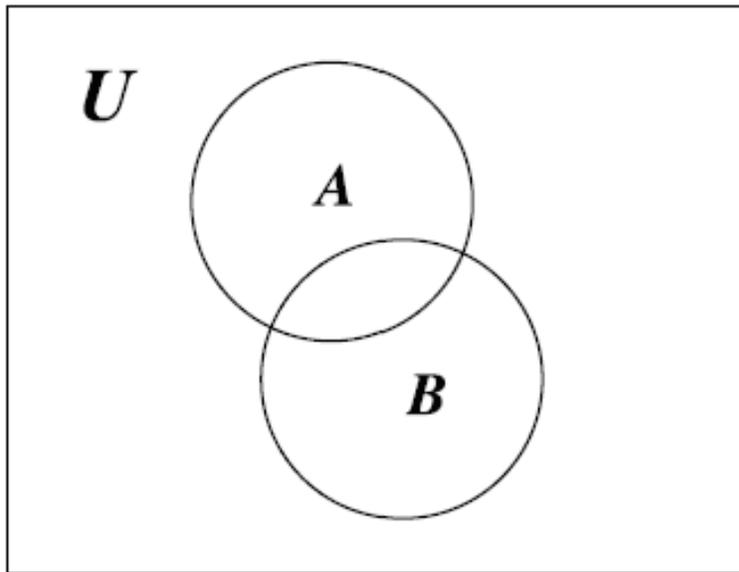


FIG. 2.4 – Deux événements A et B dans l'univers U.

2.1.3 Distinction entre événements indépendants et événements mutuellement exclusifs

La confusion est courante entre événements indépendants et événements mutuellement exclusifs. Cette confusion sémantique intervient parce que le mot indépendant a plusieurs significations. La signification primaire de quelque chose étant indépendant de quelque chose d'autre est que la deuxième chose n'a pas d'effet sur le premier. C'est la signification du mot indépendant utilisée dans la définition d'événements indépendants. L'occurrence d'un événement n'affecte pas l'occurrence ou la non occurrence des autres événements.

Il y a une autre signification du mot indépendant. C'est la signification politique d'indépendance. Quand une colonie devient indépendante du pays mère, cela devient un pays séparé distinct. Cette signification est satisfaite d'après la définition d'événements *mutuellement exclusifs* ou *disjoints* [3].

L'indépendance de deux événements n'est pas une propriété des événements eux-mêmes, c'est plutôt une propriété qui provient des probabilités des événements et de leur intersection. C'est par contre avec des événements *mutuellement exclusifs*, qui ont la propriété de n'avoir aucun élément en commun. Des événements mutuellement exclusifs avec la probabilité non négative ne peuvent pas être indépendants. Leur intersection est l'ensemble vide, donc on doit avoir une probabilité zéro, qui ne peut pas être égale au produit des probabilités des deux événements [3].

2.1.4 Probabilité marginale

La probabilité d'un des événements A , $P(A)$ est appelée sa probabilité marginale. Elle est trouvée en additionnant $P(A \cap B)$ et $P(A \cap \tilde{B})$ en utilisant les axiomes de probabilités.

- $A = (A \cap B) \cup (A \cap \tilde{B})$ Et ils sont disjoints. Donc par axiome 3.
- $P(A) = P(A \cap B) + P(A \cap \tilde{B})$ La probabilité marginale de l'événement A est trouvée en additionnant ses parties disjointes.

2.1.5 La probabilité conditionnelle

Si nous savons qu'un événement est arrivé, est ce qu'il a un effet sur la probabilité d'un autre événement arrivé? Pour y répondre, nous devons regarder la probabilité conditionnelle.

Supposons que l'événement A est arrivé. Tout ce qui en dehors de A n'est plus possible. Nous devons seulement considérer des résultats à l'intérieur de l'événement A , *L'univers réduit* $U_r = A$ est utilisée. La seule partie d'événement B qui est maintenant approprié est cette partie qui est aussi dans A ; c'est $B \cap A$.

La figure 2.5 montre qu'étant donné que l'événement A est arrivé, l'univers réduit est

maintenant l'événement A et la seule partie appropriée de l'événement B est $B \cap A$. Étant donné que l'événement A est arrivé, la probabilité totale dans l'univers réduit doit être égale à 1. La probabilité de B sachant A est la probabilité inconditionnelle de cette partie de B qui est aussi dans A, multiplié par le facteur d'échelle $\frac{1}{P(A)}$. Cela donne la probabilité conditionnelle d'événement B sachant l'événement A :

$$P(B|A) = \frac{P(A \cap B)}{P(A)} \quad (2.1)$$

Nous voyons que la probabilité conditionnelle $P(B|A)$ est proportionnelle à la probabilité commune $P(A \cap B)$ mais elle a été mise à l'échelle ainsi la probabilité de l'univers réduit est égale à 1.

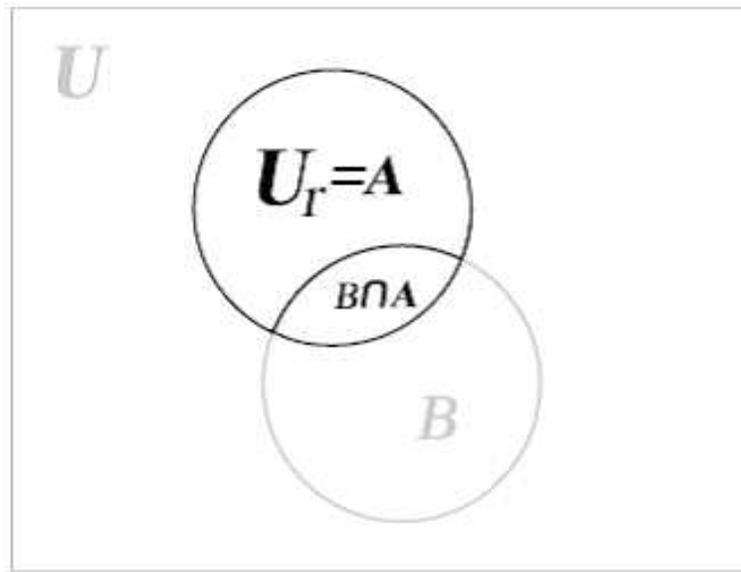


FIG. 2.5 – l'univers réduit, sachant que l'événement A est arrivé.

2.1.6 Probabilité conditionnelle pour les événements indépendants

Remarquons que quand A et B sont des événements indépendants

$$P(B|A) = P(B) \quad (2.2)$$

Puisque $P(B \cap A) = P(B) \times P(A)$ pour des événements indépendants. La connaissance de A n'a aucun effet sur la probabilité d'apparition de B quand A et B sont des événements indépendants.

2.1.7 La règle de Multiplication

Formellement, nous pourrions changer complètement les rôles des deux événements A et B. La probabilité conditionnelle de A sachant B serait :

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad (2.3)$$

Cependant, nous ne considérerons pas les deux événements de la même façon. B est un événement non observable. C'est-à-dire l'occurrence ou non occurrence de l'événement B n'est pas observée. A est un événement observable qui peut arriver avec l'événement B ou avec son complément \tilde{B} .

Cependant, les chances d'une apparition peuvent dépendre sur lequel de B ou \tilde{B} est arrivé. Autrement dit, la probabilité de l'événement A est conditionnelle sur l'occurrence ou la non occurrence de l'événement B. Quand nous purifions les fractions dans la formule de probabilité conditionnelle nous arrivons à

$$P(A \cap B) = P(A)P(B|A) \quad (2.4)$$

Ceci est connu comme la règle de multiplication pour les probabilités.

De la même façon $P(A \cap \tilde{B}) = P(\tilde{B})P(A|\tilde{B})$.

2.1.8 Le théorème de Bayes

De la définition des probabilités conditionnelles (équation 2.1), la probabilité marginale d'un événement A est trouvé en sommant les probabilités de ses parties disjointes. Puisque $A = (A \cap B) \cup (A \cap \tilde{B})$, il est claire que $(A \cap B)$ et $(A \cap \tilde{B})$ sont disjointes, on le substitue dans la définition des probabilités conditionnelles, on aura :

$$P(B|A) = \frac{P(A \cap B)}{P(A \cap B) + P(A \cap \tilde{B})} \quad (2.5)$$

Maintenant, on utilise la règle de multiplication pour trouver chacune de ces probabilités jointes, ceci donne le théorème de Bayes pour un seul événement :

$$P(B|A) = \frac{P(A|B) \times P(B)}{P(A|B) \times P(B) + P(A|\tilde{B}) \times P(\tilde{B})} \quad (2.6)$$

2.2 Les Modèles graphiques

Les modèles graphiques sont un mariage entre la théorie des probabilités et la théorie des graphes. Ils fournissent un outil naturel pour faire face à deux problèmes qui apparaissent via les mathématiques appliquées et les sciences de l'ingénieur : l'incertain et la

complexité [15].

Les modèles graphiques sont des graphes dans lesquels les noeuds représentent les variables aléatoires et les arcs représentent les indépendances conditionnelles. Les réseaux bayésiens sont un formalisme des modèles graphiques. Nombreux sont les travaux qui ont traités les réseaux bayésiens [15, 12, 21, 4, 7].

2.2.1 Les graphes

Définition 2.1. *On appelle graphe non orienté la paire $G = (V, E)$ où : $V = X_1, \dots, X_n$ représente les sommets du graphe, et $E(X_i, X_j)$ représente la famille des sous ensembles de taille 2 de V .*

Les éléments de E sont appelés arêtes du graphe.

Ici les arêtes ne sont pas orientées, il n'y a pas d'arête qui va d'un sommet à lui-même (pas de boucle). Si (X_i, X_j) appartient à l'ensemble des arêtes E , alors on représentera celle-ci par une ligne entre les deux sommets notée $X_i - X_j$ [7].

Définition 2.2. *On appelle graphe orienté la paire la paire $G = (V, E)$ où : $V = X_1, \dots, X_n$ représente les sommets (noeuds) du graphe, et $E = (e_1, \dots, e_m)$ représente une partie du produit cartésien $V \times V$, dont les éléments sont appelés arcs du graphe.*

Si (X_i, X_j) appartient à l'ensemble E , alors on appellera cet élément arc du graphe, on note $X_i \rightarrow X_j$: une flèche partant de X_i et pointant sur X_j , X_i est appelé l'origine et X_j l'extrémité [7].

2.2.2 Notion de parents et d'enfants

S'il existe un arc partant de X_i vers X_j , alors X_i s'appelle le parent de X_j et X_j l'enfant de X_i , on $Pa(X_j)$ note la suite des parents de X_j et $Ch(X_i)$ la suite des enfants de X_i . Par exemple, dans la Figure 2.6, le noeud a est le parent du noeud b et l'ensemble des parents de e est $P(e) = \{a, c, d\}$.

2.2.3 Graphe Acyclique Dirigé

On appelle cycle tout chemin d'au moins deux noeuds reliés entre eux par des flèches et pour lequel le premier et le dernier noeud sont identiques. Une classe très importante de graphes dirigés est la classe des graphes ne présentant aucun cycle. On appelle Graphes

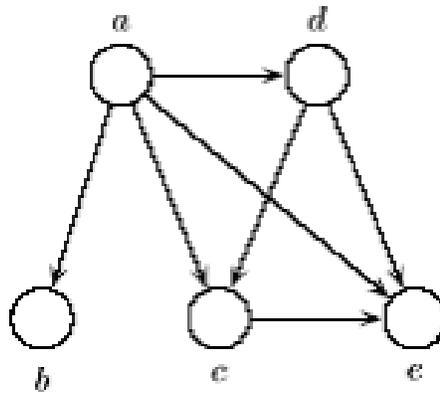


FIG. 2.6 – Exemple d'un graphe dirigé illustrant les relations parents - enfants.

Acycliques Dirigés (GAD) de tels graphes. Les GAD sont la base des modèles probabilistes appelés Réseaux Bayésiens [17](Figure 2.6).

2.2.4 Graphe Moral

Définition 2.3. *Le graphe moral G^m d'un graphe dirigé G est défini comme le graphe non dirigé avec les mêmes noeuds que G où deux noeuds X_i et X_j sont reliés dans G^m si et seulement si X_i et X_j ont un enfants en commun dans G . Dans la pratique, le graphe moral est obtenu à partir du graphe original par "mariage" des parents ayant un enfant en commun et par suppression des flèches. On construit ainsi un graphe non dirigé à partir d'un graphe dirigé.*

2.2.5 Graphe Triangulé

Dans un graphe non dirigé, la terminologie est différente de celle employée pour les graphes dirigés. Ainsi, les liens entre les noeuds sont appelés arcs au lieu de flèches, et on désigne par le terme de boucle un chemin fermé qui correspond à un cycle dans un graphe dirigé.

Un graphe triangulé est défini comme un graphe non dirigé qui a une corde¹ dans chaque boucle de longueur $n \geq 4$. La Figure 2.7 donne un exemple de graphe triangulé. La boucle $\{c, a, d, b\}$ a une corde $\{a, b\}$ qui rend cette boucle de longueur inférieure à 4. Les concepts que nous venons d'introduire et le langage employé nous sont utiles pour décrire de manière rigoureuse les techniques qui sont présentées dans les prochaines sections.

¹On appelle corde d'une boucle tout arc qui joint deux noeuds non voisins de cette boucle et qui n'est donc pas lui-même un arc de cette boucle.

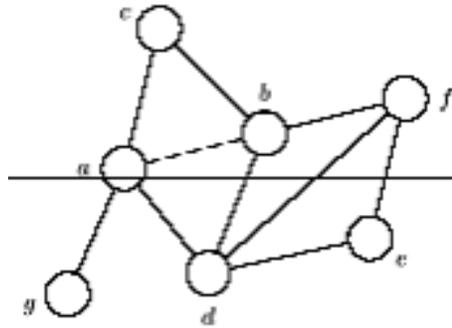


FIG. 2.7 – Exemple de graphe triangulé. La corde $\{a, b\}$ change le cycle $\{c, a, d, b\}$ en deux cycles de longueur 3.

2.3 Réseaux Bayésiens

Les modèles graphiques sont des outils théoriques très puissants. Rappelons cependant que l'objectif recherché dans ce travail est de modéliser la sûreté de fonctionnement d'un sous système à l'aide de réseaux bayésiens (RB). Un RB est un modèle graphique qui représente les indépendances conditionnelles entre un ensemble de variables aléatoires [21].

Définition 2.4. *Un réseau bayésien est défini par :*

- Une suite de variables, notée V et une suite d'arcs entre les variables, notée E .
- Chaque variable possède un nombre fini d'états exclusifs.
- Les variables et les arcs forment un graphe orienté acyclique, noté $G = (V, E)$.
- A chaque variable Y avec ses parents X_1, \dots, X_n est associée une probabilité conditionnelle. Lorsque la variable ne possède pas de parents, la dernière quantité devient une probabilité marginale $P(Y)$ [7].

Les Réseaux Bayésiens (RBs) sont une représentation de la Distribution de Probabilité Jointe (JPD) (Join Probability Distribution) [18].

Dans la figure (2.8), les noeuds représentent les variables aléatoires, les arcs indiquent les dépendances conditionnelles, la probabilité jointe $P(A,B,C,D,E)$ est factorisée en produit $P(A)P(B|A)P(C|A)P(D|B,C)P(E|D)$.

2.3.1 Indépendance Conditionnelle

Les relations d'indépendance conditionnelles permettent de simplifier le calcul de certaines lois conditionnelles et les différents réseaux bayésiens que nous serons amenés à construire auront pour fonction première de représenter les relations d'indépendance conditionnelle existantes dans le système considéré.

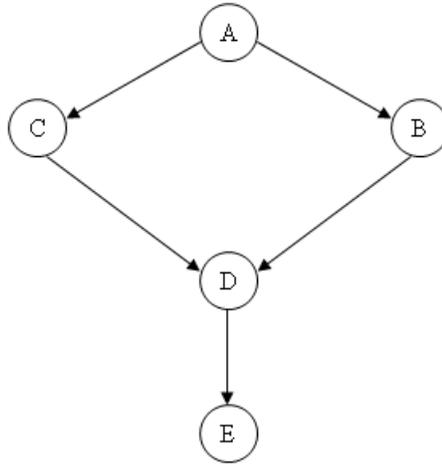


FIG. 2.8 – un exemple d'un réseau bayésien.

Pour illustrer de manière simple notre propos, considérons trois variables aléatoires A, B et C. D'après la théorie des probabilités, leur probabilité jointe s'écrit comme le produit des probabilités conditionnelles suivantes :

$$P(A, B, C) = P(A)P(B/A)P(C/A, B) \quad (2.7)$$

Si A est indépendante de B, l'équation (2.7) peut être écrite comme suit :

$$P(A, B, C) = P(A)P(B)P(C/A, B) \quad (2.8)$$

L'équation (2.8) peut être représentée par un graphe comme cela est illustré dans la Figure 2.9.

Chaque nœud correspond à une variable. Chaque flèche représente une dépendance entre

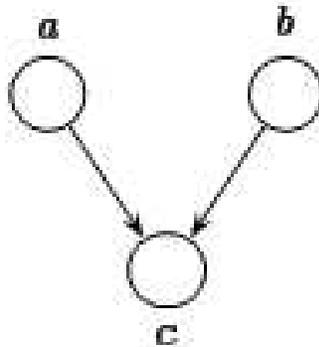


FIG. 2.9 – Représentation graphique de l'équation 2.8 avec des Réseaux Bayésiens.

les variables et est associée à la densité de probabilité du fils sachant le parent. Un réseau bayésien représente donc une distribution de probabilité qui admet la loi jointe suivante :

$$P(X) = \prod_{i=1}^N P(X_i | \text{Pa}(X_i)) \quad (2.9)$$

Cette décomposition de la loi jointe permet d'avoir des algorithmes d'inférence puissants qui font des réseaux bayésiens des outils de modélisation et de raisonnement très pratiques lorsque les situations sont incertaines ou les données incomplètes. Ils sont alors utiles pour les problèmes de classification lorsque les interactions entre les différents critères peuvent être modélisés par des relations de probabilités conditionnelles.

Lorsque la structure du réseau bayésien n'est pas fournie a priori par un expert, il est possible d'en faire l'apprentissage à partir d'une base de données. La recherche de structure de réseaux bayésiens n'est pas simple, principalement à cause du fait que l'espace de recherche est détaillé super exponentielle en fonction du nombre de variables.

Dans le RB de la figure 2.8, on a l'ensemble de noeuds $V = \{A, B, C, D, E\}$, et l'ensemble d'arcs $E = \{(A, B), (A, C), (B, D), (C, D), (D, E)\}$. Les noeuds B et C sont les fils du noeud A et sont les parents du noeud D. le noeud D, a lui-même un seul fils, c'est le noeud E. la structure graphique doit prendre la forme d'un graphe acyclique dirigé (GAD), caractérisé par l'absence de cycles dirigés.

Un RB est caractérisé par une simple et unique règle pour étendre la probabilité jointe en terme de probabilités conditionnelles simples. Soit X_1, \dots, X_N un ensemble de variables aléatoires représentées par les noeuds $i = \{1, 2, \dots, n\}$ dans le graphe, on définit $pa[i]$ comme les parents du noeud I, et soit $\chi_{pa[i]}$ l'ensemble des variables aléatoires associées avec $pa[i]$, alors :

$$P(X_1, \dots, X_n) = \prod_{i=1}^n P(X_i/\chi_{pa[i]}) \quad (2.10)$$

Par exemple, en appliquant l'équation (2.10) au RB de la figure 2.8, on obtient la factorisation suivante :

$$P(A, B, C, D, E) = P(A)P(B/A)P(C/A)P(D/B, C)P(E/D) \quad (2.11)$$

Une manière équivalente d'exprimer ces relations de dépendances est basée sur le concept de la *Couverture de Markov*². Donc, si $CM[i]$ est la couverture de Markov du noeud i, et $\chi_{CM[i]}$ est l'ensemble de variables aléatoires associées avec $CM[i]$, alors :

$$P(X_k/X_1, \dots, X_{k-1}, X_{k+1}, \dots, X_n) = P(X_k/\chi_{CM[i]}) \quad (2.12)$$

L'application de l'équation (2.11) au RB de la figure 2.6, donne :

$$P(A/B, C, D, E) = P(A/B, C) \quad (2.13)$$

$$P(B/A, C, D, E) = P(B/A, C, D) \quad (2.14)$$

$$P(C/A, B, D, E) = P(C/A, B, D) \quad (2.15)$$

$$P(D/A, B, C, E) = P(D/B, C, E) \quad (2.16)$$

$$P(E/A, B, C, D) = P(E/D) \quad (2.17)$$

Pour illustrer l'équivalence de la factorisation de l'équation (2.10), et la notion de la couverture de Markov de l'équation (2.12), on dérive l'équation (2.15) de l'équation (2.11) :

$$P(C/A, B, D, E) = \frac{P(A, B, C, D, E)}{P(A, B, D, E)}$$

²La couverture de Markov est l'ensemble des fils, des parents directs, et des ancêtres d'un noeud donné

$$\begin{aligned}
&= \frac{P(A)P(B/A)P(C/A)P(D/B, C)P(E/D)}{\sum_C P(A)P(B/A)P(C/A)P(D/B, C)P(E/D)} \\
&= \frac{P(A)P(E/D)P(B/A)P(C/A)P(D/B, C)}{P(A)P(E/D)P(B/A) \sum_C P(C/A)P(D/B, C)} \\
&= \frac{P(C/A)P(D/B, C)}{\sum_C P(C/A)P(D/B, C)}
\end{aligned}$$

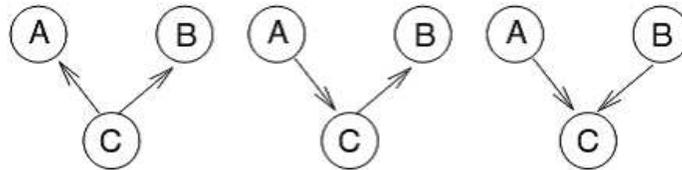


FIG. 2.10 – Trois RBs élémentaires.

Dans Figure 2.10, les RBs à gauche et au milieu ont des structures équivalentes : A et B sont conditionnellement indépendant sachant C. Le RB à droite appartient à une autre classe d'équivalence de sorte que le conditionnement sur C cause, en général, une dépendance entre A et B [9].

Considérons le RB représenté par le schéma de gauche de la figure 2.10, l'extension de la probabilité jointe suivant l'équation (2.10), donne :

$$P(A, B, C) = P(A/C)P(B/C)P(C) \quad (2.18)$$

Pour la probabilité conditionnelle $P(A, B/C)$, on obtient :

$$P(A, B/C) = \frac{P(A, B, C)}{P(C)} = P(A/C)P(B/C) \quad (2.19)$$

Donc, A et B sont *conditionnellement* indépendant sachant C. Notez, cependant, que cette indépendance ne se transforme pas aux probabilités marginales et qu'en général

$$P(A, B) \neq P(A)P(B) \quad (2.20)$$

Pour le RB représenté par le schéma de milieu de la figure 2.10, l'extension de la probabilité jointe suivant l'équation (2.11), donne :

$$P(A, B, C) = P(B/C)P(C/A)P(A) \quad (2.21)$$

Pour la probabilité conditionnelle $P(A, B/C)$, on obtient :

$$P(A, B/C) = \frac{P(A, B, C)}{P(C)} = P(B/C) \frac{P(C/A)P(A)}{P(C)} = P(B/C)P(A/C) \quad (2.22)$$

Où nous avons utilisé la règle de Bayes (2.6). Donc de nouveau, on trouve que A et B sont *conditionnellement* indépendant sachant C, tandis que, en général, cela ne se tient pas pour les probabilités marginales [9]; voir (2.20).

La situation diffère pour le RB représenté par le schéma de droit de la figure 2.10, l'extension de la probabilité jointe $P(A, B, C)$ suivant l'équation (2.11), donne :

$$P(A, B, C) = P(C/A, B)P(A)P(B) \quad (2.23)$$

La marginalisation sur C donne :

$$P(A, B) = \sum_C P(A, B, C) = P(A)P(B) \quad (2.24)$$

On a utilisé le fait qu'une fonction de probabilité est normalisée : $\sum_C P(C/A, B) = 1$. Dans ce cas et à l'opposé des deux exemples précédents que A et B sont *marginale*ment indépendants.

2.3.2 Notations et définitions

Nous utilisant les lettres majuscules pour représenter les variables ou l'ensemble des variables les lettres minuscules correspondantes sont employées pour représenter les instanciations. $P(X)$ représente la Distribution de Probabilité (DP) (Probability Distribution) de la variable aléatoire X, alors que $P(X = x)$ (abrégée en $P(x)$) représente la probabilité de l'événement $X = x$.

$Pa(X)$ est l'ensemble des parents du noeud X dans le graphe. $Pa(x)$ représente une configuration possible de ces parents.

$(X \perp Y | S)$ Signifie que X et Y sont indépendants conditionnellement à un ensemble de variables S.

$(X \perp Y)$ Signifie que X et Y sont marginalement indépendants.

Définition 2.5. Soit $G = (V, E)$ un graphe orienté sans circuit. Soit X un sommet du graphe orienté G. Soit $(\alpha_1, \dots, \alpha_q)$ une chaîne du graphe orienté G.

On dit que la chaîne $(\alpha_1, \dots, \alpha_q)$ **diverge** au sommet X si et seulement si $q > 1$ et s'il existe $r \in \{1, \dots, q\}$ tel que X est l'extrémité initiale des arcs α_r et α_{r+1} (figure 2.11).

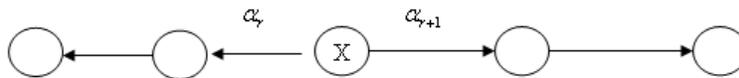


FIG. 2.11 – Exemple de chaîne divergeante en un sommet X.

Définition 2.6. On dit que la chaîne $(\alpha_1, \dots, \alpha_q)$ est **en série** au sommet X si et seulement si $q > 1$ et s'il existe $r \in \{1, \dots, q\}$ tel que X est soit l'extrémité terminale de α_r et l'extrémité initiale de α_{r+1} , soit l'extrémité initiale de α_r et l'extrémité terminale de α_{r+1} (figure 2.12).

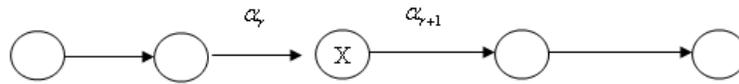


FIG. 2.12 – Exemple de chaîne en série en un sommet X.

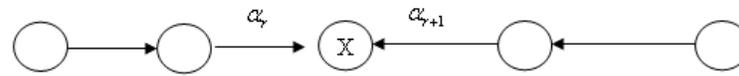


FIG. 2.13 – Exemple de chaîne convergente en un sommet X.

Définition 2.7. On dit que la chaîne $(\alpha_1, \dots, \alpha_q)$ **converge** au sommet X si et seulement si $q > 1$ et s'il existe $r \in \{1, \dots, q\}$ tel que X est l'extrémité terminale des arcs α_r et α_{r+1} (figure 2.13).

Définition 2.8. Soit $G = (V, E)$ un graphe orienté sans circuit. Soit c une chaîne du graphe orienté G . Soit Z un sous-ensemble de V . On dit que la chaîne c est **bloquée** (ou **d-séparée**) par le sous-ensemble Z si et seulement si la chaîne c vérifie au moins l'une des conditions suivantes :

1. la chaîne c diverge en un sommet de Z ;
2. la chaîne c est en série en un sommet de Z ;
3. la chaîne c converge en un sommet qui n'appartient pas à Z et dont aucun descendant propre dans le graphe G n'appartient à Z .

Dans le cas contraire, on dit que la chaîne c est **active** par rapport au sous-ensemble Z .

Dans la figure 2.14, la chaîne (α_1, α_2) est bloquée par le sous-ensemble $Z = \{P\}$. En effet, la chaîne (α_1, α_2) diverge au sommet P qui appartient au sous-ensemble $Z = \{P\}$. Dans la figure 2.15, la chaîne (β_1, β_2) est bloquée par le sous-ensemble $Z = \{P\}$. En effet, la chaîne (β_1, β_2) converge au sommet S qui n'appartient pas au sous-ensemble $Z = \{P\}$ et dont aucun descendant propre n'appartient au sous-ensemble $Z = \{P\}$.

Définition 2.9. Soit $G = (V, E)$ un graphe orienté sans circuit. Soient $X_1; X_2; X_3$ des sous-ensembles de U . On dit que les sous-ensembles X_1 et X_2 sont **d-séparés** par le sous-ensemble X_3 dans le graphe G si et seulement si les sous-ensembles $X_1; X_2; X_3$ sont deux à deux disjoints et toutes les chaînes du graphe G joignant un sommet de X_1 à un sommet de X_2 sont bloquées par le sous-ensemble X_3 . On note $\langle X_1 | X_3 | X_2 \rangle$ la **d-séparation** des sous-ensembles X_1 et X_2 par le sous-ensemble X_3 dans le graphe G

Dans la figure 2.16, Les sous-ensembles $X_1 = \{Q\}$ et $X_2 = \{R\}$ sont d-séparés par le sous-ensemble $X_3 = \{Q\}$. En effet, les chaînes (α_1, α_2) et (β_1, β_2) sont bloquées par le sous-ensemble $X_3 = \{P\}$.

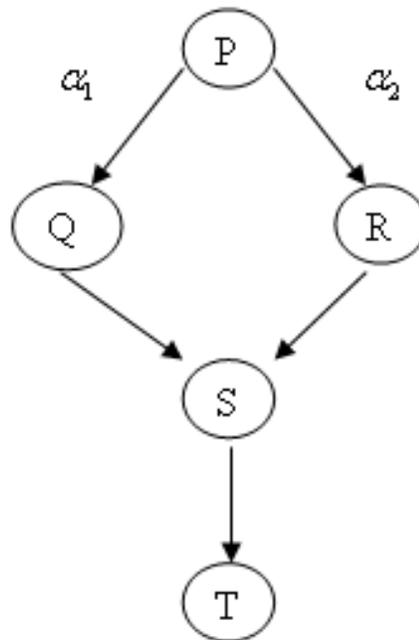


FIG. 2.14 – Exemple de chaîne bloquée par un ensemble.

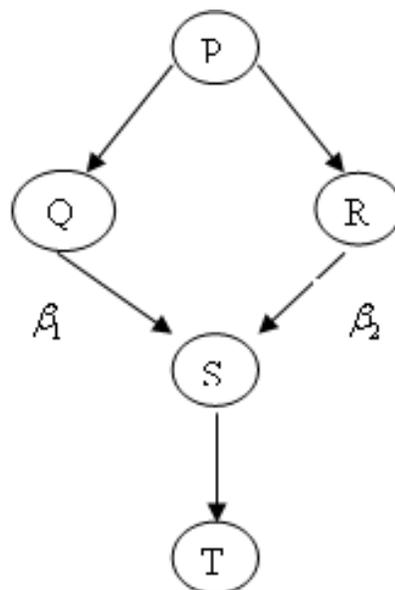


FIG. 2.15 – Autre exemple de chaîne bloquée par un ensemble.

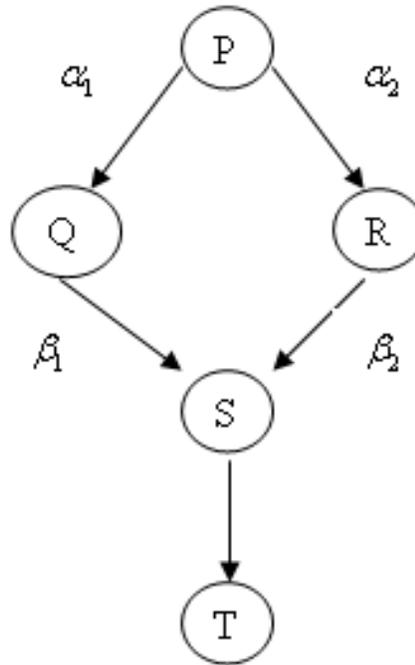


FIG. 2.16 – Exemple d'ensembles d- séparés.

2.4 Le critère de d- séparation

Considérons trois ensembles disjoints de variables X, Y et Z représentés par trois ensembles de nœuds dans un graphe acyclique dirigé G . Pour savoir si X est indépendant de Y sachant Z dans toute distribution compatible avec G , nous avons besoin de tester si des nœuds correspondants aux variables de Z **bloquent** tous les chemins allant des nœuds de X aux nœuds de Y .

Un **chemin** est une séquence consécutive d'arcs (non -dirigés) dans le graphe. Un **blocage** peut être vu comme un arrêt du flux d'informations entre les variables qui sont ainsi connectées.

Le flux d'information est dirigé par le sens des arcs et représente le flux des causalités dans le graphe, ou l'ordre dans lequel les influences vont se propager dans le graphe. Cette propagation des influences peut alors être vue comme un envoi d'information d'une variable à ses variables filles.

L'idée à la base de la d- séparation est simple quand on attribue une signification aux flèches dans le graphe. Dans la séquence $i \rightarrow m \rightarrow j$ ou dans la divergence $i \leftarrow m \rightarrow j$, si l'on conditionne m (si l'on affecte une valeur à m) alors les variables i et j qui étaient dépendantes conditionnellement à m deviennent indépendantes. Conditionner m **bloque** le flux d'information allant de i à j , c'est-à-dire qu'une nouvelle connaissance sur i ne pourra plus influencer m puisque ce dernier est maintenant connu, et donc m ne changeant plus, il n'aura plus d'influence sur j . Donc i , à travers m , n'a plus d'influence sur j non plus. Dans le cas d'une convergence $i \rightarrow m \leftarrow j$, représentant deux causes ayant le

même effet, le problème est inverse. Les deux causes sont indépendantes jusqu'à ce qu'on connaisse leur effet commun. Elles deviennent alors dépendantes.

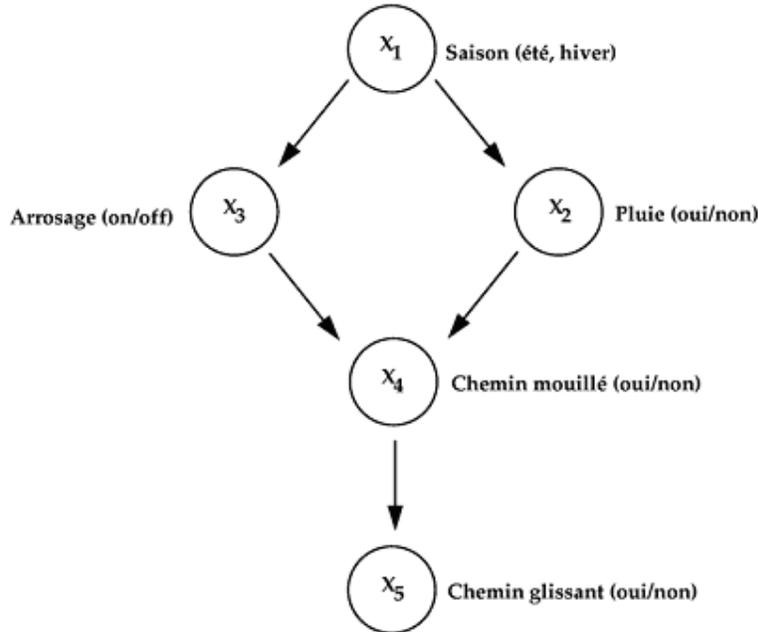


FIG. 2.17 – Un réseau bayésien représentant les dépendances entre cinq variables.

Dans la figure 2.17, si on connaît la saison X_1 alors X_2 et X_3 deviennent indépendants. Mais si on se rend compte que le chemin est glissant (X_5 est connu) ou qu'il est mouillé (X_4 est connu) alors X_2 et X_3 deviennent dépendants, car réfuter une hypothèse augmentera la probabilité de l'autre (et réciproquement).

Toujours dans la figure 2.17, $X = \{X_2\}$ et $Y = \{X_3\}$ sont d-séparés par $Z = \{X_1\}$ car les deux chemins connectant X_2 à X_3 sont bloqués par Z . Le $X_2 \leftarrow X_1 \rightarrow X_3$ chemin est bloqué car il s'agit d'une divergence dans laquelle le noeud du milieu X_1 appartient à Z . Le chemin $X_2 \rightarrow X_4 \leftarrow X_3$ est bloqué car il s'agit d'une convergence dans laquelle le noeud X_4 et tous ses descendants n'appartiennent pas à Z . Par contre, l'ensemble $Z' = \{X_1, X_5\}$, ne d-sépare pas X et Y : le chemin $X_2 \rightarrow X_4 \leftarrow X_3$ n'est pas bloqué par Z' car X_5 , qui est un descendant du noeud du milieu X_4 , appartient à Z' . On pourrait dire que le fait de connaître l'effet X_5 rend ses causes X_2 et X_3 dépendantes.

Si l'on observe une conséquence issue de deux causes indépendantes, alors les deux causes deviennent dépendantes l'une de l'autre.

Dans notre exemple, si la pluie est très forte, on pense immédiatement que c'est à cause de la pluie que le chemin est glissant. Donc on en déduit automatiquement que l'arrosage doit être mis hors de cause. De même, si le chemin est vraiment très glissant, on peut en déduire qu'il s'agit là de l'action conjuguée de la pluie et de l'arrosage. Si il est peu glissant, l'arrosage serait plutôt la cause, rendant le fait de pleuvoir quasiment improbable.

Conclusion

Selon le type d'application, l'utilisation pratique d'un réseau bayésien peut être envisagée au même titre que celle d'autres modèles : réseaux de neurones, systèmes experts, arbres de décision, etc..

Les aspects suivants des réseaux bayésiens les rendent, dans de nombreux cas, préférables à d'autres modèles : d'une part l'acquisition des connaissances, la possibilité de rassembler et de fusionner des connaissances de diverses natures dans un même modèle (retour d'expérience, expertise, observations). d'autre part ils permettent une représentation des connaissances, la représentation graphique d'un réseau bayésien est explicite, intuitive et compréhensible par un non-spécialiste.

L'inférence bayésien est basée sur l'utilisation d'énoncés probabilistes, qui dans le cas général sont trouvés par des experts étudiant un système qui leurs ait connu. Elle est particulièrement utile dans les problèmes d'induction, en se servant de la formule de Bayes. A cet effet, nous avons présenté des notions de la théorie des graphes, les propriétés des modèles graphiques, la notion de d-séparation entre ensembles de sommets d'un graphe orienté sans circuit et nous avons défini un réseau bayésien comme étant un graphe orienté sans circuit dont les sommets sont des variables aléatoires et dont les arcs permettent de représenter à l'aide de relations de d-séparation les relations d'indépendance conditionnelle.

Grâce à la définition de l'indépendance conditionnelle et aux propriétés de l'indépendance conditionnelle , nous pourrions parvenir à généraliser la définition et la construction des réseaux bayésiens au cas d'éléments aléatoires absolument quelconques.

Ainsi, nous pouvons maintenant utiliser en toute rigueur les réseaux bayésiens pour représenter et manipuler les relations d'indépendance conditionnelle.

Après avoir défini les réseaux bayésiens et étudié leur construction dans le cadre très général qui est le nôtre, nous pouvons maintenant les appliquer pour faire de l'inférence comme nous allons le voir dans le prochain chapitre.

Chapitre 3

Inférence dans les Réseaux Bayésiens

Introduction

Un réseau bayésien permet de représenter un ensemble de variables aléatoires pour lesquelles on connaît un certain nombre de relations de dépendances. Appelons $U = \{X_1, \dots, X_n\}$ l'ensemble des variables et $P(U)$ la distribution de probabilités sur cet ensemble. Si nous disposons d'une nouvelle information ε sur une ou plusieurs variables, alors on souhaiterait remettre à jour la connaissance que représente le réseau bayésien à travers $P(U)$ à la lumière de cette nouvelle information. Cette remise à jour, qui se fera bien sûr en utilisant la règle de Bayes, est appelée l'inférence. Mathématiquement parlant, l'inférence dans un réseau bayésien est le calcul de $P(U|\varepsilon)$, c'est-à-dire le calcul de la probabilité **à posteriori** du réseau sachant ε .

3.1 Spécification d'un réseau

3.1.1 Exemple

Pour illustrer notre propos, nous utiliserons un exemple issu de [2]. Il présente un modèle pour le diagnostic de l'asphyxie des nouveaux-nés. Ce domaine médical se prête bien à ce type d'analyse, car sa connaissance clinique est bonne et les données sont disponibles en grande quantité. Nous considérerons que les paramètres cliniques et le diagnostic peuvent être modélisés par des variables aléatoires, et nous aurons donc besoin de spécifier une distribution de probabilités jointe sur ces variables. Ce problème est particulièrement courant dans le domaine des systèmes experts.

La construction d'un tel modèle se décompose en trois étapes distinctes :

1. l'étape **qualitative** : on ne considère ici que les relations d'influence pouvant exister
-

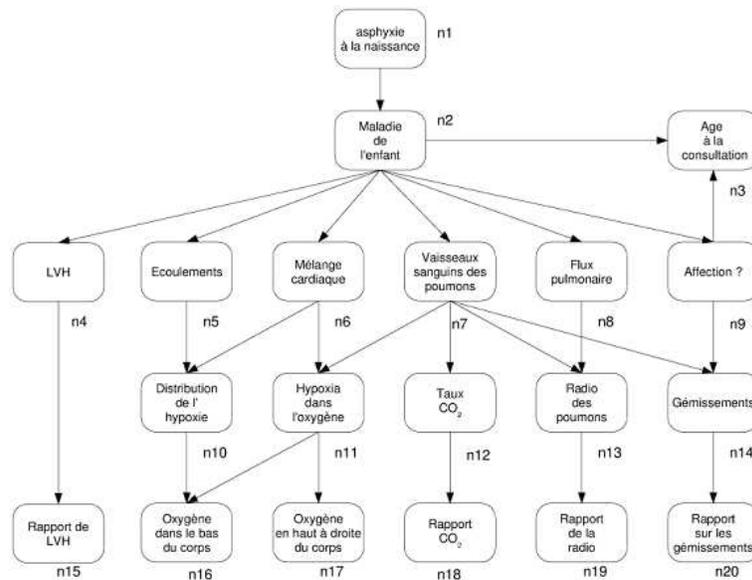


FIG. 3.1 – Représentation du problème d'asphyxie du nouveau-né.

entre les variables prises deux à deux. Ceci emmène naturellement à une représentation graphique des relations entre les variables.

2. l'étape **probabiliste** : elle introduit l'idée d'une distribution jointe définie sur les variables et fait correspondre la forme de cette distribution au graphe créé précédemment.
3. l'étape **quantitative** : elle consiste simplement à spécifier numériquement les distributions de probabilités conditionnelles.

Les maladies cardiaques congénitales peuvent être détectées à la naissance de l'enfant et sont suspectées, en général, par l'apparition de symptômes tels qu'une cyanose (le bébé devient bleu) ou un arrêt ou un dysfonctionnement cardiaque (étouffement de l'enfant). Il est alors vital que l'enfant soit transporté dans un centre spécialisé, et comme l'état de l'enfant peut se détériorer rapidement, un traitement approprié doit être administré avant le transport de l'enfant. Le diagnostic est alors fait en se basant sur les faits cliniques rapportés par le pédiatre, sur une radio, sur un ECG (électro cardiogramme) et sur une analyse sanguine.

3.1.2 Étape qualitative

L'intérêt des réseaux bayésiens est de permettre aux experts de se concentrer sur la construction d'un modèle qualitatif avant même de penser aux spécifications numériques [2].

La figure 3.1 représente le réseau bayésien modélisant ce problème de diagnostic médical. Le noeud **Maladie de l'enfant** n_2 peut prendre six valeurs correspondant aux maladies possibles dans ce cas précis de pathologie. L'arc allant du noeud n_{11} au noeud n_{16} exprime

le fait que le taux d'oxygène dans le bas du corps du patient dépend directement de l'oxygène expulsé par le patient n_{11} et de la distribution de l'hypoxie dans le corps. De même, l'oxygène expulsé n_{11} est directement influencé par l'oxygène qui est dissout dans le corps du patient n_6 et de l'état des vaisseaux sanguins dans le poumon.

3.1.3 Étape probabiliste

La spécification probabiliste du modèle passe par une représentation utilisable d'une distribution de probabilités jointe sur l'ensemble des variables. Ainsi qu'il a été montré dans la section 2.3.1 équation 2.8, la décomposition de la distribution de probabilité jointe peut se faire comme suit :

$$P(n_1, \dots, n_{20}) = \prod_{i=1}^{20} P(n_i/pa(n_i)) \quad (3.1)$$

Où les X_i sont les variables représentées par les noeuds du graphe G . La décomposition est toujours la même et permet ainsi de ne spécifier que des probabilités locales, c'est-à-dire les probabilités d'une variable sachant uniquement les variables ayant une influence directe sur elle.

3.1.4 Étape quantitative

Cette étape consiste à spécifier les tables de probabilités qui sont, pour tout i , $P(n_i/pa(n_i))$. Une table, c'est la spécification de l'ensemble des probabilités de la variable pour chacune de ses valeurs possibles sachant chacune des valeurs de ses parents. Ces probabilités sont souvent données par un expert du domaine modélisé, ou bien apprises à partir d'un corpus d'exemples. Dans notre exemple, cette étape consiste à spécifier environ 280 valeurs numériques, sachant qu'il y a 20 variables, ayant en moyenne 3 états chacune. Avec des flottants en simple précision, la mémoire nécessaire est d'environ 1,09 Ko. Si nous voulions modéliser ce problème en représentant complètement la distribution de probabilité (donc sans passer par un réseau bayésien), le nombre de valeurs numériques à spécifier serait d'environ 3^{20} , soit 3,5 milliards de valeurs. La mémoire nécessaire serait d'environ 12 Go.

3.2 Inférence

L'absence de flèches dans les RBs indique des indépendances conditionnelles qui peuvent être exploitées pour développer de meilleurs algorithmes de calcul des probabilités marginales et conditionnelles. Il y a deux problèmes de recherches principaux dans le raisonnement probabiliste utilisant des RB. Le premier problème est l'étude de l'inférence [15].

L'inférence dans un RB implique, dans une structure connue, le calcul de la probabilité marginale *a posteriori* de quelques variables sachant la valeur des variables observées. Le deuxième problème est celui de l'apprentissage de la structure qui représente les indépendances conditionnelles entre les variables.

Cette section est consacrée aux techniques d'inférence basiques. Le problème abordé consiste à savoir comment évoluent les probabilités conditionnelles d'une ou plusieurs variables X dans un réseau sachant la valeur des variables observées $Y = y$. Plus précisément, nous cherchons la probabilité $P(X/Y = y)$, où $Y = y$ est l'observation, couramment appelée *évidence*. Ces techniques sont importantes car elles servent de base pour le calcul des paramètres d'un réseau avec des variables non observées appelées *variables cachées*.

3.2.1 Les principaux algorithmes

Exact et approximatif

Les réseaux bayésiens ont été développés au début des années 1980 pour tenter de résoudre certains problèmes de prédiction et d'abduction, courants en intelligence artificielle (IA). Dans ce type de tâche, il est nécessaire de trouver une interprétation cohérente des observations avec les données connues a priori. L'inférence probabiliste signifie donc le calcul de $P(X/Y)$ où X est un ensemble d'observations et Y un ensemble de variables décrivant le problème et qui sont jugées importantes pour la prédiction ou le diagnostic. Les premiers algorithmes d'inférence pour les réseaux bayésiens ont été proposés dans [18] : il s'agissait d'une architecture à passage de messages et ils étaient limités aux arbres. Dans cette technique, à chaque noeud est associé un processeur qui peut envoyer des messages de façon asynchrone à ses voisins jusqu'à ce qu'un équilibre soit atteint, en un nombre fini d'étapes. Cette méthode a été depuis étendue aux réseaux quelconques pour donner l'algorithme de l'arbre de jonction dit JLO (de ses auteurs F.V. Jensen, S.L. Lauritzen, and K.G. Olesen).

Une autre méthode, s'appelle le **cut-set conditioning** : elle consiste à instancier un certain nombre de variables de manière à ce que le graphe restant forme un arbre. On procède à une propagation par messages sur cet arbre. Puis une nouvelle instantiation est choisie. On réitère ce processus jusqu'à ce que toutes les instantiations possibles aient été utilisées. On fait alors la moyenne des résultats. Dans la figure 2.17 du chapitre 2, si on instancie X_1 a une valeur spécifique ($X_1 = \text{hiver}$), alors le chemin entre X_2 et X_3 et passant par X_1 est **bloqué**, le réseau devient un arbre (figure 3.2). Le principal avantage de cette méthode est que son besoin en mémoire est minimal (linéaire sur la taille du réseau), alors que la méthode de l'arbre de jonction, plus rapide, a une complexité en espace exponentielle.

Bien que l'inférence dans des réseaux quelconques soit **NP difficile** [18], la complexité en temps pour chacune des méthodes citées précédemment est calculable à l'avance. Quand le résultat dépasse une limite raisonnable, on préfère alors utiliser des méthodes d'approxi-

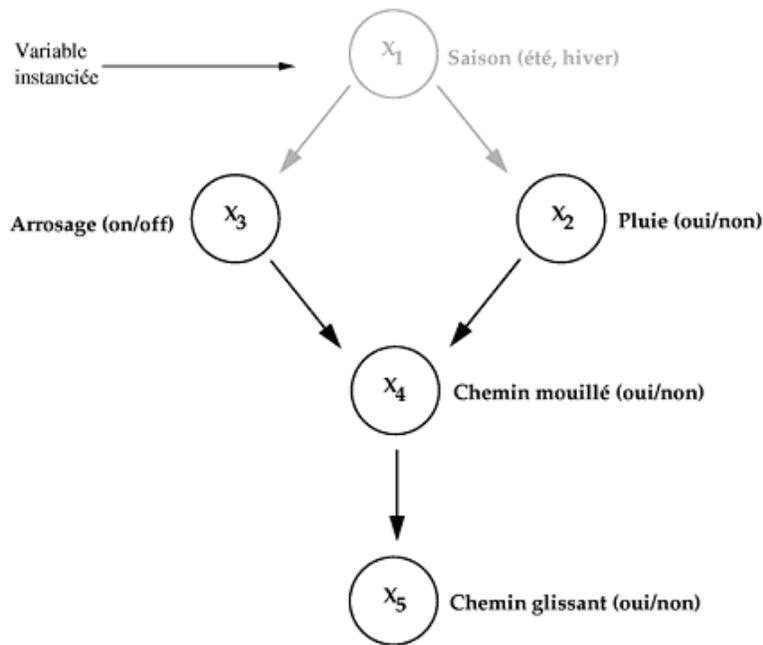


FIG. 3.2 – Graphe de la figure 2.17 transformé en arbre en instanciant X_1 .

mation. Ces méthodes exploitent la topologie du réseau et effectuent un échantillonnage de Gibbs sur des sous-ensembles locaux de variables de façon séquentielle et concurrente.

3.2.2 Approche générale de l'inférence

Soit une distribution de probabilités, le calcul de $P(Y/X)$ est trivial et nécessite une simple application de la règle de Bayes :

$$P(Y/X) = \frac{P(Y, X)}{P(X)} = \frac{\sum_{h \notin Y \cup X} P(X_H = h, Y, X)}{\sum_{h \notin Y} P(X_H = h, X)} \quad (3.2)$$

Étant donné que tout réseau bayésien définit aussi une probabilité jointe sur un ensemble de variables aléatoires, il est clair que $P(Y/X)$ peut être calculée à partir d'un DAG G . Le problème de l'inférence se réduit donc à un problème de marginalisation¹ d'une distribution de probabilités jointe. Cependant, le problème ne réside pas au niveau du calcul, mais plutôt dans son efficacité. En effet, si les variables du DAG G sont binaires, le calcul de $\sum_h P(X_1, \dots, X_n)$ prendra un temps de $O(2^N)$.

La manière la plus simple d'évaluer la probabilité cherchée est de marginaliser la fonction de probabilité jointe. Cependant, ce procédé est inefficace car quelques calculs peuvent être faits plusieurs fois. Une des techniques les plus directes est appelée *élimination de variables*.

¹Si $P(A, B, C, D)$ est une distribution de probabilités sur les variables aléatoires A, B, C et D , alors marginaliser sur D revient, pour chaque valeur de D , à faire la somme des probabilités de cette variable de manière à obtenir $P(A, B, C)$

Technique d'élimination de variables

Supposons qu'on souhaite calculer la probabilité d'une variable aléatoire X_i . L'élimination de variables consiste alors à établir un ordre dans les variables de telle façon que X_i soit la dernière. A chacune des étapes, on élimine une des variables distinctes de X_i en combinant tous les facteurs où elle est présente et en la marginalisant [6]. Le résultat final est un potentiel qui est proportionnel à la probabilité cherchée. Par exemple, considérons le réseau bayésien simple de la figure 3.3, sa probabilité jointe peut être écrite sous la forme :

$$P(A, B, C, D, F, G) = P(A) \cdot P(B/A) \cdot P(C/A) \cdot P(D/B, A) \cdot P(F/B, C) \cdot P(G/F) \quad (3.3)$$

Supposons que nous voulions marginaliser sur l'ensemble des variables sauf A afin d'obtenir $P(A)$, alors cette marginalisation serait :

$$\sum_{B,C,D,F,G} P(A, B, C, D, F, G) = \sum_B P(B/A) \cdot \sum_C P(C/A) \cdot \sum_D P(D/B, A) \cdot \sum_F P(F/B, C) \cdot \sum_G P(G/F) \quad (3.4)$$

Malgré l'apparente complexité de cette formule, le calcul de la probabilité jointe se résume à des calculs de produits très petits. En prenant la formule de droite à gauche, nous obtenons :

$$\sum_{B,C,D,F,G} P(A, B, C, D, F, G) = \sum_B P(B/A) \cdot \sum_C P(C/A) \cdot \sum_D P(D/B, A) \cdot \sum_F P(F/B, C) \cdot \lambda_{G \rightarrow F}(F) \quad (3.5)$$

Où $\lambda_{G \rightarrow F}(F) = \sum_G P(G/F)$. Dans ce cas précis, $\lambda_{G \rightarrow F}(F) = 1$ mais ce n'est pas forcément le cas à chaque fois. L'étape suivante va consister à réduire F de cette façon :

$$\sum_{B,C,D,F,G} P(A, B, C, D, F, G) = \sum_B P(B/A) \cdot \sum_C P(C/A) \cdot \lambda_{F \rightarrow C}(B, C) \cdot \sum_D P(D/B, A) \quad (3.6)$$

Où $\lambda_{F \rightarrow C}(B, C) = \sum_F P(F/B, C) \cdot \lambda_{G \rightarrow F}(F)$. En général, on essaie de calculer les termes les plus à gauche possible, de manière à minimiser le nombre de calculs nécessaires. La notation $\lambda_{F \rightarrow C}(B, C)$ signifie que l'on fait une sommation sur F et que l'on va ensuite faire une sommation sur C . Ici, C est préféré à B car il est placé plus haut dans l'ordre d'élimination des variables. Ainsi, l'ordre dans lequel les variables sont éliminées détermine la quantité de calcul nécessaire pour marginaliser la distribution de probabilités jointe. Cela influe sur la taille nécessaire pour stocker λ . Cet algorithme s'arrête quand on a marginalisé la distribution [15].

Passage de Messages en Polyarbres

L'élimination de variable est une technique qui réduit le nombre d'opérations nécessaires pour calculer la probabilité cherchée. Cependant, elle demande encore beaucoup

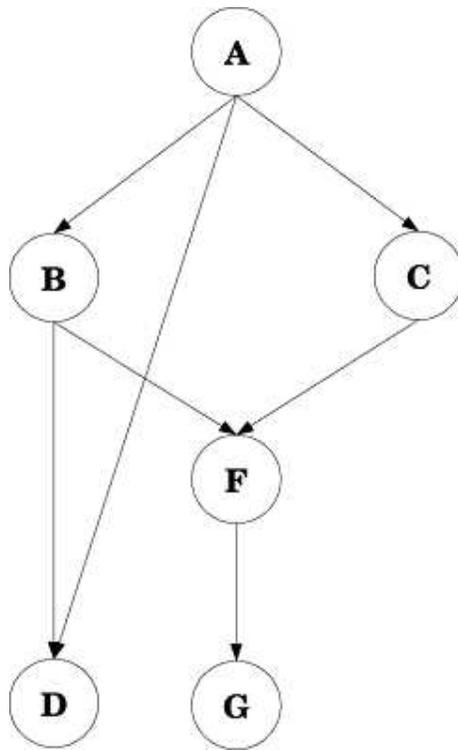


FIG. 3.3 – Un réseau bayésien simple.

des calculs. Il existe d'autres techniques plus efficaces qui tirent avantage des indépendances conditionnelles existant dans une structure en forme d'arbre [4]. Même si, dans des applications réelles, les problèmes n'induisent pas toujours une structure en forme d'arbre, les techniques appliquées à ce type de graphes sont la base pour des structures plus générales.

Le passage de message dans un polyarbre (une structure en arbre ou plusieurs racines sont permis) est une technique qui s'appuie sur la structure d'arbre du réseau et sur la d- séparation entre ses variables. En effet, dans un polyarbre, chaque noeud d sépare ses descendants de ses non- descendants. De la même façon, chaque variable est d séparée de ses frères conditionnellement à ses parents et aussi à chacun des parents de ses fils conditionnellement à ses fils.

Dans un polyarbre, il n'y a qu'un seul chemin entre deux variables. Chacun des noeuds coupe alors la structure en deux polyarbres complètement séparés. De la même façon, on coupe l'évidence en deux. On distingue alors d'une part l'évidence qui arrive à travers les parents du noeuds en question et d'autre part celle qui arrive par l'intermédiaire de ses fils. Etant donné que ce même noeud d-sépare les deux structures mentionnées, la probabilité conditionnelle cherchée est proportionnelle au produit des deux probabilités, chacune liée à une de ces structure. Ces termes qui sont appelés et peuvent être vues comme des messages envoyés par chacune des structures au noeud qui les sépare [15]. Du fait de la d-séparation, ces messages peuvent être décomposés en parties correspondant à chaque variable connectée au noeud en question. Dans le réseau, chaque variable reçoit un message de chacun de ses parentes et de chacun de ses fils. Une fois que tous les messages

ont été reçus, la variable peut envoyer, elle aussi, un message à ses voisins. A la fin de ce processus, toutes les variables ont reçues l'information provenant de l'évidence.

3.3 Algorithme de l'arbre de jonction dit JLO

Introduction

L'algorithme JLO, du nom de ses auteurs : F.V. Jensen, S.L. Lauritzen et K.G. Olesen s'applique à des réseaux ne comprenant que des variables à valeurs discrètes. L'algorithme se comporte de la façon suivante :

1. la phase de **construction** : elle nécessite un ensemble de sous-étapes permettant de transformer le graphe initial en un arbre de jonction, dont les noeuds sont des **clusters** (regroupement) de noeuds du graphe initial. Cette transformation est nécessaire, d'une part pour éliminer les boucles du graphe, et d'autre part, pour obtenir un graphe plus efficace quant au temps de calcul nécessaire à l'inférence, mais qui reste équivalent au niveau de la distribution de probabilité représentée. Cette transformation se fait en trois étapes :
 - la moralisation du graphe,
 - la triangulation du graphe et l'extraction des cliques qui formeront les noeuds du futur arbre,
 - la création d'un arbre couvrant minimal, appelé arbre de jonction ;
2. la phase de **propagation** : il s'agit de la phase de calcul probabiliste à proprement parler où les nouvelles informations concernant une ou plusieurs variables sont propagées à l'ensemble du réseau, de manière à **mettre à jour** l'ensemble des distributions de probabilités du réseau. Ceci se fait en passant des messages contenant une information de mise à jour entre les noeuds de l'arbre de jonction précédemment construit. A la fin de cette phase, l'arbre de jonction contiendra la distribution de probabilité sachant les nouvelles informations, c'est-à-dire $P(U|\varepsilon)$ où U représente l'ensemble des variables du réseau bayésien et l'ensemble des nouvelles informations sur les dites variables. ε peut, par exemple, être vu comme un ensemble d'observations faites à partir de capteurs.

Le déroulement de cet algorithme sera illustré sur l'exemple de la figure 3.1.

3.3.1 Moralisation

La première étape de transformation du graphe est la moralisation. Elle consiste à marier deux à deux les parents de chaque variable, c'est-à-dire à les relier par un arc non dirigé. Après avoir moralisé le graphe et introduit des arcs non dirigés, on finit de transformer complètement le graphe en graphe non dirigé en enlevant les directions de

chaque arc. Si G est le graphe initial, on notera G^m le graphe moralisé.

La figure 3.4 montre l'exemple de la section 3.2. Les arcs en pointillés représentent

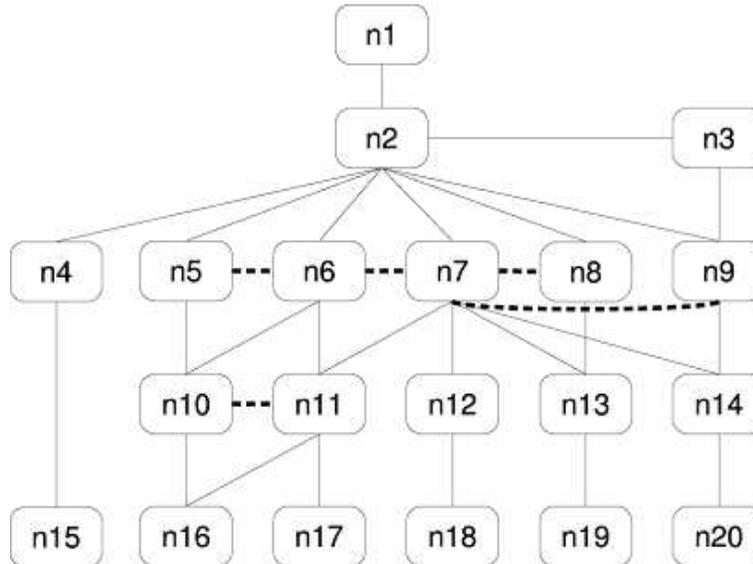


FIG. 3.4 – Graphe moralisé. Les arcs en pointillés ont été rajoutés au cours de la moralisation.

les arcs qui ont été rajoutés. La moralisation nécessite que tous les noeuds parents d'un même noeud soient reliés deux à deux.

L'idée de base est que la distribution de probabilité satisfasse aux contraintes d'indépendances conditionnelles définies par le graphe G . Cette technique de **moralisation** du graphe permet de révéler toutes les propriétés d'indépendance conditionnelle logiquement impliquées par la factorisation de la distribution jointe. Il s'agit d'une technique équivalente à celle de la d-séparation qui aboutit aux mêmes résultats. Cependant, dans le cas de la moralisation, certaines propriétés d'indépendance conditionnelle perdent leur représentation graphique dans le graphe moral G^m . Ces propriétés existent encore mais sont **cachées** dans l'ensemble des distributions de probabilités associées au graphe G .

3.3.2 Triangulation

La deuxième étape consiste à trianguler le graphe moral et à en extraire des cliques² de noeuds, qui sont des sous graphes complets de G . Ces cliques formeront les noeuds de l'arbre de jonction utilisé pour l'inférence. Il faut donc ajouter suffisamment d'arcs au graphe moral G^m afin d'obtenir un graphe triangulé G^T .

L'algorithme de triangulation opère d'une manière très simple. Un graphe est triangulé si est seulement si l'ensemble de ses noeuds peuvent être éliminés. Un noeud peut être

²une clique est un sous- graphe complètement connecté.
clique maximale : L'ajout d'un autre noeud à cette clique ne donne pas une clique.

éliminé si tous ses voisins sont connectés deux à deux. Donc un noeud peut être éliminé s'il appartient à une clique dans le graphe. Une telle clique forme un noeud pour le futur arbre de jonction qui est en train d'être construit. Ainsi, il est possible de trianguler le graphe et de construire les noeuds de l'arbre de jonction en même temps en éliminant les noeuds dans un certain ordre. Si aucun noeud n'est éliminable, il faut en choisir un parmi les noeuds restants et rajouter les arcs nécessaires entre ses voisins pour qu'il devienne éliminable. Le noeud choisi sera celui pour lequel l'espace d'état de la clique formée sera le plus petit possible. En effet, plus les cliques sont petites, plus l'espace de stockage, et a fortiori le temps de calcul, sont réduits.

L'efficacité de l'algorithme JLO reste dépendant de la qualité de la triangulation. Mais trouver une bonne triangulation dépend de l'ordre d'élimination des variables. D'une manière générale, trouver une triangulation optimale pour des graphes non dirigés reste un problème NP difficile.

Conclusion

L'inférence dans les réseaux bayésiens reste un problème NP-difficile, malgré les méthodes et techniques qu'on trouve dans la littérature qui essayent de rendre cette tâche plus ou moins à la portée de l'analyste.

On a présenté dans ce chapitre des techniques, comme la technique d'élimination de variables ou la techniques de passage de message, et l'arbre de jonction, afin de réduire le nombre de calcul.

Dans le chapitre suivant, on va procéder à l'application des réseaux bayésiens sur la chaudière.

Chapitre 4

Application pour la chaudière

Introduction

Dans ce chapitre, on va faire l'application des réseaux bayésiens sur la chaudière pour faire l'analyse du risque. Mais en premier on va faire l'application de la méthodes MADS-MOSAR pour identifier les événements non souhaités.

4.1 Identification de la chaudière

Le complexe GNL4/Z, est implanté à la zone industrielle d'Arzew, a proximité d'une population importante, le complexe dispose de 5 chaudières qui fournissent la vapeur nécessaire à l'entraînement des 9 compresseurs des lignes de liquéfaction et l'énergie électrique consommée par l'usine (puissance totale fournie : 120 000 CV). L'étude concerne l'identification des risques liés au fonctionnement de la chaudière.

Définition 4.1. *La chaudière est un dispositif permettant de chauffer l'eau et de produire de la vapeur si l'eau est chauffée au-delà de la pression atmosphérique.*

En Industrie, on utilise les chaudières pour produire la vapeur nécessaire au fonctionnement des procédés. La source de chaleur peut-être fournie par un combustible (gaz, fioul, charbon,...) ou une résistance électrique.

Les chaudières à combustible Ce type de chaudière se compose de deux compartiments distincts : l'un dans lequel brûle le combustible et un autre dans lequel l'eau est chauffée. On distingue ainsi deux types de chaudière à combustible en fonction de la circulation de l'eau à chauffer par rapport à la chaleur de combustion :

- Les chaudières à tube de fumée;
-

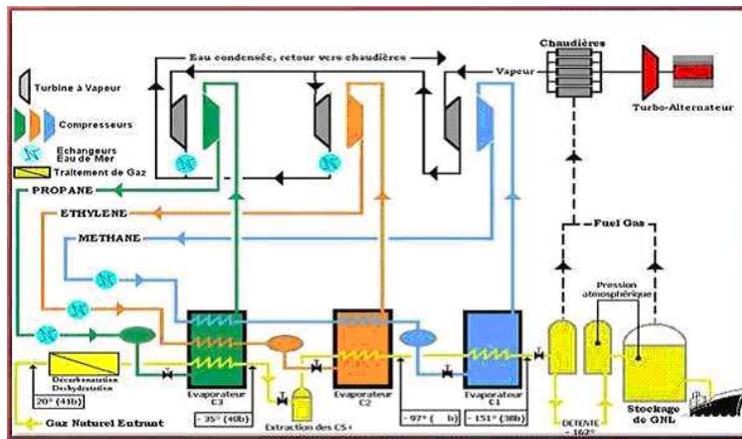


FIG. 4.1 – Schéma général retenu par le complexe GL4Z.

- Les chaudières à tube d'eau.

4.1.1 Chaudière à tube d'eau

Présentation Ce type de chaudière fournit un débit de vapeur saturée supérieur à 20 tonnes/heure, en moyenne et haute pression. Le combustible utilisé est soit du gaz, du fioul, du charbon ou déchets.

Fonctionnement Ce type de chaudière possède deux réservoirs appelés ballon distributeur (en partie inférieure) et ballon collecteur (ou encore ballon de vaporisation, en partie supérieure), reliés par un faisceau de tubes vaporisateurs, dans cet ensemble circule l'eau qui se transforme en vapeur.

Les gaz chauds produits par le brûleur sont directement en contact avec les tubes vaporisateurs, à l'intérieur de ceux-ci se produit la vaporisation. La vapeur ainsi générée est collectée dans le ballon supérieur, l'eau excédentaire est ramenée vers le ballon inférieur par des tubes de chute non soumis à la chaleur. Dans le domaine des hautes pressions, une pompe peut être installée pour faciliter cette circulation du haut vers le bas.

Caractéristiques Principales de la chaudière C'est une chaudière à tubes d'eau, du type FH 26'-24'-19' pressurisée, timbrée à 77 bars, surface tubulaire totale : $1.850 m^2$, Pression sortie surchauffeur 66.7 bars absolu, Température vapeur sèche HP 500°C, Température d'alimentation 110°C

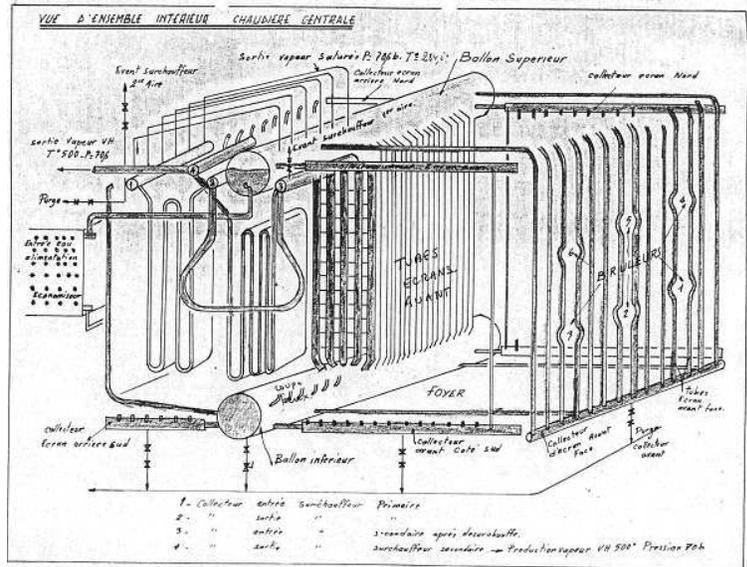


FIG. 4.2 – Vue d’ensemble intérieur chaudière centrale.

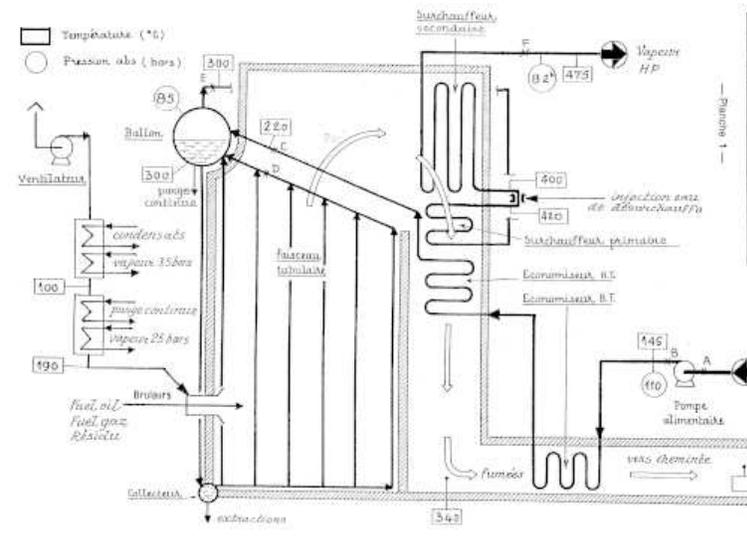


FIG. 4.3 – Condition de fonctionnement de la chaudière.

4.2 Retour d'expérience

Dans le monde les accidents des chaudières ne cesse de croître, malgré les différentes politiques mises en oeuvre pour assurer leur bon fonctionnement en toute sécurité.

On peut cité d'une manière exhaustives les accidents suivants :

En France, à l'usine de Fabrication de pâte à papier, le 03/09/2006,

Une importante émission de poussières noires se produit lors de l'ouverture du by-pass de l'électrofiltre d'une chaudière dans l'usine. Selon l'inspection, une mauvaise étanchéité de la vanne du by-pass aurait permis à de la poussière de s'accumuler derrière celle-ci.

En BANGLADESH, le 11/04/2005,

L'explosion d'une chaudière provoque l'effondrement d'une usine textile de 8 étages, faisant 47 morts et 116 disparus et blessant de nombreuses autres personnes.

RECENT HISTORY OF BOILER ACCIDENTS ACCORDING TO NATIONAL BOARD DATA					
	1996	1999	2000	2001	2002
POWER BOILERS					
SAFETY VALVE	1	1	1	4	8
LOW WATER	356	67	183	161	137
LIMIT CONTROLS	16	27	22	8	4
IMPROPER INSTALLATION	5	14	15	2	5
IMPROPER REPAIR	6	24	16	1	14
DESIGN OR FABRICATION	8	22	8	2	6
OPERATOR ERROR OR POOR MAINT.	125	140	193	82	90
BURNER FAILURE	40	27	10	29	16
HEATING BOILERS - STEAM					
SAFETY VALVE	5	2	14	2	2
LOW WATER	490	397	437	519	359
LIMIT CONTROLS	27	33	66	17	17
IMPROPER INSTALLATION	14	10	22	10	5
IMPROPER REPAIR	7	36	23	11	2
DESIGN OR FABRICATION	14	33	34	31	54
OPERATOR ERROR OR POOR MAINT.	125	258	412	406	262
BURNER FAILURE	59	20	19	29	16
HOT WATER HEATING BOILERS					
SAFETY VALVE	5	5	7	6	7
LOW WATER	112	221	258	195	96
LIMIT CONTROLS	24	68	69	19	23
IMPROPER INSTALLATION	15	31	68	13	11
IMPROPER REPAIR	3	87	28	10	2
DESIGN OR FABRICATION	20	67	40	30	60
OPERATOR ERROR OR POOR MAINT.	221	314	405	260	215
BURNER FAILURE	70	31	30	26	28

FIG. 4.4 – Statistique des pannes des chaudières dans les années antérieures [10].

BOILER ACCIDENTS, INJURIES AND DEATHS ACCORDING TO NATIONAL BOARD					
	1996	1999	2000	2001	2002
SAFETY VALVE	11	8	22	12	17
LOW WATER	958	685	878	875	592
LIMIT CONTROLS	67	128	157	44	44
IMPROPER INSTALLATION	34	55	105	25	21
IMPROPER REPAIR	16	147	67	22	18
DESIGN OR FABRICATION	42	122	82	63	120
OPERATOR ERROR OR POOR MAINT.	471	712	1011	748	567
BURNER FAILURE	169	78	59	84	60
TOTAL INCIDENTS	1768	1935	2381	1873	1439
INJURIES	56	63	24	66	16
DEATHS	4	15	8	8	3

FIG. 4.5 – Statistique des accidents, blessures et morts, entre les années 1996 et 2002 [10].

4.3 Application de la méthode MADS-MOSAR à la chaudière

La méthode MOSAR est née des travaux de P. PERILHON¹ et d'une réflexion menée par le groupe² de travail Méthodologie d'Analyse des Dysfonctionnements dans les Systèmes, ou MADS, réunissant les acteurs du Commissariat à l'Energie Atomique (ou CEA) de Grenoble, l'Institut National Supérieur des Techniques Nucléaires (ou INSTN) et l'IUT Hygiène, Sécurité, Environnement de Bordeaux. Développée au début des années 1980, elle a été appliquée avec succès dans de grandes structures telles que EDF et le CEA. La méthode MOSAR est une démarche :

- structurée, qui permet notamment, par le biais de l'analyse systémique [19], de prendre en compte les interfaces entre les éléments constituant de l'entité considérée et d'avoir ainsi une vue globale et non réductrice [19] ;
- Graduelle et progressive, "fonctionnant par paliers avec possibilité de s'arrêter provisoirement ou définitivement à chaque palier" en fonction de l'entité considérée et des objectifs visés ;
- Participative, car menée par un groupe de travail pluridisciplinaire, capitalisant les savoir-faire et faisant appel à l'imagination et à l'expérience.

Objectifs

L'objectif principal de la méthode MOSAR est d'identifier les dysfonctionnements issus de l'entité considérée ou de son environnement et dont les enchaînements peuvent conduire à un Événement Non Souhaité, ou ENS³, susceptible d'atteindre un système cible⁴. Plus largement, la méthode MOSAR cherche à "identifier, évaluer, maîtriser, gérer les processus de danger" [19] en mettant en évidence les scénarios possibles d'accidents⁵ et en déterminant les barrières de prévention et de

Démarche La méthode MOSAR est une démarche progressive composée de deux niveaux successifs :

¹Pierre PERILHON est ingénieur des Arts et Métiers, ancien responsable de l'antenne de Grenoble de l'INSTN, expert au CEA en sécurité non nucléaire et en sûreté nucléaire et concepteur de la méthode MOSAR.

²Le groupe de travail "Méthodologie d'Analyse des Dysfonctionnements dans les Systèmes" ou MADS, a été constitué en 1989 avec pour objectif de rassembler les expériences et les compétences développées depuis une vingtaine d'années en matière de maîtrise et de gestion des risques. Ce groupe compte de nombreux acteurs, au premier rang duquel se trouvent J.L. ERMINE et P. PERILHON

³Un événement est défini comme un "changement affectant la stabilité ou la forme du système et conduisant à une modification de son comportement". Un Événement Non Souhaité est défini comme l'ensemble de dysfonctionnements susceptibles de provoquer des effets non souhaités sur des systèmes cible. L'événement non souhaité est l'enchaînement de l'événement initiateur (événement à l'origine du changement d'état), événement initial (événement qui caractérise le changement d'état), événement principal (ou flux de danger) et de l'événement final (événement généré par l'impact de l'événement principal)

⁴Un système cible est un système sensible au flux de danger et subissant des dommages. La méthode MOSAR distingue quatre systèmes cible : les matériels, les individus, les populations et les écosystèmes

⁵Un scénario d'accident possible est défini par une suite d'événements qui peut logiquement survenir

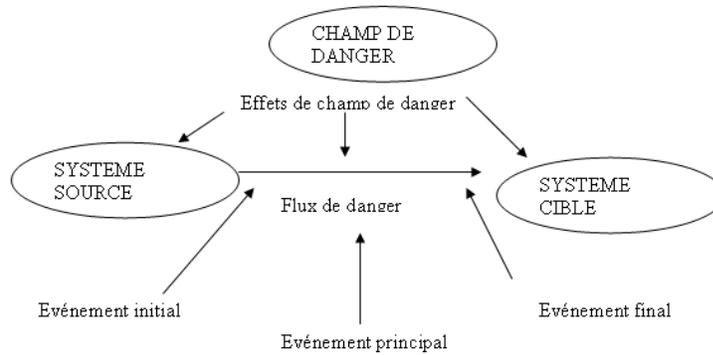


FIG. 4.6 – Schéma de processus de danger [19].

1. Un niveau macroscopique, décrit par le module A de la méthode et dont le résultat est une étude de sécurité principale de portée limitée à l'analyse des risques principaux ou "de proximité" [19] ;
2. Un niveau microscopique, décrit par le module B de la méthode, et dont la portée étend l'analyse macroscopique à l'ensemble des risques de l'entité considérée, en ayant en particulier recours à des méthodes telles que l'AMDEC pour une analyse détaillée, orientée sûreté de fonctionnement, des dysfonctionnements opératoires ou techniques.

4.3.1 Module A

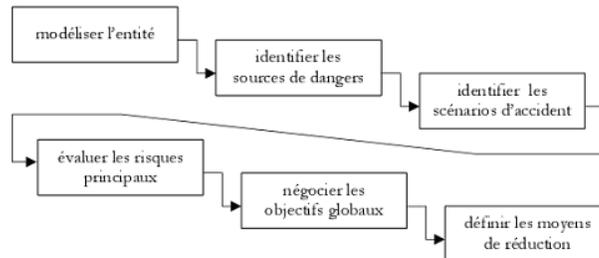


FIG. 4.7 – Niveau macroscopique de MADS-MOSAR, décrit par le module A.

4.3.2 Décomposition du système en sous systèmes

On peut décomposer la chaudière en : un ensemble de sous systèmes matériels et humains.

SS1 : Réservoir/ Ballon supérieur : composé de dispositifs d'alimentations, tubes, Level glace, le réservoir, vanne prise échantillon vapeur.

SS2 : Réservoir inférieur/ Ballon Inférieur : composé de Ballon, tube écran, collecteur écran, dispositif de purge.

SS3 : système de surchauffeur composé de collecteur de sortie surchauffeur primaire, faisceau tubulaire du surchauffeur primaire, collecteur de entrée surchauffeur primaire, collecteur de sortie surchauffeur secondaire, faisceau tubulaire du surchauffeur secondaire, collecteur de entrée surchauffeur secondaire, désurchauffeur, tubes liaison en les deux ballon.

SS4 : Brûleur composé de : détecteur de flamme, flexible entrée air, filtre à gaz, robinet à gaz.

SS5 : Equipement de chauffe composé de : ventilateur de secours, circuit alimentation de gaz, Event Inter robinet (EIR), trou d'homme chambre combustion.

SS6 : Environnement (Humain : population, le matériel : reste des installations du complexe).

SS7 : le personnel opérateur.

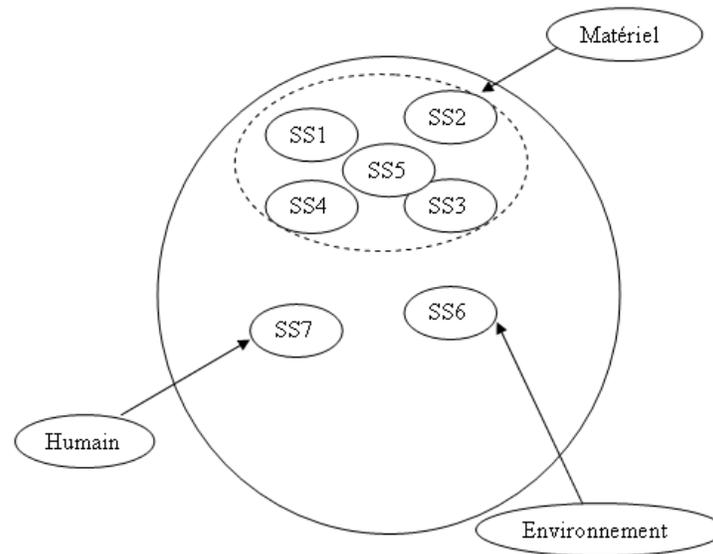


FIG. 4.8 – découpage du système en sous système.

Type de sous système source de danger	Evènements initiateurs	Evènements initiaux	Evènements Principaux
SS1 : Réservoir supérieur			
Tube alimentation eau	Corrosion Bouchage	Perçage de tube Eclat de tube	Fuite d'eau Niveau ballon sup. très bas
Vanne de régulation	Bloquée fermé	Manque d'eau	Niveau ballon sup. très bas
Level Glace	Cassure de glace	Fuite d'eau	Risque de brûlure sur l'opérateur Difficulté de lecture de niveau d'eau
Joint porte autoclave	Usure	Fuite d'eau	Risque de brûlure sur l'opérateur Niveau ballon sup. très bas
Tube sortie vapeur saturée	Corrosion	Fuite de vapeur Eclat de tube	Déstabilisation de la quantité d'énergie Risque de brûlure sur l'opérateur
Vanne prise échantillon vapeur	Usure joint	Fuite de vapeur	Risque de brûlure sur l'opérateur

TAB. 4.1 – Module A pour le SS1.

Type de sous système source de danger	Evènements initiateurs	Evènements initiaux	Evènements Principaux
SS2 : Réservoir Inférieur			
Purge	Corrosion Choc	Fuite d'eau Bouchage	Brûlure Accumulation d'impureté
Joint porte autoclave	Usure	Fuite d'eau	Risque de brûlure sur l'opérateur
Tube écran	Corrosion	Fuite d'eau	Déstabilisation de niveau réservoir inférieur
Collecteur écran	Corrosion	Fuite d'eau	Risque de brûlure sur l'opérateur

TAB. 4.2 – Module A pour le SS2.

Type de sous système source de danger	Evènements initiateurs	Evènements initiaux	Evènements Principaux
SS3 : système de surchauffe			
Collecteur entrée surchauffeur primaire	Corrosion	Perçage	Fuite de vapeur
Faisceau tubulaire surchauffeur primaire	corrosion	Perçage, Eclat Augmentation de pression	Fuite de vapeur explosion
Collecteur entrée surchauffeur secondaire	Corrosion	Perçage	Fuite de vapeur
Faisceau tubulaire surchauffeur secondaire	corrosion	Perçage, Eclat Augmentation de pression	Fuite de vapeur explosion
Désurchauffeur	Dysfonctionnement vanne de pulvérisation	Augmentation de température	Eclatement de faisceau tubulaire
 Tubes liaison	Corrosion	Fuite d'eau	explosion

TAB. 4.3 – Module A pour le SS3.

Type de sous système source de danger	Evènements initiateurs	Evènements initiaux	Evènements Principaux
SS4 : Brûleur			
Détecteur de Flamme	Poussière Usure	Pas de flamme	Brûleur éteint
Flexible entrée air	Colmate	Fuite air	Etouffement flamme
Filtre à gaz	bouchage	Pas de gaz	Brûleur éteint
Robinet à gaz	Blocage usure	Pas de gaz Trop de gaz Fuite de gaz	Min gaz Min gaz Maxi gaz incendie

TAB. 4.4 – Module A pour le SS4.

Type de sous système source de danger	Evènements initiateurs	Evènements initiaux	Evènements Principaux
SS5 : Equipement de chauffe			
Ventilateur de secours	Défaillance du moteur	Pas d'air	Maxi gaz
Circuit alimentation de gaz	Corrosion	Fuite de gaz	Incendie
Event Inter robinet (EIR)	Usure joint Corrosion	Fuite de gaz Blocage	Risque incendie Maxi gaz Min gaz
Trou d'homme chambre combustion	Usure joint	Apparition flamme	Risque brûlure sur l'opérateur

TAB. 4.5 – Module A pour le SS5.

Type de sous système source de danger	Evènements initiateurs	Evènements initiaux	Evènements Principaux
SS6 : Environnement			
Population	Incendie Emission polluants	Explosion Rejet de débris Pollution	Blessures Dégât matériel Morts Maladies respiratoires
Installations	Défaillance chaudière	Rejet de débris	explosion

TAB. 4.6 – Module A pour le SS6.

Type de sous système source de danger	Evènements initiateurs	Evènements initiaux	Evènements Principaux
SS7 : Personnel			
Opérateur	Mal intension Manque de formation Stress Excès de charge de travail	Chute Défaillance Equipement Manque de concentration	Blessures Déstabilisation du système Perte de contrôle

TAB. 4.7 – Module A pour le SS7.

4.4 Etude Statistique

4.4.1 L'analyse de la variance

L'analyse de la variance (ANOVA) a pour objectif d'étudier l'influence d'un ou plusieurs facteurs sur une variable quantitative. L'ANOVA se résume à une comparaison multiple de moyennes de différents échantillons constitués par les différentes modalités des facteurs. On présente un tableau récapitulatif des différentes étapes d'une ANOVA

Procédure de calcul

Effectifs : n_1, \dots, n_k ;

Total des effectifs : $N = \sum_{i=1}^k n_i$;

Somme des pannes : $T_i = \sum_{i=1}^k X_i$;

Total Général : $T_G = \sum_{i=1}^k T_i$;

Moyennes des pannes : $\bar{X}_T = \frac{T_G}{N}$;

Variance Totale : $SC_T = \sum_{i=1}^n (X_i - \bar{X}_T)^2 = \sum_{i=1}^N X_i^2 - \frac{T_G^2}{N}$;

Variance entre les classes (inter -groupe) : $SC_g = \sum_{j=1}^k \frac{T_j^2}{n_j} - \frac{T_G^2}{N}$;

Variance à l'intérieur de chaque classe (intra -groupe) : $SC_r = \sum_{i=1}^N X_i^2 - \sum_{j=1}^k \frac{T_j^2}{n_j}$;

Variance totale (Carrés Moyens Totaux) : $CM_T = \frac{SC_T}{N-1}$;

Variance inter- groupe (k groupe) : $CM_g = \frac{SC_g}{k-1}$;

Variance inter- groupe : $CM_r = \frac{SC_r}{N-k}$;

Statistique de Décision : $F_{cal} = \frac{CM_g}{CM_r}$;

Si $F_{cal} < F_{lu}(k-1, N-k)$ ddl et $\alpha = 0.05$ alors accepter H_0 .

Si $F_{cal} > F_{lu}(k-1, N-k)$ ddl et $\alpha = 0.05$ alors rejeter H_0 .

Dans notre cas H_0 est " le systeme fonctionne correctement "

Pour des raison de simplification, on a adopté la codification suivant la Legendre qui suit :

Composant		Anomalie	
A	Automate	b	bouchage
B	Brûleur	d	défaillance
C	Chaudière	f	fuite
F	Foyer	p	pression haute
I	Indicateur de mesure		
J	Joint		
L	Level Glace		
P	Pompe		
T	Tube		
V	Vanne		

TAB. 4.8 – Legendre des pannes par composant.

Exemple :Cf : fuite de chaudière

Le tableau 4.9 donne les statistiques issues de notre enquête sur les pannes de la chaudière au complexe GLN4/Z :

Dans le tableau 4.10, X1 ; X2 ; X3 ; X4 ; X5 ; X6 ; X7 ; X8 ; X9 ; X10 ; X11 ; X12 correspondent respectivement au panne suivantes Cf ; Vf ; Tf ; Vd ; Tb ; Ad ; Fp ; Ld ; Bd ; Id ; Pd ; Jd.

Groupe	Cf	Vf	Tf	Vd	Tb	Ad	Fp	Ld	Bd	Id	Pd	Jd
Nombres												
1	6	7	12	2	10	2	1	3	7	3	3	10
2	6	4	4	1	8	1	1	2	3	3	1	8
3	3	3	3	1	7	0	0	2	3	3	1	3
4	2	2	3	1	6	0	0	2	2	2	1	2
5	2	2	2	1	6	0	0	2	2	2	0	1
6	1	2	1	1	6	0	0	2	2	2	0	1
7	1	1	1	0	3	0	0	2	1	1	0	1
8	1	1	1	0	3	0	0	1	1	1	0	0
9	1	0	1	0	2	0	0	1	0	0	0	0
10	0	0	0	0	1	0	0	0	0	0	0	0
11	0	0	0	0	1	0	0	0	0	0	0	0
12	0	0	0	0	1	0	0	0	0	0	0	0
effectifs	12	12	12	12	12	12	12	12	12	12	12	12
total x_i	23	22	28	7	54	3	2	17	21	17	6	26
moyenne	1,92	1,83	2,33	0,58	4,50	0,25	0,17	1,42	1,75	1,42	0,50	2,17

TAB. 4.9 – données statistiques des pannes de la chaudière.

Groupe	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12
Nombres												
1	6	7	12	2	10	2	1	3	7	3	3	10
2	6	4	4	1	8	1	1	2	3	3	1	8
3	3	3	3	1	7	0	0	2	3	3	1	3
4	2	2	3	1	6	0	0	2	2	2	1	2
5	2	2	2	1	6	0	0	2	2	2	0	1
6	1	2	1	1	6	0	0	2	2	2	0	1
7	1	1	1	0	3	0	0	2	1	1	0	1
8	1	1	1	0	3	0	0	1	1	1	0	0
9	1	0	1	0	2	0	0	1	0	0	0	0
10	0	0	0	0	1	0	0	0	0	0	0	0
11	0	0	0	0	1	0	0	0	0	0	0	0
12	0	0	0	0	1	0	0	0	0	0	0	0
effectifs	12	12	12	12	12	12	12	12	12	12	12	12
total x_i	23	22	28	7	54	3	2	17	21	17	6	26
moyenne	1,92	1,83	2,33	0,58	4,50	0,25	0,17	1,42	1,75	1,42	0,50	2,17

TAB. 4.10 – représentation du tableau 4.9 en procédant à un changement de variables.

$SC_T = 723$	$CM_T = 5,058$
$SC_g = 1878$	$CM_g = 15,62$
$SC_r = 536$	$CM_r = 4,09$
$F_{cal} = \frac{CM_g}{CM_r} = 3,82$	
$(k - 1, N - k) ddl$	$(11, 132) ddl, \alpha = 0.05$
$F_{lu} = 1,86$	

TAB. 4.11 – Resultat de l'ANOVA.

$F_{cal} > F_{lu}(11, 132) ddl$ et $\alpha = 0.05$ alors rejeter H_0

4.4.2 Analyse en Composante principale

Le but de l'analyse en composante principale est d'expliquer la quantité maximale de variance avec le moins nombre de composantes principales.

Composante principale :

Les composantes principales sont des combinaisons linéaires des variables originales calculées avec le critère de variance maximale. Les composantes principales sont centrées, non corrélées et ordonnées de la plus grande variance à la plus petite.

L'analyse en composants principale est généralement utilisée comme une étape dans une série d'analyses. Par exemple, vous pouvez utiliser les composantes Principales pour réduire vos données et éviter la multi - co linéarité ou quand vous avez trop de prédictions par rapport au nombre d'observations.

Avec l'analyse en composante principale, vous pouvez découvrir souvent des relations non soupçonnées, vous permettant d'interpréter les données d'une nouvelle façon.

Analyse en Composante Principale de : X1 ; X2 ; X3 ; X4 ; X5 ; X6 ; X7 ; X8 ; X9 ; X10 ; X11 ; X12

Eigenanalysis of the Correlation Matrix									
Eigenvalue	10,201	1,070	0,406	0,163	0,093	0,043	0,020	0,004	0,001
Proportion	0,850	0,089	0,034	0,014	0,008	0,004	0,002	0,000	0,000
Cumulative	0,850	0,939	0,973	0,987	0,994	0,998	1,000	1,000	1,000
Eigenvalue	0,000	0,000	-0,000						
Proportion	0,000	0,000	-0,000						
Cumulative	1,000	1,000	1,000						

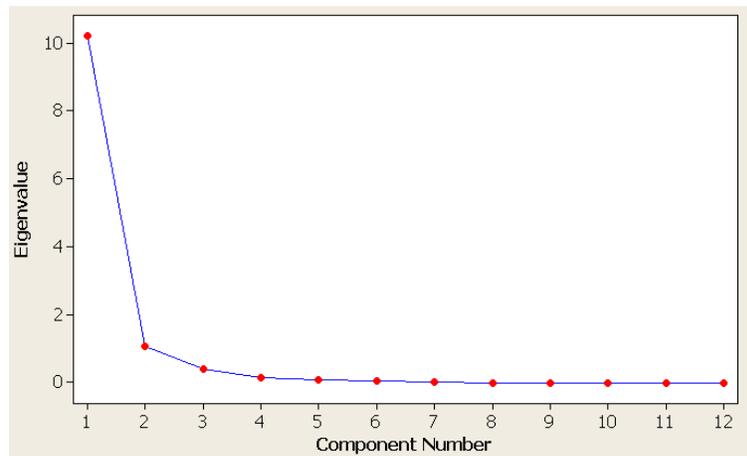


FIG. 4.9 – Valeurs propres en rapport avec les composantes principales.

Interprétation

La première composante principale a une variance 10,202 (égale à la plus grande valeur propre) et représente 0.850 (85 %) de la variation totale dans les données.

La deuxième composante principale (variance 1.070) représente 0.939 (93.9 %) de la variation de données totale.

La première composante principale avec une variance égale aux valeurs propres plus grandes que 1 représentent 0.939 (93.9 %) de la variabilité totale, suggérant que 2 composantes principales expliquent en juste proportion la variation dans les données.

4.4.3 Construction du réseau Bayésien

La construction du RB passe par trois étapes :

1. Identification des variables et de leur espace ;
2. Définition de la structure du RB ;
3. Définition de la loi de probabilité jointe des variables.

Les variables sont identifiées moyennant une discussion entre les experts et le modélisateur.

La deuxième étape consiste à identifier les liens entre variables c'est-à-dire : répondre à la question : " pour quels couples (i,j) la variable X_i influence t-elle la variable X_j ? ".

Généralement, cette étape s'effectue par l'interrogation d'experts et on se basant sur un retour d'expérience, bien sûr s'il existe.

La dernière étape consiste à renseigner les tables de probabilités des noeuds (TPN) associées aux différentes variables.

On distingue 2 cas selon la position d'une variable X_i dans le RB :

- La variable X_i n'a pas de parents, les experts doivent exprimer la loi de probabilité marginale de X_i ;

- La variable X_i possède des parents les experts doivent exprimer la dépendance de X_i en fonction des variables parents.

Le recueil de lois de probabilités auprès des experts est une étape délicate dans la construction du RB. Cependant une discussion approfondie avec les experts peut aboutir à une appréciation qualitative, c'est-à-dire lorsqu'un événement est clairement défini, les experts peuvent généralement exprimer si celui-ci est : " **probable** ", " **peu probable** ", " **hautement improbable** ", etc.

Dans ce contexte il est possible d'utiliser une table de conversion d'appréciations qualitatives en probabilités, comme l'échelle de Lichtenstein et Newman [16],(voir Figure 4.10). Il peut y avoir diverses manières de déterminer les probabilités des Noeuds. Par exemple,

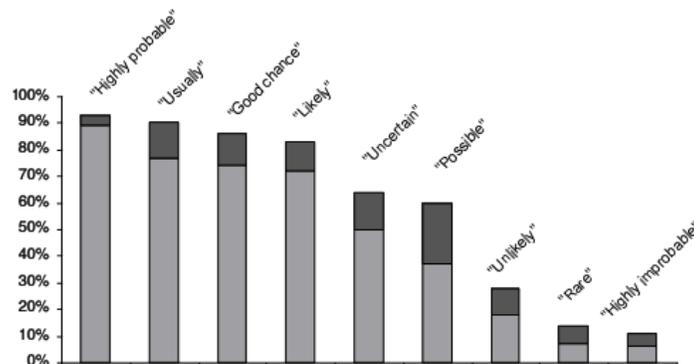


FIG. 4.10 – Correspondance entre appréciations qualitatives et probabilités.

nous pourrions tirer $P(Dysf = 'Vrai')$ d'un historique sur le dysfonctionnement de la chaudière. 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.

Un questionnaire a été donné à 2 experts pour nous aider à extraire quelques probabilités, et en s'appuyant sur l'échelle décrit dans la figure 4.10, on en déduit les valeurs possibles de différents probabilités. Les autres probabilités sont extraites du retour d'expérience.

Vu la complexité exponentielle des réseaux bayésiens, et en s'appuyant sur la méthode de l'analyse de en composante principale, et avec l'aide de deux experts, on a pu identifier le réseau bayésien correspondant aux évènements pouvant altérer le bon fonctionnement de la chaudière, et il se présente comme décrit dans la figure 4.11

Structure du réseau bayésien

On a pris la notation suivante :

Tf : variable fuite de tube ;

Vd : variable défaillance de la vanne ;

L : variable Niveau Ballon Supérieur très bas ;

Id : variable défaillance des indicateurs de mesure ;

Pcc : variable pression de la chambre de combustion très haute ;

Dysf : variable dysfonctionnement de la chaudière ;

Expl : variable risque d'explosion.

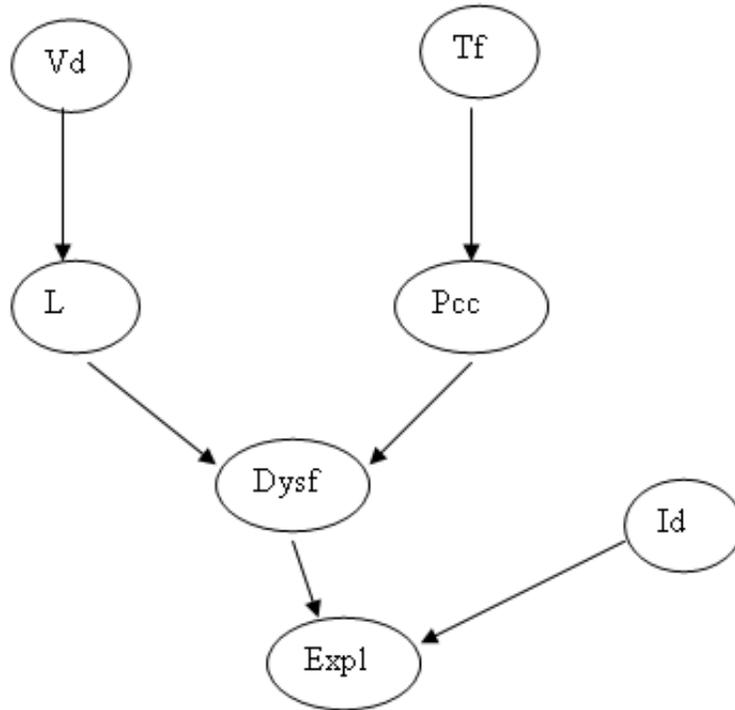


FIG. 4.11 – RB correspondant à l'état de fonctionnement de la chaudière.

Elicitation des probabilités

Vd	Tf	Id
0.6	0.75	0.55

TAB. 4.12 – Probabilités marginales des variables Vd,Tf,Id.

On a juste élicité la probabilité dans le cas où la variable prend l'état "**Vrai**", la probabilité pour l'état où la variable est fautive se déduit par complémentarité car par exemple : $P(Vd = 'Faux') = 1 - P(Vd = 'Vrai')$

Pour les probabilités conjointes, on a :

Vd	Vrai	Faux
L	0.65	0.5

TAB. 4.13 – Probabilité de L sachant Vd.

Tf	Vrai	Faux
Pcc	0.7	0.25

TAB. 4.14 – Probabilité de Pcc sachant Tf.

Pcc	Vrai		Faux	
L	Vrai	Faux	Vrai	Faux
Dysf	1	1	1	0.2

TAB. 4.15 – Probabilité de Dysf sachant L et Pcc.

Id	Vrai		Faux	
Dysf	Vrai	Faux	Vrai	Faux
Expl	0.85	0.15	0.1	0

TAB. 4.16 – Probabilité de Expl sachant Id et Dysf.

pour des raison de simplification des calculs de probabilités, on a adopté un changement de variable : soit $X_1 = Vd$, $X_2 = Tf$, $X_3 = L$, $X_4 = Pcc$, $X_5 = Dysf$, $X_6 = Id$, $X_7 = Expl$.

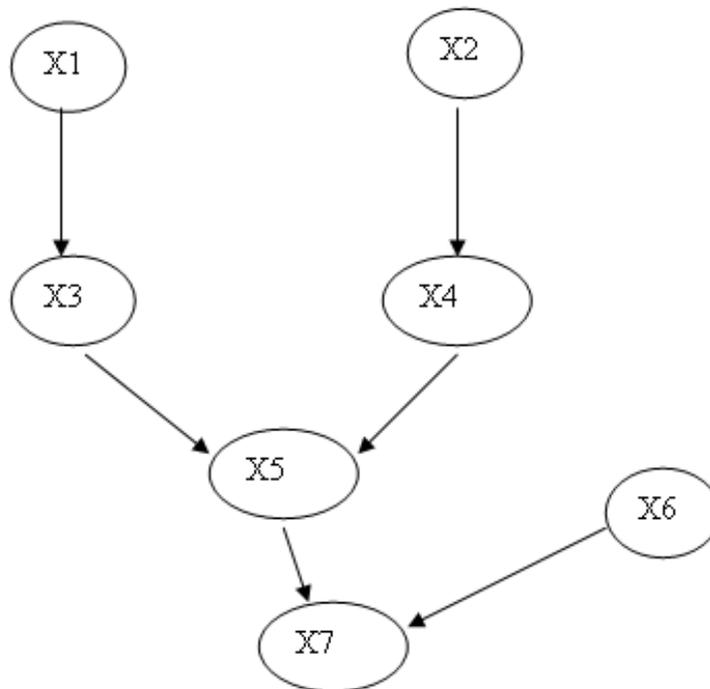


FIG. 4.12 – simplification du RB par changement de variable de la figure 4.11.

Inférence

La tâche la plus intéressante qu'on veut résoudre avec les réseaux bayésiens est *l'inférence probabiliste* ; par exemple, considérons l'état de la variable " Pression de la chambre de combustion très haute " et la variable " niveau du ballon supérieur très bas ", et supposant qu'on observe le fait qu'il y a un dysfonctionnement de la chaudière (*l'évidence*), alors selon notre RB deux possibilités se présentent : soit qu'il y a une forte pression au niveau de la chambre de combustion , soit il y a une diminution rapide du niveau du

ballon supérieur. On appliquant l'équation 2.9 sur le RB de la figure 4.12, on aura :

$$P(X_1, X_2, X_3, X_4, X_5, X_6, X_7) = \prod_{i=1}^7 P(X_i/Pa(X_i)) \quad (4.1)$$

Pour notre exemple, on a :

$$P(X_1, X_2, X_3, X_4, X_5, X_6, X_7) = P(X_1)P(X_2)P(X_3/X_1)P(X_4/X_2) \\ P(X_5/X_3, X_4)P(X_6)P(X_7/X_5, X_6) \quad (4.2)$$

On peut utiliser la règle de Bayes pour calculer la probabilité à posteriori de chaque explication (on prend 0 == Faux, 1 == Vrai), donc :

$$P(X_3 = 1/X_5 = 1) = \frac{P(X_3 = 1, X_5 = 1)}{P(X_5 = 1)} \quad (4.3)$$

$$P(X_4 = 1/X_5 = 1) = \frac{P(X_4 = 1, X_5 = 1)}{P(X_5 = 1)} \quad (4.4)$$

Le modèle peut répondre sur la question "*quelle est la vraisemblance qu'il y a une défaillance de vanne, sachant qu'il y a un dysfonctionnement de la chaudière ?*"

$$P(X_3 = 1/X_5 = 1) = \frac{\sum_{x_1, x_2, x_4 \in \{0,1\}} P(X_1 = x_1, X_2 = x_2, X_3 = 1, X_4 = x_4, X_5 = 1)}{P(X_5 = 1)} \quad (4.5)$$

$$P(X_3 = 1/X_5 = 1) = \frac{P(X_1=0, X_2=0, X_3=1, X_4=0, X_5=1) + P(X_1=0, X_2=0, X_3=1, X_4=1, X_5=1)}{P(X_5=1)} \\ + \frac{P(X_1=0, X_2=1, X_3=1, X_4=0, X_5=1) + P(X_1=0, X_2=1, X_3=1, X_4=1, X_5=1)}{P(X_5=1)} \\ + \frac{P(X_1=1, X_2=0, X_3=1, X_4=0, X_5=1) + P(X_1=1, X_2=0, X_3=1, X_4=1, X_5=1)}{P(X_5=1)} \\ + \frac{P(X_1=1, X_2=1, X_3=1, X_4=0, X_5=1) + P(X_1=1, X_2=1, X_3=1, X_4=1, X_5=1)}{P(X_5=1)} \quad (4.6)$$

Et donc, $P(X_3 = 1/X_5 = 1) = \frac{0,563}{0,8647} = 0,651$

Sachant que :

$$P(X_5 = 1) = \sum_{X_1, X_2, X_3, X_4 \in \{0,1\}} P(X_1, X_2, X_3, X_4, X_5 = 1) = 0,8647$$

4.4.4 Discussions

La probabilité de la défaillance de la vanne sachant qu'il y a un dysfonctionnement de la chaudière est plus de 50% (65,1%). Ceci nous permet d'actualiser la probabilité a priori de la variable "défaillance de la vanne", en plus il nous permet de prendre des décisions concernant la gestion des pièces de rechanges, ainsi que les précaution et les mesures à prendre afin d'éviter des accidents et incidents dont leurs conséquences seront majeurs

sur l'entreprise et l'écosystème.

En fait, l'utilisation la plus importante des réseaux bayésiens est la révision des probabilités chaque fois qu'on observe des événements. Les réseaux bayésiens offrent un formalisme intéressant lorsqu'il faut mesurer l'incertitude des scénarios qui décrivent des changements d'états d'un modèle.

En situation d'incertitude, les états du système susceptible d'arriver sont inconnus et les enchaînements causaux permettent de prévoir dans quelles conditions un scénario se réalise. Le réseau bayésien tente d'extraire de la connaissance afin d'arriver à une décision.

L'intérêt particulier des réseaux bayésiens est de tenir compte simultanément de connaissance a priori d'experts (dans le graphe) et de l'expérience contenue dans les données. L'intérêt aussi de l'utilisation des réseaux bayésiens dans l'analyse des événements incertains, c'est qu'ils permettent une description qualitative entre les différentes variables représentant ces événements (le graphe causal) et la description quantitative des dépendances entre les événements (les probabilités conditionnelles).

Le principe d'évaluation des risques adopté par GL4Z et illustré dans la figure 4.13 considère que l'indice de risque provient uniquement du produit de deux facteurs : la probabilité de son occurrence et sa gravité.

Hors pour un gestionnaire de risques, il n'est pas raisonnable d'assigner une probabilité

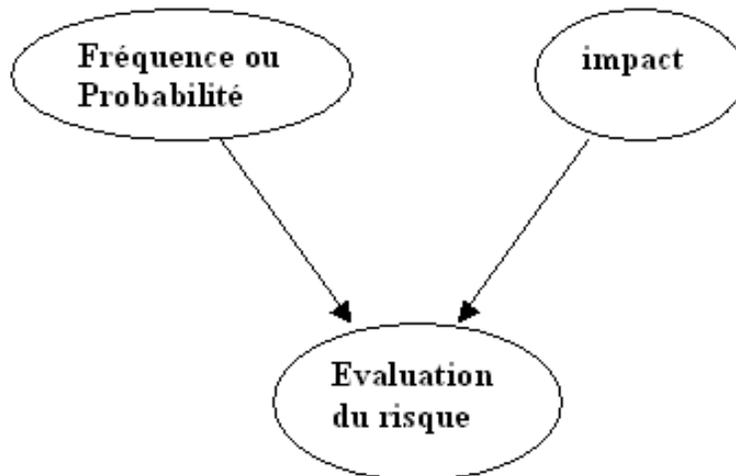


FIG. 4.13 – Principe de l'évaluation des risques adopté par GL4Z.

ou une fréquence directe de la survenue d'un incident sans considérer les événements dont il est conditionnel, par exemple les événements déclenchants et ceux permettant de réduire l'impact de l'incident.

De même, on ne peut pas estimer l'impact de l'incident sans tenir compte des facteurs ou événements autre que l'incident et qui ont une incidence directe sur la gravité du sinistre (par exemple, un incident survenu à un opérateur en heurtant un robinet de vanne, peut être de gravité différente selon l'intervention du personnel du complexe, qualifié pour as-

sister le blessé).

Certes, la mesure du risque selon la pratique de GL4Z, sert à hiérarchiser les incidents mais il reste insignifiant car il ne permet pas de nous renseigner, dans une situation donnée, des pertes humaines ou matérielles engendrées par chaque incident.

Nous concluons que la méthodes d'évaluations du risque pratiquée par GL4Z (impact x fréquence) jusqu'à présent, reste limitée et ne permet pas une amélioration continue des indicateurs de performances de la gestion la maintenance de la chaudière, et par conséquent de sa sinistralité en cas d'accident. En effet , il est difficile de quantifier les évaluations en termes de pertes humaines et de dégâts matériels d'une part, et d'autre part justifier la pertinence des actions menées et des dépenses engagées suite à l'identification des risques supposés majeurs.

Avec les réseaux bayésiens, cette évaluation sera reprise dans une approche causale intégrant à la fois :

- Le risque lui-même,
- Le ou les événements en conséquence,
- Le ou les événements de déclenchement,
- Les paramètres qui permettent de réduire l'impact du risque.

Avec cette approche de causalité, assurée à travers la construction du réseau bayésien, l'évaluation du risque causé par le dysfonctionnement de la chaudière ne tient pas compte uniquement de sa probabilité et de son impact direct, mais aussi des événements déclencheurs (fuite de tube ou défaillance de vanne), ainsi que les instruments de mesures.

En exploitant donc toutes ces informations relatives à ces éléments en relations, et dont une grande partie sont déterminées par le retour d'expérience, chaque évaluation de risque aura un sens par rapport à l'environnement où il peut arriver.

4.4.5 Conclusion

L'application des réseaux bayésiens sur la chaudière nous a permis de faire une prévision des pannes pouvant la rendre en dysfonctionnement, et ceci en faisant une analyse ascendante ou à postériori dans le réseau bayésien. Ainsi on peut désormais par le biais de cette analyse appliquer une politique de maintenance pour garantir une meilleur disponibilité par la gestion de pièce de rechange, en optant pour des pièces de durée de vie la plus maximale possible et la plus normalisée, afin de garantir une sécurité suffisante.

Conclusions et perspectives

Conclusions

Les accidents des installations industrielles notamment dans le secteur de la pétrochimie, engendrent des pertes humaines, matérielles et une atteinte à l'écosystème, malgré leur diminution à travers le monde.

La gestion des risques via la Sécurité de Fonctionnement a suscité un progrès incontestable par le recours des théories de la mathématique appliquée.

Technique mathématique combinant statistiques et intelligence artificielle, les réseaux bayésiens permettent d'analyser de grandes quantités de données pour en extraire des connaissances utiles à la prise de décision, contrôler ou prévoir le comportement d'un système, diagnostiquer les causes d'un phénomène, etc.

Les réseaux bayésiens qui tirent leur profil de la théorie des graphes et des probabilités en particulier par l'intégration de la formule de Bayes ont la capacité de faire l'actualisation des données en combinant observation et retour d'expérience.

Par le biais des Réseaux Bayésiens, on a modélisé les pannes les plus fréquentes et les plus pertinentes causant le dysfonctionnement de la chaudière par la création de liens type cause-effet, rendant ainsi possible la prédiction d'un phénomène engendrant un risque potentiel sachant des observations. La modélisation des pannes de la chaudière par les Réseaux Bayésiens est une tâche difficile malgré le recours des experts du domaine, car le regard envers l'état du système diffère d'un expert à un autre.

Les variables utilisées sont toutes de type discret à valeur binaire $\{0, 1\}$, et ceci pour limiter l'espace utilisé par les tables de probabilités des noeuds, rappelons le que cet espace augmente en exponentielle (2^N).

L'intérêt majeur des Réseaux Bayésiens est le calcul des probabilités à posteriori des variables, en se basant sur les connaissances des experts et l'observation (l'évidence) et le retour d'expérience. En effet ceci permet de capitaliser les connaissances et d'enrichir la base de données qui va servir plus tard comme données à priori pour le système.

Ainsi, au fur et à mesure on reçoit des informations sur l'état du système, on aura la capacité d'actualiser les probabilités des variables cibles. Aussi nous pouvons prévoir les probabilités des variables parentes en donnant des hypothèses sur l'état des variables descendantes.

Ce processus va permet de prévoir les mesures à prendre sur le plan d'ingénierie ou sur le plan de management sur les différents niveaux de l'entreprise, afin de prendre des dé-

cisions permettant de limiter l'apparition de risques dont leur conséquence va nuire l'état de l'entreprise que ce soit sur le stade organisationnel, structurel, et par conséquent sur sa situation financière.

Néanmoins les Réseaux Bayésiens présentent les limites suivantes :

Lisibilité des graphes : En effet, même si la connaissance manipulée dans les réseaux bayésiens, ou extraites des données par les algorithmes d'apprentissage associé est lisible puisque représentée sous forme de graphes, elle reste moins lisible que celle représentée par un arbre de décision, par exemple, surtout si ce graphe présente un grand nombre de noeuds.

La communauté de recherche travaillant sur les réseaux bayésiens n'a pas encore vraiment intégré le problème des variables continues : l'essentiel des algorithmes développés pour l'inférence et l'apprentissage dans les réseaux bayésiens, aussi bien que les outils disponibles sur le marché pour mettre en oeuvre ces algorithmes utilisent des variables discrètes.

Perspectives

L'utilisation des réseaux bayésiens est nouvelle, on compte donner à notre travail un aspect plus détaillé et plus développé en intégrant le concept de la logique floue dans la création des tables de probabilités des noeuds. Par ce biais, on espère surmonter le problème de quantification des probabilités.

Annexe A

Réglementation

Décret exécutif n° 90-245 du 18 août 1990 portant réglementation des appareils à pression de gaz, (JORA N° 36 du 22-08-1990)

Décret exécutif n° 90-246 du 18 août 1990 portant réglementation des appareils à pression de vapeur, (JORA N° 36 du 22-08-1990)

Décret exécutif n° 98-339 du 13 Rajab 1419 correspondant au 3 novembre 1998 définissant la réglementation applicable aux installations classées et fixant leur nomenclature (JORA N° 82 du 04-11-1998)

Décret exécutif n° 99-253 du 28 Rajab 1420 correspondant au 7 novembre 1999 portant composition, organisation et fonctionnement de la commission de surveillance et de contrôle des installations classées, (JORA N° 79 du 10-11-1999)

Loi n° 04-20 du 13 Dhou El Kaada 1425 correspondant au 25 décembre 2004 relative à la prévention des risques majeurs et à la gestion des catastrophes dans le cadre du développement durable, (JORA N° 84 du 29-12-2004)

Décret exécutif n° 2006-198 du 4 Joumada El Oula 1427 correspondant au 31 mai 2006 définissant la réglementation applicable aux établissements classés pour la protection de l'environnement,(JORA N° 37 du 04-06-2006)

Bibliographie

- [1] K. BACLAWSKI. Bayesian network development. *International Workshop on Software Methodologies, Tools and Techniques*, pages 18–48, September 2004.
 - [2] D. BELLOT. *Fusion de données avec des réseaux bayésiens pour la modélisation des systèmes dynamiques et son application en télémédecine*. Thèse de Doctorat, Univ. Henri Poincaré-Nancy1, Novembre 2002.
 - [3] W.M. BOLSTAD. *Introduction to BAYESIAN STATISTICS*. John Willey, wiley-interscience edition, 2004.
 - [4] CHAN P. CHEUNG, MON. *Réseaux bayésiens et filtres particulières pour l'égalisation adaptative et le décodage conjoints*. Thèse de Doctorat, Ecole Nationale Supérieure des Télécommunications de Bretagne, Decembre 2003.
 - [5] M. CHEVALIER. La sûreté de fonctionnement (sdf). *Magazine InterSections*, 2004.
 - [6] G. COZMAN, F. Generalizing variable elimination in bayesian network. In *Workshop on Probabilistic Reasoning in Artificial Intelligence*, November 2000.
 - [7] F. CROSET. *Aide à l'optimisation de la maintenance à partir des Réseaux Bayésiens et Fiabilité dans un contexte doublement censuré*. Thèse de Doctorat, Université de Joseph Fourier, Janvier 2003.
 - [8] V. DELCROIX. Les réseaux bayésiens versus d'autres modèles probabilistes pour le diagnostic multiple de systèmes complexes. *5èmes journées d'extraction et gestion des connaissances, Université René Descartes*, 2005.
 - [9] Richard DYBOWSKI Dirk HUSMEIER and Stephen ROBERTS. *Probabilistic Modeling in Bioinformatics and Medical Informatics*. Springer edition, 2005.
 - [10] K. HESELTON. *Boiler Operator's Handbook*. The fairmont press inc edition, 2005.
 - [11] M. HIBOU. Réseaux bayésiens pour la modélisation de l'apprenant en eiah : modèles multiples versus modèle unique. In *1ères Rencontres Jeunes Chercheurs en EIAH, RJC- EIAH'2006*, pages 40–47, 2006.
 - [12] P. LERAY. *Réseaux Bayésiens : Apprentissage et modélisations de systèmes complexes*. Thèse de Doctorat, Université de Rouen, 2006.
 - [13] K. MARTI. *Coping with uncertain, modeling and strategy issues*. Springer edition, 2006.
 - [14] Y. MORTUREUX. La sûreté de fonctionnement : méthodes pour maîtriser les risques. *Technique de l'Ingénieur*, 2000.
-

-
- [15] K.P. MURPHY. *Dynamic Bayesian Networks : Representation, Inference and Learning*. PhD thesis, University of California, Berkeley, 2002.
- [16] Patrick NAÏM. *Réseaux Bayésiens*. Eyrolles edition, 2004.
- [17] J. PEARL. *Probabilistic reasoning in intelligent system*. Morgan kaufman edition, 1988.
- [18] J. PEARL. Bayesian networks. In *M.A. Arbib Edition, Handbook of Brain Theory and Neural Networks*, pages 157–160, November 2000.
- [19] P. PERILHON. Proposition d'une méthode "générique" d'analyse de risques évolution d'une approche analytique déterministe et probabiliste vers une vision systémique, sociétale et réglementaire. *une Serie des Annales des Mines*, pages 53–60, Mai 2003.
- [20] G. SHAFER. *A mathematical theory of evidence*. Priceton university press edition, 1976.
- [21] E.S. SOTO. *Réseaux Bayésiens Dynamiques pour la Vérification du Locuteur*. Thèse de Doctorat, Ecole Nationale Supérieure des Télécommunications de Bretagne, Mai 2005.
- [22] SPIEGELHALTER. Bayesian analysis in expert systems. *Statistical Sciences*, (8) :219–283, 1993.
- [23] C. STARR. An introduction to bayesian belief networks and their applications to land operations. *Defence Operations Analysis Symposium*, January 2004.
- [24] A. VILLEMEUR. *Sûreté de Fonctionnement des systèmes industriels*. Dunod edition, 1988.
- [25] P. WEBER. Modélisation de processus industriels par réseaux bayésiens orientés objets - application à l'analyse des performances d'un processus industriel. *Intelligence Artificielle*, (18) :299–326, 2004.
- [26] Y. XIANG. *Probabilistic Reasoning in Multiagent Systems : A Graphical Models Approach*. Cambridge University Press, 2002.
- [27] H. ZIV. Constructing bayesian - network models of software testing and maintenance uncertainties. In *ICSM'97 International Conference on Software Maintenance*, 1997.
- [28] G. ZWINGELSTEIN. *Sûreté de fonctionnement des systèmes industriels complexes*. *Technique de l'Ingénieur*, 2000.
-

Résumé

Notre travail s'intègre dans le cadre de la prise de décision pour la stratégie de maintenance d'une chaudière, par l'application des réseaux bayésiens dans une finalité de sûreté de Fonctionnement, tout en veillant sur une meilleure disponibilité, et en préservant la sécurité des opérateurs et de l'installation dans sa totalité dans un contexte de gestion de risque.

Les relations de causalités entre les différents événements (type cause-effet) pouvant mettre la chaudière en dysfonctionnement sont prises en compte, tout en intégrant les probabilités conditionnelles, en s'appuyant sur les avis des experts du domaine et le retour d'expérience.

Les réseaux bayésiens sont devenus un outil de raisonnement sous l'incertain, le suivi de tâches tel que le diagnostic, la prédiction, et la prise de décision, font d'eux une discipline de recherche de l'intelligence artificielle en plein essor.

Le traitement des données via l'inférence nous a permis de faire une analyse montante et descendante et d'enrichir la base du retour d'expérience par l'acquisition d'observations (*évidences*).

Mots clé : Sûreté de fonctionnement, réseaux bayésiens, Chaudière, prise de décision, risque, maintenabilité, disponibilité, sécurité
