Département de formation doctorale en informatique **École doctorale IAE+M Lorraine**
UFR STMIA

Fusion de données avec des réseaux bayésiens pour la modélisation des systèmes dynamiques et son application en télémédecine

THÈSE

présentée et soutenue publiquement le 26 Novembre 2002

pour l'obtention du

Doctorat de l'université Henri Poincaré – Nancy 1

(spécialité informatique)

par

David Bellot

Composition du jury

Président	Jacques Duchêne	PR, Univ. de technologie de Troyes, LM2S
Rapporteurs	Pierre Bessière	CR CNRS, INRIA Rhône-Alpes
	Paul Rubel	PR INSA Lyon 1, Laboratoire LISI
Rapporteur interne	Jean-Pierre Thomesse	PR INPL, LORIA
Directeur de thèse	Jean-Paul Haton	PR Univ. Henri-Poincaré Nancy I, LORIA
Examinateur	Anne Boyer	MC Univ. Nancy 2, LORIA
Invité	François Charpillet	DR INRIA, LORIA

Laboratoire Lorrain de Recherche en Informatique et ses Applications - INRIA Lorraine - UMR 7503

Mis en page avec la classe thloria.

A Brigitte

A ma famille

Remerciements

Respectons la tradition et dédions cette première page aux remerciements. Bien sûr, ils ne peuvent être que nombreux, hétérogènes et incomplets. Et pourtant, ces deux pages me serviront à fusionner et à combiner toute la gratitude et l'amitié que je porte à tous ceux grâce à qui ce travail a été possible.

Tout d'abord, je tiens à remercier Jean-Paul Haton, mon directeur de thèse. Du petit livre vert que je lisais dans la cours du lycée aux réseaux bayésiens, des avions mystérieux aux patients *on-line*, je lui dois ma passion dévorante pour l'Intelligence Artificielle et toutes ses merveilleuses applications. Merci Jean-Paul pour vos précieux conseils, votre aide et cette façon si magique que vous avez de donner, à chaque fois, un éclairage nouveau à nos recherches, alors même que l'on croyait être définitivement tombé dans une impasse. Ce fut un réel plaisir et un immense honneur d'être votre étudiant.

Cette thèse a été co-encadrée par Anne Boyer et François « le chef » Charpillet : il aime bien qu'on l'apelle comme ça.

Je remercie Anne, qui a co-encadré cette thèse, pour la patience et la disponibilité dont elle a su faire preuve tout au long de nos interminables séances de travail. Elle a su écouter avec beaucoup d'attention mes folles réflexions bayésiennes. Elle m'a enseigné la précision et la rigueur dans la rédaction des textes scientifiques. Son expérience m'a permis de donner à cette thèse l'homogénéité nécessaire et son immuable gentillesse a été pour moi un réconfort et une aide précieuse durant ces années de dur labeur où nous nous sommes cotoyés.

François, mon autre directeur de thèse, m'a fait découvrir le monde de la recherche. Sa passion autant que sa patience m'ont permis d'aborder des mondes inconnus et fait connaître la vie secrète des agents intelligents. Parfois nos vues furent divergentes, mais chaque fois nos discussions furent très enrichissantes et me permirrent d'avancer dans l'univers étonnant de l'incertitude et de la perception des environnements dynamiques. Je le remercie pour son indulgence et sa patience à mon égard, il a du mérite ! Ses intuitions et ses visions ont été pour moi une très riche source d'inspiration et je pense qu'il faudra encore de nombreuses générations de thésards pour arriver, enfin, au bout de ses idées.

Jean-Pierre Thomesse a été mon rapporteur interne. Il a toujours su nous faire profiter de son extraordinaire bonne humeur et de sa terrible efficacité dans les réunions de travail TIISSAD. Il est passé maître dans cet art et je lui dois beaucoup.

Je remercie aussi Paul Rubel et Pierre Bessière d'avoir accepté de rapporter sur mon travail thèse. Paul Rubel m'a fait comprendre un point fondamental dans notre domaine : *en médecine la vérité*

n'est jamais intégralement connue. Pierre Bessière m'a donné l'occasion de poursuivre, après ma thèse, mes travaux dans le domaine des réseaux bayésiens. Je lui en suis très reconnaissant.

Monsieur le Professeur Jacques Duchêne a présidé ma soutenance et je lui dois toute ma gratitude.

La recherche n'a pas été mon seul travail : je me suis aussi initié au passionnant mais Ô combien épuisant métier d'enseignant à l'Université. Mes maîtres en ce domaine sont nombreux. Je tiens tout particulièrement à remercier Odile Mella, Brigitte Wrobel-Dautcourt et Laurent Vigneron. Durant ma dernière année de thèse, j'ai eu la chance de pouvoir être seul responsable d'un cours. Ceci a été possible grâce au soutien permanent d'Odile Thierry et de Kamel Smaïli.

Je me dois aussi de remercier tous mes collègues et amis du LORIA. Ils sont nombreux et je risque d'en oublier. Ils m'ont fait l'honneur de m'accompagner et de me supporter durant ces quatres longues années. Merci à mes collègues de bureau, Olivier, Mr OpenGL, et Romaric, l'homme aux schémas, pour leur formidable compagnie et leur bonne humeur malgré les viscicitudes de la vie d'un thésard à Nancy. Nous avons quand même passé de sacrés moments ensemble !

Je remercie Christophe A. de m'avoir fait profiter de son savoir encyclopédique sur le C++. Mes amitiés vont aussi à tous les membres de l'équipe MAIA (et des autres équipes aussi) : Eric « le Toulousain » a qui le soleil manque cruellement (et je le comprends), Rédimé, Irina, Iadine, Frank, Alain, Vincent et Vincent, Christine, Alexis, Régis, Laurent P., Luc, et plein d'autres encore...

Les rendez-vous du jeudi soir ont été épiques autant qu'inoubliables, et je garderai un formidable souvenir des terribles batailles de T.A. (et autres jeux pleins de poésie) que nous avons menées, en particulier avec Dominique, celui qui *déchnargue* les *schmolls* à coup de *Berta* (comprenne qui pourra). J'espère que dans 20 ans, tu parleras encore avec la même passion de Mhz et de gaussiennes.

Je dois un grand merci à Christophe C. pour nos longues et tardives soirées philosophiques sur l'Amour, la Vie et le devenir de Motorola ! Et un autre à mon ami Laurent, l'auteur du système 2. Je ne désespère pas de te faire aimer, un jour, Linux et les réseaux bayésiens.

Et, bien sûr, je remercie Loïc pour son amitié, le soutien qu'il m'a apporté et sa perpétuelle bonne humeur : en espérant que nos soirées gastronomiques à Nancy ne soient que les premières d'une interminable série.

Je ne saurais terminer ces deux pages sans une pensée pour ma famille et pour Brigitte, qui a la bonne idée de partager ma vie. Sans eux je n'aurais jamais pu arriver jusque-là et ils m'ont permis de réaliser ce rêve qui me tenait tant à coeur.

Table des matières

Table des figures

Introduction

L'intelligence artificielle est un artefact qui se propose d'émuler sur une machine, des comportements qui sont réputés intelligents quand on les rencontre chez un animal supérieur ou chez l'homme. Ces comportements peuvent être cognitifs, des comportements de prise de décision, de raisonnement, mais aussi de perception. Il en existe un certain nombres qui sont abordés par l'intelligence artificielle. On pourrait en particulier citer la résolution de problèmes, la représentation des connaissances et le raisonnement, l'apprentissage et l'adaptation. Mais l'intelligence artificielle ne se réduit pas seulement à ces domaines assez généraux, mais s'intéresse aussi aux phénomènes de perception de l'environnement, comme la vision, la reconnaissance de la parole, le traitement du langage naturel ou encore la compréhension de signaux issus de capteurs.

Placée dans le cadre de l'intelligence artificielle, cette thèse présente des travaux portant sur un domaine plus restreint, celui de la perception d'environnements évolutifs et incertains. Dans cette problématique, un système intelligent doit utiliser un ensemble de capteurs dirigés vers un environnement et se construire une représentation propre des entités, objets et événements de l'environnement. Le but d'un tel système intelligent est toujours la prise de décision : elle peut avoir de multiples formes comme le choix d'une action ou encore la découverte d'une information cachée. Le notion de décision est prise ici au sens le plus large. La décision servira à l'accomplissement d'une tâche ou d'une application spécifique.

La télémédecine est littéralement la médecine à distance. Parmi toutes les formes de celle-ci, nous nous intéresserons, dans cette thèse, au diagnostic à distance. Il consiste à observer un patient, qui sera considéré comme un environnement pour le système intelligent, et à estimer l'état de santé de ce patient, par rapport à la pathologie dont il souffre. Cette estimation est un diagnostic : les capteurs observent les symptômes du patient et le système intelligent en déduit l'affection dont souffre le patient. Dans le cadre particulier de la télémédecine, ce type de diagnostic nous conduira à avoir une représentation permanente de l'état du patient et à la remettre à jour chaque fois que de nouvelles observations sont disponibles. Il s'agit de *monitoring*. Cette tâche est rendue d'autant plus difficile que le patient est loin du système intelligent et du médecin, que les capteurs fournissent des données incertaines et bruitées, et enfin, qu'il est nécessaire, pour qu'un tel type de système soit utile, de pouvoir détecter les tendances à l'aggravation avant qu'un réel problème ne survienne. Ce type de détection doit permettre au médecin de pouvoir, toujours à distance, adapter la thérapie chaque fois que cela est nécessaire, et ainsi d'assurer au patient un confort et une sécurité accrue.

Cette thèse présente dans un premier temps la notion de télémédecine intelligente où le patient et le médecin sont en relation à travers un système intelligent. Ce système est là pour surveiller le

patient à travers des capteurs multiples et hétérogènes et fournir une estimation régulière sur l'état du patient. En cas de problème, il aura aussi la charge d'alerter le médecin.

Dans un deuxième temps, la problématique de la télémédecine est placée dans un cadre plus large : celui de la fusion de données pour la modélisation des systèmes dynamiques. En effet, la multiplication des capteurs, et la variété de leurs caractéristiques nous a emmené à nous intéresser au problème de la fusion d'ensembles des données dans le but de fournir un résultat cohérent et utile. Dans le cas d'un système intelligent destiné à la télémédecine, il s'agira de fusionner des données physiologiques issues de capteurs observant le patient, afin de fournir un diagnostic utilisable par le médecin pour réguler au mieux l'état de santé de son patient. Une étude complète et un cadre original pour la fusion de données seront proposés au cours du deuxième chapitre, et notre contribution se situera au niveau de la définition d'une notion de gain qualifié dans un processus de fusion de données.

Le troisième chapitre, quant à lui, se focalisera sur les modèles graphiques et plus particulièrement sur les réseaux bayésiens, qui sont une solution proposée au problème de la modélisation d'un patient, de l'utilisation et de la fusion de données issues de capteurs hétérogènes pour faire un diagnostic en environnement incertain. Une synthèse du domaine est proposée, ainsi qu'une étude de l'inférence exacte dans les réseaux bayésiens.

Ce cadre formel servira alors de base au dernier chapitre pour la mise en place d'une application de télémédecine : Diatelic^{TM}, un projet de suivi à distance de l'état d'hydratation de patients subissant une dialyse à domicile. Ce chapitre s'intéressera surtout à la modélisation de systèmes dynamiques pour faire du *monitoring*, et en particulier à l'utilisation des réseaux bayésiens dynamiques, qui permettent de raisonner sur des séquences d'observations. Ce formalisme sera appliqué au projet Diatelic^{TM}, afin de modéliser un patient et de diagnostiquer, chaque fois que de nouvelles observations sont disponibles, son état d'hydratation. Des résultats seront commentés, et l'ensemble des résultats sera proposé dans l'annexe A. Ce chapitre mettra en valeur la capacité du modèle des réseaux bayésiens à traiter ce type de problème de diagnostic en environnement incertain et la nécessité de fusionner des données hétérogènes pour obtenir un diagnostic utile au médecin.

Chapitre 1

Introduction à la télémédecine intelligente

Résumé :

L'accroissement de la durée de vie, et la réduction des durées d'hospitalisation contribuent à développer les soins à domicile. Dans le cadre de la *télémédecine intelligente*, nous nous intéressons à la construction et l'intégration de systèmes dits coopératifs où le personnel médical occupe un rôle à part entière : il s'agit de systèmes où l'être humain intervient dans le fonctionnement du système autant comme consommateur de connaissances et d'informations que comme pourvoyeur d'autres informations. Les deux types de résultats recherchés sont la détection de situations d'urgences ou d'aggravation subite de l'état du patient (et la détermination de tendances d'aggravation) et la prédiction basée sur l'analyse de données quotidiennes.

Ce travail de thèse s'est déroulé en grande partie au sein d'un projet national du nom de TIISSAD pour *Technologies de l'Information Intégrées aux Services de Soins Á Domicile*. Le projet TIISSAD s'est intéressé à la fois à la prévention du risque et à la gestion des risques avérés, de façon à améliorer la qualité de vie de ses bénéficiaires, leur sécurité, leur sérénité, et celles de ceux qui y sont attachés. L'un des objectifs du projet TIISSAD consiste à *introduire de l'intelligence dans les systèmes de télé-assistance afin de maximiser leurs fonctionnalités et leur robustesse, et d'accroître leur capacité d'aide au pronostic, au diagnostic et à la décision thérapeutique*.

Enfin, ce chapitre termine par une présentation de l'application qui nous servira de plate-forme expérimentale tout au long de cette thèse : Diatelic^{TM}. Le projet Diatelic^{TM}vise à l'utilisation de l'intelligence artificielle pour la surveillance en continu de patients souffrant d'insuffisance rénale. Les patients sont en général en attente d'une greffe de rein et doivent subir un traitement par Dialyse Péritonéale Continue Ambulatoire (DPCA). Ainsi, il est clairement apparu que le suivi de l'état d'hydratation est primordial pour assurer la survie du patient dans le meilleur état de santé possible. Pour terminer, nous présenterons succinctement les deux premiers systèmes Diatelic^{TM}(v1 et v2) développés et justifierons le besoin d'une nouvelle modélisation pouvant améliorer le suivi des patients.

> *« La télémédecine est l'apport de services de santé et de soins dans des situations où la distance est un facteur critique. Les technologies de l'information et de la communication sont utilisées pour l'échange de données dans le but de faire un diagnostic, un acte préventif ou thérapeutique, de la recherche, de l'évaluation ou encore de l'enseignement, en ayant pour souci permanent l'amélioration de l'état de santé des individus ou des communautés* [Organization, 1997]. »

1 La télémédecine intelligente

1.1 Un enjeu social et médical

Cette définition, donnée par l'Organisation Mondiale de la Santé, vise à promouvoir l'émergence d'une nouvelle forme de médecine, où les distances, autrefois facteurs critiques, sont désormais abolies pour permettre un échange et une inter-opérabilité accrue entre les professionnels de santé et les patients. Cette nouvelle notion de service vise principalement à la transmission d'informations médicales du patient vers le médecin et du médecin vers le patient, quelle que soit la distance séparant les deux protagonistes. Ce besoin d'intégrer une nouvelle forme de médecine à distance, dite *télémédecine*, est apparu progressivement avec l'avènement des nouvelles technologies de transmission de l'information numérique parallèlement à un besoin croissant de spécialistes médicaux et dans certains cas à cause d'une pénurie de ces mêmes spécialistes.

Littéralement, télémédecine signifie *médecine à distance*, mais la généralisation de l'usage du mot a coïncidé avec l'utilisation massive des technologies de communication pour le transfert d'images médicales avec une qualité rendant possible un diagnostic à distance. Ainsi la télémédecine apparaît comme étant très dépendante de l'évolution technologique, et peut être considérée comme une technique offrant une nouvelle forme d'accès aux soins médicaux et à la formation médicale. Cette évolution technologique de la médecine s'articule autour de deux aspects principaux qui sont l'aide à la décision médicale par le biais de l'expertise médicale, tout au long de la chaîne de soins, et la formation des personnels de santé, grâce à une inter-communication largement accrue entre les divers acteurs du domaine médical.

La télémédecine s'oriente vers plusieurs modalités de fonctionnement et d'action, comme l'utilisation à distance d'un robot par le biais d'échanges de données multimédia, l'échange d'informations médicales entre plusieurs praticiens (dossiers médicaux) à des fins de consultations et d'expertises, la consultation à distance de données recueillies en temps réel sur le terrain, généralement associée à une situation d'urgence, ou encore l'aide à la décision et l'interaction avec une tierce partie informatique comme un intermédiaire expert entre patients et professionnels de santé.

Cette approche pluri-modale met en avant la nécessité d'une coopération entre de nombreux acteurs venant des mondes scientifiques, médicaux, techniques et administratifs, mais aussi et surtout des vrais bénéficiaires de cette nouvelle technologie : les patients. Il s'agit alors de faire profiter, directement ou indirectement, les différents acteurs, des compétences d'une institution hospitalière, par le biais d'une demande d'expertise. Mais il peut aussi s'agir de faciliter, grâce à la technologie, un mode de prise en charge déjà existant tel qu'une demande d'expertise par envoi de

dossiers médicaux entre professionnels de santé, de l'amélioration de la surveillance à domicile par l'utilisation de capteurs et le déclenchement d'alarmes à destination du médecin traitant, de la télé-imagerie médicale dans des régions où les distances sont particulièrement grandes (cas des pays nordiques ou de l'Australie, par exemple).

L'accroissement de la durée de vie et la réduction des durées d'hospitalisation contribuent, en effet, à développer les soins à domicile tels que le Maintien À Domicile (MAD) et l'Hospitalisation À Domicile (HAD). La demande sociétale dans l'amélioration des conditions de vie, en particulier des personnes âgées mais aussi des personnes atteintes de maladies chroniques, est sans cesse croissante et appelle à la proposition de solutions nouvelles pour répondre à cette problématique.

Pour les bénéficiaires de l'HAD, le passage à la technologie de la télémédecine ne se fait pas sans intrusion anxiogène *a fortiori* pour les personnes seules. Une telle intrusion se traduit la plupart du temps par une instrumentation du patient, comme un émetteur porté en médaillon autour du cou pour les personnes âgées. L'hospitalisation à domicile doit apporter aux patients un niveau de confort et de sécurité au moins égal à celui proposé en milieu hospitalier, et si possible faire percevoir l'HAD aux patients comme une alternative optimale. Les intervenants médicaux doivent recevoir, quant à eux, un niveau de service au moins équivalent et une capacité d'action comparable à celle qu'ils sont capables de délivrer en milieu hospitalier. La sécurité est donc un facteur majeur pour la mise en place de services de télémédecine [Thomesse et al., 2002]. Les patients ont besoin d'une écoute et d'échanges continuels avec le personnel soignant qui est confronté à cette demande permanente dans un laps de temps non extensible. Il est nécessaire que le système de télémédecine puisse prendre en compte tous les facteurs et percevoir toutes les informations concernant le patient et son entourage de manière à produire un diagnostic sans faille qui ne mettra jamais en danger le patient.

Les technologies de l'information et de la communication jouent un rôle important pour atteindre les objectifs de sécurité, de fiabilité, et aussi d'économie que se donne la télémédecine. Elles peuvent servir de medium efficace pour la transmission d'informations visant à évaluer l'état de santé des patients, à détecter des anomalies ou des incidents et à informer avec pertinence le responsable médical du patient, en vue d'une intervention. Il s'agit alors de développer des systèmes de télé-assistance à domicile ou en milieu hospitalier visant à une bonne compréhension du patient et de l'évolution de son état et permettant une interaction privilégiée et immédiate avec le médecin ou le personnel de secours.

Un système de télé-assistance ou plus simplement de télémédecine pourra regrouper les fonctions suivantes :

– acquisition de données à travers toute forme de capteurs physiques ou virtuels (tels que la saisie par un patient d'une fiche de renseignements) ;
– transport des données entre le patient ou le personnel médical et le système de télémédecine. Ce dernier pourra faire office de simple routeur ou au contraire ajouter de la valeur aux informations qu'il traite. De plus, la sécurité de transfert des données doit être assurée tout au long de la chaîne de traitement des informations ;
– interaction homme-machine permettant un dialogue évolué entre la partie informatique et la partie humaine. Il s'agit non seulement d'offrir des accès basiques d'interaction avec le système de télémédecine (consultation, saisie de données, interrogation régulière de l'état de patients), mais aussi des fonctionnalités plus évoluées telles que la possibilité pour le médecin d'intervenir sur les capacités et les connaissances des systèmes d'aide à la décision ;

– stockage des données : une attention particulière doit être portée sur l'accessibilité aux données quelque soit le lieu, la méthode et/ou l'outil de consultation ;
– aide à la décision et perception : rejoignant en partie la problématique de l'interaction homme-machine, il s'agit surtout de doter le système des capacités nécessaires pour percevoir, analyser et comprendre l'état de chaque patient, en fonction de toutes les données disponibles au cours du temps. Il s'agit en particulier de la problématique générale de cette thèse : aider le médecin, au jour le jour, à décider quelle est la meilleure thérapie pour le patient, en observant le patient.

Les systèmes d'information, les bases de données et les dossiers informatisés facilitent la prise de décision en améliorant l'accès aux données pertinentes et leur mise en perspective. Mais il ne s'agit que d'une aide indirecte présentant des faits sur lesquels le décideur doit appliquer un raisonnement. Les systèmes d'aide à la décision ont l'ambition d'assister l'homme, voire de le suppléer, en remplaçant ou en reproduisant le raisonnement humain.

Une décision suppose l'utilisation de connaissances et de modèles que l'on confrontera au monde réel dans le but d'effectuer un choix. Les trois types d'informations suivantes entrent en jeu :

– les faits observés qui sont souvent entachés d'erreur et/ou d'incertitude,
– les connaissances théoriques, souvent à la base de la modélisation des phénomènes étudiés,
– l'expérience, ou connaissance pratique, qui s'acquiert au cours du temps souvent à partir de mécanismes d'apprentissage.

Une des difficultés majeures de la prise de décision, en particulier en médecine et en santé publique, vient de l'incertitude associée aux informations (connaissances, observations) qui entrent en jeu dans le processus de décision. L'incertitude intervient à plusieurs niveaux :

– sur les connaissances : certaines connaissances sont d'ordre statistique (fréquence des maladies ou des signes) et sont associées par nature à un risque d'erreur, mais d'autres connaissances sont incomplètes, par défaut d'exploration ou par insuffisance de conceptualisation,
– sur les faits : la description de l'état présent n'est jamais parfaite, soit par manque de moyens ou de temps (urgence), soit par défaut de mesure ou mauvaise interprétation d'un symptôme, d'un signe ou d'un résultat,
– au niveau du langage : certains termes utilisés sont intrinsèquement ambigus et incertains, surtout dans le cadre d'une consultation médicale : *je me sens mal, j'ai une douleur ici, etc...*

Bien que portant sur un objet précis dans le cadre d'un domaine scientifique déterminé, la décision ne peut s'abstraire de l'environnement (psychologique, social, culturel, économique) de l'objet d'étude ou de l'observateur. Les systèmes de décision peuvent s'appliquer à plusieurs types de problèmes tels que le classement ou le diagnostic, en tenant compte de l'incertitude sur la situation réelle de l'objet d'étude (patient, organe, population, etc...), ou encore le problème d'optimisation ou de planification visant à indiquer la démarche la plus efficace compte tenu de l'objectif et des contraintes pour aboutir, par exemple, à une stratégie thérapeutique.

Plusieurs modes de fonctionnement de tels systèmes sont envisageables. Dans le mode passif, par exemple, l'utilisateur interroge explicitement le système en lui fournissant les données nécessaires et attend une réponse, au moins partielle. On peut aussi envisager un mode semi-actif, où le système joue surtout un rôle de garde-fou en aidant en temps réel[1] l'opérateur humain dans sa tâche thérapeutique : rappels des règles de sécurité, contrôle des interactions médicamenteuses, alerte

[1] La notion de temps réel dépend de l'application et de la pathologie. La réaction d'un système de télémédecine peut être très variable dans le temps (seconde, jour, etc...).

sur le changement d'état du patient. Enfin on peut finalement distinguer un mode actif où le système dispose d'une complète autonomie d'action et peut intervenir autant dans la surveillance et le diagnostic que dans la thérapie. On dira alors que le système fonctionne en *boucle fermée*. Cependant, un système réel ne sera jamais complètement actif, mais intéressera l'opérateur humain au processus décisionnel et thérapeutique. Il nous semble que le médecin peut et doit être un composant actif du système, au même titre que les autres parties intelligentes. Il s'agit donc plus d'un mode de fonctionnement *coopératif* que passif ou semi-actif.

La télémédecine vise en général à mettre en relation, grâce à divers moyens de télécommunications modernes, du personnel médical et des patients, ou bien du personnel médical et des systèmes d'informations. Il s'agit dans ce cas de fournir à un patient éloigné, l'accès à une qualité de soin qu'il ne peut trouver près de chez lui (voir qu'il ne peut trouver du tout, si il n'y a pas de médecins ou d'hôpital près de son lieu de résidence). La télémédecine peut être aussi le moyen de permettre aux personnels soignants de pouvoir accéder en tout temps et en tout lieu à des informations vitales sur le patient, de manière à optimiser les soins.

Dans le cadre de la *télémédecine intelligente*, nous nous intéressons plus particulièrement à la construction et l'intégration de système experts dits *coopératifs* où le personnel médical occupe un rôle à part entière et coopère avec des machines *douées* d'un certain type de raisonnement médical. Il s'agit de systèmes où l'être humain intervient dans le fonctionnement du système autant comme consommateur de connaissances et d'informations que comme pourvoyeur d'autres informations. Le domaine visé par la télémédecine intelligente est avant tout celui du diagnostic en temps réel où le patient constitue la source principale des informations et où le médecin représente le consommateur principal. Dans de tels systèmes, le patient est observé en continu par un nombre fini de capteurs et senseurs en relation avec la pathologie concernée. Mais dans un souci d'intégration de chacun des acteurs, l'opérateur humain intervient aussi au niveau du contrôle des diagnostics et de l'apprentissage du domaine par le système. Il joue ainsi un rôle d'utilisateur et de professeur pour le système de télémédecine intelligente. Le but visé sera de nature décisionnelle : observer le patient pour évaluer et/ou prédire son état afin de prendre une décision d'ordre thérapeutique, celle-ci pouvant être prise par le système seul, ou par le médecin seul, ou par les deux conjointement.

La télémédecine intelligente se donne donc pour objectif la régulation voire la guérison du patient, mais en opérant à distance et en utilisant un intermédiaire informatique capable de suppléer le médecin dans sa tâche de surveillance médicale. Dans le cas de maladies chroniques où il n'existe pas encore de traitement curatif, le but sera de réguler au mieux l'état du patient. L'intérêt est alors de proposer au malade la formule de l'hospitalisation à domicile qui lui permettra, grâce au système de télémédecine, de vivre dans un contexte familier tout en ayant un niveau de prestation et de sécurité commun à celui d'un centre hospitalier. Ce niveau de sécurité est atteint grâce à l'effort permanent de surveillance et d'observation prodigué par le système intelligent. Les buts à atteindre peuvent être résumés de la façon suivante :

– permettre la relation patient-médecin malgré l'éloignement et la criticité de la ressource médicale,
– améliorer le confort et la sécurité du patient et du médecin,
– mettre le patient au centre du dispositif de santé et le responsabiliser,
– assurer un suivi médical *en continu*, selon trois actions : en recueillant des données dans le milieu de vie du patient, au moment pertinent (fixé par le protocole ou lors de la survenu d'un

problème), en sauvegardant ces données et en les mettant à la disposition du médecin et du patient, enfin, en suscitant des consultations entre les visites prévues par le protocole afin de prévenir un événement grave.

1.2 Un enjeu scientifique

L'objectif général de la démarche médicale à visée décisionnelle est de réduire au maximum l'incertitude quant à l'état actuel ou futur du patient par l'acquisition et le traitement d'informations et l'utilisation de connaissances sur le domaine d'intérêt. Pour parvenir à cet objectif, plusieurs moyens doivent être mis en œuvre : au niveau de l'observation, au niveau de la communication, au niveau du traitement des données et au niveau de l'interaction avec le personnel médical. Nous nous intéressons ici particulièrement au problème de l'observation ou plutôt de la perception et à celui du traitement des données et de l'interaction avec les opérateurs humains. Il s'agit donc d'introduire de l'intelligence dans les systèmes de télémédecine afin de maximiser leurs fonctionnalités et leur robustesse, et d'accroître leur capacité d'aide au pronostic, au diagnostic et à la décision thérapeutique.

La télémédecine est un vaste domaine, intégrant aussi tout ce qui touche à la médecine à distance, en particulier les robots chirurgicaux pouvant être manipulés à plusieurs milliers de kilomètres par un chirurgien. Dans le cadre de cette thèse, nous restreindrons notre étude à un problème de perception et d'aide à la décision en nous intéressant plus particulièrement aux pathologies chroniques qui relèvent d'un protocole de suivi bien identifié. On peut citer dans cette catégorie :

- *l'insuffisance rénale,* les patients dont le rein ne fonctionne plus subissent un traitement par dialyse pouvant prendre différentes formes (péritonéale à domicile, hémodialyse, etc...). Ce traitement vise à pourvoir un mécanisme de substitution aux fonctions rénales,
- *l'insuffisance cardiaque,* les patients peuvent être sujet à tout moment à une attaque cardiaque. Un tel type de pathologie est néanmoins prévisible et peut être évité ou contré si l'aggravation de cas est détecté suffisamment tôt,
- *les personnes âgées vivant éventuellement seules.* Même s'il ne s'agit pas ici d'une pathologie chronique, les problèmes et aggravations recherchés sont clairement identifiés et le problème de la surveillance de telles personnes s'apparente fortement aux deux pathologies chroniques citées précédemment. Les problèmes rencontrés sont la chute et l'aberrance ou la dégradation du comportement (fugue, marche irrégulière, malaise, changement subit des habitudes quotidiennes, ...).

Dans ces contextes, la détection de situations d'urgences ou d'aggravations subites de l'état du patient, et la détermination de tendances et la prédiction basées sur l'analyse de données quotidiennes sont les deux types de résultats recherchés. La découverte de telles situations engendre automatiquement le génération d'une alerte à l'attention du personnel médical. L'utilisation de telles données sur l'évolution du patient permet alors la constitution d'un dossier patient retraçant son historique clinique. Il s'agit donc, non seulement de détecter les situations à risque, mais aussi d'augmenter la connaissance dont on dispose concernant le patient de manière à améliorer sans cesse les performances du système intelligent. Ce but ne peut être atteint qu'en utilisant au mieux l'ensemble des données disponibles. Il s'agit donc de les combiner efficacement de manière à apporter une plus-value au flux d'informations que le système reçoit, sachant que la plupart du temps, les données reçues seront entachées d'erreurs, bruitées, incertaines voire subjectives et qu'on ne

pourra leur accorder qu'une confiance limitée. En effet, les capteurs et senseurs disponibles ne sont en aucun cas capables de fournir une information parfaite, mais plutôt une information incertaine qu'il faudra utiliser malgré tout. C'est pourquoi, le système intelligent doit être à même de *fusionner* les données et informations, quelque soit leur nature pour parvenir au comportement recherché. L'enjeu scientifique qui nous intéresse se situe alors dans le cadre de la fusion de données hétérogènes et incertaines dans le but de faire un diagnostic sur un domaine partiellement connu. Le domaine d'étude n'est lui-même que partiellement connu car la connaissance médicale est encore limitée et ne rend pas compte actuellement de la totalité de la physiologie du patient. De plus, l'état réel du patient ne peut jamais être sûr. Un diagnostic ne rend compte que d'une vérité partielle. Pour disposer d'un diagnostic complet, il faudrait pouvoir prendre en compte la totalité des paramètres dirigeant la physiologie du patient ainsi que ses antécédents, son historique et son environnement. Il est donc quasiment impossible d'obtenir un diagnostic complet. Cependant, une connaissance partielle de l'état du patient est souvent suffisante pour réguler son état ou le guérir. On peut résumer ce problème par l'assertion suivante : *dans le cadre du diagnostic médical, la vérité n'est jamais intégralement connue*.

Dans le cas de la télémédecine, le diagnostic est d'autant plus difficile à réaliser qu'il y a peu de retour de la part du personnel de santé. Le système est en effet destiné à suppléer ou assister les médecins dans leur action de diagnostic. Du fait même de la précarité de la ressource médicale, il serait donc aberrant de vouloir un retour permanent et complet sur chaque diagnostic effectué par le système intelligent. Cependant, le personnel médical étant de fait impliqué dans le processus décisionnel, il existe un retour, même partiel, qu'il faut utiliser et fusionner avec les autres informations de manière à confirmer, infirmer ou apprendre à mieux diagnostiquer.

Dans la suite du document, nous emploierons indifféremment les termes « agent intelligent », « assistant intelligent » ou « système intelligent » pour évoquer un système de télémédecine intelligent. Nous rappelons alors qu'un assistant intelligent est caractérisé par sa capacité à prendre l'initiative d'une action de façon autonome. Pour être qualifié d'*intelligent*, un assistant doit posséder des facultés de perception et de modélisation de son environnement, de raisonnement, et de prise de décision rationnelle. Dans l'objet de notre étude, l'assistant intelligent est sujet à un processus d'apprentissage continu ou régulier visant les tâches suivantes :

– fusion et présentation de données et/ou d'informations,
– surveillance automatique ou assistée du patient,
– routage intelligent des alarmes,
– aide à la décision,
– adaptation au couple médecin-patient par la création de profils patient relatifs à une pratique médicale.

Les données à traiter sont de natures différentes. En particulier, on peut citer les données issues de capteurs physiologiques (poids, température, tension artérielle, ...), ou encore des indications subjectives fournies par le patient lui-même ou par une tierce personne telles que « je ne me sens pas bien », « le patient a l'air déprimé », « le patient a une douleur à tel endroit ». On peut également mentionner des connaissances médicales d'ordre plus général telles que celles décrivant les relations entre symptômes et causes d'une maladie.

Nous utiliserons alors le terme capteur (ou l'anglicisme *senseur*) pour désigner toute source d'information destinée à être fusionnée avec les autres informations dans le but de produire un diagnostic. Un capteur peut donc être physique, réel, virtuel, mais il délivrera toujours des informations

sur le patient et son environnement. Les informations peuvent provenir directement d'un capteur physique ou être issues d'un processus de raisonnement particulier. Ce peut être aussi une consigne fixée par le médecin, ou encore une connaissance médicale ou une information issue du dossier du patient.

L'objectif de la *fusion* de telle données est donc d'améliorer la robustesse et la qualité de la *décision*, c'est-à-dire d'un diagnostic pouvant engendrer une décision thérapeutique. Il apparaît donc clairement que plus on utilise de données, plus le diagnostic se fera de façon sûre et complète. Cependant, les données étant de nature très différentes, nous sommes confrontés directement au problème de leur modélisation au sein d'un modèle unique.

Ce modèle a pour but de modéliser le processus de fusion de données et nous incite à se poser des questions de base sur la fusion de données. Pourquoi fusionner ? Que doit-on fusionner ? Comment fusionner ? Est-ce vraiment utile de fusionner ? Dans le cadre de cette thèse, ces questions d'ordre général peuvent être traduites de la façon suivante :

– quelle est la nature des données à fusionner (numériques/symboliques, continues/discrètes, certaines/incertaines, temporelles/non-temporelles, ...) ?
– quelles sont les propriétés des processus à modéliser (observables, partiellement observables, markoviens, non-markoviens, ...) ?
– quelles sont les exigences en termes de résultats (sensibilité, spécificité du système) ?

1.3 Spécificités de la télémédecine

La télémédecine est donc la médecine à distance. Dans le cadre que je viens de présenter et que je continuerai à développer plus tard avec la présentation du projet TIISSAD, certaines questions spécifiques se posent. Elles sont en relation avec l'éloignement entre le patient et le médecin, la limitation des moyens d'action à domicile, la fiabilité des mesures prises à domicile, le caractère partiel et évolutif des observations et des connaissances ou encore la variabilité inter et intra-sujet.

L'éloignement entre le médecin et le patient est bien sûr la caractéristique principale qui a donné son nom à la télémédecine. Avec la télémédecine, l'éloignement doit devenir un avantage pour le patient. En effet, la distance le séparant de son lieu de consultation médicale est abolie car, par le biais d'un système intelligent, l'expertise et la connaissance du médecin lui sont apportées à domicile. Ainsi, non seulement le patient n'est plus contraint par les déplacements, mais en plus il peut accéder à une assistance médicale de qualité, car le système de télémédecine doit être opérationnel 24h sur 24 et fournir en permanence une assistance au moins du niveau de celle que peut donner le médecin. L'interaction nécessaire entre le médecin et le système intelligent permet un transfert de connaissance et de savoir-faire dans la machine. Ainsi, chaque patient peut recevoir des soins de qualité. Ainsi, par cette approche originale de la télémédecine dite intelligente, la contrainte de l'éloignement se transforme en une forme d'ubiquité de l'expertise médicale permettant de soigner avec la même qualité et la même sécurité, un nombre plus grand de patients.

Cependant, ce type de télémédecine ne peut remplacer une véritable consultation en centre spécialisé. Dans le cas de nombreuses pathologies, il est difficile ou du moins couteux d'équiper médicalement le logement d'un patient. L'utilisation de moyens d'observation intelligents tels que ceux préconisés dans le cadre de la télémédecine intelligente (capteurs et système expert) peuvent permettre un suivi régulier et ainsi prévenir les situations à risque pour le patient. Il s'agit

donc d'une solution intermédiaire entre l'hospitalisation à domicile complète (avec les soins) et le simple suivi périodique par le médecin (consultation dans le centre médical). Pour les pathologies qui s'y prêtent, cette solution permet d'atteindre un niveau de confort et de sécurité accru pour un coût moindre qu'une hospitalisation, même à domicile.

Cependant, la télémédecine intelligente se heurte à un problème qui est toutefois courant dans d'autres disciplines : la qualité et la fiabilité des observations. Dans le cadre qui nous intéresse, les données recueillis chez le patient peuvent être sujettes à de nombreuses altérations dû à de multiples causes : défaillance d'un capteur, saisie erronée par le patient, oubli du patient de se connecter, erreur volontaire (pour dissimuler une information, comme un poids trop important par exemple), etc... Il faut donc pouvoir prendre en compte ce type de données, et partir de l'hypothèse que les données que l'on traite seront toujours incertaines. Elles sont incertaines car elles peuvent être altérées, mais elles sont aussi incertaines car dans de nombreux cas la connaissance médicale n'est pas encore suffisante pour pouvoir appréhender dans sa totalité la phénomène observé. Le type de raisonnement associé aux systèmes experts de télémédecine intelligente doit donc être capable de travailler avec des connaissances incertaines.

Cette incertitude dans les connaissances est fortement liée au caractère évolutif des phénomènes observés. La patient est un être vivant, donc en perpétuelle évolution. Assister médicalement à distance un patient doit donc être un processus répétitif dans le temps, souvent sous forme d'un cycle observation-diagnostic. Les observations faites sur le patient ne sont valables que pendant une période donnée ; le système expert doit donc remettre régulièrement à jour sa connaissance sur le sujet. Cette remise à jour doit permettre de gérer correctement l'évolution et les transformations physiologiques que subit le patient. Cette approche doit même aider à adapter, ou à laisser s'adapter, le système intelligent à de multiples patients. Il s'agit là d'une contrainte forte de la télémédecine : la réalisation d'un système intelligent doit se faire en gardant à l'esprit que ce système sera en charge de l'observation de nombreux patients aux caractéristiques les plus diverses et qui chacun ont leur propre façon d'évoluer. Ce problème sera souvent résolu par l'utilisation d'un profil patient, qui permet de résumer au sein d'un corpus de données, les caractéristiques propres du patient, et ainsi de pouvoir n'utiliser qu'un seul type de système. Le problème de la création et de l'entretien d'un profil patient reste, à son tour, dépendant de l'application (et donc, ici, de la pathologie).

2 Le projet TIISSAD

Ce travail de thèse s'est déroulé en grande partie au sein d'un projet national du nom de TIISSAD pour *Technologies de l'Information Intégrées aux Services de Soins Á Domicile*. L'objectif principal de ce projet est de définir une architecture ouverte, générique, modulaire et inter-opérable pour la construction de systèmes intelligents de monitorage à distance adapté à de multiples pathologies. Ce but implique la définition de composants élémentaires inter-opérables s'intéressant à chacun des aspects d'un système de télémédecine. On trouvera une description complète de ce projet dans [Thomesse et al., 2001]. L'un des objectifs du projet TIISSAD consiste à *introduire de l'intelligence dans les systèmes de télé-assistance afin de maximiser leurs fonctionnalités et leur robustesse, et d'accroître leur capacité d'aide au pronostic, au diagnostic et à la décision thérapeutique*. L'approche développée concerne les pathologies chroniques qui relèvent d'un protocole de suivi bien identifié. Dans ce contexte, deux types de résultats sont recherchés :

– la détection de situations d'urgences,
– la détermination de tendances et la prédiction basées sur l'analyse de données quotidiennes.

2.1 Organisation du projet TIISSAD

Ce projet a débuté au début de l'année 2000 et s'est terminé fin 2001. Le travail s'est organisé autour d'une collaboration entre plusieurs laboratoires :

– le Laboratoire Lorrain de Recherche en Informatique et ses Applications (LORIA) à Nancy,
– le TIM-C à Grenoble,
– le LAAS-CNRS à Toulouse,
– l'INSERM U558 à Toulouse et
– l'INSERM ERM107 à Lyon

ainsi que des centres médicaux :

– l'Association Lorraine pour le Traitement de l'Insuffisance Rénale (ALTIR) de Nancy,
– le CHU de Toulouse et l'Hôpital du Muret (gériatrie),
– le CHU de Grenoble (gériatrie) et
– le CHU de Lyon (cardiologie).

Le projet TIISSAD s'est intéressé à la fois à la prévention du risque et à la gestion des risques avérés, de façon à améliorer la qualité de vie de ses bénéficiaires, leur sécurité, leur sérénité, et celles de ceux qui y sont attachés.

Les objectifs de ce projet ont été d'étudier le développement de systèmes de télé-assistance à domicile de personnes âgées et de patients souffrant de l'une des deux pathologies suivantes : l'insuffisance rénale ou l'insuffisance cardiaque. Il s'agissait aussi de développer de nouvelles méthodes, de nouveaux modèles et des outils génériques les plus indépendants possibles des pathologies envisagées.

2.2 Contenu scientifique du projet

Les différents laboratoires engagés dans ce projet se sont intéressés chacun à un aspect du problème de la télémédecine, de manière à pouvoir proposer un modèle générique de système intelligent de télémédecine. En particulier, les points suivants ont été abordés à travers les travaux des équipes de recherche :

– identification, organisation et interaction des différents acteurs intervenant dans un système de télémédecine intelligente ;
– analyse et proposition de méthodes et outils pour la représentation des données et de la connaissance médicale et para-médicale, et pour la communication entre les différents acteurs quelque soit leur position géographique ;
– modélisation et création de capteurs dits intelligents : ce sont des capteurs qui fournissent des données de plus haut niveau que les seules mesures physiques, donc des données déjà traitées. Un travail de réflexion a été engagé pour déterminer les familles adéquates de capteurs pour les applications de télémédecine ;
– modélisation UML d'un système de télémédecine ;
– interfaces homme-machine ;

– analyse du principe de la fusion de données pour la création d'assistants intelligents.

Le projet a été divisé en plusieurs groupes de travail portant sur les systèmes intelligents pour l'assistance à domicile, sur la création d'un système domotique patient et sur les interfaces homme-machine. De plus, des travaux déjà en cours, réalisés par les équipes participantes, ont été intégrés au projet pour servir de base de réflexion. Ces travaux portaient sur :

– la télé-surveillance de personnes atteintes de la maladie d'Alzheimer, menée par le LAAS à Toulouse
– la télé-surveillance de certains patients insuffisants cardiaques, menée par l'INSERM de Lyon
– la télé-surveillance de dialysés à domicile, qui est le projet Diatelic^{TM}mené au LORIA.

Les résultats du projet TIISSAD ont permis la spécification d'un système générique de télé-surveillance, en particulier au travers d'une modélisation en UML. Les acteurs et leurs rôles respectifs ont pu être clairement identifiés. Ils ont aussi permis la confrontation des méthodes de traitement des données dans chaque application et la comparaison des méthodes de fusion de données, de modélisation d'environnement incertain et d'aide à la décision. Un cahier des charges pour la création d'un système domotique de santé a été réalisé. Enfin, le projet a montré l'intérêt des technologies telles que XML pour la représentation et l'échange de données médicales, quelles qu'elles soient, et aussi pour la spécification d'interfaces homme-machine.

Les principales fonctionnalités d'un système TIISSAD sont de signaler l'occurrence d'un événement normal en *fusionnant* les valeurs prises par plusieurs capteurs, et à partir des données observées et de la connaissance médicale relative à la pathologie considérée, de représenter la tendance d'évolution clinique du patient et, éventuellement, de prévoir l'état futur du patient. Le système TIISSAD doit donc permettre de faire un diagnostic, mais aussi un pronostic sur l'état du patient, ce qui permettra au personnel médical en charge d'adapter la thérapie. Le terme de *fusion de données* s'est imposé au cours du projet, car les pathologies étudiées mettent en œuvre de nombreuses sources d'informations de type très différents : mesures physiologiques (poids, tension, etc...), indications subjectives (« je vais bien », « le patient a l'air fatigué », etc...) et recommandations thérapeutiques (changement dans le traitement, ajout ou retrait d'un médicament, changement de posologie, etc...). En plus des problèmes classiques du domaine du diagnostic et du pronostic, des difficultés spécifiques ont été mises en évidence, au sein du projet, telles que :

– l'éloignement médecin/patient,
– la limitation des moyens d'action à domicile,
– la fiabilité et la confiance accordées aux mesures prises à domicile,
– le caractère partiel et évolutif des observations et des connaissances,
– la variabilité inter et intra-sujet comme les renseignements fournis par le patient, ou encore les contre-indications thérapeutiques (résistances aux médicaments, allergies, etc...).

L'objectif essentiel a donc été d'améliorer la robustesse et la qualité de la décision prise par le système intelligent intégré à un système TIISSAD. Les différentes applications ont pu être confrontées et ont montré l'apport des différents modèles utilisés, tels que les modèles à seuil, les modèles markoviens ou encore les modèles graphiques et réseaux bayésiens.

Parmi les publications faites sur les activités du projet TIISSAD, [Thomesse et al., 2001] donne un aperçu général et complet de l'ensemble du projet, [Jeanpierre and Charpillet, 2002] présente certains résultats de la version 2 du projet Diatelic^{TM}(surveillance des dialysés) et [Bellot et al., 2002a] est relatif à la version 3 du même système Diatelic^{TM}, présenté dans cette thèse.

3 Le projet Diatelic^{TM}

Basé sur une collaboration avec l'ALTIR du Centre Hospitalier Universitaire de Nancy, le projet Diatelic^{TM}vise à l'utilisation de l'intelligence artificielle pour la surveillance en continu de patients souffrant d'insuffisance rénale. Les patients sont soit en attente d'une greffe de rein, soit victime d'un disfonctionnement du rein (souvent dû à leur âge) : ils doivent alors subir un traitement par Dialyse Péritonéale Continue Ambulatoire (DPCA). Dans une première partie, nous exhiberons les principaux problèmes de la DPCA : l'hyperhydratation, la déshydratation et les péritonites (infections du péritoine). Ainsi, il est couramment admis que le suivi de l'état d'hydratation d'un patient est primordial pour assurer au patient une meilleure survie en maintenant le meilleur état de santé possible. Dans un second temps, nous présenterons le projet Diatelic^{TM}, dont l'objectif est d'assurer le suivi automatique de l'état d'hydratation des patients sous DPCA. Cette application nous servira de support expérimental tout au long de cette thèse.

3.1 La dialyse péritonéale à domicile

3.1.1 Les principales fonctions du rein

Situé de part et d'autre de la colonne vertébrale à la hauteur des fausses côtes, le rein est un organe dédié à l'élimination des déchets du sang et du retour de ce dernier sous forme purifiée dans l'organisme. Il s'agit aussi d'un organe essentiel dans le maintien du volume et de la composition ionique des fluides dans l'organisme (homéostasie). Il est capable de s'adapter à de nombreuses situations, ce qui implique que les débits et la composition des urines sont très variables et fonction du contexte. Il n'y a donc pas de composition *normale* de l'urine, mais peut-être une moyenne physiologique.

Chaque rein est constitué de plus d'un million de très petites unités appelées néphrons contenant un filtre appelé glomérule. Ces filtres débarrassent le sang de l'eau et des déchets. L'eau est renvoyée dans le corps, et les déchets sont concentrés dans l'urine.

Le rein est la principale voie d'excrétion des déchets métaboliques non volatils, dont certains sont potentiellement toxiques. C'est le cas en particulier de l'urée, l'acide urique, la créatinine ou encore l'acide oxalique. De plus le rein élimine un grand nombre de produits chimiques exogènes tels que les toxines ou les médicaments et leur métabolites : il a une action de détoxification et d'élimination de ces produits. Il participe à la régulation endocrine des volumes extra-cellulaires et de la pression artérielle. Il intervient dans le catabolisme des protéines de petit poids moléculaire, des hormones polypeptidiques (insuline, glucagon, hormone de croissance, etc...) et la régulation de la composition des fluides biologiques. Le rein est aussi le site de production de nombreuses hormones, la cible et l'effecteur endocrine d'hormones fabriquées ailleurs dans l'organisme ou dans le rein lui-même. Enfin, il assure et maintien le bilan (quantité) et la composition (concentration) ionique d'un grand nombre d'ions mono- ou divalents tel que Na^+, K^+, Ca^{2+}, Mg^{2+}, Cl^-, Li^+, H^+, $(CO_3)^{2-}$, $(PO_4)^{3+2}$

Chez l'homme adulte le rein pèse environ 150 grammes. Il comporte deux régions : le cortex et la médullaire. Dans le cortex se trouvent les glomérules qui sont des pelotons de capillaires, faisant

[2] Na^+ : sodium, K^+ : potassium, Ca^{2+} : calcium, Mg^{2+} : magnésium, Cl^- : chlore, Li^+ : litium, H^+ : hydrogène, $(CO_3)^{2-}$, $(PO_4)^{3+}$

partie d'un néphron. Ils sont utilisés par le rein pour la filtration du sang. La médullaire a une extrémité qui se projette dans la cavité excrétice (petit calice) qui fait office de sortie du rein.

Le rein est un organe richement vascularisé qui reçoit environ $\frac{1}{4}$ du débit cardiaque. Un réseau d'artères et artérioles spécifiques parcourt cet organe et participent à son fonctionnement. Les glomérules participent à l'interconnexion des artérioles et ainsi à l'échange de fluides entre le cortex et le médullaire.

L'unité fonctionnelle du rein est le néphron. Chaque rein en compte environ 1,2 million avec des variations de 0,7 à 1,5 millions déterminées génétiquement et qui pourraient expliquer la susceptibilité à certaines maladies rénales. Chaque néphron comporte un glomérule et le tubule attenant. Le glomérule assure une fonction de filtration en séparant l'eau plasmatique et ses constituants non protéiques des protéines.

De plus, le rein étant un véritable organe endocrine, il est capable de synthétiser et sécréter un grand nombre d'hormones destinées à la régulation du volume extra-cellulaire, la régulation de la pression artérielle, diverses étapes d'excrétion du sodium et du potassium, la régulation de la masse globulaire, la régulation du métabolisme minéral, etc...

3.1.2 L'insuffisance rénale chronique

La qualité de la fonction rénale est appréciée par le débit de filtration glomérulaire. La mesure de la clairance[3] de l'insuline constitue la méthode de référence de détermination du débit de filtration glomérulaire. La créatinine est une substance azotée provenant de la dégradation de la créatine. La créatine, présente dans la plupart des tissus, est synthétisée à partir d'acides aminés puis transformée dans le tissu musculaire par une enzyme : la créatine kinase. Cette enzyme sert à constituer des réserves d'énergie pour l'organisme ; on la trouve essentiellement dans les muscles. Quand le muscle est au repos, la créatine est utilisée comme source d'énergie à moyen terme. En cas de besoin, l'organisme puise dans cette réserve de créatine, qui est transformée de façon inverse pour obtenir de nouveau l'un de ses composants de base : l'adénosine triphosphate. Cette dernière molécule constitue une source d'énergie immédiatement disponible pour une activité musculaire. La créatinine est donc un déchet issue de cette production d'énergie, et elle doit être éliminée de l'organisme par le rein à travers les urines. Si son taux augmente anormalement, cela signifie que la fonction rénale est insuffisante. La clairance de la créatinine se détériore alors jusqu'à ce que le patient se trouve en situation d'insuffisance rénale : les reins ne jouent plus leur rôle normal de filtres. La clairance de la créatinine traduit la capacité que possèdent les reins à filtrer le sang pour en extraire et évacuer la créatinine. Cette capacité varie en fonction de l'âge, de la quantité d'eau bue, de l'état de santé, mais surtout de la capacité d'épuration des glomérules.

L'insuffisance rénale chronique est l'aboutissement d'un processus de destruction progressive des néphrons (et donc des glomérules qu'ils contiennent). Ceux restants, supposés fonctionnels, vont tenter de s'adapter, mais cette adaptation s'avère délétère et on assiste alors à un processus d'auto-aggravation. Lorsque la réduction néphronique[4] atteint 90 %, l'insuffisance rénale chronique est dite terminale et nécessite une dialyse voire une transplantation de rein. L'insuffisance rénale est

[3] coefficient d'épuration qui correspond à l'aptitude à éliminer.

[4] Réduction néphronique : destruction des néphrons

donc une insuffisance excrétoire aboutissant à la rétention de substances normalement éliminées dans l'urine.

La *dialyse* est un traitement pour les personnes à un stade avancé de l'insuffisance rénale chronique. Ce traitement vise à remplacer la fonction de filtrage et d'épuration du sang, fonctions normalement assurées par le rein, en éliminant de l'organisme les déchets et l'excès d'eau. La dialyse peut être un traitement temporaire, mais lorsque les reins cessent complètement de fonctionner, la dialyse doit être effectuée à intervalles réguliers. Le seul autre traitement connu à l'heure actuelle est la greffe de rein.

Actuellement, il existe deux types de dialyses, que nous allons maintenant présenter : l'*hémodialyse* où le sang est purifié en passant à travers un rein artificiel (machine de dialyse), la *dialyse péritonéale* où la membrane péritonéale ou péritoine fait office de filtre (le sang est donc filtré à l'intérieur du corps).

L'hémodialyse

L'hémodialyse nécessite un rein artificiel, le dialyseur, qui contient la membrane d'épuration. L'épuration se fait par échange entre le sang et un bain de dialyse fabriqué et contrôlé par un générateur. Cette technique nécessite un abord vasculaire[5] fournissant un débit élevé, facile d'accès. La fistule artério-veineuse créée par un chirurgien au niveau de l'avant-bras est le meilleur système. Elle sera piquée avec deux aiguilles pour le départ et le retour du sang. La séance dure en moyenne 4 heures. Trois séances hebdomadaires nécessaires ont lieu en centre spécialisé, en centre d'auto-dialyse ou à domicile après apprentissage. La figure 1.1 montre le circuit simplifié d'un hémodialyseur.

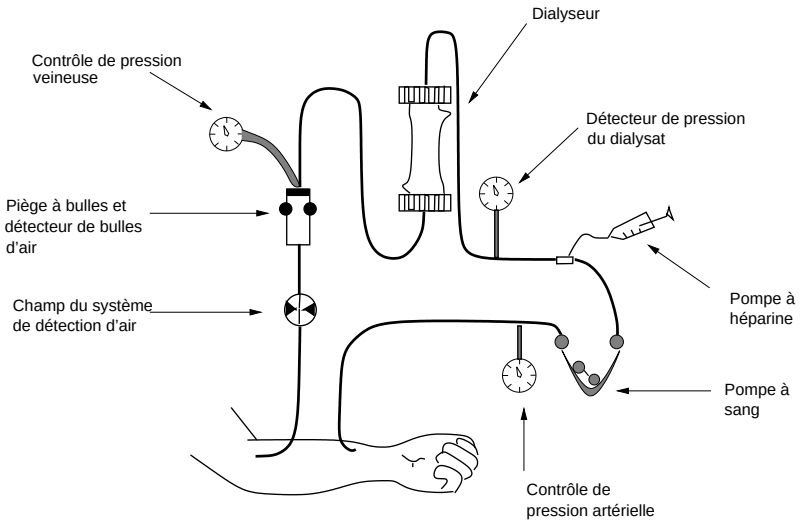


FIG. 1.1 – Circuit d'hémodialyse

[5] lieu d'échange avec le système sanguin

La dialyse péritonéale

La membrane du péritoine tapisse la cavité abdominale dans laquelle se trouvent l'estomac, la rate, le foie et les intestins. La dialyse péritonéale est effectuée en remplissant la cavité abdominale de liquide stérile (dialysat) via un cathéter installé lors d'une opération préparatoire (figure 1.2). La membrane péritonéale sert alors de membrane dialysante. Après un temps d'échange à travers cette membrane, le liquide est évacué par un tube prolongateur reliant le cathéter à la poche externe de dialysat. Le patient remplit à nouveau le péritoine avec un liquide propre de dialysat. On utilise des poches de 2 litres, et ce trois à quatre fois par jour, tous les jours. Le remplissage est effectué manuellement par gravité ou avec une machine automatisée. Les séances ont lieu à domicile après apprentissage par le patient. Environ 10 % des dialysés sont traités par cette méthode. Contrairement à l'hémodialyse, dans cette forme de thérapie, le sang reste à l'intérieur du corps. La solution de la dialyse entre dans la cavité péritonéale par un cathéter inséré dans l'abdomen par un médecin. Les liquides en trop et les déchets passent à travers la membrane péritonéale dans la solution de dialyse qui est alors évacuée hors de l'abdomen.

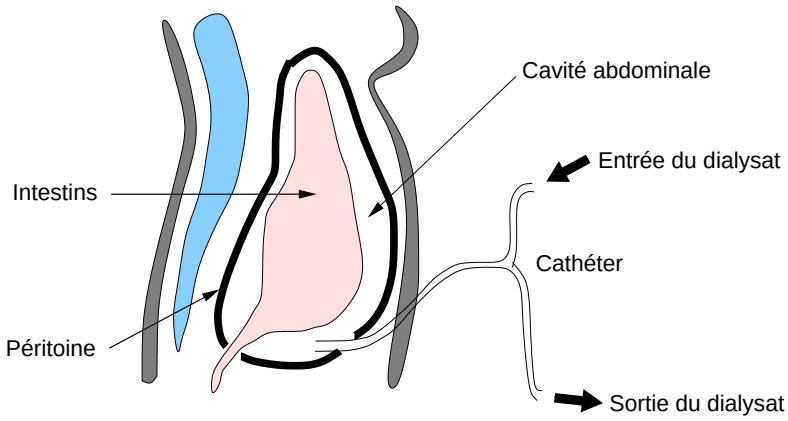


FIG. 1.2 – Le péritoine équipé d'un cathéter

Il existe deux formes de dialyses péritonéales :

1. la **Dialyse Péritonéale Continue Ambulatoire** (DPCA) est effectuée par le patient. Elle fonctionne plus ou moins comme un rein sain puisqu'il s'agit d'un processus de dialyse continue 24 heures sur 24, sept jours par semaine. Comme les dialysés ne sont pas hospitalisés ni attachés à une machine, ils peuvent se déplacer durant la thérapie. Ils peuvent ainsi faire le traitement à la maison ou au travail. Dans cette dialyse, le patient porte sur lui une poche de dialysat qui est vidée dans le péritoine, puis de nouveau remplie progressivement avec les déchets filtrés par le péritoine.
2. la **Dialyse Péritonéale Automatisée** (DPA) se fait à la maison à l'aide d'un cycleur (machine reproduisant le fonctionnement d'un rein) qui effectue les échanges automatiquement, habituellement pendant que le patient dort. La plupart du temps, le dialysat reste dans le péritoine pendant la journée pour filtrer et extraire les déchets et les fluides en trop. Il est ensuite récupérer progressivement par le cycleur.

3.2 Améliorer les conditions de traitement : la dialyse péritonéale adéquate

Un traitement efficace est celui qui assure au patient une durée de vie normale avec un état clinique satisfaisant, lui permettant une bonne insertion socio-professionnelle et familiale. Cependant les techniques de dialyse sont encore loin de suppléer le rein normal : l'épuration des solutés obtenue par dialyse est beaucoup plus faible que celle obtenue par les reins normaux.

La notion de dialyse *adéquate* résulte de la définition de *normes* biologiques propres aux dialysés, compatibles avec une survie de bonne durée et de bonne qualité. L'adéquation d'un traitement par dialyse comporte de multiples dimensions, qui ont chacune pour but de prévenir les complications de l'insuffisance rénale chronique :

– qualité de vie,
– épuration des toxines,
– prévention de l'arthropathie, de la neuropathie, de la dénutrition,
– prévention cardio-vasculaire,
– équilibre acido-basique, phospho-calcique, hydro-sodé,
– contrôle de la tension artérielle et de l'anémie.

À travers de nombreuses études réalisées depuis 1989 [Durand and Kessler, 1998], un certain nombre de cibles à atteindre et de techniques de diagnostic ont pu être définies pour réaliser l'adéquation d'une dialyse et ainsi répondre au critère primordial : la bonne survie du patient. Il existe un certain nombre d'effets indésirables provoqués par un traitement par dialyse qui peuvent hypothéquer l'adéquation. La plupart sont difficilement ou pas du tout mesurables et beaucoup n'ont pas été cliniquement évalués sur le long terme. En particulier, il n'existe aucun indicateur précis pour les problèmes suivants :

– degré de bio-compatibilité des solutés injectés au patient,
– fréquence des infections péritonéales,
– degré d'hyperhydratation acceptable par jour,
– nombre de manipulations tolérables par le patient chaque jour (échange de poches de dialysat).

Parmi les indicateurs importants à surveiller, nous noterons en particulier celui de la tension artérielle, celui du poids idéal du patient et celui du taux d'hydratation. Ces trois facteurs sont liés. Le poids idéal devra être en permanence ré-évalué par le médecin de manière à éviter des épisodes de trouble de la tension, ou encore l'augmentation du taux des protéines (due à une déshydratation excessive). Le taux d'hydratation est aussi un des facteurs les plus importants et surtout, à l'heure actuelle, le plus difficile à estimer. Il est donc nécessaire de l'évaluer par des méthodes indirectes de manière à pouvoir détecter des anomalies : déshydratation ou hyperhydratation excessive. Une déshydratation peut entraîner une altération de l'état général du patient. Une hyperhydratation peut avoir des conséquences importantes sur le système respiratoire, ou encore provoquer une élévation de la tension artérielle. En particulier, on notera qu'une accumulation excessive d'eau entre deux dialyses exposerait le patient au risque d'œdème pulmonaire et obligerait donc un traitement par dialyse plus lourd, en général mal toléré par le patient. Le volume des boissons et, d'une manière générale, l'alimentation du patient doivent être très contrôlés.

La régulation du taux d'hydratation est donc un problème délicat et difficile pouvant donc entraîner des complications importantes sur l'état de santé du patient. Ce taux d'hydratation doit être

maintenu à un état stable et un état d'hyperhydratation ou de déshydratation ne peut être toléré que sur une période très courte (1 à 3 jours).

3.3 Objectifs de Diatelic^{TM}

Le projet Diatelic^{TM}a été lancé en 1997 pour assurer le suivi automatique de l'état d'hydratation des patients sous DPCA et détecter les états d'hyperhydratation et de déshydratation : son but est de démontrer que l'on peut assurer aux patients qui utilisent DIATELIC^{TM}, une meilleure qualité de vie, une meilleure survie, mais aussi une moindre morbidité et une réduction des coûts. De plus l'équipe médicale de l'ALTIR a souhaité avoir un outil d'aide à la décision concernant en particulier le réglage de certains paramètres utiles à la régulation de l'état d'hydratation du patient, et en particulier le poids idéal du patient. Ce poids étant directement en relation avec le taux d'hydratation, une bonne gestion de ce paramètre permet une meilleure régulation de l'état d'hydratation du patient, et offre ainsi l'assurance de le garder en bonne santé.

Le projet Diatelic^{TM}s'inscrit donc clairement dans le cadre du diagnostic médical, et plus généralement de l'évaluation de l'état d'un système à partir d'observations faites sur ce système, mais aussi de l'aide à la décision. Le but est de maintenir le système dans un état donné via des interventions externes (faites par l'utilisateur). Un premier brevet a été déposé sur ce système [Chanliau et al., 2001]. On pourra aussi consulter la référence suivante [Jeanpierre and Charpillet, 2002], portant sur le fonctionnement du système Diatelic^{TM}v2 que je vais brièvement présenter dans la section suivante.

3.4 Solutions proposées dans Diatelic^{TM}v2

Deux versions du système ont été réalisées et ce projet a donné lieu à la création d'une start-up destinée à commercialiser ce service d'assistance et de surveillance à domicile des patients dialysés. La première version, nommée Diatelic^{TM}1, était basée sur un système expert utilisant le cadre des modèles à base de règles de production. La description des règles a été rendue possible grâce à un transfert d'expertise de la part des médecins. Ce modèle reflétait correctement les règles de base utilisées par les médecins pour faire leur diagnostic. Des cœfficients de pondérations sont associés à chaque règle pour prendre en compte plus efficacement l'incertitude due aux mesures. Aucune publication significative n'a été faite sur ce premier système mais les expériences préliminaires faites avec lui ont permis aux auteurs de ce système, de montrer son inadéquation à traiter le problème de Diatelic^{TM}. Bien que ce premier système était constitué des règles empiriques utilisées par les médecins, deux problèmes majeurs sont apparus :

– lors de la création d'une nouvelle règle, l'ensemble des règles devait être ré-évalué afin de vérifier la cohérence du système expert. Ce problème est très courant dans les systèmes à base de règles de production. Le système peut donc difficilement évoluer.
– la prise en compte des spécificités de chaque patient ne peut pas se faire très facilement, car la plupart du temps il faut modifier les pondérations associées à chaque règle, ce qui peut apporter une ou plusieurs incohérences dans l'ensemble des règles. De plus, les méthodes d'apprentissage pour les systèmes à base de règles de production ne sont ni forcément adaptées, ni véritablement efficaces pour un problème tels que Diatelic^{TM}, où les données d'observations sont essentiellement numériques et souvent incertaines et entachées d'erreur.

Ces problèmes ont conduit à la réalisation d'une deuxième version du système expert, nommée Diatelic^{TM}v2 [Jeanpierre and Charpillet, 2002], utilisant cette fois-ci un modèle de décision Markovien partiellement observable (POMDP)[Rabiner, 1989],[Koenig and Simmons, 1996], pour modéliser l'évolution d'un patient. Ce modèle fait tout d'abord une hypothèse markovienne d'ordre 1 : l'état d'un patient dépend des observations faites le jour même du diagnostic et de son état au jour précédent. Le POMDP utilise 5 états représentant trois états-diagnostic possibles et deux états d'aide à la décision :

– état *normal* : probabilité que le patient aille bien,
– état *hyper-hydraté* et état *deshydraté* : probabilité que le patient soit dans un état d'hydratation grave,
– état *poids idéal trop haut*, état *poids idéal trop bas* : probabilité que le poids idéal soit mal adapté,

Les cinq états sont mutuellement exclusifs et vérifient à chaque instant : $\sum_i P(etat = i) = 1$ où i varie de 1 à 5. Les données fournies par le patient sont normalisées par des opérateurs flous, avant d'être intégrées dans le POMDP. L'algorithme de Forward-Backward [Rabiner, 1989] est utilisé pour estimer la nouvelle distribution de probabilités sur l'ensembles des 5 états. Cette estimation est fournie au médecin par le biais d'un interface Web. Si un état dépasse un seuil critique, alors une alarme est générée permettant d'attirer l'attention du médecin sur l'état d'un patient particulier. Ce système est en expérimentation depuis deux ans et surveille 15 patients à distance. Ce modèle montre des performances intéressantes autant pour la prise en compte de l'historique des observations que de sa réactivité par rapport aux aggravations de l'état des patients. Il est à noter cependant que les 5 états étant mutuellement exclusifs, il n'est alors pas possible d'estimer en même temps l'état d'hydratation du patient et si le poids idéal est correctement réglé pour le patient en cours.

Notons enfin que ce système introduit une méthode originale de fuzzication des observations permettant de passer d'un espace continu à un espace discret d'observations[Chanliau et al., 2001]. Cette méthode sera en partie ré-utilisée dans le système à base de réseaux bayésiens dynamiques présenté dans cette thèse.

3.5 Expérimentations et premiers résultats du système Diatelic^{TM}v2

La période de recrutement des patients a débuté en juin 1999 et s'est terminée en août 2000. L'expérimentation médicale, qui se déroule sur une période de deux ans, se termine en août 2002. Les patients sont pris *normalement* en charge par l'ALTIR puis formés spécifiquement à l'utilisation du système DIATELIC^{TM}v2. Ainsi, 30 malades ont été recrutés, dont 15 ont été équipés du système DIATELIC^{TM}v2, et les 15 autres ont suivi la procédure classique de DPCA. Les résultats obtenus après un an d'utilisation sont intéressants et montrent que :

– le nombre de jours d'hospitalisation a été réduit de 40 % sur une durée normale moyenne de 3 semaines par an,
– la consommation de médicaments est considérablement réduite (en particulier les anti-hypertenseurs),
– la tension artérielle et le poids des malades suivis sont plus réguliers et plus conformes aux normes,
– les économies sont estimées à 10 000 euros par an et par patient.

Le système DIATELIC^{TM}v2, bien que présentant des résultats encourageants, peut être grandement amélioré. Les états d'un POMDP étant exclusifs, le système ne donne pas toujours un diagnostic complet : en effet, si le patient a un poids idéal trop haut, aucune information n'est accessible sur l'état d'hydratation. Or un système tel que DIATELIC^{TM}doit être en mesure de fournir un diagnostic et une aide à la décision en même temps : ici le diagnostic est l'estimation de l'état d'hydratation du patient et l'aide à la décision est l'estimation de la justesse du poids idéal du patient, ce poids étant déterminé par le médecin. De plus, la modélisation et l'extension de ce modèle à la lumière de nouvelles données médicales est assez difficile et nécessite une restructuration complète du système. Enfin, le système, dans son état actuel, ne permet pas de gérer les situations critiques où certaines données sont manquantes, ou particulièment erronées. Dans ce cas précis, des données neutres (moyenne sur les jours précédents) sont introduites dans les observations.

4 Conclusion : télémédecine et raisonnement dans l'incertain

Ce premier chapitre a posé les bases du problème de cette thèse : le diagnostic en télémédecine ne peut être fait qu'à partir d'informations incertaines et variables dans le temps, et la vérité sur l'état du patient est imposible à connaître. Souvent les données disponibles sont particulièrement hétérogènes, mais il est nécessaire de les utiliser conjointement dans le but d'optimiser la qualité du diagnostic produit par un système intelligent.

Après une présentation du cadre de la télémédecine et du problème particulier de l'assistance intelligente à distance, le système Diatelic^{TM}v2 et le problème du diagnostic à distance ont été présentés. Diatelic^{TM}servira de base, dans la suite de cette thèse, à l'ensemble des expérimentations avec le modèle dit des réseaux bayésiens dynamiques.

Le problème fondamental posé par Diatelic^{TM}est donc le suivant : comment fournir un diagnostic fiable quand les données dont on dispose sont hétérogènes, incertaines et partiellement bruitées ? Comment modéliser un système qui répondra à la première question et qui fournira un diagnostic, y compris dans des situations critiques telles que celles posées par le manque ou l'incohérence des données ? Mais ce problème en entraîne immédiatement un autre : comment évaluer la pertinence et l'efficacité du système. En fait, le travail présenté dans cette thèse exhibe un problème majeur des systèmes d'aide au diagnostic : ils sont très difficiles à évaluer et nécessitent en général la mise en place d'un protocole expérimental lourd et fastidieux. De plus, une telle évaluation prend souvent plusieurs années et donc les résultats définitifs ne peuvent être connus dans des temps raisonnables (souvent supérieurs à plusieurs années) : mettre en place une expérimentation médicale est un processus difficile.

L'approche développée dans cette thèse sera présentée au cours des trois chapitres suivants. Elle montrera progressivement comment on est passé de la spécification du problème Diatelic^{TM}a une approche plus générale de la fusion de données et comment nous avons convergé vers une solution particulière en nous basant sur le formalisme des réseaux bayésiens. Il permettra en particulier d'apporter une contribution à une des questions fondamentales de la télémédecine et du diagnostic médical : l'incertitude sur les données et les diagnostics.

L'intérêt du problème posé par la télémédecine est donc bien celui de l'incertitude et du raisonnement dans l'incertain. Les données sont incertaines, peu fiables et bruitées. Mais les modèles sont

aussi incertains et pas toujours adéquats. Néanmoins, il faut donner un diagnostic ou en tout cas une indication suffisamment précise pour que le médecin puisse prendre la bonne décision.

Chapitre 2

Processus de fusion de données

Résumé :

La fusion de données correspond à la volonté d'utiliser simultanément plusieurs sources de données, ayant des caractéristiques éventuellement différentes, afin d'obtenir une nouvelle information de meilleure qualité.

Ce chapitre présente une approche originale de la fusion de données et applique ces notions à la modélisation et au diagnostic probabiliste dans le cadre du projet Diatelic^{TM}. Notre contribution se situe au niveau de la définition d'une notion de gain dans un processus de fusion de données.

Une première partie présente la notion de fusion de données et montre les principales techniques utilisées dans le cadre de la perception d'environnements incertains. Ensuite, j'aborde une présentation plus générale de la fusion de données pour introduire la notion de gain qualifié. Le chapitre se termine sur une application de cette approche à la modélisation du problème Diatelic^{TM}en terme de fusion de données.

« La fusion de données vise à l'association, la combinaison, l'intégration et le mélange de multiples sources de données représentant des connaissances et des informations diverses dans le but de fournir une meilleure décision par rapport à l'utilisation séparée des sources de données. Trois éléments sont toujours considérés dans le cadre de la fusion de données : les sources, les algorithmes et le résultat. Le résultat sera composé de données ayant une valeur ajoutée par rapport aux données initiales en provenance de la source. Cette nature de cette valeur ajoutée dépendra de l'application. La fusion de données forme alors un cadre original pour le problème de la perception d'environnements lorsque plusieurs capteurs sont utilisés, éventuellement au cours du temps. »

1 La fusion de données : introduction et buts

La fusion de données correspond à la volonté d'utiliser simultanément plusieurs sources de données différentes, ou de grouper des informations hétéroclites afin d'obtenir une nouvelle information de meilleure qualité. Cette nouvelle information est souvent destinée à la prise de décision. Ce chapitre présente la fusion de données sous l'angle du problème de la perception d'un environnement quelconque.

Le problème de la combinaison et de l'utilisation simultanée de données et d'informations à partir de plusieurs sources se rencontre dans de nombreux champs d'application souvent liés au besoin de vouloir percevoir un environnement à partir de capteurs plus ou moins fiables, plus ou moins précis, plus ou moins efficaces. Mais le terme de fusion de données s'est finalement étendu à de plus vastes domaines et se rencontre maintenant fréquemment en représentation logique de la connaissance, ou encore en fouille de données [Gebhardt and Kruse, 1998]. De plus, nous assistons à un engouement notable pour la fusion de données pour des applications telles que le diagnostic médical, le commerce, la robotique, les finances, le traitement du signal ou encore la compréhension automatique de documents scientifiques. Tous ces domaines ont en commun le fait de devoir manipuler de grandes quantités de données de nature et de type très variés, afin d'obtenir une information de meilleure qualité. Cette notion de qualité, très importante pour la fusion de données, dépend alors de l'application visée. Par exemple, la fusion de plusieurs documents scientifiques permettra d'obtenir un document composite et décrivant avec précision la structure et les relations qui existent entre les documents analysés. Le résultat sera une information facile à utiliser pour faire, par exemple, une recherche thématique dans un grand ensemble de documents scientifiques. La qualité de l'information obtenue tient ici à la facilité qu'elle apportera au processus de recherche de documents.

Cependant, les différentes approches de la fusion de données, les applications nombreuses et portant sur des domaines très différents les uns des autres, nous ont emmené au constat qu'il n'existe pas de formalisme précis pour décrire un processus de fusion de données [Abidi et al., 1992, Hall and Llinas, 1997].

1.1 Approche classique de la fusion de données

1.1.1 Une nécessité

Le domaine de la fusion de données devient particulièrement important autant d'un point de vue fondamental que pour des applications pratiques. La fusion de données trouve ses applications dans un grand ensemble de domaines tels que la robotique, le contrôle et le commandement militaire, la médecine, la vision robotique, l'interprétation d'images (fusion d'images satellites ou médicales), le contrôle et le monitoring de processus, l'extraction de connaissances dans de grandes banques de données, etc...

De nombreuses approches de la fusion de données ont été considérées, allant de l'utilisation de plusieurs capteurs complémentaires et/ou concurrents au raisonnement symbolique destiné à l'interprétation de données. Le but reste souvent l'utilisation optimale des données disponibles pour prendre une décision, quelle que soit cette décision : un diagnostic, l'interprétation d'un signal ou d'une scène, une planification d'actions, etc... Ainsi, un élément crucial dans des systèmes souhaitant aboutir à une telle prise de décision est l'existence d'un mécanisme capable de modéliser, de *fusionner* et d'interpréter les informations disponibles. Les données fusionnées reflètent non seulement l'information générée par chaque source de données, mais encore l'information qui n'aurait pu être inférée par aucune des sources prises séparément !

1.1.2 Niveaux de fusion

On rencontre généralement trois principales méthodes de fusion dans les systèmes complexes. Souvent elles sont utilisées simultanément. La fusion de données basée sur des modèles physiques (ou biologiques) représente la première méthode. Le modèle le plus connu est le filtre de Kalman [Maybeck, 1990] qui est largement utilisé pour le contrôle de systèmes dynamiques où il est nécessaire de fusionner des données redondantes et de bas niveau (numériques) au cours du temps. Le système est décrit par un modèle linéaire. L'erreur associée aux capteurs et au système est modélisée par un bruit blanc gaussien. Dans ce cas là, le filtre de Kalman fournira un estimateur statistique optimal pour les données fusionnées [Luo and Kay, 1989]. Ce type d'approche s'accorde aux cas où le système observé à un comportement linéaire connu à l'avance. Il s'agit d'une procédure classique de prédiction-mise à jour.

La seconde méthode est basée sur l'utilisation de techniques de classification et plus généralement d'intelligence artificielle : modèles statistiques et probabilistes, et théorie de l'information. De nombreuses techniques existent et nous pouvons citer le raisonnement bayésien, la théorie de l'évidence ou encore les opérateurs récursifs [Abidi and Gonzalez, 1992]. Ce type d'approche mêle des connaissances de plus haut niveau avec des données numériques. L'évolution de l'environnement observé n'a pas besoin d'être connu à l'avance et peut souvent être appris.

Enfin la troisième méthode s'inspire des modèles cognitifs. Dans cette dernière grande famille, nous pouvons aussi inclure les réseaux de neurones ou encore les techniques de vision artificielle. Ce type de modèle s'inspire de la réalité (psychologie, biologie,...) pour construire des modèles de raisonnement efficace. Il mêle souvent connaissances qualitatives et quantitatives. Ils sont souvent caractérisés par la nécessité d'apprendre leur comportement avant de pouvoir fonc-

tionner. Ils s'accorde autant avec des environnements au comportement linéaire que non-linéaire [Hertz and Krogh, 1991].

La fusion de données est souvent décrite comme un processus à plusieurs niveaux, dans lequel trois niveaux fonctionnels ont été clairement identifiés [Haton et al., 1998] :

– *la fusion de bas niveau* : extraction d'information directement depuis les capteurs pour former des hypothèses partielles, principalement à partir de méthodes numériques ;
– *la fusion de haut niveau* : amélioration et intégration des hypothèses partielles produites par le premier niveau, combinées avec des informations symboliques (comme des connaissances déclaratives), dans le but d'abstraire les données et de fournir une interprétation symbolique de la situation actuelle ;
– *la décision finale* : souvent basée sur une interprétation symbolique des résultats, afin de remplir une tâche spécifique (c'est-à-dire de donner la *décision*). Dans une application militaire ce sera la désignation de la menace. Dans une application médicale, ce sera le diagnostic sur l'état du patient considéré. Dans le cas de la robotique, la décision est représentée par un plan ou une politique d'actions à effectuer. Cette dernière activité peut faire partie du processus de fusion de données, en particulier dans le cas de systèmes multi-modèles où la décision finale se fait à partir de multiples décisions intermédiaires [Barret, 1990]. Dans ce cas le système peut procéder à un vote, à un tirage aléatoire ou encore à un choix selon un critère discriminant.

Cependant, ce modèle à trois niveaux nécessite d'être affiné. Par exemple, dans le cas de la vision robotique, il est utile de distinguer le niveau où l'on fusionne : les pixels, les régions (ensembles de pixels), la représentation globale de la scène et l'interprétation finale. De nombreux modèles à plusieurs niveaux (souvent supérieur à 3) ont été proposés en vision robotique dans lesquels les informations numériques et symboliques sont fusionnées à chaque niveau. Dans ce domaine, on pourra se référer en particulier aux travaux de Henri Maître présentés dans [Bloch and Maître, 1994].

La nature hypothétique de la fusion de données, due en partie à l'imprécision associée aux données initiales, rend nécessaire l'utilisation de mécanismes de contrôle sophistiqués basés sur l'intégration d'hypothèses multiples. En particulier, la fusion de décisions multiples en reconnaissance des formes est un cas important à considérer. De nombreuses approches ont été proposées dans ce sens, autant en vision qu'en reconnaissance de l'écriture et l'hypothèse sous-jacente à toutes ces méthodes est qu'en mélangeant les résultats donnés par plusieurs méthodes de reconnaissance selon certaines règles de combinaisons, les performances peuvent être considérablement améliorées et apporter des résultats bien supérieurs à toutes les méthodes prises chacune isolément. Bien sûr, il y a là aussi de nombreuses façons de combiner ces méthodes : certaines se basent sur des heuristiques comme le vote majoritaire sur l'ensemble des méthodes, d'autres encore introduisent une mesure de fiabilité à chaque méthode et utilisent un vote pondéré. D'autres méthodes encore préfèrent l'utilisation de la théorie de Dempster-Shafer ou encore des règles de combinaison bayésienne. Enfin, les réseaux de neurones peuvent aussi être à la base de modèles de combinaison.

1.1.3 Le besoin de connaissances supplémentaires

Dans de nombreux problèmes de fusion de données, l'idée est de reconstruire un modèle de l'environnement observé, sans connaissance préalable, à partir de signaux, d'images ou d'autres types d'informations issues de capteurs divers. Par exemple, la cartographie 3D topographique et géographique de la surface de la Terre utilise uniquement des techniques radar et de photogrammétrie.

Aucune connaissance sur l'aspect même de la planète n'est utilisée. Cependant, dans de nombreux cas, une certaine quantité de connaissances *a priori* est disponible. Ces connaissances risquent dans certains cas d'être soit dépassées, soit partielles. Mais si la structure des informations a été correctement définie, la fusion de données peut se ramener à un problème de mise à jour d'un ensemble de connaissances (remplacement des anciennes informations par les nouvelles, ajout des informations manquantes pour compléter la description de l'environnement, etc...).

Dans le cas de problèmes plus complexes tels que l'analyse de situation (militaire, financière,...), les aspects pertinents du problème peuvent ne pas être directement en relation avec les données brutes fournies par les capteurs. Dans certains cas, on peut même ne pas avoir de modèle structurant les données. Ce cas se rencontre souvent en diagnostic médical, où la donnée des symptômes et les observations physiques du patient sont souvent insuffisantes pour déterminer directement la cause, c'est-à-dire la maladie ou le trouble dont souffre le patient. En effet, l'état de santé d'un patient est aussi dépendant de l'état d'un nombre important d'organes qui sont souvent peu ou pas observables. Mais d'autres facteurs entrent aussi en jeu tels que l'état moral du patient, son cadre de vie, ses antécédents médicaux ou encore son hygiène de vie.

Le cas le plus complexe est sans doute l'analyse de situation, la prédiction et la décision dans des systèmes sociaux dynamiques souvent rencontrés en analyse géopolitique ou économique, ou dans le domaine de la prise de décision militaire sur un champ de bataille [Waltz and Llinas, 1990, Abidi and Gonzalez, 1992, McMichael et al., 1996]. Dans ce type de système, la masse de connaissance à traiter est particulièrement grande et en constante augmentation et les entités observées sont souvent très complexes (comme une foule, ou un ensemble militaire comprenant des hommes, des véhicules, des aéronefs). De plus, les données sont souvent complexes et parfois assez abstraites : il peut s'agir simplement d'images et de vues aériennes, mais il peut aussi s'agir de données recueillies lors d'un sondage d'opinion ou d'une étude démographique.

1.2 Un état de l'art des techniques classiques

En me basant sur les considérations générales présentées dans les sections précédentes, je vais introduire dans cette nouvelle section un panel représentatif bien que non exhaustif des techniques de fusion de données les plus courantes. Ces techniques sont souvent issues d'autres domaines mais sont ici appliquées à la combinaison et l'intégration de données hétérogènes provenant de plusieurs sources. Ces pourquoi nous les appelleront *techniques classiques*.

Si l'on se restreint au domaine de l'intelligence artificielle et de la robotique, on peut identifier principalement deux grandes classes de techniques de fusion de données :

– celle basée sur des modèles probabilistes, possibilistes ou de la théorie de l'évidence,
– celle basée sur les techniques des moindres carrés.

Cependant, cette classification, datant d'une dizaine d'années [Abidi and Gonzalez, 1992], a légèrement évolué depuis les dernières années, car nous pouvons y inclure aussi les techniques basées sur le raisonnement en logique [Gebhardt and Kruse, 1998] ou encore les techniques basées sur les systèmes multi-agents [Haton et al., 1998]. De plus, il est intéressant de noter que ces techniques servent en particulier à combiner différents systèmes de fusion de données entre eux afin de bâtir un nouveau système de fusion de données, opérant, par exemple, à plusieurs niveaux d'abstraction.

Les techniques que je vais maintenant présenter diffèrent par les caractéristiques de l'environnement dans lesquels elles opèrent, par le type d'informations délivrées par les capteurs, par le mode de représentation de l'information que l'on adopte, ou encore par la manière de représenter l'incertitude associée à la technique de fusion. Dans [Luo and Kay, 1989], on peut trouver une étude comparative des techniques suivantes : moyenne pondérée, filtre de Kalman, estimation bayésienne, théorie de la décision statistique, théorie de l'évidence de Dempster-Shafer, logique floue et règles de production avec coefficients de confiance. Je vais maintenant donner un bref aperçu de certaines de ces techniques.

1.2.1 Méthodes d'estimation bayésienne

Dans cette approche, les capteurs sont considérés comme un ensemble d'entités capables de fournir une décision à tout instant. Chaque capteur est alors vu comme un estimateur bayésien. Ainsi, les distributions de probabilités associées à chaque capteur sont combinées dans une seule fonction de distribution de probabilités jointes *a posteriori*, en utilisant la règle de Bayes :

$$P(A|B) = \frac{P(B|A).P(A)}{P(B)}$$

La vraisemblance de cette fonction est maximisée pour obtenir la fusion finale de l'ensemble des capteurs. On trouvera dans [Xu et al., 1992] les détails mathématiques ainsi qu'une application à la reconnaissance de l'écriture.

Cette approche est relativement simple, mais elle ne permet pas encore de représenter des modèles complexes évoluant dans le temps. De plus, l'intégration de connaissances qualitatives est relativement difficiles, mais cette approche forme une base pour d'autres méthodes plus récentes, telles que les Modèles de Markov Cachés (HMM, Hidden Markov Models), ou encore les réseaux bayésiens, qui eux permettront justement l'intégration de telles connaissances et la modélisation de systèmes dynamiques.

1.2.2 Cartes d'évidence

Faire naviguer un robot dans un environnement inconnu ou partiellement connu nécessite l'utilisation de nombreux capteurs souvent de même nature et disposés sur le robot de telle façon qu'ils puissent observer au maximum l'environnement dans lequel est plongé le robot. Les cartes d'évidence représentent une approche aujourd'hui classique pour la modélisation des objets se trouvant dans l'espace dans lequel évolue un robot et sert de base pour la navigation et la planification des mouvements du robot. Cette technique, initialement développée par H. Moravec [Moravec, 1987] et A. Elfes [Elfes, 1989] a été depuis utilisée sous diverses formes, notamment pour la localisation d'un robot [Thrun and Bücken, 1996] et pour diverses tâches de navigation en conjonction avec d'autres modèles comme les cartes topologiques [Thrun et al., 1998] (qui ne représentent que certains points intéressants de l'environnement du robot). Cette technique de fusion est particulièrement adaptée lorsque l'on a de nombreux capteurs de même type et qu'il est nécessaire de mettre à jour, de façon bayesienne, une connaissance sur un environnement physique. Cette approche est particulièrement adaptée à la navigation robotique. L'utilisation de connaissances qualitatives est plus difficile cependant, ainsi que l'utilisation de capteurs hétérogènes.

Dans l'approche de H. Moravec, l'environnement dans lequel évolue le robot est totalement inconnu et le robot le découvre au fur et à mesure de ses évolutions. Classiquement, le robot est équipé d'une couronne de sonars lui conférant une capacité de perception à 360°. La surface globale dans laquelle évolue le robot est discrétisée dans une représentation à 2 ou 3 dimensions. A chaque cellule de cette carte discrétisée est associée la probabilité de présence d'un objet : un meuble, un mur, etc... Lors de ses déplacements, le robot utilise ses capteurs pour accumuler des *preuves* et re-estime à chaque instant la probabilité pour chaque case de contenir un objet. Cette réestimation est faite à partir d'une application de la règle de Bayes. Au départ, chaque case contient une probabilité de 0,5, ce qui correspond à une incertitude totale. Chaque écho de chaque sonar sert ainsi à confirmer ou infirmer la présence d'un objet dans chaque case. Au bout d'un certain temps, le robot a une connaissance plus approfondie de son environnement. En effet, les cases correspondant à un objet (meuble ou mur par exemple) verront leur probabilité tendre vers 1 alors que les espaces vides verront leur probabilité tendre vers 0.

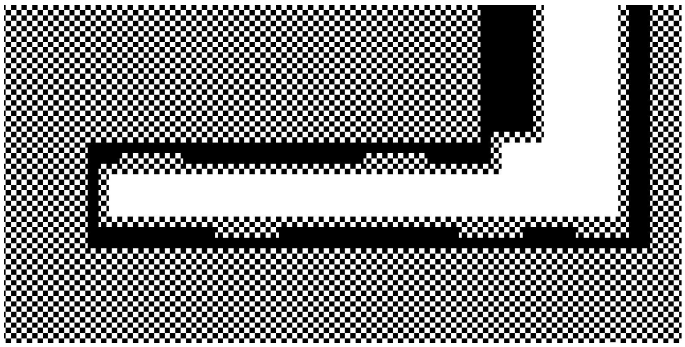


FIG. 2.1 – Carte d'évidence idéale

La figure 2.1 représente la carte idéale d'un couloir. Les murs ont une probabilité de présence d'un objet de 1 et l'espace vide de 0. Au-delà du mur, rien n'est connu, donc les probabilités restent à 0,5 (en gris sur la figure).

FIG. 2.2 – Carte d'évidence après parcours du robot

La figure 2.2 représente une carte obtenue après plusieurs passages du robot. On voit nettement apparaître les murs, l'espace vide, ainsi que les zones de doute. Cependant, due à l'imprécision et aux échos parasites, la cartographie n'est pas parfaite. Elle permet cependant de planifier correctement une trajectoire. Ces deux cartes sont extraites de [Martin and Moravec, 1996].

Ainsi, chaque écho est transformé en une preuve supplémentaire de présence ou d'absence d'un objet et sert à renforcer ou diminuer la probabilité de présence d'un objet dans chaque case de la carte. L'ensemble des signaux sonars émis et reçus est donc fusionné au sein d'un même modèle. Sa construction nécessite cependant, dans un premier temps, un parcours en aveugle de l'environnement. La fusion d'information de plus haut niveau, comme un plan ou des balises, permet d'améliorer la fiabilité et la précision du modèle, et d'accélérer le processus de cartographie [Thrun and Bücken, 1996, Thrun et al., 1998, Yamauchi, 1995].

D'autres applications de ce modèle concernent son utilisation pour localiser un robot dans un espace donné. Dans ce cas, chaque case contient la probabilité de présence du robot, et la somme des probabilités des cases fait 1. Cette méthode est moins sensible aux problèmes d'imprécision des capteurs et des problèmes de glissement des roues des robots sur le sol [Thrun and Bücken, 1996].

1.2.3 Les modèles de Markov cachés

Les modèles de Markov cachés, que nous appelleront par la suite HMM, sont des modèles statistiques de données séquentielles, qui ont été largement utilisés dans des applications telles que la reconnaissance de la parole [Rabiner, 1989], la reconnaissance de forme, l'intelligence artificielle ou encore la modélisation de séquences biologiques. Leur succès tient principalement à l'algorithme d'apprentissage de Baum-Welch, qui est un cas particulier de la procédure EM (*expectation-maximization*) pour l'estimation du maximum de vraisemblance. Cet algorithme d'apprentissage permet en effet d'apprendre efficacement un modèle à partir de données observées. L'intérêt en fusion de données est que ce modèle est particulièrement adapté à la modélisation de processus stochastiques et gèrent donc bien les flux de données. Cependant, la gestion de multiples capteurs n'est possible qu'avec un pré-traitement de l'ensemble des données issues des capteurs pour les transformer en une observation unique, qui sera l'observation utilisée à chaque pas de temps par le HMM. Donc un HMM est plutôt adapté à la fusion de données au cours du temps, plutôt qu'à un instant donné. Dans ce dernier cas, il sera nécessaire d'utiliser une autre méthode de fusion en amont du HMM.

Les HMM peuvent être appliqués à des ensembles de données qui ont une certaine propriété dite propriété de Markov : l'état d'un système peut être totalement déterminé à condition de connaître *(a)* l'état dans lequel il était à l'instant précédent et *(b)* les observations faîtes sur le système à l'instant courant. La distribution de probabilités jointes d'une séquence d'observations $y_1^T = \{y_1, y_2, \ldots, y_T\}$ peut toujours être factorisée de la façon suivante :

$$p(y_1^T) = p(y_1) \prod_{t=2}^{T} p(y_t | y_1^{t-1})$$

Cependant, il est matériellement impossible de modéliser une séquence de données, où la distribution conditionnelle $p(y_t|y_1^{t-1})$ d'une variable observée y_t à l'instant t dépend de l'ensemble des valeurs précédentes prises par la variable y_i où $1 \leq i \leq t-1$. Cependant, la modélisation devient

possible, si l'on fait l'hypothèse qu'une séquence d'observation passée peut être *résumée* au sein d'une variable aléatoire, dite *variable d'état*, qui contient toutes les informations nécessaires de la séquence y_1^t nécessaire à l'évaluation de la distribution de la prochaine observation y_t.

Un HMM d'ordre k est une distribution de probabilités sur une séquence d'observations $q_1^t = \{q_1, q_2, \ldots, q_t\}$ possédant la propriété d'indépendance conditionnelle suivante :

$$p(q_t|q_1^{t-1}) = p(q_t|q_{t-k}^{t-1})$$

Et puisque q_{t-k}^{t-1} résume toutes les informations du passé, alors q_t est généralement appelée une variable d'état. Ainsi, cette propriété de Markov est interprétée en disant que seule la connaissance des k instants précédents est suffisante pour connaître l'état courant. Souvent, on utilise seulement un modèle de Markov à l'ordre 1 où seul l'instant précédent est nécessaire pour connaître l'instant courant. Ceci est contenu dans la distribution suivante :

$$p(q_1^T) = p(q_1) \prod_{t=2}^{T} p(q_t|q_{t-1})$$

Un tel modèle est complètement spécifié si l'on connaît l'*état initial* $p(q_1)$ et les *probabilités de transition* (d'un état à l'état suivant) $p(q_t|q_{t-1})$. Souvent, ces probabilités de transition sont supposées homogènes, c'est-à-dire les mêmes pour chaque pas de temps. Néanmoins, lorsque l'ordre k devient grand, ces modèles sont particulièrement lourd à manipuler. Par exemple, pour une variable d'état multinomiale $q_t \in \{1, \ldots, n\}$, le nombre requis de paramètres pour représenter les probabilités de transition est de l'ordre de $O(n^{k+1})$. Dans le cas d'un modèle de Markov caché, on ne modélise pas directement les observations, mais plutôt l'état du système à travers une *variable d'état* : on ne fait pas l'hypothèse que les observations ont la propriété de Markov, mais plutôt l'hypothèse qu'il existe une variable, non observée, mais en relation avec les observations, et qui a la propriété de Markov. Les relations entre la variable d'état et les observations sont définies par les deux hypothèses d'indépendance conditionnelle suivantes :

$$p(y_t|q_1^t, y_1^{t-1}) = p(y_t|q_t)$$

et

$$p(q_{t+1}|q_1^t, y_1^t) = p(q_{t+1}|q_t)$$

En d'autres termes, l'information contenue dans la variable q_t d'état suffit à décrire l'historique des observations et des états jusqu'à l'instant t quand on veut prédire la valeur de la variable d'observation y_t ou du prochain état q_{t+1}.

Dans de nombreuses applications, la variable d'état est discrète et les *états* du HMM représentent les différentes valeurs que peut prendre la variable d'état. Ainsi, le HMM est représenté grâce à un graphe orienté où chaque noeud représente un état possible et les arcs une probabilité de transition non-nulle $p(q_t|q_{t-1})$ entre deux états, donc un automate d'états finis probabiliste. La figure 2.3 représente un HMM courant en reconnaissance de la parole, servant à modéliser une distribution de séquences acoustiques associées à un unique phonème.

De nombreux algorithmes ont été développés pour les HMM. Il existe en particulier l'algorithme de Viterbi [Viterbi, 1967] qui permet, à partir d'une séquence d'observations y_1^T d'inférer la séquence d'états q_1^T la plus vraisemblable lui correspondant pour un HMM donné. Cet algorithme

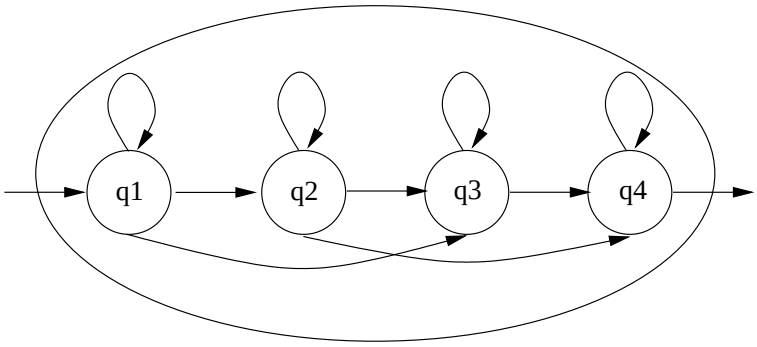


FIG. 2.3 – Exemple d'un HMM *gauche-droite* utilisé en reconnaissance de la parole pour représenter la distribution de séquences acoustiques associées à un unique phonème.

est largement appliqué dans les problèmes de reconnaissance de la parole [Rabiner, 1989] ou de biologie moléculaire [Baldi and Brunak, 1998] (par exemple) où chaque état correspond à un label de classification et chaque séquence d'état forme une séquence de labels. La figure 2.4 représente une partie d'un HMM où chaque séquence d'état (entourée par un ovale) a un sens particulier : elles correspondent chacune à un mot. Ce HMM permet de reconnaître des séquences précises de mots connectés (parole continue). Il est ainsi possible grâce à cet algorithme et au HMM correspondant, de transformer les séquences d'observations (le signal sonore au cours du temps) en une séquence de mots, correspondant donc à une information de plus haut niveau d'abstraction.

Les probabilités de transition sont souvent apprises en cherchant le modèle reflétant au mieux un ensemble de séquences d'observations pour lesquelles on connaît la séquence d'états correspondante [Rabiner, 1989].

1.2.4 Théorie de l'évidence de Dempster-Shafer

Cette théorie [Dempster, 1967, Shafer, 1976] est dérivée de l'approche bayésienne, mais utilise deux mesures pour qualifier le degré de croyance que l'on a sur une hypothèse, calculé à partir d'indices la confirmant ou l'infirmant. La théorie peut assigner une mesure de certitude à des ensembles d'hypothèses autant qu'à des hypothèses seules. Cette approche permet de raisonner sur des ensembles d'hypothèses dans un premier temps, et de se restreindre petit à petit aux hypothèses plausibles, au fur et à mesure que de nouvelles évidences apparaissent. Cette approche de la fusion de données est adaptée à la fusion de multiples capteurs. Sa mise en oeuvre pour la fusion de données au cours du temps reste plus difficile qu'avec un HMM. C'est pourquoi cette théorie est plus adaptée à la fusion de multiples sources à un instant donné. La manipulation de données qualitatives est aussi plus aisée que les modèles présentés précédemment, comme nous allons le voir maintenant.

Supposons que Ω soit un ensemble d'hypothèses exhaustives et mutuellement exclusives et $P(\Omega)$ l'ensemble de toutes les disjonctions possibles sur Ω. Supposons qu'il existe une source de données en relation d'une certaine manière avec Ω. Dans la théorie de Dempster-Shafer, on assignera aux éléments de $P(\Omega)$ une mesure de certitude, grâce à l'affectation probabiliste m suivante :

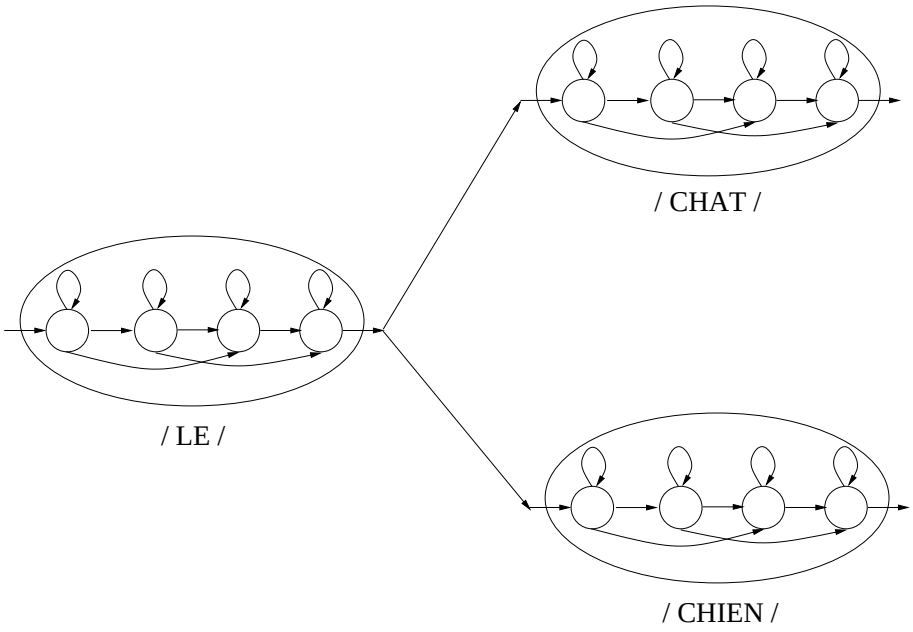


FIG. 2.4 – Partie d'un HMM servant à la reconnaissance de mots connectés, avec des groupes d'états ayant un sens particulier (ici un mot). Une séquence d'états correspond à une séquence de mots.

– $m(\emptyset) = 0$,
– $\sum_{E/E \in P(\Omega)} m(E) = 1$

L'ensemble des hypothèses E telles que $m(E)$ est positif est appelé ensemble focal de m. Cette affectation permet de calculer deux fonctions Cr (Croyance ou Crédibilité) et Pl (Plausibilité) qui sont des mesures de la certitude accordée à chaque hypothèse :

$$\forall E, F \in P(\Omega),\ Cr(F) = \sum_{E/E \subseteq F} m(E)$$

Ainsi, $Cr(F)$ est la somme de toutes les masses de croyance assignées aux ensembles qui font que F est certain. C'est donc une mesure de la confiance totale accordée à chaque hypothèse ou sous-ensemble d'hypothèses. La fonction de plausibilité Pl est définie par :

$$\forall F \in P(\Omega),\ Pl(F) = 1 - Cr(\neg F)$$

Ainsi, $Pl(F)$ est la somme des masses assignées aux ensembles qui ne sont pas inconsistants avec F, c'est-à-dire qui rendent F plausible.

Ainsi pour une source unique de données se focalisant sur des éléments de $P(\Omega)$, il est possible de calculer un intervalle $[Cr(E), Pl(E)]$ pour chaque hypothèse E, simple ou non. L'intervalle $[0, 1]$ correspondra alors à une ignorance totale. Mais considérons maintenant deux sources de données, et prenons un exemple tiré de l'analyse d'images [Haton et al., 1998]. La première source de données fournit des informations sur les niveaux de gris des pixels et la seconde source, sur la texture d'une zone dans l'image. Supposons que quatre sortes de zones puissent exister : A, B, C

et D. L'algorithme d'interprétation cherche à étiqueter une zone délimitée par un contour connu. Les deux sources distribuent leurs masses de probabilité, respectivement m_1 et m_2 sur l'ensemble des hypothèses possibles. Bien que nous puissions calculer des intervalles de certitude pour les quatre catégories, en prenant en compte séparément chaque source de données, il est intéressant ici d'utiliser la règle de combinaison de Dempster-Shafer afin de fusionner les deux *opinions* que sont m_1 et m_2. La masse affectée à tout Z appartenant à $P(\Omega)$ et différent de l'ensemble vide peut être calculée en combinant m_1 et m_2 avec une somme orthogonale :

$$m_{12}(Z) = m_1 \oplus m_2(Z) = K. \sum_{X_i, Y_j / X_i \cap Y_j = Z} m_1(X_i) \times m_2(Y_j)$$

où X_i et Y_j appartiennent à $P(\Omega)$ et avec $\frac{1}{K} = 1 - \sum_{X_i, Y_j / X_i \cap Y_j = \emptyset} m_1(X_i) \times m_2(Y_j)$. K est nécessaire dans la formule pour éliminer la masse résiduelle non-nulle qui aurait été assignée à $\emptyset$ et pour ainsi avoir $m_{12}(\emptyset) = 0$ tout en ayant la masse totale égale à 1. Intuitivement, K est une mesure du degré de conflit entre les deux sources, qui peut ne pas être pertinente dans le cas où les deux sources sont discordantes, car K masquera alors le conflit entre les deux sources [Smets, 1988].

Étant donné que la somme orthogonale est commutative et associative, il est possible de prendre en compte les données en provenance des deux sources dans n'importe quel ordre. Cependant, il est à noter que la règle de combinaison de Dempster-Shafer ne permet de fusionner que deux sources indépendantes l'une de l'autre. Des extensions ont été proposées afin de prendre en compte des *opinions* dépendantes dans [Ling and Rudd, 1988].

1.2.5 Les modèles graphiques probabilistes

Les modèles graphiques sont un mariage entre la théorie des probabilités et la théorie des graphes. Ils fournissent un outil simple et efficace pour traiter des problèmes souvent rencontrés en intelligence artificielle, en mathématiques appliquées ou en ingénierie : l'incertitude et la complexité [Jordan, 1999]. Les modèles graphiques sont aussi connus sous le nom de réseaux de croyance, réseaux d'indépendance probabiliste ou encore réseaux bayésiens [Pearl, 1988, Cowell et al., 1999].

La notion qui est à la base des réseaux bayésiens est celle de la modularité : un système complexe est construit en combinant des parties plus simples. Et la théorie des probabilité scelle l'ensemble, c'est-à-dire la façon dont les parties sont combinées. Elle assure que l'ensemble est consistant. La théorie des graphes fournit des outils intuitifs permettant la modélisation par l'humain de systèmes complexes et la création d'algorithmes efficaces de raisonnement et d'apprentissage. De nombreux problèmes, tels que les modèles de Markov cachés, les filtres des Kalman, les modèles d'Ising sont des cas particuliers des modèles graphiques et ces derniers apportent ainsi un cadre et une formulation unique à l'ensemble de ces approches [Smyth et al., 1996].

Les modèles graphiques forment un cadre particulièrement efficace pour la fusion de données. Ils permettent de modéliser un problème simplement et d'utiliser des données en provenance de multiples sources, quelles soient qualitatives ou quantitatives. Ces données servent à mettre à jour la connaissance et les croyances que l'on a sur le problème, de façon bayesienne, à partir d'un unique formalisme d'inférence et d'avoir ainsi une nouvelle vue du problème sachant les nouvelles données.

Ces modèles seront largement abordés dans le chapitre 3 qui est entièrement consacré au raisonnement causal probabiliste et plus particulièrement aux réseaux bayésiens.

1.2.6 Techniques des moindres carrés et filtres de Kalman

Les techniques dites des moindres carrés regroupent en particulier des modèles tels que les filtres de Kalman ou encore les méthodes d'optimisation et de régulation. Ces techniques trouvent de très nombreuses applications, en particulier dans le domaine du contrôle de processus dynamique mais surtout dans le domaine de la poursuite de cibles en mouvement. Ces techniques ont été appliquées en premier lieu à la poursuite de cibles aériennes autant dans le domaine militaire que civil mais ont trouvé leurs applications dans de nombreux autres domaines, comme la navigation robotique et la fusion multi-capteurs pour le contrôle et la perception d'un robot. De nombreuses déclinaisons du filtre de Kalman existent, et servent dans des problèmes d'initialisation de piste (trouver la cible que l'on va ensuite poursuivre) ou encore de poursuites multi-cibles (gestion du trafic aérien près d'un aéroport) [Barret, 1990, Bar-Shalom and Li, 1995].

Certaines techniques des moindres carrés on été développées pour la localisation d'un objet connu. Dans ce contexte, les données récupérées à partir des capteurs et les erreurs de mesures associées sont interprétées comme un ensemble de contraintes sur l'espace des solutions. Chaque point de l'espace représente une position possible pour l'objet observé. Au fur et à mesure que les données arrivent, l'espace des solutions est réduit itérativement, jusqu'à un point de convergence qui représente alors la position réelle de l'objet [Abidi and Gonzalez, 1992].

Filtre de Kalman

Le filtre de Kalman est l'une des méthodes utilisées en environnement dynamique [Maybeck, 1990] quand il est nécessaire de fusionner des données redondantes de bas niveau au cours du temps. De nombreuses extensions de ce filtre ont été proposées tels que les filtres $\alpha\beta$ et $\alpha\beta\gamma$. Ces filtres sont une simplification du filtre de Kalman destinés à alléger les calculs. Les filtres multi-modèles combinent plusieurs modèles d'évolution afin de choisir le meilleur à chaque instant. Ils peuvent aussi fusionner les résultats de chacun des modèles au cours du temps [Barret, 1990].

Dans le cas où le système peut être décrit par un modèle linéaire et que l'erreur associée autant au système qu'aux capteurs peut être modélisée par un bruit blanc Gaussien, le filtre de Kalman fournira d'uniques estimations statistiquement optimales pour les données fusionnées [Luo and Kay, 1989].

Nous proposons de diviser le filtrage de Kalman en 5 étapes [Crowley and Demazeau, 1993] :

– représentation d'état : un modèle dynamique de l'environnement est une liste de primitives décrivant une partie de l'environnement à l'instant t. Chaque primitive représente une estimation de l'état local de l'environnement comme une conjonction de propriétés estimées $\widehat{x}(t) \equiv \{\widehat{x}_1(t), \widehat{x}_2(t), \ldots, \widehat{x}_n(t)\}$. L'état actuel de l'environnement est estimé par un processus d'observation qui projette l'environnement sur un vecteur d'observations $Y(t)$ en prenant en compte le bruit qui peut perturber le processus d'observation. $\widehat{x}(t)$ et $Y(t)$ doivent être accompagnés d'une estimation de leur incertitude. Ainsi des observations successives feront varier le facteur de confiance au cours du temps ;

– prédiction : cette étape permet de projeter le vecteur estimé $\widehat{x}(t)$ sur une valeur prédite $X^*(t + \Delta t)$ et aussi de projeter l'incertitude estimée sur l'instant $t + \Delta t$;
– correspondance entre l'observation et la prédiction : cette étape suppose une continuité temporelle et calcule une distance de Mahalanobis entre les propriétés prédites et observées. Un seuil de rejet permet de séparer les bonnes occurrences des fausses alarmes ;
– mise à jour : quand on a vérifié qu'une observation correspond à la prédiction, le filtre de Kalman procède à une estimation de l'ensemble des propriétés et leurs dérivées à partir de l'association de l'ensemble de propriétés prédites et observées. Le point intéressant en fusion de données, est que le filtre de Kalman fournit aussi une estimation de la précision associée aux éléments de ces ensembles ;
– élimination de primitives incertaines et ajout de nouvelles primitives au modèle : cette étape utilise les facteurs de confiance précédemment dérivés.

Dans [Crowley and Demazeau, 1993], une discussion complète est donnée sur l'utilisation de propriétés symboliques. Dans [Bar-Shalom and Li, 1995], des nombreuses applications à la poursuite multi-cibles sont données. [Murphy, 1999] montre une reformulation des filtres de Kalman sous forme de réseaux bayésiens.

1.2.7 Fusion multi-agents

Cette section n'entre dans aucun des deux cadres précédents, car il s'agit plus ici d'architecturer un système de fusion de données, que de proposer une technique ou des algorithmes permettant de faire de la fusion de données. Cependant, dans le cas où l'on utilise plusieurs algorithmes simples, l'architecture d'un système de fusion de données devient primordiale et participe à la réussite du système.

Les techniques multi-agents sont actuellement en plein essor. Parmi les raisons de leur grande acception, nous pourrions citer les suivantes :

– ces modèles permettent d'incorporer les différents acteurs du système de fusion d'une façon modulaire et souvent indépendante,
– la fusion de données est souvent hiérarchique, et de nombreux modèles multi-agents ont cette structure même, ce qui les rend adéquats pour faire de la fusion,
– le problème du contrôle de la fusion peut être rendu indépendant du problème de la fusion lui-même. Dans les modèle multi-agents, la connaissance associée au contrôle peut être facilement séparée pour chaque processus opérationnel. En outre, des stratégies sophistiquées de contrôle sont possibles, avec, par exemple, des négociations complexes entre algorithmes de fusion et agents.

Dans un premier temps, les techniques multi-agents s'appliquent surtout à donner une architecture plus performante au processus de fusion. Chaque agent est libre d'implémenter un algorithme ou une technique de fusion différente. L'intérêt est donc de pouvoir fusionner des données grâce à un mécanisme approprié d'interaction entre les différents agents. Le modèle le plus courant est celui du tableau noir (blackboard) [Llinas and Antony, 1993] où la communication entre agents est implémentée à travers une mémoire commune. Ses applications sont l'interprétation d'images, la navigation robotique [Ayari, 1996] ou encore l'évaluation de situations [Brogi et al., 1988].

Dans un deuxième temps, de nouvelles approches sont proposées où la fusion est réalisée de manière implicite, par un ensemble d'agents coopérants et interagissant ensembles. Ces techniques, dites émergeantistes, ont pour but d'arriver à fusionner des données, en utilisant des agents au comportement simple [Ayari, 1996]. Elles sont généralement appliquées à la fusion de capteurs, ou encore à l'analyse et au traitement d'images.

La particularité de ces techniques tient essentiellement à la granularité des agents. En effet, l'originalité de l'approche émergeantiste est l'utilisation d'un grand nombre d'agents, souvent simples (voire des agents purement réactifs), mais dont l'ensemble de leurs interactions permet l'émergence d'un comportement complexe, qui n'est pas codé au départ dans le fonctionnement interne des agents. De plus, ce type d'approche se distingue des modèles de fusion du type tableau noir où les agents sont tous différents et le fonctionnement global du système est connu *a priori*.

1.2.8 Conclusion sur l'état de l'art

Nous venons de présenter un panel non exhaustif des techniques de fusion de données. Cette présentation a néanmoins l'avantage de proposer un certain nombre de solutions particulièrement adaptées à la fusion multi-capteurs dans des environnements dynamiques où il est nécessaire de prendre en compte une certitude et une précision fluctuante des capteurs au cours du temps. Ces approches sont adaptées et adaptables à de nombreux problèmes, mais n'abordent le problème de la fusion que d'un point de vue technique. Dans la section qui suit, je vais m'attacher à présenter une vue plus globale de la notion de fusion de données, en introduisant le concept de processus de fusion de données, ainsi que celle de gain dans un processus de fusion de données. Un exemple détaillé sera présenté et appliqué au projet Diatelic^{TM}décrit dans le chapitre 1.

2 Une approche générique de la fusion de données dans les systèmes dynamiques

La conception d'un système présentant un comportement intelligent implique souvent l'utilisation d'une grande quantité de données, souvent riches et complexes, de nature et de fiabilité différentes, bruitées, imprécises. Ces données proviennent de capteurs ou de sources d'informations diverses et hétérogènes. Ces capteurs sont souvent à la base d'un processus de perception d'un environnement et les données fournies ne permettent pas une description complète et explicite de l'environnement (environnement est partiellement observable). Dans le cadre des systèmes dynamiques, nous sommes confrontés à une quantité importante de données, en provenance de multiples sources hétérogènes, de niveaux d'abstraction différents, bruitées, incertaines, redondantes et d'une durée de vie quelconque. Au final, nous souhaitons obtenir une information de meilleure qualité au sens de l'application. Cette notion de qualité peut être rattachée à des notions plus classiques telles que la précision, la confiance, la complétude. Il peut aussi s'agir d'obtenir une information plus synthétique ou résumée, ou en d'autres termes une information d'un plus haut niveau d'abstraction.

La première partie de ce chapitre a fait un état de l'art des techniques existantes dans le domaine de la fusion de données, appliquées essentiellement à la perception d'environnement. Puisque le but d'un système de fusion de données est que l'agent réalisant la fusion puisse fournir un résultat

de meilleure qualité qui servira à prendre une décision, ce dernier aura alors besoin d'observer efficacement son environnement pour rendre cette décision la plus pertinente possible. Il devra bien sûr utiliser au mieux l'ensemble des données dont il dispose.

Si la décision prise en combinant et en mélangeant l'ensemble des informations est de meilleure qualité que si l'on avait utilisé les informations séparément, alors la fusion de données aura été utile pour cet agent.

L'intérêt de la fusion réside alors dans son utilité par rapport aux buts fixés à l'agent et amène les questions suivantes :

– est-t-il nécessaire de fusionner les données ?
– doit-on fusionner l'ensemble des sources ou simplement les décisions issues du traitement des sources ?
– quel est l'algorithme le plus adapté ?
– peut-on utiliser plusieurs algorithmes ? Si oui, quel est la meilleure architecture permettant de connecter ces différents algorithmes ?

La question principale reste donc : *fusionner des données, est-ce-utile et à quel point ?*

Dans la suite de ce chapitre, je vais tenter d'apporter les premiers éléments de réponse, en présentant tout d'abord une définition d'un processus de fusion de données, puis une typologie associée à la notion de processus de fusion de données. Ceci permettra de poser les bases nécessaires à l'introduction de la notion de gain qualifié dans un processus de fusion de données. Un exemple, basé sur l'étude du système Diatelic^{TM}, sera ensuite présenté.

2.1 Définition et typologie d'un processus de fusion de données

Les sources de données d'un système basé sur la fusion de données, ont chacune des caractéristiques différentes, mais toutes doivent être utilisées conjointement, afin d'obtenir un bénéfice. Ainsi le choix d'un ou de plusieurs algorithmes de fusion dépendra de l'application, des caractéristiques des sources de données et enfin du bénéfice que l'on souhaite obtenir. Parmi les buts à atteindre, on peut citer :

– l'utilisation exhaustive des données : l'idée est d'extraire des données disponibles un maximum d'information afin d'obtenir une vue la plus globale possible de l'environnement observé. Par exemple, dans le cas de la poursuite de cibles aériennes, on voudra avoir une vue la plus exhaustive possible de l'ensemble des aéronefs en présence, ainsi que le type et la trajectoire de chacun. Ceci est nécessaire afin de planifier les atterrissages et les décollages (dans le cas d'une application civile) ou encore d'estimer le nombre et la menace représentés par les forces ennemies (dans le cas d'une application militaire) ;
– le mélange et la combinaison les plus pertinentes, dans le but d'extraire de nouvelles informations inaccessibles auparavant et de renforcer la certitude que l'on a sur les informations déjà connues. Dans le cas d'un monitoring médical (diagnostic d'un état physiologique du patient remis à jour en permanence), lorsque le système de monitoring découvre un état alarmant, celui-ci devra attendre d'être confirmé avant de déclencher une alarme et de prévenir un médecin. Ceci est nécessaire pour éviter les fausses alarmes qui rendraient, le cas échéant, le système de monitoring peu crédible d'un point de vue médical. Un cas particulier est l'abstraction et la synthèse de données où l'idée est d'utiliser plusieurs ensembles de données de bas niveau, afin d'obte-

nir un ensemble unique final de haut niveau d'abstraction. Typiquement, il s'agit de l'analyse d'images, où plusieurs images sont utilisées afin d'extraire et de reconnaître les objets présents dans cette image. Le résultat est souvent une description de l'image [Bloch and Maître, 1994].
– le renforcement du signal utile et l'atténuation du bruit afin d'obtenir une information sinon digne de confiance, en tout cas utilisable. Dans le cas de la poursuite d'avions militaires grâce à un radar, le signal est souvent brouillé. Ceci est dû à de mauvaises conditions météo ou à un brouillage ennemi volontaire. L'utilisation de plusieurs radar permet d'éliminer une partie du bruit et d'obtenir ainsi un signal dans lequel on peut avoir confiance [Barret, 1990].

2.1.1 Définition

Ainsi, en première approche, la fusion de données peut être vue comme un processus dont le but est de combiner l'information dans le but d'améliorer le processus de prise de décision.

A la vue de ce qui a été dit jusqu'à maintenant, cette définition peut être étendue de la façon suivante :

Définition d'un processus de fusion de données *Un processus de fusion de données vise à l'association, la combinaison, l'intégration et le mélange de données fournies par au moins deux ensembles de données issues d'une même source au cours du temps ou de plusieurs sources, éventuellement au cours du temps. Le but d'un processus de fusion de données est de fournir un autre ensemble de données de meilleure qualité que si l'on avait utilisé chaque ensemble initial séparément. La notion de qualité dépend de l'application.*

Ainsi, trois éléments fondamentaux ont été identifiés : les sources de données qui fournissent les ensembles initiaux de données qui seront fusionnés, le processus de fusion de données et les algorithmes qui lui sont associés, et enfin le résultat, qui est un ensemble de données lui-même.

Si le processus de fusion de données agit comme un filtre, alors il peut aussi être considéré comme une source de données pour un autre processus de fusion de données. La définition est donc récursive.

2.1.2 Sources de données

En dépit du fait que chaque processus utilisant au moins deux sources d'information pourrait être appelé un processus de fusion de données, je me restreindrai ici aux problèmes où les sources contiennent des données incertaines ou bruitées. Ce type de sources de données est particulièrement fréquent dans les problèmes de perception de l'environnement : les capteurs électroniques sont en effet sujets au bruit et ne fournissent jamais une information sûre. Par exemple, pour augmenter ses capacités de perception, un robot recevra en général plusieurs types de capteurs tels que des capteurs infra-rouges, une ou plusieurs caméras, des sonars ou encore des capteurs de contact. Ceci lui permettra de percevoir son environnement selon plusieurs modalités et ainsi d'augmenter sa capacité de détection et de compréhension des objets qui l'entourent.

On peut ainsi définir les caractéristiques essentielles suivantes pour une source de données. Une source peut donc fournir des données :

– *numériques/symboliques* : la source peut fournir des données numériques ou des connaissances uniquement qualitatives.
– *objective*s : l'information que donne la source est vérifiable et éventuellement démontrable. Il n'y a aucune incertitude associée à cette information. Il peut s'agir, par exemple, de faits exacts dans une base de connaissances.
– *subjective*s : les informations sont incertaines et bruitées. C'est le cas le plus courant. En effet, il y a une différence notable entre les données réelles et les données observées (et donc fournies pas la source de données). Ceci peut provenir :
  – soit de la qualité du capteur servant de source de données,
  – soit d'un agent intermédiaire agissant comme filtre entre la donnée réelle et la donnée observée. Dans le cas d'un système de télé-médecine, et en particulier dans le cas de Diatelic^{TM}, le patient est lui-même responsable de la saisie de ses données. Parfois, lorsqu'un patient a pris trop de poids, il ne souhaite pas que le médecin le sache, souvent par peur des réprimandes. Il fournit donc au système expert Diatelic^{TM}, un poids inférieur à son poids réel. La donnée fournie n'est donc plus digne de confiance. Néanmoins elle devra quand même être utilisée, car il s'agit de la seule disponible.
– *temporelles ou causales* : il s'agit plus ici d'une caractéristique sur les relations entre les données que sur les données elles-mêmes. En effet, une source peut aussi fournir des données qui sont toutes corrélées de façon temporelle, causale (éventuellement une autre modalité). Il existe donc un ordre temporel entre les données, ou encore une relation de cause à effet entre les données.
– *spatiales* : dans l'espace de représentation des données, il existe une relation entre les données, mesurable par une métrique. L'ordre des données n'est pas forcément strict, comme ce serait le cas dans une source temporelle.

Cette première approche ne concerne bien sûr que les données à l'intérieur d'une source. Dans le paragraphe qui suit, je vais m'intéresser aux relations entre les sources de données.

2.1.3 Relations entre les sources de données

Connaître les relations entre les sources est important car ceci permettra de caractériser, d'ordonner ou encore de corréler les différentes sources de données disponibles. On pourra alors décider qu'elle est la meilleure stratégie ou le meilleur algorithme pour fusionner les données. Nous avons identifié les relations suivantes entre les sources [Bellot et al., 2002b] :

– *distribution* : les sources de données donnent des informations sur le même environnement mais chacune en ayant soit un point de vue différent, soit la capacité de n'observer qu'une partie de l'environnement. Par exemple, un véhicule robotisé utilisera deux caméras, une pour l'avant, une pour l'arrière, lui permettant ainsi de voir l'ensemble de son environnement. Un autre exemple est celui des radars longue et courte portée fréquemment utilisés sur les bâtiments de la Marine. En effet, chaque bateau possède un radar ayant des capacités différentes de celui des autres bateaux. Mais l'ensemble des données est fusionnée afin de fournir une image de la même zone : celle où se trouvent les bateaux.
– *complémentarité* : chaque source perçoit uniquement un sous-ensemble de l'environnement global. La fusion de l'ensemble de ces sources donnera une vue élargie de l'environnement. Les sources, dans ce cas, ne sont pas redondantes.

– *hétérogénéité* : chaque source fournit des informations dont les caractéristiques sont complètement différentes des autres sources. Il s'agit là d'un cas très classique. Par exemple, dans le système Diatelic^{TM}, chacun des capteurs fournit des données physiologiques s'intéressant à différentes caractéristiques du corps humain : poids, tension artérielle ou encore température.
– *redondance* : les sources de données décrivent le même environnement avec des informations de même caractéristique. C'est le cas lorsqu'un capteur est doublé afin de prévenir une panne ou simplement pour augmenter la précision ou la confiance accordée aux mesures. Si des différences apparaissent entre les mesures, alors un processus de fusion de données aura intérêt à utiliser ces différences pour améliorer la qualité du traitement de données et extraire de nouvelles informations. La redondance est nécessaire pour sécuriser ou rendre plus robuste un système percevant un environnement.
– *contradiction* : l'espace de définition de deux sources est le même mais elles fournissent des informations totalement différentes l'une de l'autre, tout en observant le même environnement. Par exemple, une source pourra fournir comme mesure l'intervalle de $[0.1, 0.3]$ alors qu'une autre source fournira l'intervalle $[0.31, 0.41]$.
– *concordance* : les données fournies par deux sources sont compatibles entre elles et se corroborent l'une l'autre. Par exemple, une première mesure issue d'une source de données fournira l'intervalle $[1, 5]$ et une seconde mesure issue d'une deuxième source de données fournira la mesure (plus précise) $[2, 4]$ pour une observation faite sur le même environnement dans les mêmes conditions.
– *discordance* : les informations fournies par deux sources sont incompatibles entre elles. La différence entre la contradiction et la discordance est qu'ici une source ne va pas invalider l'information donnée par l'autre source, mais simplement donnera un résultat qui est aussi acceptable (dans le cadre de l'application) que celui de la première source. Par exemple, une source de données peut fournir l'intervalle $[-1, 2]$ et une autre source, l'intervalle $[3, 5]$. Si le domaine des capteurs se situe au moins dans l'intervalle $[-1, 5]$ alors les deux réponses, bien que discordantes, sont valides.
– *différence de granularité* : les sources de données fournissent des données redondantes mais chaque source observe l'environnement à une échelle différente.
– *synchronisation* : les données fournies par les sources sont concordantes au niveau du temps. Les sources de données doivent donc fournir un flux de données et non pas un ensemble entièrement disponible. La notion de temps et de date de délivrance d'une donnée est nécessaire. De même, deux sources de données peuvent être asynchrones entre elles. Par exemple, une source peut fournir un flux de données, alors qu'une autre source fournira simplement des informations à la demande.

Cette liste présente donc les caractéristiques des relations pouvant exister entre des sources de données. Cependant, il apparaît clairement que dans une application réelle, les sources en présence pourront être munies de plusieurs de ces relations.

Par exemple, si deux caméras observent une pièce selon un angle de vue différent, ces deux sources de données seront à la fois distribuées et complémentaires. Elles observent la même pièce (complémentaires) mais selon un point de vue différent (distribuées). Elles ne sont pas redondantes, car l'information fournie par chacune d'elles est différente, car ici encore, leurs points de vue sont différents. Elles ne sont pas discordantes, sauf dans le cas d'un dysfonctionnement de l'une des deux caméras. Elles ont la même granularité et sont synchrones. Elles sont normalement concordantes sauf, encore une fois, dans le cas d'un dysfonctionnement de l'une des caméras.

Cependant, l'interprétation des données avec un même algorithme d'analyse d'images peut donner des informations différentes. Ainsi, elles peuvent devenir discordantes et dans certains cas contradictoires (par exemple, si l'une des caméras observe une personne mais l'autre non, car cette personne est cachée par un objet pour l'autre caméra).

2.1.4 Qualité des données

Fusionner des données est intéressant seulement si cela permet d'augmenter la qualité des données en sortie du processus de fusion par rapport aux données en entrée. Cette notion de qualité est intéressante à plusieurs points de vue, car elle permet de caractériser le contenu d'un ensemble selon les besoins de l'application. Des données de mauvaises qualité peuvent avoir un impact négatif sur les résultats attendues d'une application. Si la fusion de données tend à améliorer la qualité des données en sortie, il est cependant légitime de s'interroger sur la qualité des données en entrée. En effet, la recherche d'un gain particulier dans un processus de fusion peut aussi passer par l'estimation de la qualité des données avant et après le processus de fusion.

En général, le point de vue du consommateur de données tend à définir des données de bonne qualité comme des données qui correspondent exactement à ses besoins et à l'utilisation qu'il en fait. La qualité des données est alors mesurée selon trois approches : intuitive, théorique et empirique. L'approche intuitive est préférée quand la sélection des attributs de qualité des données est basée sur l'expérience ou une compréhension intuitive de ce qui est un *attribut important*. La plupart des études sur la qualité des données se basent sur cette approche. Ainsi, si de nombreux attributs peuvent être utilisés, seuls un petit nombre d'entre eux reviennent régulièrement telles que la précision ou la fiabilité. Dans les systèmes d'informations (base de données, sites web, etc...), la satisfaction de l'utilisateur est aussi utilisée comme critère de qualité [Delone and McLean, 1992]. D'autres critères seront l'opportunité, la complétude, la pertinence, l'accessibilité, ou encore l'interprétabilité [Wang et al., 1996], [Naumann, 1998].

L'approche théorique pose le problème de savoir comment les données peuvent devenir déficientes durant le processus de fabrication de ces données. Parmi les diverses études faîtes sur le sujet, l'une d'entre elles utilise une approche ontologique dans laquelle les attributs de qualité des données sont dérivés sur la base des déficiences des données, qui sont définies comme les inconsistances entre une vue de l'environnement inférée à partir de sa représentation dans une base de connaissances et la vue que l'on obtient en observant directement l'environnement réel [Wand and Wang, 1996].

La mesure de la qualité des données reste cependant une tâche difficile et souvent subjective, car la qualité des données dépend fortement de l'application visée. Une approche empirique de mesure de la qualité des données a été proposée par R.Y.Wang et D.M. Strong dans *[Wang et al., 1996]*. Cette étude est basée sur une étude expérimentale telle que celle utilisée dans les prospections marketing. Leur approche consiste à dépasser le concept de qualité des données à travers seulement la précision et à proposer une classification hiérarchique de la notion de qualité des données. Ils ont ainsi identifié quatre dimensions de la qualité de l'information, regroupant quinze attributs mesurables :

– la qualité intrinsèque,
– la qualité contextuelle,
– la qualité de la représentation,
– la qualité d'accessibilité.

La qualité *intrinsèque* est caractérisée par la précision, l'objectivité, la confiance et la réputation. La qualité *contextuelle* est définie par la valeur ajoutée, la complétude, le volume approprié de données, la pertinence et la continuité. La qualité *en représentation* est déterminée par la consistance, la concision, l'interprétabilité et la facilité de compréhension. La qualité *d'accessibilité* est repérée par la facilité de l'accès et la sécurité. Cependant, malgré cette classification, la mesure de la qualité d'une information à travers la comparaison ou le classement de sources de données reste difficile, il s'agit par ailleurs d'un problème de décision multi-attributs ([Naumann, 1998]). Ainsi, la mesure de la qualité des données d'une source, et son application au calcul d'un gain apporté par un processus de fusion de données passe par les questions suivantes :

1. existe-t-il une mesure de la qualité globale de l'ensemble des sources ?
2. existe-t-il une mesure de la qualité des données en sortie du processus de fusion ?
3. ces mesures sont-elles compatibles avec le gain souhaité ?
4. la stratégie de fusion choisie permet-elle d'obtenir un gain positif ? de le maximiser ?

En se basant simplement sur la définition commune de la qualité des données qui est l'aptitude des données pour une intention, et si l'utilisation des données a pour but la prise de décision, la qualité des données est dépendante des décisions considérées. Des données de grande qualité peuvent alors être utilisées pour déterminer quelle est la meilleure parmi un ensemble de décisions possibles. Ainsi des données de mauvaise qualité entraîneront la plupart du temps une décision de mauvaise qualité, et des données de bonne qualité feront tendre vers la découverte de la meilleure décision dans une situation donnée [Arnborg et al., 2000].

Dans le cadre de la théorie de la décision, l'utilisation de l'information est le mieux décrite par la prise de décision basée sur l'utilité espérée (de la décision). Ce point nous permet de relier la recherche d'une meilleure qualité à la notion plus vaste de la conception des agents à rationalité limitée [Zilberstein, 1995].

En effet, imaginons que le processus de fusion définisse un modèle de comportement pour un agent, alors le but de cet agent serait de maximiser la qualité des données en sortie ou, autrement dit, de maximiser le gain obtenu par le processus de fusion. Par exemple, pour un robot, le but sera d'augmenter la précision et la complétude de sa connaissance sur l'environnement dans lequel il évolue. Dans le cas d'un diagnostic médical, la confiance accordée aux diagnostics représentera le point important à maximiser. Mais dans le cas d'un système de monitoring médical, l'opportunité des alarmes sera toute aussi importante. Plus le système fournit des alarmes pertinentes, plus la fusion apporte des données de qualité en sortie. Une alarme pertinente est une alarme justifiant réellement une intervention médicale ou une modification de la thérapie appliquée au patient. Une alarme non pertinente ne pourra être qu'ignorée par le personnel médical.

Un agent rationnel souhaite accomplir une tâche de façon optimale dans le cadre d'une architecture d'agent donnée. La limite d'un processus de fusion tient dans la quantité d'information globale : le résultat de la fusion n'est qu'une autre façon d'exprimer les données issues de l'ensemble des sources de données. L'agent cherche à améliorer le gain (éventuellement en améliorant la qualité des données), et non la quantité d'informations.

En se basant sur ces considérations, je vais définir, dans la section suivante, la notion de gain dans un processus de fusion de données et plus particulièrement la notion de gain qualifié, qui

me permettra de donner une vue unifiée des notions de recherche d'une meilleure qualité et d'une utilité optimale en fusion de données.

2.1.5 Notion de gain qualifié

A partir des précédentes considérations, je vais maintenant présenter la notion de gain qualifié, qui s'inspire directement des notions de qualité sur les données et de l'idée qu'un processus de fusion de données peut être vu comme un agent qui va tenter d'atteindre un but unique : obtenir le gain en qualité le plus élevé. Ce gain dépend fortement de l'application. Néanmoins, en se basant sur les divers travaux sur la qualité des données présentés dans la section précédente, il est possible de déterminer quatre qualificatifs de gain englobant toutes les situations possibles.

Dans le but de simplifier le problème de l'estimation de la qualité d'un processus de fusion de données, une qualification de la notion de gain est ici proposée. Ce point de vue original permet de déterminer facilement les exigences d'un processus de fusion de données.

Gain en représentation

Le gain en représentation est conditionné par le niveau d'abstraction des données en entrée du processus de fusion. Il est obtenu lorsque l'ensemble des données en sortie du processus de fusion possède un niveau d'abstraction plus élevé ou une plus grande granularité que chaque ensemble de données fourni par les différentes sources. Le nouveau niveau d'abstraction ou la nouvelle granularité doivent apporter une sémantique plus riche sur l'ensemble des données en sortie par rapport aux ensembles de données de chaque source.

Par exemple, un algorithme de classification, en regroupant et abstrayant des données, définit des classes jusqu'alors inconnues et apporte ainsi un gain en représentation positif et non nul.

Gain en certitude

Le gain en certitude est lié à la confiance que l'on accorde aux données. Un gain en certitude intervient sur un fait déjà connu (au moins partiellement) et pour lequel on a une confirmation. Par exemple, dans un modèle de langage en reconnaissance de la parole, un gain en perplexité sans gain en reconnaissance peut donc être vu comme un gain en certitude ; c'est-à-dire que les mots prononcés par le locuteur sont *mieux* reconnus, mais le système n'en a pas plus reconnu pour autant [Bigi, 2000].

Gain en précision

Le gain en précision est lié à la variabilité des données obtenues en sortie par rapport aux données initiales du processus de fusion. Si cette variabilité (mesurée, par exemple, par rapport à une valeur moyenne) tend à diminuer alors le gain en précision augmentera. Si les données initiales sont bruitées et/ou entachées d'erreur, alors pour obtenir un gain en précision, le processus de fusion de données tentera de réduire le bruit et d'éliminer les erreurs. En général, le gain en précision et le gain en certitude sont corrélés. Cependant la détermination d'un gain en précision nécessite une connaissance *a priori* sur la valeur moyenne des données en sortie du processus de fusion.

Dans les applications de poursuite de cibles aériennes, il est possible d'améliorer la précision de la position estimée de la cible en utilisant plusieurs radars. Il y a gain en précision à partir du moment où la cible est déjà connue. Si la cible n'est pas certaine, alors le gain en précision s'accompagne d'un gain en certitude.

Gain en complétude

Le gain en complétude renseigne sur l'étendue des connaissances qu'un agent a sur son environnement, sachant que le modèle de perception est basé sur un processus de fusion de données. Ainsi, l'apport de nouvelles connaissances sur l'environnement complètera la vue que l'agent a déjà sur cet environnement même. De plus, si les données sont redondantes et/ou concordantes, le gain en complétude peut être accompagné d'un gain en certitude et en précision. Cependant, un fort gain en complétude peut aussi se faire au détriment du gain en précision. En effet, l'agent peut percevoir très rapidement son environnement, mais sans vérifier l'exactitude de ses nouvelles connaissances. Ainsi, la certitude sur ses connaissances n'augmente que très peu. Par exemple, dans le cadre d'un apprentissage non-supervisé, un gain en complétude peut être observé au fur et à mesure que l'algorithme d'apprentissage découvre de nouveaux concepts ou de nouvelles classes au sein des ensembles de données qu'il fusionne.

Les quatre qualifications, que je propose, englobent l'ensemble des situations et d'autres qualifications ne seront alors qu'une composition des ces quatre qualificatifs initiaux. Cette approche permet donc de déterminer les points intéressants à étudier dans un processus de fusion de manière à évaluer son utilité et, éventuellement, de manière à lui donner le moyen de maximiser cette utilité en essayant de maximiser en permanence un ou plusieurs gains.

Dans la section suivante, je vais présenter le projet Diatelic^{TM}sous un nouveau point de vue : celui de la fusion de données et ainsi illustrer la notion de processus de fusion de données et la notion de gain qualifié avec ce problème de télé-médecine intelligente.

2.2 Diatelic^{TM} : un processus de fusion de données

2.2.1 Situation du problème

L'intérêt porté à Diatelic^{TM}tient au fait que, dans ce problème, on cherche à monitorer une pathologie chronique (l'insuffisance rénale) et plus particulièrement à détecter les problèmes d'hydratation. Puisque les seules données dont nous disposons sont celles fournies par le patient (température, tension, poids, type de poche utilisé), nous sommes obligés de les fusionner afin d'en extraire l'information pertinente : l'estimation du taux d'hydratation. L'idéal sera bien sûr de détecter un tel incident pathologique avant qu'il n'arrive de manière à réguler au mieux l'état du patient. La conséquence directe d'une telle régulation, au niveau du traitement des données, est que l'on risque de ne jamais atteindre un état pathologique réel. On n'aurait donc jamais d'indication qu'un état pathologique a été correctement détecté. En contre partie, si le patient ne se retrouve jamais dans une situation critique, alors le système de télé-médecine aura parfaitement fonctionné et atteint son but principal : la sécurité et le confort maximal pour le patient.

La fusion des données fournies chaque jour ainsi que la fusion des états des jours précédents va donc permettre d'estimer avec une plus grande certitude l'état du patient et ainsi de voir son évolution au cours du temps. Cette évolution déterminera les tendances du patient et permettra la détection des cas critiques. Dans un cas idéal, on devrait prendre en compte l'ensemble des informations concernant le patient. Cependant, si l'historique des diagnostics et des observations est disponible, il n'en est pas de même d'autres données qui seraient pourtant très riches d'informations telles que le régime alimentaire du patient, ou encore le type et la quantité d'exercices physiques qu'il fait chaque jour.

Il est donc clair que les informations disponibles doivent être fusionnées afin d'obtenir les éléments recherchés. Dans la suite de cet exemple, je vais reprendre chaque élément constituant un processus de fusion de données et l'appliquer au projet Diatelic^{TM}.

2.2.2 Détermination des sources

La première étape est de déterminer les sources de données dont nous disposons. Précisons un point : une source ne délivrera que des données du même type. Ainsi, nous pouvons en premier lieu identifier 6 sources :

- la température du patient, relevée chaque jour ;
- la tension artérielle, relevée chaque jour, qui est en fait composée de quatre valeurs : les tensions systoliques et diastoliques prises en position debout et en position couchée ;
- le poids, relevé chaque jour ;
- le poids idéal (parfois appelé *poids sec*) qui peut être modifié chaque jour par le médecin, et qui est pour cette raison considéré comme une source de données synchrones avec les trois précédentes ;
- le type de poche de dialyse utilisé, relevé chaque jour, et qui peut être changé quotidiennement sur ordre du médecin ;
- l'ultra-filtration, relevée chaque jour, qui correspond à la différence de volume en entrée et en sortie de dialyse.

A ces 6 sources, nous pouvons rajouter les connaissances médicales qui doivent être explicitement intégrées dans le processus de fusion. En effet, si nous n'utilisons que ces 6 sources-là, nous ne pouvons déterminer avec précision l'état d'hydratation du patient. Cependant si nous prenons en compte le fait que le patient souffre d'insuffisance rénale, nous devons alors associer au processus de fusion l'ensemble des connaissances médicales disponibles pour le traitement de telles pathologies. Bien sûr, ceci peut paraître trivial. La connaissance sur le domaine d'expertise est souvent une connaissance *a priori*. C'est aussi le cas dans Diatelic^{TM}. Mais ma remarque se justifie par le fait que dans certains cas, la connaissance sur le domaine d'expertise peut évoluer pendant le déroulement du processus de fusion et doit alors être prise en compte. Cette connaissance devient alors une source de données (vraisemblablement asynchrone par rapport aux sources de données capteurs) qu'il faudra fusionner avec les autres.

Enfin, une 8ème source existe et il s'agit de l'état quotidien du patient et en particulier de son état d'hydratation. L'existence de cette 8ème source sera justifiée dans le paragraphe suivant sur les caractéristiques et les relations entre les sources.

2.2.3 Caractéristiques des sources

En supposant que la connaissance médicale n'évolue pas, seules les 6 premières sources sont intéressantes. Tout d'abord, les valeurs mesurées sont toutes *subjectives* car, dans le cas de Diatelic^{TM}, le patient est responsable de la saisie des données. Il peut donc les modifier à volonté. L'expérience du médecin a montré que le patient ne triche que rarement sur sa température ou sur l'ultra-filtration, mais qu'il est plus enclin à, parfois, modifier légèrement la valeur du poids ou de la tension artérielle. De la même manière, le poids idéal est considéré comme émanant d'une source subjective, car il est laissé à l'appréciation du médecin. En revanche, les types de poche utilisés sont considérés comme des valeurs objectives, car il n'y a pas possibilité d'interpréter la valeur réelle !

Les sources de données sont considérées comme temporelles, car chaque jour le patient fournit un ensemble de valeurs. Cependant, les relations causales entre les données au sein de chaque source ne sont pas clairement définies. En considérant le fait qu'il s'agit d'un même et unique patient, et que l'état de ce patient est fortement dépendant de son état le jour précédent, il apparaît clairement que les mesures relevées par le patient sont dépendantes de leur valeur de la veille. Cependant, le lien de causalité est particulièrement difficile à établir et, dans le meilleur des cas, ne pourra être établi que statistiquement. Dans ce cas, l'approche la plus classique est de considérer que le problème est markovien et de créer un état *résumant* l'historique des observations. Dans le cas de la télé-médecine, cette approche est satisfaisante mais pas parfaite. En effet, il existe de nombreux autres paramètres à prendre en considération si l'on veut correctement déterminer les liens causaux entre les observations :

– est-ce-que le patient s'est reposé aujourd'hui ? A-t-il beaucoup dormi ?
– a-t-il mangé ce matin ? A-t-il peu bu ?
– a-t-il fait des exercices physiques ?
– sa nourriture d'aujourd'hui est-elle riche ou pauvre en eau ? et celle d'hier, ou plus ?

Il existe encore de nombreuses autres questions, trop nombreuses pour être correctement intégrées. De plus, certains événements plus anciens (remontant à 2, 3 ou 4 jours) peuvent encore avoir de l'influence sur le patient. Enfin, dans le cadre de la télé-médecine, un problème supplémentaire apparaît : celui du manque de données. En effet, si, un jour, le patient ne peut avoir accès à un terminal Diatelic^{TM}, alors il ne pourra renseigner la base de données sur ses valeurs du jour. Dans ce cas, le système devra tout de même fournir une information sur l'état d'hydratation du patient.

Ces considérations sur la nécessité de *résumer* l'état du patient nous permettent de justifier l'existence de la 8ème source : l'état d'hydratation du patient. Cette 8ème source est aussi subjective car ses informations sont calculées au sein de l'algorithme de fusion lui-même. Néanmoins, il apparaît que l'état d'hydratation de la veille doit être fusionné avec les données quotidiennes. Remarquons que ce type précis de fusion est aussi une mise à jour de la connaissance que l'on a sur l'état du patient. Et réciproquement, la mise à jour de cette connaissance (à l'instant t) est donc une fusion des données issues des capteurs et de la connaissance que l'on avait jusqu'à l'instant $t-1$.

2.2.4 Relations entre les sources

Chaque source s'intéresse à un aspect particulier du patient ; elles sont donc distribuées. Chacune d'entre elles apporte des informations différentes. On considère alors qu'elles sont hétérogènes

plutôt que complémentaires. Ce cas est particulièrement fréquent en diagnostic médical. Les données sont fournies par le patient chaque jour, ou presque. Aucun autre capteur que ceux définis précédemment n'intervient. Donc les sources ne sont pas redondantes entre elles, ceci étant dû à leur hétérogénéité.

Enfin les sources sont concordantes. En effet, ceci est particulier au problème de Diatelic™. L'évolution de la tension artérielle d'un patient permet en première approche de déterminer les cas d'hyperhydratation ou de déshydratation. Mais, la tension n'est pas suffisante. L'évolution du poids permet aussi de deviner les cas critiques, car une chute de poids est souvent la cause d'une déshydratation, et une prise de poids peut être due à une hyper-hydratation. Cependant, la fusion de ces deux mesures est nécessaire pour obtenir un diagnostic complet sur l'état d'hydratation du patient. Cette nécessité est issue de l'expertise des médecins néphrologues. Ils ont observé que les changements dans l'état d'hydratation d'un patient sont la plupart du temps accompagnés de modifications de la tension et du poids du patient : chute ou augmentation brutale de la tension et/ou du poids.

Enfin, nous noterons que les sources sont parfaitement synchronisées entre elles, car le patient fournit l'ensemble des valeurs en même temps. Par contre, l'état d'hydratation du patient (la 8ème source) ne sera pas synchronisé avec les 6 autres dans le cas où les données ne seraient pas saisies un jour par le patient.

La dernière remarque porte sur la contradiction. Nous faisons ici l'hypothèse que les sources de données ne sont pas contradictoires. Dans le cas où le patient ne respecterait pas le protocole médical, cette hypothèse serait invalidée. Néanmoins, nous supposons que les données reflètent à peu près bien les valeurs physiologiques réelles, et par conséquent, une source ne pourra en contredire une autre, car elles observent le même et unique patient.

2.2.5 Conséquences de cette étude préliminaire

Cette première étude nous permet de mieux cerner le problème de Diatelic™, et plus généralement celui du diagnostic en télé-médecine. Nous avons identifié 8 sources de données, dont une qui est apparue après avoir considéré le problème des liens entre les observations d'un jour à l'autre. Ensuite, nous avons choisi de résoudre le problème posé par Diatelic™en utilisant un modèle *résumant* l'historique des observations au sein d'un même état. Ainsi, nous pouvons modéliser Diatelic™avec des techniques markoviennes. Enfin, ce modèle devra être en mesure de pouvoir prédire ou estimer un état même lorsque les données ne sont pas présentes. Pour terminer, même si cela parait évident, il faut noter le fait que les données sont incertaines et bruitées et que le ou les algorithmes nécessaires au processus de fusion devront raisonner en environnement incertain.

Plusieurs modèles répondent à un ou plusieurs de ces critères. Par exemple, les HMM assurent que les observations à un instant t dépendront uniquement d'un état caché X_t et cet état caché dépendra uniquement de l'état X_{t-1} à l'instant précédent.

Les filtres de Kalman permettent d'estimer un état et de *prolonger* cette estimation même si aucune observation n'est disponible. Cependant il est nécessaire d'avoir un modèle linéaire décrivant l'évolution du système. Enfin, les réseaux bayésiens permettent de raisonner en environnement incertain, mais aussi de gérer plusieurs sources de données hétérogènes et aussi les dépendances

causales directes ou indirectes entre plusieurs variables. Nous avons en effet vu que le poids et la tension sont fortement liés. Ce dernier modèle est donc particulièrement intéressant car :

1. il permet de modéliser un processus markovien, hypothèse faite pour Diatelic^{TM},
2. il permet de représenter des influences causales entre des variables aléatoires, ce qui nous permet de modéliser la connaissance médicale mise en oeuvre dans Diatelic^{TM},
3. il permet de gérer des connaissances incertaines grâce à l'utilisation d'une représentation probabiliste de la connaissance et de fusionner des observations avec des historiques de diagnostic, par l'utilisation de mécanismes bayésiens de mise à jour de la connaissance.

Ce modèle parait donc être particulièrement adéquat à la modélisation d'un problème comme celui posé par Diatelic^{TM}.

2.2.6 Gain qualifié appliqué à Diatelic^{TM}

Pour compléter cette étude, j'applique dans cette section la notion de gain qualifié à Diatelic^{TM}, de manière à exhiber les points importants à analyser dans le processus de fusion de données de Diatelic^{TM}.

Gain en représentation

Sachant que le diagnostic est toujours fait sur la variable (le taux d'hydratation), il n'y a pas de gain en représentation pendant le calcul du diagnostic. En effet, il y a toujours la même différence de niveau d'abstraction entre les sources de données et le résultat. Cependant, si nous utilisons les données issues des sources, ainsi que l'historique des diagnostics pour apprendre un profil de patient, nous pouvons espérer un gain en représentation pendant le processus d'apprentissage. Un profil de patient est l'ensemble des données caractérisant le patient et intervenant dans le calcul du diagnostic. Dans le cas d'un modèle tel qu'un HMM, ce profil de patient serait les matrices d'observations et de transition. Dans le cas d'un réseau bayésien, ce serait les paramètres du réseau (probabilités *a priori*).

Gain en certitude

Le gain en certitude est obtenu dans un cas précis : quand un problème sur l'état d'hydratation est détecté, si de nouvelles données viennent confirmer ce problème, alors la confiance accordée à la précédente découverte grandit. Si au contraire, les nouvelles données ne confirment pas cette découverte, alors il peut s'agir d'une fausse alarme, et il n'y aura pas de gain en certitude.

Cette étude du gain en certitude, dans le cas de Diatelic^{TM}, est très importante, car elle pourrait permettre d'éliminer les fausses alarmes et de ne fournir qu'une information toujours pertinente au médecin.

Gain en précision

Le gain en précision est en relation avec le gain en certitude. En effet, on obtient un gain en précision lorsque le système est capable de détecter le bon état pathologique (hyper-hydratation

ou déshydratation). Si le système donne le bon taux d'hydratation, alors le médecin pourra ajuster la thérapie et ainsi réguler au mieux l'état du patient.

Gain en complétude

Le gain en complétude est obtenu seulement dans le cas de l'apprentissage du profil du patient. Les paramètres représentant le profil du patient peuvent être appris au début ou bien adapté en cours de traitement (par exemple sur demande du médecin). En effet, si le médecin n'est pas en accord avec le diagnostic donné par le système, il peut soit attendre le lendemain pour voir si le système confirme ou infirme son diagnostic de la veille, soit modifier les paramètres du patient de manière à ce que le système donne un autre diagnostic, avec lequel le médecin sera, cette fois-ci, d'accord. Les informations que le médecin apporte pourraient être, dans ce cas, vues comme une nouvelle source d'information.

Ainsi, le gain en complétude sera obtenu lorsque le système apprendra un meilleur profil de patient et connaîtra donc mieux (ou plus largement) le patient.

3 Conclusion

Ce second chapitre a présenté une approche originale de la fusion de données et en particulier, j'ai introduit la notion de processus de fusion de données et de gain qualifié. Cette approche systématique de la fusion de données permet d'analyser avec une plus grande rigueur des problèmes faisant intervenir de multiples sources de données fournissant des données au cours du temps (flux de données). Le point important est l'introduction de la notion de gain qualifié, permettant d'étudier et de contrôler la qualité des données en sortie du processus de fusion. Cette étude du gain obtenu au cours du processus de fusion peut aussi nous renseigner sur l'efficacité des algorithmes de fusion utilisés et en particulier estimer si les algorithmes atteignent réellement les buts proposés par l'application. Cette approche permet donc d'estimer si il est intéressant de fusionner les données, et dans le cas contraire, elle peut donner une indication pour modifier le processus de fusion afin que ce dernier obtienne les résultats escomptés.

Cette vue de la fusion de données a été illustrée par un exemple basé sur le projet Diatelic^{TM}. Cet exemple est aussi une étude sur la conception du système Diatelic^{TM}en tant que processus de fusion de données. Ceci nous a permis d'extraire les problèmes suscités par les données particulières à Diatelic^{TM}et d'entrevoir certaines solutions avant même de choisir un modèle particulier. Le choix de ce modèle va donc être dicté par cette étude.

Dans la suite de cette thèse, je vais présenter le modèle utilisé par Diatelic^{TM}pour fusionner les données et obtenir un diagnostic. Ce modèle est basé sur les réseaux bayésiens dynamiques suite à l'étude faite dans ce chapitre. Nous montrerons en effet que les réseaux bayésiens nous permettent de répondre à un maximum des exigences du processus de fusion de données que j'ai décris dans ce chapitre. Dans le chapitre suivant, je présente une synthèse sur le raisonnement bayésien et causal utilisé pour faire du diagnostic.

Chapitre 3

Réseaux bayésiens et inférence

Résumé :

Les réseaux bayésiens font partie de la famille des modèles graphiques. Ils regroupent au sein d'un même formalisme la théorie des graphes et celle des probabilités afin de fournir des outils efficaces autant qu'intuitifs pour représenter une distribution de probabilités jointe sur un ensemble de variables aléatoires. Ce formalisme très puissant permet une représentation intuitive de la connaissance sur un domaine d'application donné et facilite la mise en place de modèles performants et clairs. La représentation de la connaissance se base sur la description, par des graphes, des relations de causalité existant entre des variables décrivant le domaine d'étude. A chaque variable est associée une distribution de probabilités locale quantifiant la relation causale.

Ce chapitre est une synthèse sur les réseaux bayésiens : il présente les principales définitions et les principaux théorèmes formant la base de ce domaine. Dans un premier temps, je présente le formalisme des réseaux bayésiens en abordant aussi les notions d'indépendance conditionnelle et de *d-séparation*. Ensuite, je présente l'algorithme JLO servant à l'inférence exacte dans les réseaux bayésiens, et je termine le chapitre en proposant un exemple complet d'inférence. Ce chapitre se veut donc une introduction aux réseaux bayésiens, et, sans être exhaustif, il aborde néanmoins les points nécessaires pour comprendre ce domaine.

1 Les modèles graphiques

1.1 Introduction

Les modèles graphiques portent de nombreux noms : réseaux de croyance, réseaux probabilistes, réseaux d'indépendance probabiliste ou encore réseaux bayésiens. Il s'agit d'un formalisme pour représenter de façon factorisée une distribution jointe de probabilités sur un ensemble de variables aléatoires. Ils ont révolutionné le développement des systèmes intelligents dans de nombreux domaines.

Ils sont le mariage entre la théorie des probabilités et la théorie des graphes. Ils apportent des outils naturels permettant de traiter deux grands problèmes couramment rencontrés en intelligence artificielle, en mathématiques appliquées ou en ingénierie : l'incertitude et la complexité. Ils jouent en particulier un rôle grandissant dans la conception et l'analyse d'algorithmes liés au raisonnement ou à l'apprentissage [Jordan, 1999], [Becker and Naïm, 1999], [Dawid, 1992].

A la base des modèles graphiques se trouve la notion fondamentale de la *modularité* : un système complexe est construit par la combinaison de parties simples !

La théorie des probabilités fournit le ciment permettant la combinaison de ces parties, tout en assurant que le modèle est et reste consistant. Elle fournit un moyen d'interfacer modèles et données. La théorie des graphes apporte d'une part une interface intuitive grâce à laquelle un humain peut modéliser un problème comportant des variables interagissant entre elles, d'autre part un moyen de structurer les données. Ceci mène alors vers une conception naturelle d'algorithmes génériques efficaces.

Beaucoup de systèmes probabilistes classiques issus de domaines tels que les statistiques, la théorie de l'information, la reconnaissance des formes ou encore la mécanique statistique, sont en fait des cas particuliers du formalisme plus général que constituent les modèles graphiques. Dans ce domaine, on peut citer les modèles de Markov cachés, les filtres de Kalman ou encore les modèles d'Ising [Jordan, 1999]. Ainsi, les modèles graphiques sont un moyen efficace de voir tous ces systèmes comme des instances d'un formalisme commun sous-jacent.

L'avantage immédiat réside dans le fait que les techniques développées pour certains domaines peuvent alors être aisément transférées à un autre domaine et être exploitées plus facilement. Les modèles graphiques fournissent ainsi un formalisme naturel pour la conception de nouveaux systèmes [Jordan, 1999].

J'ai voulu que ce chapitre soit une synthèse du domaine. Pour l'écrire, je me suis largement inspiré des ouvrages et articles suivants :

- Causality de Judea Pearl [Pearl, 2001],
- Probabilistic Networks and Expert Systems de R.G. Cowell, A.P. Dawid, S.L. Lauritzen et D.J. Spiegelhalter [Cowell et al., 1999],
- Les réseaux bayésiens de A. Becker et P. Naïm [Becker and Naïm, 1999],
- Probabilistic Independance Networks for Hidden Markov Probability Models de P. Smyth, D. Heckerman et M.I. Jordan [Smyth et al., 1996].

Cette synthèse a pour but de clarifier le domaine et de présenter les derniers résultats au niveau de la théorie sur les réseaux bayésiens. Elle est nécessaire en particulier pour comprendre le fonctionnement de l'inférence exacte et des algorithmes associés.

Dans un premier temps, je vais présenter les définitions et les théorèmes de base sur les réseaux bayésiens. Ensuite je parlerai de la notion d'indépendance conditionnelle, qui est à la base du processus de représentation de la connaissance dans les réseaux bayésiens. La seconde partie portera sur l'inférence et en particulier sur l'algorithme JLO dit algorithme de l'arbre de jonction. Il s'agit d'un algorithme d'inférence exacte. Je terminerai sur une conclusion et sur quelques perspectives.

1.2 Les réseaux bayésiens

1.2.1 Introduction

Dans cette section, nous allons nous focaliser sur un modèle particulier de la famille des modèles graphiques : les réseaux bayésiens, qui utilisent des graphes dirigés acycliques.

Le rôle des graphes dans les modèles probabilistes et statistiques est triple :

1. fournir un moyen efficace d'exprimer des hypothèses,
2. donner une représentation économique des fonctions de probabilité jointe,
3. faciliter l'inférence à partir d'observations.

Supposons que nous disposions d'un ensemble U de variables aléatoires discrètes toutes binaires, tel que $U = \{x_1, x_2, \ldots, x_n\}$. Pour être stockée, la probabilité jointe $P(U)$ de cet ensemble nécessitera un tableau comprenant 2^n entrées : une taille particulièrement grande, quelque soit le système utilisé. Par contre, si nous savons que certaines variables ne dépendent en fait que d'un certain nombre d'autres variables, alors nous pouvons faire une économie substantielle en mémoire et par conséquent en temps de traitement. De telles dépendances vont nous permettre de décomposer cette très large distribution en un ensemble de distributions locales beaucoup plus petites, chacune ne s'intéressant qu'à un petit nombre de variables. Les dépendances vont aussi nous permettre de relier ces petites distributions en un grand ensemble décrivant le problème que l'on veut modéliser. On pourra ainsi répondre de façon cohérente et efficace à diverses questions que l'on pourrait se poser sur cette distribution de probabilités. Dans un graphe, il est possible de représenter chaque variable du problème par un noeud et chaque dépendance entre les variables par un arc.

Les graphes dirigés et non dirigés sont abondamment utilisés pour faciliter une telle décomposition des connaissances. Les modèles à base de graphes non dirigés sont souvent appelés *champs de Markov* [Pearl, 1988] et ont servi initialement à représenter des relations temporelles (ou spatiales) symétriques [Isham, 1981, Cox and Wermuth, 1996, Lauritzen, 1996]. Les graphes dirigés acycliques sont utilisés pour représenter des relations temporelles ou causales, en particulier dans [Lauritzen, 1982], [Wermuth and Lauritzebn, 1983] et [Kiiveri et al., 1984]. Judea Pearl les a nommés *Réseaux Bayésiens* en 1985 pour mettre en évidence trois aspects :

1. la nature subjective des informations,
2. l'utilisation de la règle de Bayes comme principe de base pour la mise à jour des informations,
3. la distinction entre les modes de raisonnement causal et fondé (basé sur des évidences).

Cette dernière distinction émane directement de l'article de Thomas Bayes en 1763 [Bayes, 1763].

1.2.2 Décomposition d'une distribution de probabilités

Le schéma de base de la décomposition des graphes dirigés acycliques peut être illustré de la façon suivante. Supposons que nous ayons une distribution P définie sur un ensemble de n variables discrètes, ordonnées arbitrairement de cette façon : $X_1, X_2, \ldots, X_n$. La règle de la chaîne permet d'obtenir la décomposition suivante :

$$P(X_1, X_2, \ldots, X_n) = P(X_n|X_{n-1}, \ldots, X_2, X_1) \ldots P(X_2|X_1)P(X_1) \quad (3.1)$$

$$= P(X_1) \prod_{j=n}^{2} P(X_j|X_{j-1}, \ldots, X_1) \quad (3.2)$$

Supposons maintenant que les probabilités conditionnelles de certaines variables X_j ne soient pas dépendantes de tous les prédécesseurs de X_j (c'est-à-dire $X_1, X_2, \ldots, X_{j-1}$) mais seulement de certains d'entre eux. En d'autres termes, supposons que X_j soit indépendante de tous ses autres prédécesseurs sauf d'un certain nombre d'entre eux : ceux qui ont une influence directe sur X_j. Nous appellerons cet ensemble restreint $pa(X_j)$. Alors nous pouvons écrire :

$$P(X_j|X_1, \ldots, X_{j-1}) = P(X_j|pa(X_j))$$

et la décomposition 3.2 devient alors :

$$P(X_1, X_2, \ldots, X_n) = \prod_j P(X_j|pa(X_j))$$

Cette formule permet de simplifier énormément les informations nécessaires pour le calcul de la probabilité jointe de l'ensemble $\{X_1, \ldots, X_n\}$. Ainsi, au lieu de spécifier la probabilité de X_j conditionnellement à toutes les réalisations de ses prédécesseurs $X_1, \ldots, X_{j-1}$, seules celles qui sont conditionnées par les éléments de $pa(X_j)$ doivent être précisées. Cet ensemble est appelé *les parents Markoviens de* X_j (ou simplement les *parents*).

Définition 1.1 (Parents Markoviens) *Soit $V = \{X_1, \ldots, X_n\}$ un ensemble ordonné de variables et $P(V)$ la distribution de probabilité jointe sur ces variables. Si $pa(X_j)$ est un ensemble minimal de prédécesseurs de X_j qui rendent X_j indépendant de tous ses autres prédécesseurs, alors $pa(X_j)$ sont les* parents markoviens *de X_j. En d'autre termes, $pa(X_j)$ est tout sous-ensemble de $\{X_1, \ldots, X_{j-1}\}$ qui satisfait l'équation*

$$P(X_j|pa(X_j)) = P(X_j|X_1, \ldots, X_{j-1}) \quad (3.3)$$

et tel qu'aucun sous-ensemble propre de $pa(X_j)$ ne satisfasse l'équation 3.3.

La définition 1.1 affecte à chaque variable X_j un ensemble $pa(X_j)$ d'autres variables qui sont suffisantes pour déterminer la probabilité de X_j. La connaissance des autres variables est redondante une fois que l'on connaît les valeurs des variables de l'ensemble $pa(X_j)$. Cette affectation de variables peut être représentée par un graphe, dans lequel les noeuds sont les variables et les arcs dirigés dénotent l'influence directe qu'ont les parents sur leurs enfants. Le résultat d'une telle construction est appelé un *réseau bayésien* dans lequel un arc dirigé allant de X_i à X_j définit X_i comme étant un parent markovien de X_j, en accord avec la définition 1.1.

La définition des parents markoviens pose donc la base théorique de la notion de relation modale entre des connaissances dans un réseau bayésien. En effet, il est possible de représenter toute sorte de modalité entre les variables dans un réseau bayésien. Elles peuvent être d'ordre causal, temporel, hiérarchique, etc... En général, une seule modalité est utilisée dans un même réseau et la plupart du temps, il s'agit de la causalité. Ceci permet de représenter l'influence directe d'une variable sur une autre : si il existe un arc dirigé allant d'une variable A à une variable B, alors A est une des causes possibles de B, ou encore A a une influence causale directe sur B.

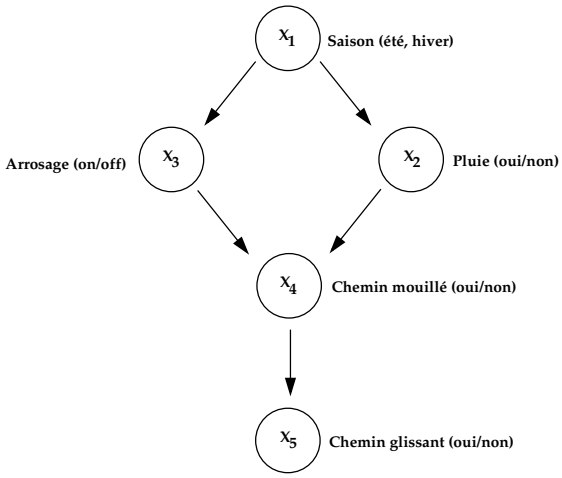


FIG. 3.1 – Un réseau bayésien représentant les dépendances entre cinq variables

La figure 3.1 représente un réseau bayésien simple contenant cinq variables. Il décrit parmi les saisons de l'année (X_1) si la pluie tombe (X_2), si un arrosage est en marche (X_3), si le chemin est mouillé (X_4) et si le chemin est glissant (X_5). Toutes les variables sont binaires. Par exemple, l'absence d'un arc allant de X_1 à X_5 signifie que la saison n'a pas une influence directe sur l'état glissant ou pas du chemin. Autrement dit, le fait que le chemin soit glissant est conditionné par le fait qu'il soit ou non mouillé, rendant inutile la connaissance sur la météo ou l'allumage (ou non) des arrosages, et *a fortiori* sur la saison. Enfin, en reprenant la définition 1.1 (parents markoviens), le graphe de la figure 3.1 se décompose de la façon suivante :

$$P(X_1, \ldots, X_5) = P(X_1).P(X_2|X_1).P(X_3|X_1).P(X_4|X_2, X_3).P(X_5|X_4)$$

Sachant un graphe G et une distribution de probabilités P, alors la décomposition de la définition 1.1 ne nécessite plus d'ordre sur les variables. Nous pouvons conclure qu'une condition nécessaire pour qu'un graphe G soit un réseau bayésien d'une distribution de probabilités P, est que P admette une décomposition sous forme d'un produit dirigé par G tel que [Pearl, 2001] :

$$P(X_1, \ldots, X_n) = \prod_i P(X_i|pa(X_i)) \tag{3.4}$$

Définition 1.2 (Compatibilité de Markov) *Si une fonction de probabilité P admet une factorisation sous la forme de l'équation 3.4 relativement à un graphe acyclique dirigé (GAD) G, on dira que G représente P et que P est compatible ou P est Markov-relatif à G.*

Assurer la compatibilité entre des graphes acycliques dirigés et des probabilités est important en modélisation statistique car la compatibilité est une condition nécessaire et suffisante pour qu'un GAD G puisse expliquer un corpus de données empiriques représentées par P, c'est-à-dire pour décrire un processus stochastique permettant de générer P [Pearl, 1988, Pearl, 2001].

Un moyen facile de caractériser un ensemble de distributions compatible avec un GAD G est de lister l'ensemble des indépendances (conditionnelles) que chacune des distributions devront satisfaire. Ces indépendances peuvent être aisément lues à partir du GAD en utilisant un critère graphique appelé la *d-séparation* (dans [Pearl, 1988], le *d* signifie *directionnel*).

1.2.3 Le critère de *d-séparation*

Considérons trois ensembles disjoints de variables X, Y et Z représentés par trois ensembles de noeuds 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 noeuds correspondants aux variables de Z *bloquent* tous les chemins allant des noeuds de X aux noeuds 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.

Définition 1.3 (d-Séparation) *Un chemin p est dit* d-séparé *(ou* bloqué*) par un ensemble Z de noeuds si et seulement si :*

1. *p contient une séquence $i \to m \to j$ ou une* divergence *$i \leftarrow m \to j$ tel que $m \in Z$, ou*
2. *p contient une* convergence *$i \to m \leftarrow j$ telle que $m \notin Z$ et tel qu'aucun descendant de m n'appartienne à Z.*

Un ensemble Z d-sépare *X de Y si et seulement si Z bloque chaque chemin partant d'un noeud quelconque de X à un noeud quelconque de Y.*

L'idée à la base de la d-séparation est simple quand on attribue une signification aux flèches dans le graphe. Dans la séquence $i \to m \to j$ ou dans la divergence $i \leftarrow m \to j$, 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 influer 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 $i \to m \leftarrow j$, 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.

Dans la figure 3.1, 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 3.1, $X = \{X_2\}$ 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 chemin $X_2 \leftarrow X_1 \rightarrow X_3$ est bloqué car il s'agit d'une convergence 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.

Ce schéma de raisonnement est plus connu sous le nom du *paradoxe de Berkson* en statistique [Berkson, 1946].

1.2.4 Quelques propriétés de la *d-séparation*

La connexion entre la *d-séparation* et l'indépendance conditionnelle est établie grâce au théorème suivant que l'on doit à Verma et Pearl dans [Verma and Pearl, 1988] et dans [Geiger et al., 1988] :

Théorème 1.1 (Implications probabilistes de la d-séparation) *Si les ensembles X et Y sont d-séparés par Z dans un GAD G, alors X est indépendant de Y conditionnellement à Z dans chaque distribution compatible (déf. 1.2) avec G. Réciproquement, si X et Y ne sont pas d-séparés par Z dans un GAD G, alors X et Y sont dépendants conditionnellement à Z dans au moins une distribution compatible avec G.*

On notera à présent par $(X \perp\!\!\!\perp Y|Z)_P$ la notion d'indépendance conditionnelle et par $(X \perp\!\!\!\perp Y|Z)_G$ la notion graphique de *d-séparation*. Le théorème peut être réécrit de la façon suivante [Pearl, 2001] :

Théorème 1.2 *Pour tous les ensembles disjoints de noeuds (X, Y, Z) dans un GAD G et pour toute fonction de probabilités P on a :*

1. $(X \perp\!\!\!\perp Y|Z)_G \Longrightarrow (X \perp\!\!\!\perp Y|Z)_P$ *toutes les fois que G et P sont compatibles ; et*
2. *si $(X \perp\!\!\!\perp Y|Z)_P$ est vérifié dans toute distribution compatible avec G, alors $(X \perp\!\!\!\perp Y|Z)_G$.*

Enfin, un autre test de *d-séparation* a été donné dans [Lauritzen et al., 1990]. Il est basé sur la notion de graphes ancestraux. Pour tester si $(X \perp\!\!\!\perp Y|Z)_G$, on efface tous les noeuds de G sauf ceux qui sont dans $\{X, Y, Z\}$ et leurs ancêtres, puis on connecte par un arc chaque paire de noeuds qui a un enfant commun (*mariage* des noeuds) et on transforme le graphe dirigé en graphe non-dirigé. Alors $(X \perp\!\!\!\perp Y|Z)_G$ est vérifié si et seulement si Z intercepte tout chemin allant d'un noeud de X à un noeud de Y dans le graphe non-dirigé résultant. Ici seule la topologie du graphe compte, et non plus l'ordre dans lequel le graphe G avait été construit initialement. Cette approche sera importante lors de la présentation de l'algorithme d'inférence dans les réseaux bayésiens.

2 Modélisation et inférence dans les réseaux bayésiens

2.1 Introduction

Un réseau bayésien permet donc de représenter un ensemble de variables aléatoires pour lesquelles on connaît un certain nombre de relations de dépendances. Appelons U 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 *a posteriori* du réseau sachant ε.

2.2 Spécification d'un réseau

2.2.1 Exemple

Pour illustrer notre propos, nous utiliserons un exemple issu de [Cowell et al., 1999]. 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 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.

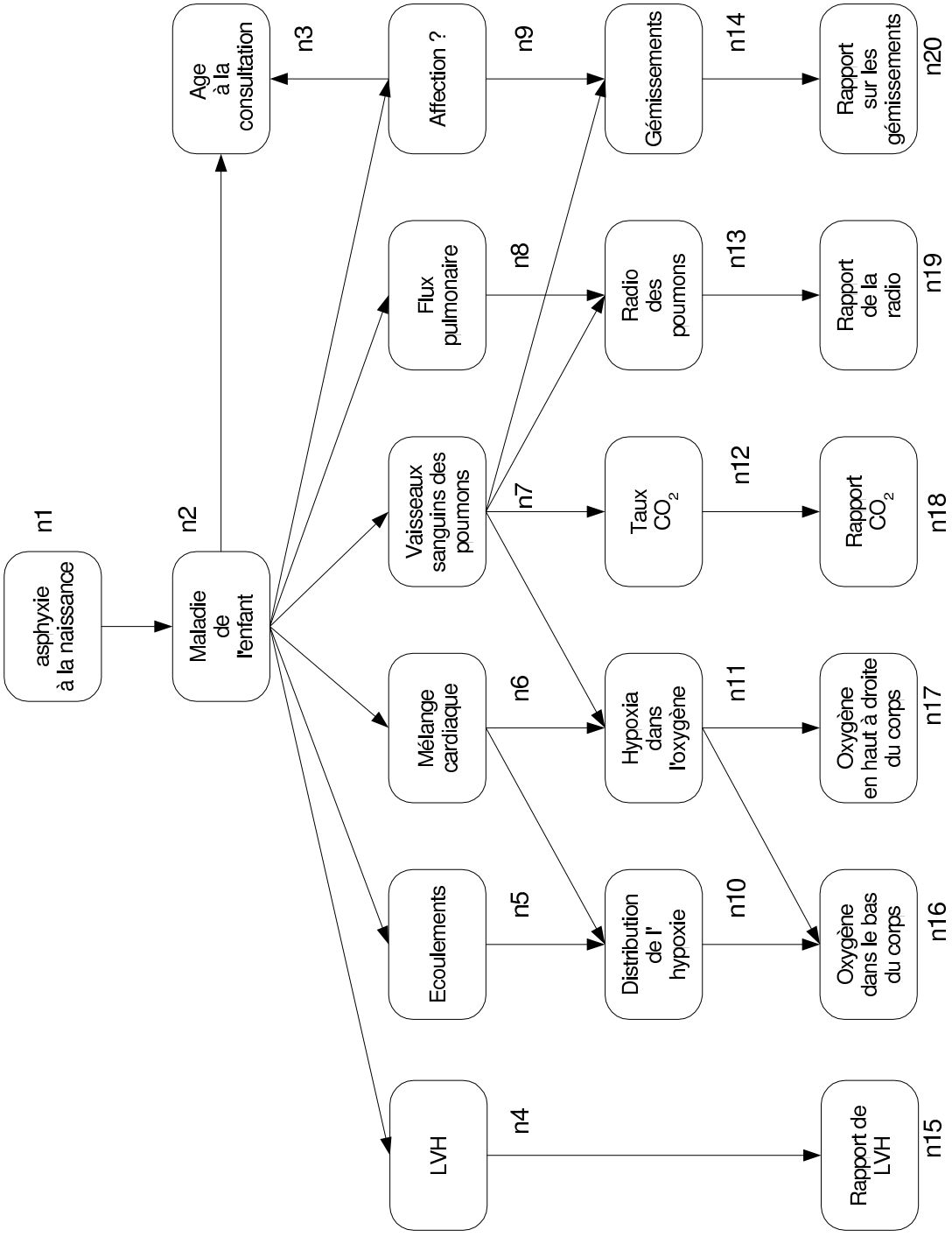


FIG. 3.2 – Représentation du problème d'asphyxie du nouveau-né

2.2.2 Étape qualitative

Ainsi que J. Pearl le montre dans [Pearl, 1988], 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.

La figure 3.2 [Cowell et al., 1999] 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.

Ce graphe illustre donc la première étape consistant à modéliser qualitativement le problème et à déterminer les influences existant entre les variables.

2.2.3 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 1.2, la décomposition de la distribution de probabilité jointe peut se faire comme dans l'équation 3.4 :

$$P(n_1, \ldots, n_{20}) = \prod_{i=1}^{20} P(n_i | pa(n_i))$$

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.

2.2.4 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 apprise 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 12Go !

2.3 Les principaux algorithmes

2.3.1 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 $P(Y|X)$ 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 [Pearl, 1982] et dans [Kim and Pearl, 1983] : 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 JLO qui sera l'objet de notre étude. Cette méthode est aussi appelée *algorithme de l'arbre de jonction* et a été développée dans [Lauritzen, 1988] et [Jensen et al., 1990].

Une autre méthode, développée dans [Pearl, 1988] et dans [Jensen, 1996], s'appelle le *cut-set conditionning* : 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 instanciation est choisie. On réitère ce processus jusqu'à ce que toutes les instanciations possibles aient été utilisées. On fait alors la moyenne des résultats. Dans la figure 3.1, si on instancie X_1 a une valeur spécifique ($X_1 = hiver$, par exemple), alors le chemin entre X_2 et X_3 et passant par X_1 est *bloqué*, et le réseau devient un arbre (cf. figure 3.3) . Le principal avantage de

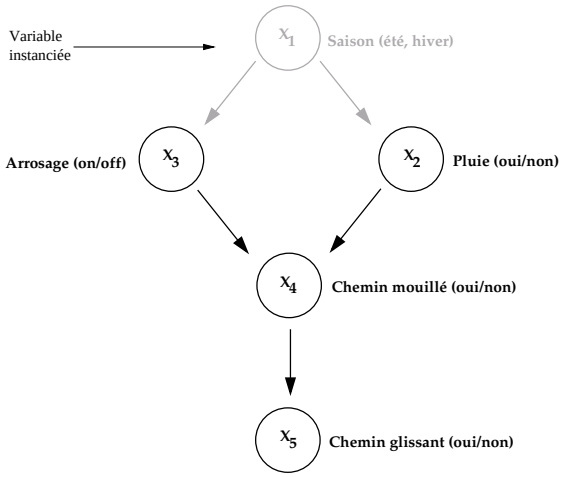


FIG. 3.3 – Graphe de la figure 3.1 transformé en arbre en instanciant X_1.

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. Des méthodes hybrides ont été proposées pour tenter de concilier complexité en temps et en espace, dans [Shachter et al., 1994] et dans [Dechter, 1996b].

Bien que l'inférence dans des réseaux quelconques soit *NP-difficile* [Cooper, 1990], 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 une méthode d'approximation [Pearl, 1988]. Ces méthodes exploitent la topologie du réseau et effectuent un échantillonage de Gibbs sur des sous-ensembles locaux de variables de façon séquentielle et concurrente [Jaakkola and Jordan, 1999], [Jordan et al., 1999].

2.3.2 Approche générale de l'inférence

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

$$P(Y|X) = \frac{P(Y,X)}{P(X)} = \frac{\sum_{h \not\subset Y \cup X} P(X_H = h, Y, X)}{\sum_{h \not\subset Y} P(X_H = h, X)}$$

Étant donné que tout réseau bayésien défini aussi une probabilité jointe sur un ensemble de variables aléatoires, il est clair que $P(Y|X)$ peut être calculée à partir d'un GAD G. Le problème de l'inférence se réduit donc à un problème de marginalisation[1] d'une distribution de probabilités jointe. Cependant, le problème ne réside pas tant au niveau du calcul, mais plutôt de son efficacité. En effet, si les variables du GAD G sont binaires, le calcul de $\sum_h P(X_1, \ldots, X_N)$ prendra un temps de $O(2^N)$.

Considérons le réseau bayésien simple de la figure 3.4. Sa probabilité jointe peut être écrite sous la forme :

$$P(A,B,C,D,F,G) = P(A) \cdot P(B|A) \cdot P(C|A) \cdot P(D|B,A) \cdot P(F|B,C) \cdot P(G|F)$$

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

$$\sum_{B,C,D,F,G} P(A,B,C,D,F,G) =$$

$$\sum_B P(B|A) \cdot \sum_C P(C|A) \cdot \sum_D P(D|B,A) \cdot \sum_F P(F|B,C) \cdot \sum_G P(G|F)$$

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 :

$$\sum_{B,C,D,F,G} P(A,B,C,D,F,G) =$$

$$\sum_B P(B|A) \cdot \sum_C P(C|A) \cdot \sum_D P(D|B,A) \cdot \sum_F P(F|B,C) \cdot \lambda_{G \to F}(F)$$

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

$$\sum_{B,C,D,F,G} P(A,B,C,D,F,G) =$$

$$\sum_B P(B|A) \cdot \sum_C P(C|A) \cdot \lambda_{F \to C}(B,C) \cdot \sum_D P(D|B,A)$$

[1] Si $P(A,B,C,D)$ 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 $P(A,B,C)$.

où $\lambda_{F\to C}(B,C) = \sum_F P(F|B,C)\cdot\lambda_{G\to F}(F)$. En général, on essaie de calculer les termes les plus à gauche possible, de manière à minimiser le nombre de calculs nécessaires. La notation $\lambda_{F\to C}(B,C)$ 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.

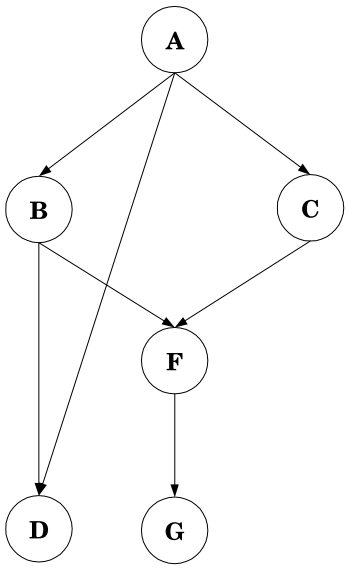


FIG. 3.4 – Un réseau bayésien simple [Kask et al., 2001]

2.4 Algorithme de l'arbre de jonction dit JLO

2.4.1 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 [Lauritzen, 1988],[Jensen et al., 1990]. Des extensions pour des distributions gaussiennes et des mixtures de gaussiennes ont été proposées dans [Lauritzen and Wermuth, 1989] et dans [Cowell et al., 1999]. Un algorithme similaire à été développé par Dawid dans [Dawid, 1992]. Il résout le problème de l'identification du maximum *a posteriori* (MAP) avec une complexité en temps équivalente à celle de l'algorithme JLO. Cet algorithme sera présenté plus tard. De plus, il y a de nombreuses variantes de ces deux algorithmes, mais on peut montrer [Shachter et al., 1994] que tous les algorithmes d'inférence exacte sur les réseaux bayésiens sont équivalents ou peuvent être dérivés de l'algorithme JLO ou de l'algorithme de Dawid. Ainsi, ces algorithmes subsument les autres algorithmes d'inférence exacte dans les modèles graphiques.

L'algorithme se comporte de la façon suivante :

– 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 ;
– 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 ce 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 à la distribution de probabilité sachant les nouvelles informations, c'est-à-dire $P(U|\varepsilon)$ où $\underline{U}$ représente l'ensemble des variables du réseau bayésien et ε l'ensemble des nouvelles informations sur lesdites variables. ε peut, par exemple, être vu comme un ensemble d'observations faîtes à partir de capteurs.

Le déroulement de cet algorithme sera illustré sur l'exemple de la figure 3.2.

2.4.2 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 3.5 montre l'exemple de la section 2.2. Les arcs en pointillés représentent 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. De plus, [Cowell et al., 1999] montre que le graphe moral G^m satisfait aux mêmes propriétés que 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^m. [Cowell et al., 1999] présente une justification de l'équivalence de la distribution de probabilités jointe issue du graphe initial G et de la distribution issue du graphe moral G^m.

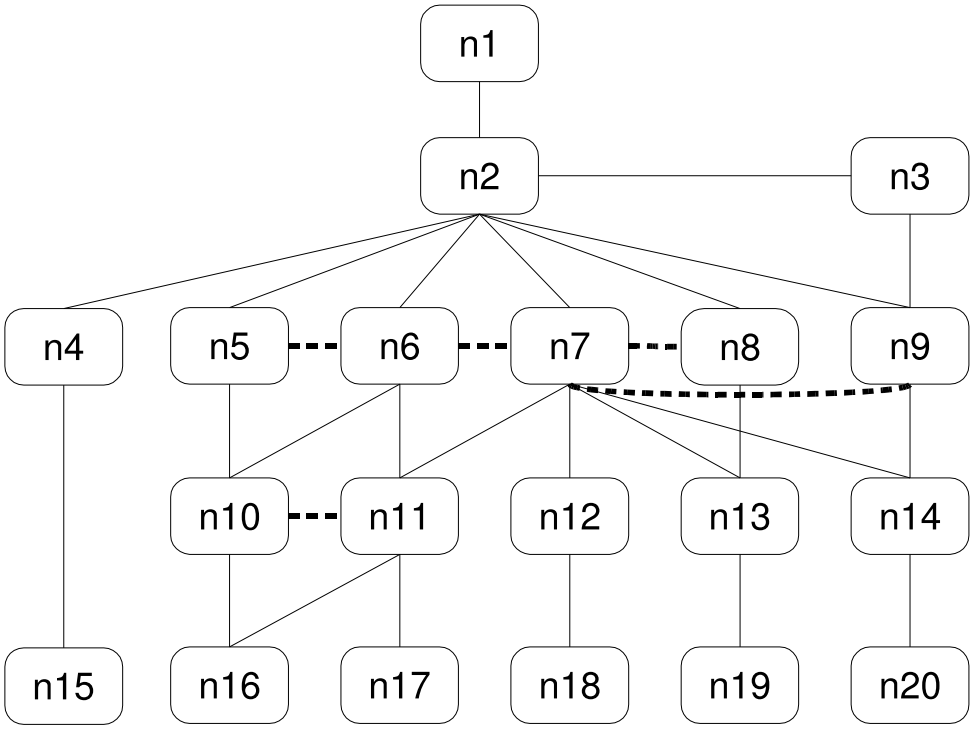


FIG. 3.5 – Graphe moralisé. Les arcs en pointillés ont été rajoutés au cours de la moralisation

2.4.3 Triangulation

La deuxième étape consiste à trianguler le graphe moral G^m et à en extraire des cliques 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 éliminé si tous ses voisins sont connectés deux à deux. Donc un noeud peut être éliminé si 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 [Yannakakis, 1981]. Dans [Kjaerulff, 1992], Kjaerulff donne un aperçu de plusieurs algorithmes de triangulation pour des graphes acyliques dirigés. Pour des problèmes où les cliques de grande

taille sont inévitables, la méthode présentée par Kjaerulff (*simulated annealing*) donne de bons résultats, bien qu'elle nécessite un temps de calcul assez long. Cependant, pour l'algorithme JLO, le calcul de l'arbre de jonction n'est nécessaire qu'une seule fois. Il s'agit là d'un compromis acceptable. Kjaerulff a aussi présenté un algorithme utilisant un *critère d'optimalité* $c(v)$ en fonction d'un noeud v. Par exemple, on peut vouloir maximiser (ou minimiser) une fonction d'utilité ou de coût associée à la sélection d'un noeud du graphe non-dirigé. [Olmsted, 1983] et [Kong, 1986] donnent l'algorithme suivant sur un graphe G ayant k noeuds :

Algorithme 2.1 (Triangulation avec critère d'optimalité) – *Aucun noeud n'est numéroté,* $i = k$.
– *Tant qu'il y a des noeuds non-numérotés faire*
  – *Sélectionner un noeud* v *non-numéroté optimisant le critère* $c(v)$.
  – *Donner à* v *le numéro* i.
  – *Former l'ensemble* C_i *avec le noeud sélectionné et ses voisins non-numérotés.*
  – *Connecter deux à deux tous les noeuds de* C_i *s'ils ne sont pas encore connectés.*
  – *Éliminer le noeud* v *sélectionné et décrémenter* i *de 1.*

Bien sûr, cet algorithme dépend de la qualité du critère d'optimalité $c(v)$ pour sélectionner les noeuds. Ce critère peut tenter de minimiser la taille de l'espace d'états joints de la clique C_i. Ce critère donne en général de bons résultats. On peut aussi tenter de minimiser le nombre d'arcs à ajouter dans une clique C_i si le noeud était sélectionné. L'idée est toujours de privilégier la taille des cliques la plus petite possible afin d'optimiser au mieux le temps de calcul nécessaire aux traitements des tables de probabilités conditionnelles.

La figure 3.6 représente le graphe triangulé de l'exemple de la section 2.2. Dans cet exemple, il a été nécessaire de rajouter deux arcs supplémentaires entre n_5 et n_7 et entre n_5 et n_{11}. Les nombres situés à droite de chaque noeud représentent l'ordre d'élimination des variables au cours de la triangulation. Cet ordre d'élimination va aussi nous permettre d'extraire les cliques. Nous obtenons les cliques suivantes, au cours de la triangulation du graphe G^m :

n°	Contenu
1	n_{12}, n_{18}
2	n_7, n_{12}
3	n_2, n_5, n_6, n_7
4	n_5, n_6, n_7, n_{11}
5	n_5, n_6, n_{10}, n_{11}
6	n_2, n_7, n_9
7	n_2, n_3, n_9
8	n_2, n_7, n_8
9	n_7, n_9, n_{14}
10	n_7, n_8, n_{13}
11	n_{10}, n_{11}, n_{16}
12	n_1, n_2
13	n_2, n_4

14	n_4, n_{15}
15	n_{11}, n_{17}
16	n_{14}, n_{20}
17	n_{13}, n_{19}

Il apparaît clairement que pour toute clique C_i, il existe une clique C_j telle que $C_i \cap C_j \neq \emptyset$. Cette propriété est intéressante et permettra la construction de l'arbre de jonction.

2.4.4 Arbre de jonction

La construction de l'arbre de jonction est la dernière partie avant de procéder à l'inférence proprement dite. Nous rappelons que pour un réseau bayésien donné, l'arbre de jonction est construit une et une seule fois. Les calculs probabilistes auront lieu dans l'arbre de jonction autant de fois que nécessaire. Cependant, pour un réseau bayésien donné, il existe plusieurs arbres de jonction possibles : il sont fonction de l'algorithme de triangulation et de l'algorithme de construction utilisé.

Nous commençons par deux définitions importantes : la décomposition et le graphe décomposable.

Définition 2.1 (Décomposition) *Un triplet (A, B, C) de sous-ensembles disjoints d'un ensemble de noeuds V d'un graphe non-dirigé G forme une* décomposition *de G (ou* décompose *G), si $V = A \cup B \cup C$ et si les conditions suivantes sont satisfaites :*

- *C sépare A de B,*
- *C est un sous-ensemble complet de V.*

A, B ou C peuvent être vides, mais si A et B ne sont pas vides, alors ont dira que l'on a une *décomposition propre* de G.

Définition 2.2 (Graphe décomposable) *Un graphe non-dirigé G est* décomposable *si :*

- *soit il est complet,*
- *soit il possède une décomposition propre (A, B, C) telle que les deux sous-graphes $G_{A \cup C}$ et $G_{B \cup C}$ soit décomposables.*

Ces deux définitions seront utilisées par la suite pour prouver l'existence d'un arbre de jonction. Voici d'abord la définition d'un tel arbre :

Définition 2.3 (Arbre de jonction) *Soit $\mathcal{C}$ une collection de sous-ensembles d'un ensemble fini V de noeuds et soit $\mathcal{T}$ un arbre avec $\mathcal{C}$ comme ensemble de ses noeuds, alors $\mathcal{T}$ est un* arbre de jonction *si toute intersection $C_1 \cap C_2$ d'une paire (C_1, C_2) d'ensembles dans $\mathcal{C}$ est contenue dans chaque noeud sur le chemin unique allant de C_1 à C_2 dans $\mathcal{T}$.*

Si G est un graphe non-dirigé, $\mathcal{C}$ est l'ensemble de ses cliques et $\mathcal{T}$ est un arbre de jonction avec $\mathcal{C}$ son ensemble de noeuds, alors $\mathcal{T}$ est un arbre de jonction (de cliques) pour le graphe G. On a alors le théorème suivant [Cowell et al., 1999] :

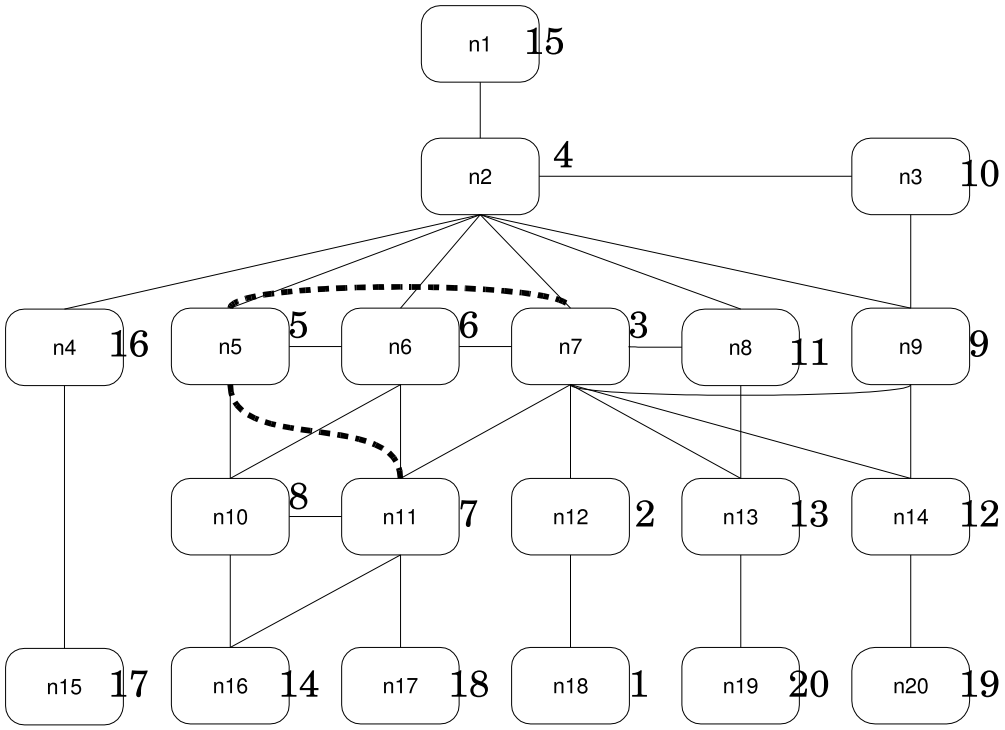


FIG. 3.6 – Graphe triangulé. Les nombres à droite des noeuds représentent l'ordre d'élimination des noeuds. Les lignes en pointillés sont les arcs qu'il a été nécessaire de rajouter

Théorème 2.1 *Il existe un arbre de jonction $\mathcal{T}$ de cliques pour le graphe G si et seulement si G est décomposable.*

De plus, l'intersection $S = C_1 \cap C_2$ entre deux voisins dans l'arbre de jonction est aussi un séparateur dans le graphe non-dirigé G des ensembles de noeuds C_1 et C_2 (en fait, il s'agit même d'un séparateur minimal). On appelle S le *séparateur* des noeuds C_1 et C_2 dans l'arbre de jonction. On notera $\mathcal{S}$, l'ensemble des séparateurs. Quand un graphe G admet plusieurs arbres de jonction, on peut alors montrer que $\mathcal{S}$ reste le même, quel que soit l'arbre de jonction. La définition qui suit est utile pour la construction de l'arbre de jonction : il s'agit de la propriété de *running-intersection*.

Définition 2.4 (Propriété de running-intersection) *Une séquence $(C_1, C_2, \ldots, C_k)$ d'ensembles de noeuds a la propriété de* running-intersection *si pour tout $1 < j \leq k$, il existe un $i < j$ tel que $C_j \cap (C_1 \cup \cdots \cup C_{j-1}) \subseteq C_i$.*

Il existe un classement très simple des cliques d'un graphe décomposable : ce classement permet de construire un arbre de jonction possédant la propriété de *running-intersection*. Et inversement, si les cliques ont été ordonnées pour satisfaire cette propriété, alors on peut construire un arbre de jonction avec l'algorithme suivant :

Algorithme 2.2 (Construction de l'arbre de jonction) *Soit un ensemble $(C_1, \ldots, C_p)$ de cliques ordonnées de manière à avoir la propriété de* running-intersection.

- *Associer un noeud de l'arbre de jonction à chaque clique C_i.*
- *Pour $i = 2, \ldots, p$*
  - *ajouter un arc entre C_i et C_j où j est une valeur prise dans l'ensemble $\{1, \ldots, i-1\}$ et tel que :*

$$C_i \cap (C_1 \cup \cdots \cup C_{i-1}) \subseteq C_j$$

L'application de cet algorithme à l'exemple de la section 2.2 nous permet d'obtenir l'arbre de jonction de la figure 3.7.

On notera dans cet exemple, que la variable n_2 est contenue dans les cliques C_3, C_6, C_7, C_{12} et C_{13}, avec un particulier $C_3 \cap C_{13} = \{n_2\}$. L'ensemble $\{C_3, C_6, C_7, C_{12}, C_{13}\}$ forme un sous-arbre connecté dans l'arbre de jonction.

Il est aussi possible de construire l'arbre simplement en utilisant l'algorithme de Kruskal [Cormen et al., 1990]. L'algorithme est sensiblement équivalent au précédent. On définit un poids w pour chaque lien entre deux cliques C_i et C_j et tel que

$$w = |C_i \cap C_j|$$

pour tous les couples (i, j). Alors un arbre de cliques satisfera la propriété de *running-intersection* si et seulement si c'est un arbre couvrant de poids maximal [Smyth et al., 1996]. L'algorithme construit un arbre de jonction en choisissant successivement chaque couple de cliques dont le poids est maximal sauf si ce lien crée un cycle.

Ceci termine la construction de l'arbre de jonction. Il est à noter que la complexité dans le pire des cas de l'heuristique de la triangulation est de l'ordre de $O(N^3)$ et que la création de l'arbre (c'est-à-dire le calcul de l'arbre couvrant) est de l'ordre de $O(N^2 \log N)$.

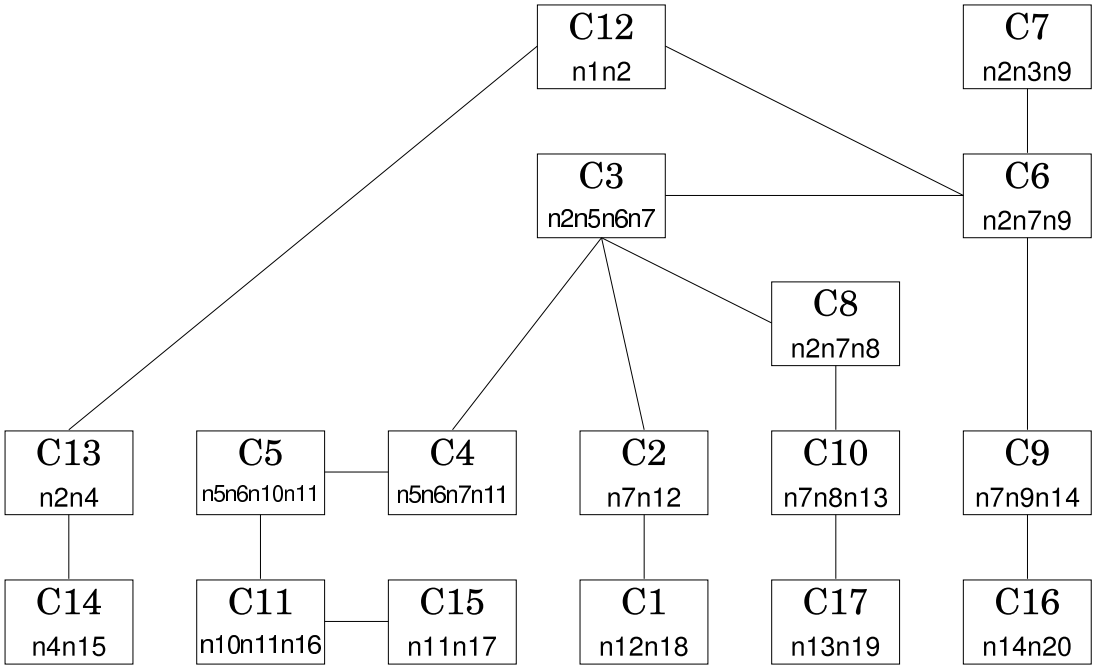


FIG. 3.7 – Arbre de jonction. Les C_i représentent les numéros des cliques, les $n_i \dots n_j$ sont les noeuds contenus dans chaque clique.

2.4.5 Initialisation de l'arbre de jonction

A partir de cet instant, on considère que l'arbre de jonction est correctement construit. Cette première étape ne doit être faîte qu'une seule fois. Les étapes qui vont maintenant suivre concernent le calcul sur les probabilités dans les réseaux bayésiens. Elles doivent être répétées autant de fois que nécessaire, c'est-à-dire chaque fois que les spécifications numériques changent ou qu'une nouvelle observation est disponible.

Cette étape consiste à utiliser les spécifications numériques du graphe initial G afin de calculer une spécification numérique équivalente pour l'arbre de jonction.

La distribution jointe de probabilité pour un graphe non-dirigé G^u peut être exprimée sous forme d'une simple factorisation :

$$P(X_1, \dots, X_n) = \prod_{C \in \mathcal{C}} a_C(x_C) \tag{3.5}$$

où $\mathcal{C}$ est l'ensemble des cliques de G^u, x_C est une affectation de valeurs aux variables de la clique C et les fonctions a_C sont des fonctions non-négatives, prenant leurs valeurs dans l'ensemble des affectations possibles des valeurs des variables de la clique C et rendent une valeur dans l'intervalle $[0, \infty[$. L'ensemble des fonctions de cliques associées à un graphe non-dirigé G^u représentent la spécification numérique de ce graphe [Smyth et al., 1996]. Dans la littérature sur les champs de Markov, une telle fonction $a_c(x_C)$ est souvent appelée une fonction de potentiel.

Nous savons d'ores et déjà que les cliques du graphe triangulé ont été arrangées sous forme d'un arbre de jonction. En considérant à présent l'ensemble des séparateurs associés à chaque couple de cliques adjacentes dans l'arbre de jonction, on donne à chaque séparateur une fonction de potentiel b_S définie de façon équivalente aux fonctions de potentiel a_C. Comme un séparateur est égal à l'intersection de deux cliques adjacentes, la distribution de probabilités jointe du réseau bayésien initial peut se factoriser de la façon suivante :

$$P(X_1 \dots X_n) = \frac{\prod_{C \in \mathcal{C}} P(x_C)}{\prod_{S \in \mathcal{S}} P(x_S)} \tag{3.6}$$

où $P(x_C)$ et $P(x_S)$ sont les distributions de probabilités marginales jointes des variables de la clique C (respectivement du séparateur S). Ce résultat est très important et permet de justifier le calcul d'inférence dans les réseaux bayésiens avec l'algorithme de l'arbre de jonction.

A partir de cette nouvelle formulation de la distribution de probabilité d'un réseau bayésien, l'initialisation se fait très simplement. Comme l'arbre de jonction vérifie la propriété de *running-intersection,* on sait alors que chaque variable X_i se trouve dans au moins une clique. On affecte alors, de façon unique, chaque X_i à une et une seule clique de l'arbre. Certaines cliques risquent de n'avoir aucune variable affectée. Après avoir affecté l'ensemble des variables, chacune à une clique particulière, on définit les fonctions de potentiel de la façon suivante :

$$a_C(x_C) = \begin{cases} \prod_i P(X_i|pa(X_i)) \; si \; X_i \; est \; mis \; dans \; C \\ 1 \; si \; aucune \; variable \; n'est \; dans \; C \end{cases}$$

Les fonctions de potentiel b_S des séparateurs sont initialisées à 1.

A cet instant, l'arbre de jonction est initialisé et consistant avec le réseau bayésien initial. On peut donc propager une observation et calculer la probabilité *a posteriori* de chaque variable du réseau sachant des observations.

2.4.6 Propagation par passages de messages locaux

Le principe de l'algorithme est de passer l'information nouvelle d'une clique à ses voisines dans l'arbre de jonction, et de mettre à jour les voisines et les séparateurs avec cette information locale. Le point important dans l'algorithme JLO est que la représentation $P(X_1 \dots X_n)$ de l'équation 3.6 reste vraie après chaque passage de messages d'une clique à une autre. Une fois que tous les messages locaux ont été transmis, la propagation convergera vers une représentation marginale de la distribution de probabilités sachant le modèle initial (c'est-à-dire les paramètres du réseau bayésien qui nous ont permis d'initialiser l'arbre de jonction) et sachant les évidences observées.

Quelques définitions

On appelle table de probabilités conditionnelles (CPT : conditional probability table), le tableau contenant l'ensemble des probabilités d'un ensemble de variables aléatoires discrètes pour chacune des valeurs de ces variables. Un tel tableau forme un *potentiel discret* ou simplement un *potentiel*. De plus, la table de probabilités conditionnelles représentant $P(x|pa(x))$ forme aussi un *potentiel* avec la propriété additionnelle de sommer à 1 pour chaque configuration des parents.

Une *évidence* correspond à une information nous donnant avec une certitude absolue la valeur d'une variable. Dans le cas normal, un noeud X binaire d'un réseau bayésien contiendra l'information *« je pense que $X = x_1$ avec une certitude de $P(X = x_1)$ et je pense que $X = x_2$ avec une certitude de $P(X = x_2)$. »* Dans le cas d'une évidence, le noeud contiendra *« je sais que $X \neq x_2$. »* Une deuxième forme d'évidence est appelée évidence vraisemblable ou simplement vraisemblance (*likelihood findings*), et permet d'apporter une observation avec simplement une distribution de vraisemblance sur l'ensemble des états possibles de l'observation, relatant l'incertitude que l'on a sur l'observation. La somme des valeurs doit être égale à 1.

Flux d'information entre les cliques

L'équation 3.6 nous permet de spécifier numériquement les paramètres de l'arbre de jonction :

$$P(U) = \frac{\prod_{C \in \mathcal{C}} a_C(x_C)}{\prod_{S \in \mathcal{S}} b_S(x_S)}$$

où a_C et b_S sont des fonctions de potentiel non-négatives.

Le passage de messages procède de la façon suivante. On définit le *flux* d'une clique C_i à une clique adjacente C_j de la manière suivante. Soit S_k le séparateur de ces deux cliques, alors

$$b^*_{S_k}(x_{S_k}) = \sum_{C_i \setminus S_k} a_{C_i}(x_{C_i}) \tag{3.7}$$

est la marginalisation sur l'ensemble des états des variables qui sont dans la clique C_i mais pas dans le séparateur S_k. La clique C_j est mise à jour avec le potentiel suivant :

$$a^*_{C_j}(x_{C_j}) = a_{C_j}(x_{C_j})\lambda_{S_k}(x_{S_k})$$

où

$$\lambda_{S_k}(x_{S_k}) = \frac{b^*_{S_k}(x_{S_k})}{b_{S_k}(x_{S_k})}$$

Le terme $\lambda_{S_k}(x_{S_k})$ est appelé le *facteur de mise à jour*. L'idée du message est de transférer la nouvelle information que C_i a reçu (l'évidence) et que C_j ne connaissait pas encore. Cette nouvelle information est alors *résumée* dans le séparateur S_k qui est le seul point commun entre C_i et C_j. Un flux correspond à l'envoi de messages depuis C_i vers tous ses voisins dans l'arbre de jonction. Ce flux introduit alors une nouvelle représentation de $P(U)$ probabiliste de l'arbre de jonction telle que :

$$K^* = (\{a^*_C : C \in \mathcal{C}\}, \{b^*_S : s \in \mathcal{S}\})$$

Enfin, pour compléter l'algorithme, il faut un ordonnancement du passage des messages. Cet ordonnancement est dicté par l'arbre de jonction. Un ordonnancement est l'ordre dans lequel les messages sont passés d'une clique à une autre de telle manière que toutes les cliques reçoivent une information de chacune de leur voisine.

L'ordonnancement le plus direct opère en deux temps. On choisit une clique comme racine de l'arbre de jonction. Tout noeud de l'arbre peut être choisi comme racine. Ensuite la première phase dite de *collection* consiste à passer les messages depuis les feuilles de l'arbre jusqu'à la racine. Si un noeud doit recevoir plusieurs messages, alors les messages sont envoyés séquentiellement.

L'ordre n'est pas important. Une fois la phase de collection complétée, commence la phase de *distribution* qui consiste à opérer de manière inverse : les messages sont transmis depuis la racine jusqu'aux feuilles. Il y a au plus deux messages parcourant chaque arc de l'arbre : celui du fils et celui du père. On notera en plus que le flux des messages dans l'arbre de jonction n'a aucun lien avec les arcs du réseau bayésien initial et ne reflètent pas la structure du réseau.

A la fin, le réseau a atteint un *état d'équilibre*, ce qui signifie que si aucune information nouvelle n'est introduite dans le réseau, alors un nouveau passage de messages dans l'arbre de jonction ne modifiera pas les fonctions de potentiel. Ceci est cohérent avec le fait que le passage d'un message correspond à la transmission d'une nouveauté (en terme d'information probabiliste) d'une clique à une autre. Une fois que toutes les nouvelles informations ont été propagées à l'ensemble du réseau, l'ensemble des noeuds a eu connaissance de l'ensemble des nouveautés. L'état d'équilibre est atteint.

Si ε est l'ensemble des évidences introduites dans le réseau, et si $P(U)$ est la distribution de probabilité *a priori* du réseau, alors l'algorithme JLO, tel qu'il a été présenté, donne après propagation, $P(U|\varepsilon)$.

Entrer une évidence dans le réseau

Classiquement, un réseau bayésien est utilisé de façon dynamique. A chaque fois qu'une nouvelle information est obtenue, elle est insérée dans le réseau et on la propage à l'ensemble du réseau. Avant une phase de propagation, JLO permet d'insérer autant d'évidences qu'il y a de variables.

D'un point de vue plus formel, une *évidence* est une fonction $\varepsilon : \chi \rightarrow \{0, 1\}$. Cette évidence représente le fait que certains états χ de la variable aléatoire sont *impossibles*. Si tous les états sauf un sont déclarés impossibles, alors après propagation de l'évidence, la variable ayant reçu l'évidence aura une probabilité $P(X_\varepsilon = x_i) = 1$ où x_i représente le seul état possible. Une évidence vraisemblable est une fonction $\varepsilon : \chi \rightarrow [0, 1]$ et telle que la somme des affectations à chaque état de la variable soit égale à 1.

Pour insérer une évidence dans un réseau bayésien, on procède de la façon suivante : pour chaque clique contenant la variable v recevant l'évidence, la fonction de potentiel ϕ_C de la clique (autrement dit sa table de probabilités) est multipliée par l'évidence. En pratique, s'il s'agit d'une évidence simple (au contraire d'une évidence semblable), alors on met à 0 les entrées de la table correspondant aux valeurs impossibles décrétées par l'évidence. Le fonction de potentiel modifiée correspond alors à la représentation de $P^\varepsilon(x) \equiv P(x\&\varepsilon) \equiv P(x)\varepsilon(x) \propto P(x|\varepsilon)$.

Après avoir entré une (ou plusieurs) évidence(s) dans le réseau, on procède à une propagation complète (collection et distribution) de manière à ce que le réseau atteigne un état d'équilibre. Ensuite l'ensemble des tables de probabilités des cliques de l'arbre de jonction sont normalisées. On obtient la formule suivant pour la probabilité jointe *a posteriori* :

$$P(x\&\varepsilon) = \frac{\prod_{C \in \mathcal{C}} P(x_C\&\varepsilon)}{\prod_{S \in \mathcal{S}} P(x_S\&\varepsilon)}$$

et en normalisant on obtient finalement :

$$P(x|\varepsilon) = \frac{\prod_{C \in \mathcal{C}} P(x_C|\varepsilon)}{\prod_{S \in \mathcal{S}} P(x_S|\varepsilon)}$$

A ce moment, obtenir la probabilité *a posteriori* sachant l'évidence d'une variable quelconque, revient à prendre une clique contenant ladite variable et à marginaliser la table de probabilités afin d'obtenir la table de probabilité de la variable seule. Toute clique contenant la variable d'intérêt est un candidat adéquat.

2.4.7 Un exemple de propagation

Pour illustrer l'algorithme de propagation, nous allons réutiliser l'exemple de la figure 3.2. Considérons deux cliques adjacentes $C_{13} = \{n_2, n_4\}$ et $C_{14} = \{n_4, n_{15}\}$ et leur séparateur $S = \{n_4\}$.

L'arbre de jonction (dont on ne représente qu'une petite partie ici) est initialisé avec $P(n_{15}|n_4)$ pour C_{14} et $P(n_4|n_2)$ pour C_{13} (voir figure 3.8(a)). La variable n_2 servira, quant à elle, à initialiser la clique C_{12} avec $P(n_2|n_1)$. Les séparateurs sont initialisés à 1.

Ensuite on incorpore l'évidence $rapport\,LVH\ =\ oui$, et la clique C_{14} contient alors des 0 correspondant aux états impossibles. Après passage du message, le séparateur et la clique C_{13} sont mis à jour, ce qui donne les potentiels de la figure 3.8(b). Deux étapes ont été nécessaires :

– on marginalise C_{14} sur toutes les variables non contenues dans S (en fait toutes sauf n_{15}) pour obtenir b^*_S, le nouveau potentiel du séparateur (3.8(b)) ;
– on calcule le ratio λ qui sert à modifier le potentiel de la clique C_{13}, en multipliant chaque terme de ce potentiel avec λ.

Dans la figure 3.8(c), on retrouve C_{13} modifié par les messages en provenance du reste de l'arbre de jonction (phase de distribution). Dans cette figure, le séparateur et C_{14} n'ont pas encore reçu le message de C_{13}. On calcule donc, de la même manière, le message de C_{13} à C_{14} et on modifie le séparateur, puis on remet à jour C_{14}. La figure 3.8(d) nous donne les nouveaux potentiels pour C_{14} et le séparateur. A ce moment là, la propagation est complètement finie en terme d'envoi de messages entre les cliques. Le potentiel de chaque clique est maintenant égal à $P(C_i, \varepsilon)$, c'est-à-dire, la probabilité de la clique *et* de l'évidence. En normalisant la clique sur ε, on obtient $P(\varepsilon) = 0.284$. Ceci est vrai pour toute clique. On peut alors normaliser sur l'ensemble des cliques afin d'obtenir la probabilité du réseau sachant l'évidence (et non plus en conjonction avec l'évidence). Ceci nous donne finalement la figure 3.8(e), et termine complètement l'algorithme JLO.

On peut bien sûr insérer plus d'une évidence à la fois, et propager après pour calculer $P(U|\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)$. On peut aussi incorporer les évidences une à la fois et propager à chaque fois pour voir l'évolution de la probabilité du réseau. Dans les deux cas, on aura le même résultat à la fin. De plus, si une variable contient n états discrets, il est possible de donner au plus $n - 2$ états impossibles, laissant ainsi un certain *degré de liberté* sur les états probables de la variable.

2.4.8 Complexité de l'étape de propagation

La complexité de l'algorithme JLO au moment de la propagation des messages est de $O(\sum_{i=1}^{N_C} n_e(C_i))$ où N_C est le nombre de cliques de l'arbre de jonction et $n_e(C_i)$ est le nombre d'états de la clique C_i. Ainsi, pour réduire cette complexité, il est nécessaire de construire des cliques ayant un petit nombre de variables (et si possible, avec des variables ayant un petit nombre d'états). Cependant, le problème de trouver un arbre de jonction optimal, avec des cliques les plus petites possible, reste

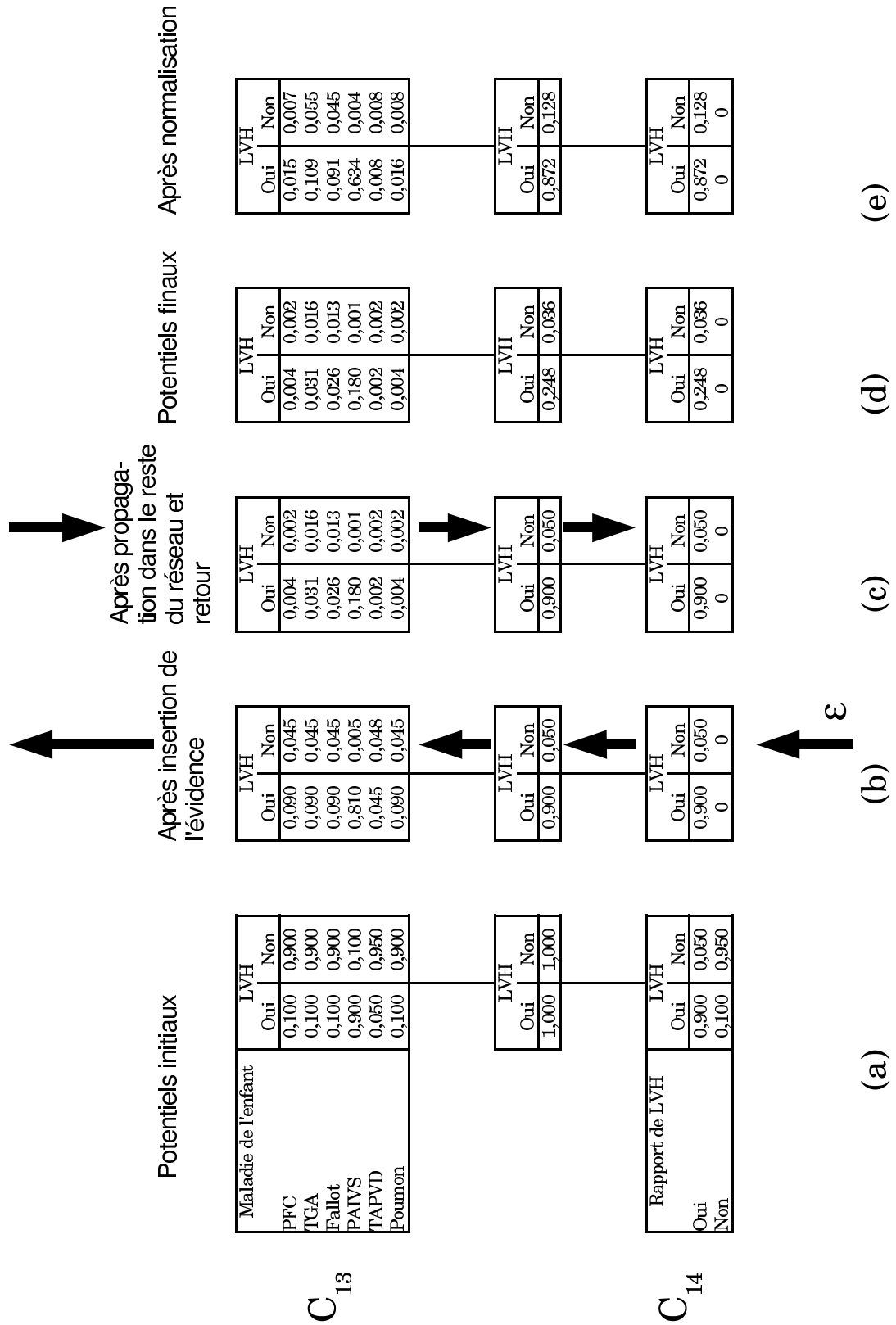


FIG. 3.8 – Exemple d'une propagation entre les cliques C_{14} et C_{13} (et le reste du réseau)[Cowell et al., 1999]

un problème NP-difficile. En pratique, l'heuristique proposée dans [Jensen et al., 1990] donne de bons résultats.

3 Conclusion

3.1 Max-propagation

L'algorithme JLO est à la base d'une famille plus complète d'algorithmes permettant de faire de l'inférence sachant un ensemble d'évidences sur certaines variables du réseau. En particulier, l'algorithme de Dawid sert à trouver la configuration la plus probable de toutes les variables sachant une représentation sous forme de réseau bayésien d'une fonction p de probabilité [Dawid, 1992]. L'algorithme JLO est utilisé en modifiant simplement la routine principale de propagation. Pour créer les messages inter-cliques, au lieu de faire une marginalisation avec une somme , on peut utiliser une fonction de maximisation telle que, si $W \subseteq U \subseteq V$ sont des sous-ensembles des noeuds du réseau et ϕ est un potentiel sur U, alors l'expression $M_{U\setminus W}\phi$ dénote la *max-marginalisation* de ϕ sur W, définie par

$$\left(M_{U\setminus W}\phi\right)(x) = \max_{z\in\mathcal{U}\setminus\mathcal{W}} \phi(z.x)$$

Les flux de messages sont alors calculés de la même façon qu'avec l'algorithme JLO, en remplaçant simplement la marginalisation de la formule 3.7 par celle donnée précédemment. Le schéma de propagation est ensuite exactement le même. Les mêmes résultats et les mêmes propriétés que ceux de l'algorithme JLO s'appliquent à l'algorithme de Dawid. Après une propagation complète, le réseau contient la configuration jointe la plus probable de l'ensemble des variables du réseau. Bien sûr, si une variable a été *instanciée* par une évidence, elle gardera sa valeur, après propagation. Cet algorithme peut être vue comme donnant *la meilleure explication* de l'évidence.

3.2 Perspectives

Même si le problème de l'inférence est résolu, un algorithme tel que JLO n'est applicable que sur des applications raisonnables. Il est possible de traiter de très grands réseaux, à condition que les cliques gardent une taille acceptable[2]. Les recherches actuelles s'orientent vers l'amélioration des divers algorithmes de propagation. On trouvera des références dans [Kjaerulff, 1998], [Shenoy, 1997] ou encore [Madsen and Jensen, 1998].

L'algorithme de Viterbi pour les modèles de Markov cachés [Viterbi, 1967] est un cas particulier de l'algorithme de max-propagation [Smyth et al., 1996], comme d'autres algorithmes de décodage tels que BCJR [Bahl et al., 1974] (dans [Frey, 1998] on trouvera plus de détails sur cette méthode). On peut aussi transformer d'autres algorithmes en instance des algorithmes de propagation [Murphy, 2002] tels que les transformations de Hadamard, les FFT [Kschischang et al., 2001], les algorithmes de satisfaction en logique propositionnelle (Davis et Putnam) [Dechter, 1998] ou encore certains algorithmes de parsing de grammaires [Parsing, 1999].

[2]La *taille acceptable* dépend du nombre de variables et du nombre de valeurs discrètes que peut prendre chaque variable. Par exemple, si toutes les variables sont binaires, alors une taille acceptable ne dépassera pas 20 à 25 variables (entre 4Mo et 128Mo par table de probabilités).

L'algorithme de propagation (JLO, Dawid, etc...) a été généralisé par R. Dechter en 1996. Il porte le nom d'algorithme de *bucket élimination* [Dechter, 1996a].

Cependant, les recherches ne se limitent pas à la découverte d'algorithmes de propagations plus performants, mais s'orientent aujourd'hui vers la modélisation efficace et intuitive de réseaux bayésiens toujours plus grands :

- apprentissage de la structure du réseau [Friedman, 1998],
- adaptation en ligne de la structure d'un réseau [Chaodhury et al., 2002],
- évaluation de la pertinence d'un réseau par rapport à un problème donné,
- modélisation intuitive avec de nouveaux modèles comme les réseaux bayésiens orientés objet [Koller and Pfeffer, 1997], les réseaux bayésiens hiérarchiques [Murphy and Paskin, 2001], etc...
- incorporation de la notion de temps dans un réseau bayésien [Aliferis and Cooper, 1995], [Jr. and Young, 1999].

Bien que des solutions partielles existent déjà, ces problèmes, encore aujourd'hui, restent largement ouverts et dignes d'intérêt.

Chapitre 4

Modélisation des systèmes dynamiques : application à Diatelic^{TM}

Résumé :

Ce chapitre présente une contribution au domaine de la télémédecine et de la fusion de données à travers une utilisation des réseaux bayésiens dynamiques pour la modélisation d'un problème de diagnostic médical à distance, dans le cadre du projet Diatelic^{TM}. Scindé en trois parties, il aborde, dans un premier temps, les réseaux bayésiens dynamiques, qui sont une solution intéressante à la modélisation des processus stochastiques. Puis, dans un deuxième temps, le système Diatelic^{TM}v3 est présenté, son architecture est expliquée et la modélisation du phénomène avec des réseaux bayésiens dynamiques est détaillée. La troisième partie s'intéresse, quant à elle, aux expérimentations réalisées à partir d'un prototype développé dans le cadre de cette thèse et montre des résultats encourageants, tout en insistant sur la nécessité, à plus long terme, d'une évaluation et d'une validation médicale rigoureuse. Des résultats seront présentés et mettrons aussi en évidence la nécessité de fusionner les données pour obtenir un diagnostic de bonne qualité, utile au médecin.

1 Introduction : le diagnostic

Le diagnostic est l'opération qui consiste à trouver les causes d'un phénomène connaissant un ensemble d'observations, c'est-à-dire l'opération qui consiste à répondre à la question : *pourquoi a-t-on obtenu le résultat que l'on vient d'observer ?* Ceci est l'acception la plus générale de la notion de diagnostic. Dans le cas du diagnostic médical, diagnostiquer revient à trouver les causes qui ont engendrées un certain nombre de symptômes chez le patient. Les causes sont dans ce cas la maladie ou l'affection dont souffre le patient.

Les programmes de diagnostic ont été sans doute les premiers systèmes développés dans le cadre de l'intelligence artificielle appliquée. Ces systèmes sont apparus dans les années 1970. Bien que les problèmes abordés soient souvent similaires, les approches proposées pour les résoudre étaient souvent radicalement différentes. De nombreuses approches ont été proposées dans la littérature pour modéliser la notion de diagnostic [Lucas, 1998]. Dans les systèmes utilisant une connaissance empirique, la classification à base d'heuristiques est une méthode commune pour faire du diagnostic. Dans le cadre de systèmes basés sur des modèles de domaines détaillés, le diagnostic à base de modèles est aussi une pratique courante. Typiquement, ces types de systèmes incorporent des connaissances structurées, causales et des descriptions des interactions entres les différents objets et entités modélisés.

Par exemple, dans [Chandrasekaran, 1988], on trouve une analyse du processus de diagnostic sous forme d'un ensemble de petites *tâches génériques pour la résolution de problèmes*. D'autres études ont été faites, en particulier sur la représentation de connaissances en diagnostic avec, dans un premier temps, la proposition de systèmes à base de règles empiriques [Buchanan and Shortliffe, 1984]. Plus tard, on verra apparaître les systèmes basés sur des modèles physiques et/ou biologiques : ils connaîtront une certaine popularité au niveau des applications industrielles [Beschta et al., 1993, Dague, 1994] ou médicales [Downing, 1993]. L'idée est de construire un système à base de connaissances utilisant un modèle explicité du domaine concerné.

Les aspects plus formels du diagnostic ont été étudiés, surtout ceux utilisant des connaissances causales, avec en particulier l'introduction de la logique comme base formelle [Poole et al., 1987, Poole, 1994]. Dans la théorie logique du *diagnostic abductif*, ce dernier est formalisé comme un raisonnement des effets vers les causes, avec la connaissance causale représentée comme des implications logiques de la forme causes $\rightarrow$ effets.

Les causes sont en général des défauts ou des anomalies, mais elles peuvent aussi inclure des situations normales. De nombreuses approches de la causalité en logique ont été proposées ; on citera en particulier [Lin, 1996] et [Cain and Turner, 1995]. La plupart de ces approches sont orientées vers le qualitatif, mais dans un grand nombre de problèmes, les systèmes de diagnostic doivent gérer des connaissances incertaines, et raisonner en environnement incertain en utilisant des mesures particulières de l'incertitude. Les réseaux bayésiens représentent un modèle particulièrement efficace, car ils permettent une représentation au sein d'un même modèle de connaissances qualitatives causales et de connaissances quantitatives exprimant l'incertitude que l'on a sur les connaissances qualitatives. En effet, de nombreuses applications sont basées sur l'utilisation d'un réseau bayésien en diagnostic médical et essentiellement sur le diagnostic à partir d'observations incertaines. Les modèles utilisés sont soit construits à partir d'une expertise humaine, soit construits par

différents algorithmes qui exhibent[1] les relations causales contenues de façon intrinsèque dans les données. On peut citer en référence [Cooper and Herskovitz, 1992],[Heckerman and et al., 1995] et [Cheng and Bell, 1997]. Ce dernier cas est possible et efficace lorsque les données sont présentes en grande quantité et que le corpus de données est suffisamment complet (corpus sans *trous*) [Wu et al., 2001]. Dans le cadre de cette thèse, l'expertise médicale étant plus exploitable que les simples données d'observations des patients, nous avons choisi de modéliser le système Diatelic^{TM}à partir des connaissances des médecins, et non en utilisant une méthode d'apprentissage automatique.

Dans le cadre d'environnements évoluant dans le temps, le diagnostic vise à fournir une estimation de l'état dans lequel se trouve un système que l'on observe, et de remettre à jour cette estimation périodiquement ou chaque fois que de nouvelles observations sont disponibles. Cependant, l'état de ce système se modifie régulièrement pour des raisons souvent inconnues ou du moins non-observables (*cachées*). Il est donc nécessaire de pouvoir observer, au moins partiellement, l'état de ce système et après chaque observation de pouvoir ré-estimer l'état dans lequel se trouve ce système. Dans le cas du diagnostic médical, les observations sont la plupart du temps les symptômes que présente le patient et l'état à estimer est son état physiologique, à partir duquel un diagnostic peut être plus facilement déduit (soit par le système expert, soit par un médecin dans le cas de l'aide à la décision).

Ce problème d'estimation en continu (ou de façon régulière) de l'état d'un environnement ou d'un système est plus connu sous le nom de problème de *monitoring* ou encore *monitorage*. Il s'agit d'estimer l'état sachant un ensemble d'observations, donc dans le cas d'une approche probabiliste, d'estimer $P(E|O_1 \dots O_n)$ où E est l'état du système et $\{O_1 \dots O_n\}$ est un historique des observations. Dans le cas du monitoring à horizon fini, le problème se réduira à l'estimation de $P(E|O_i \dots O_n)$ où $(n-i)$ est la taille de la séquence d'observations nécessaires pour obtenir une bonne estimation de l'état de l'environnement observé.

La section qui suit va donc s'intéresser à la modélisation des systèmes dynamiques et en particulier à l'utilisation des réseaux bayésiens pour l'estimation de l'état d'un système dynamique sur lequel on peut faire des observations au cours du temps.

2 Le monitoring

Le monitoring est un processus visant à observer le monde et à noter aussi les événements importants pour une tâche spécifique afin d'estimer l'état dans lequel se trouve le monde que l'on observe. Les événements importants sont inattendus, inhabituels et relatent un changement de l'état du monde qui l'observe. Le monitoring est effectué en utilisant des capteurs qui fournissent des données d'observations sur le monde. A partir de ces données, un système de traitement de l'information fourni une représentation adéquate du monde observé en vue d'une prise de décision. La décision peut à son tour avoir des conséquences sur le monde observé et entraîner la venue de nouveaux événements qui devront être pris en compte. Le monitoring est donc un processus répétitif voire cyclique d'une phase d'observation d'un monde (ou environnement) suivie d'une phase d'estimation de l'état de cet environnement. Dans de nombreux cas, l'estimation se focalise sur des paramètres *cachés*, c'est-à-dire qui ne sont pas directement observables.

[1] Cette méthode est couramment appelée l'*apprentissage de structure*.

Quand les paramètres sont cachés, la première idée est d'utiliser plus de capteurs et plus de sortes de capteurs (un exemple est présenté dans [Rao, 1991]). Cependant, il n'est jamais possible de tout observer : soit le système ne dispose pas assez de capteurs, soit les capteurs ne sont pas assez discriminants. En effet, il existe toujours une limite à la vitesse d'échantillonage des capteurs, et tous les capteurs sont toujours sujet au bruit et aux erreurs. Pour reprendre l'exemple du système Diatelic^{TM}présenté dans les chapitres 1 et 2, la mesure de l'état d'hydratation d'un corps vivant n'est possible que de deux façons :

– dans le cas d'un être humain, il existe des pèse-personnes à impédance électrique permettant d'estimer la masse graisseuse d'un patient en faisant passer un courant électrique dans le corps du patient. Cette méthode est approximative. C'est pourquoi elle ne peut pas être utilisée dans le cas de la surveillance d'un patient sous dialyse péritonéale ;
– dans le cas général d'un corps vivant, on peut mesurer la quantité d'eau présente dans le corps en brûlant ce corps puis en faisant la différence du poids avant et après la combustion. Après la combustion, seules les matières solides sont présentes, l'eau s'étant évaporée.

Dans la plupart des cas, le but d'un système de monitoring est de détecter les changements. Ceux-ci peuvent survenir brutalement (changement de direction d'un mobile, malaise subit chez un patient,...) ou de façon plus continue (tendances, amélioration ou aggravation de l'état d'un patient,...).

Parmi les modèles de monitoring les plus connus, on trouve les filtres de Kalman, les HMM et les réseaux bayésiens dynamiques. Nous allons voir par la suite que les réseaux bayésiens dynamiques (RBD) peuvent généraliser de nombreux modèles tels que les filtres de Kalman ou encore diverses sortes de HMM [Smyth et al., 1996, Murphy, 2002].

3 Les modèles d'espace d'états

3.1 Introduction

Dans de nombreux domaines, l'observation d'un environnement se fait à travers un flux d'informations appelé données séquentielles. Ces données peuvent autant être issues d'un système dynamique que générées par un processus spatial à une dimension, comme des séquences de données biologiques. Il est alors possible d'analyser ces données en ligne (analyse en temps réel) ou après leur collecte. Dans le cas d'une analyse en ligne, il est courant de vouloir prédire les observations futures, ou les états futurs (si ceux-ci sont cachés) de l'environnement que l'on est en train d'observer, connaissant l'ensemble des observations jusqu'au temps courant. Le futur est un phénomène incertain, et si l'on prédit un état futur, on ne peut jamais être sûr qu'il sera tel qu'il a été prédit. Il est donc intéressant d'avoir une mesure de certitude associée à la prédiction, ce qui permettra d'estimer la confiance que l'on peut accorder à cette prédiction. En supposant que t est l'instant courant, on appelle $y_{1:t} = (y_1, \ldots, y_t)$ la séquence d'observations jusqu'au temps t (en supposant le temps discret). Prédire l'état futur de l'environnement revient donc à estimer la distribution de probabilité $P(y_{t+h}|y_{1:t})$ à l'instant $t + h$ où h est appelé l'horizon, c'est-à-dire la distance à laquelle on veut prédire dans le futur.

Dans le cas où l'observateur peut agir sur l'environnement, ses actions peuvent être prises en compte sous forme d'une commande u (aussi appelée *entrée*). Appelons alors $u_{1:t}$ les actions ef-

fectuées sur l'environnement depuis l'instant initial jusqu'à l'instant courant, et $u_{t+1:t+h}$ les h prochaines actions. Alors, prédire revient à calculer la distribution de probabilité $P(y_{t+h}|u_{1:t+h}, y_{1:t})$ sur l'ensemble des états que peut prendre l'environnement. De nombreuses approches permettent de résoudre ce problème dans de nombreuses situations. On peut citer en particulier les réseaux de neurones ou encore les arbres de décisions [Meek et al., 2002]. Pour les données discrètes (comme les données textuelles par exemple), les modèles n-grammes sont couramment utilisés [Jelinek, 1997], $(n-1)$ étant la taille maximale de la séquence d'observation $1:t$ (celle-ci ne dépassant que rarement 2).

Dans ces modèles classiques, le problème est que la prédiction doit être basée sur une fenêtre de temps finie $y_{t-l:t}$ où $l \geq 0$ est la taille de la fenêtre temporelle. Si le système modélisé est markovien avec un ordre inférieur à l, alors la prédiction avec une fenêtre $y_{t-l:t}$ plus petite que l'historique $y_{1:t}$ complet donnera les mêmes résultats. Mais dans de nombreux cas, l'ordre est grand et inconnu.

3.2 Définition d'un modèle d'espace d'états

Pour résoudre les problèmes que nous venons d'évoquer et en particulier celui du monitoring, il existe aussi une grande famille de modèles qui sont les modèles d'espace d'états. Ils sont connus pour être efficaces dans la résolution des problèmes de prédiction et de monitoring. Avec ce type de modèle, on suppose que l'environnement observé possède un certain nombre d'états cachés qui génèrent lesdites observations, et que ces états cachés évoluent au cours du temps, souvent comme une fonction des entrées (les actions sur l'environnement). Dans le cas d'un monitoring en ligne, le but est d'inférer les états cachés sachant les observations jusqu'au temps présent. Si X_t est l'état caché à l'instant t, alors l'inférence sera l'estimation de la distribution de probabilité $P(X_t|y_{1:t}, u_{1:t})$ sur l'ensemble des états cachés. C'est ce que l'on appelle un état de croyance (*belief state*). L'état de croyance peut être mis à jour récursivement en utilisant la règle de Bayes. Ainsi un monde peut être décrit par un ensemble d'états de croyance, que l'on ne peut observer et un ensemble d'observations possibles à partir desquelles on infère les états de croyance.

Pour définir un modèle d'espace d'états, il faut d'abord connaître l'état initial $P(X_1)$ qui est une distribution de probabilités initiale sur l'ensemble des états. Ensuite, on définit une fonction de transition d'état à état $P(X_t|X_{t-1})$. Cette fonction permet de connaître la probabilité de se retrouver dans un état particulier à l'instant t sachant que l'on était dans un autre état particulier à l'instant précédent. On définit de même une fonction d'observation $P(y_t|X_t)$. Dans le cas où l'observateur a un contrôle particulier sur l'environnement, la fonction de transition devient $P(X_t|X_{t-1}, u_t)$ et la fonction d'observation est $P(y_t|X_t, u_t)$. Ceci permet en particulier de modéliser la perception active [Murphy, 2002]. Dans la plupart des cas, on supposera que le système modélisé est markovien d'ordre 1, c'est-à-dire que $P(X_t|X_{1:t-1}) = P(X_t|X_{t-1})$. De plus, on supposera que le système est homogène (la fonction de transition ne change pas au cours du temps), ce qui permet de modéliser des séquences d'observations de longueur infinie.

3.3 Inférence

Un modèle d'espace d'état est un modèle qui décrit comment X_t génère (ou cause) Y_t et X_{t+1}. L'inférence dans de ce type de modèle consiste simplement à inverser ce processus, c'est-à-dire à

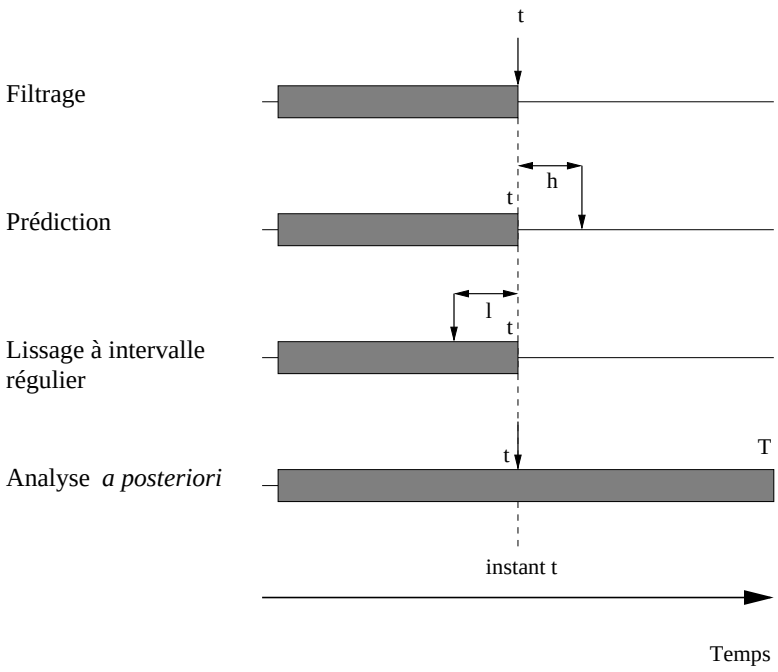


FIG. 4.1 – Principales façon d'inférer dans les modèles d'espace d'états. La partie grisée représente la période pour laquelle on dispose d'observations, les flèches verticales donnent l'instant de l'inférence (t)

calculer $X_{1:t}$ sachant $Y_{1:t}$. Ceci revient en fait à estimer les états cachés en utilisant les observations seulement. Il existe plusieurs sortes d'inférences possibles dans ce type de modèle ainsi qu'on peut le voir sur le figure 4.1 .

Filtrage

Il s'agit de l'inférence la plus courante. Elle consiste à estimer l'état de croyance en utilisant la règle de Bayes, à partir des observations jusqu'à l'instant *t*, ce qui revient à calculer :

$$P(X_t|y_{1:t}) \propto P(y_t|X_t, y_{1:t-1})P(X_t|y_{1:t-1})$$

Comme on fait une hypothèse de Markov sur le système modélisé, on peut écrire que $P(y_t|X_t, y_{1:t-1}) = P(y_t|X_t)$. Cette tâche est aussi appelée *monitoring*.

Lissage à intervalle régulier

Connu aussi sous le terme de *hindsight*, cette opération consiste à estimer $P(X_{t-l}|y_{1:t})$ où l est un intervalle de temps. Il s'agit donc d'estimer un des états précédents à la lumière des nouvelles observations. C'est une révision des états de croyance passés. Le lissage est utile pour l'apprentissage des paramètres d'un modèle.

Prédiction

Il s'agit de l'opération inverse : estimer un état futur sachant les observations jusqu'à l'instant courant. Ceci revient à calculer $P(X_{t+h}|y_{1:t})$ où $h > 0$. En marginalisant sur X_{t+h}, on peut aussi prédire les observations futures.

Analyse a posteriori

Il s'agit là encore d'un lissage, mais que l'on effectue après avoir obtenu l'ensemble des observations. Ceci revient donc à calculer $P(X_t|y_{1:T})$ pour tout $1 \leq t \leq T$.

4 Les réseaux bayésiens dynamiques

4.1 Introduction

Parmi les différentes approches des modèles d'espace d'états, les réseaux bayésiens dynamiques sont un cas intéressants car ils généralisent de nombreuses autres approches (filtres de Kalman, HMM, etc...). Ils permettent de spécifier rapidement des modèles complexes, mais aussi de combiner ou d'étendre d'autres modèles. Les HMM sous forme de réseaux bayésiens dynamiques ont été particulièrement étudiés et ont permis l'élaboration de modèles dérivés très simplement, comme les HMM semi-continus utilisés en reconnaissance de la parole pour réduire l'espace de stockage nécessaire au traitement des vecteurs d'observations [Gales, 1999], les HMM auto-régressif où

les observations y_t ont une influence directe sur les observations à l'instant y_{t+1} permettant d'affaiblir l'hypothèse classique que l'état X_t ne dépend que de l'état précédent X_{t+1} et des observations courantes y_t [Hamilton, 1994], ou encore les HMM hiérarchiques qui sont une extension des HMM pour modéliser des domaines ayant une structure hiérarchique et des dépendances à différentes échelles de temps ou différents niveaux d'abstraction [Fine et al., 1998].

En plus de la possibilité de modéliser des processus dynamiques, l'intérêt des réseaux bayésiens dynamiques réside dans le fait qu'il est possible d'élaborer de nouveaux modèles en modifiant la structure du réseau et de combiner différents réseaux entre eux afin d'obtenir un nouveau modèle. Il est ainsi possible de connecter plusieurs HMM entre eux simplement en reliant de façon approprié deux réseaux bayésiens dynamiques équivalents aux HMM.

4.2 Définition

Un réseau bayésien est un modèle statique. En fait, la notion de temps n'intervient pas dans un réseau bayésien *classique.* Mais pour modéliser un processus stochastique, on utilise un réseau bayésien dynamique qui est créé en répétant dans le temps un réseau classique et en reliant ces réseaux classiques par des liens causaux d'un pas de temps à l'autre. *Un réseau bayésien dynamique* est donc une chaîne du même réseau bayésien répété autant de fois que nécessaire (suivant la longueur de la séquence d'observations). Chaque répétition est un pas de temps permettant de représenter l'évolution d'un processus stochastique. Ils contiennent chacun un certain nombre de variables aléatoires représentant les observations et les états (cachés) du processus.

Plus formellement, un réseau bayésien dynamique est une extension des réseaux bayésiens pour modéliser des distributions de probabilités sur un ensemble semi-infini d'ensembles de variables aléatoires $Z_1, Z_2, \ldots$.[2] Dans la plupart des applications, on souhaite décomposer Z_t en trois sous-ensembles : X_t, les variables d'états (non-observables), Y_t, les variables observées, et U_t les variables d'entrées (ou de commandes). On ne s'intéressera qu'aux processus à temps discret. Les processus à temps continus ne peuvent être représentés avec un tel modèle. Les processus à temps continu ne peuvent être représentés par ce modèle. Il existe un autre modèle, appelé réseau bayésien temporel [Jr. and Young, 1999] permettant de représenter le temps au sein même de la fonction de distribution de probabilité conditionnelle associée à chaque noeud du réseau. Cette fonction devient alors dépendante du temps. Cependant, ce modèle ne permet pas de prendre en compte efficacement un historique d'observations ou un historique des états (ou les deux en même temps), sauf, bien sûr, à créer des variables représentant le passé, le présent ou encore le futur au sein du même réseau. Ceci revient alors à créer un réseau bayésien dynamique.

Dans un RBD, la variable t représentant le temps s'incrémentera de un à chaque fois qu'une nouvelle observation est disponible. Il est aussi possible de considérer les processus à événements discrets dans lesquels la variable t ne sera incrémentée que lorsqu'une nouvelle observation est disponible. Ainsi le temps réel entre t et $t+1$ ne sera pas le même à chaque valeur que prendra t.

Définition 4.1 (Réseau bayésien dynamique) *Un réseau bayésien dynamique est une paire $D = (B_1, B_{\rightarrow})$ où B_1 est un réseau bayésien qui définit un état initial $P(Z_1)$ et $B_{\rightarrow}$ est un réseau bayésien temporel à deux pas de temps (2TBN) qui définit $P(Z_t|Z_{t-1})$ à l'aide d'un graphe acyclique*

[2] Généralement, l'ensemble est fini $Z_1, \ldots, Z_n$ mais il peut grandir autant que l' on veut ($Z_1, \ldots, Z_k$ où $k > n$).

de la façon suivante :

$$P(Z_t|Z_{t-1}) = \prod_{i=1}^{N} P(Z_t^i|pa(Z_t^i))$$

où Z_t^i est le i-ème noeud à l'instant t (qui peut être un composant de X_t, Y_t ou U_t) et $pa(Z_t^i)$ sont les parents de Z_t^i dans le graphe.

Les noeuds dans le premier pas de temps du 2TBN n'ont pas de paramètres associés, mais chaque noeud du deuxième pas de temps du 2TBN a une distribution de probabilités conditionnelles qui lui est associée et qui définit $P(Z_t^i|pa(Z_t^i))$ pour tout $t > 1$. Les distributions ont une forme arbitraire (tables, mixtures de gaussiennes, etc...). Les parents d'un noeud Z_t^i peuvent aussi bien être dans le même pas de temps que dans le pas de temps précédent. On suppose que le modèle est markovien d'ordre un. Les arcs reliant les noeuds entre les différents pas de temps vont de la gauche vers la droite, représentant ainsi le déroulement continu du temps. Si il existe un tel arc d'un noeud Z_{t-1}^i à un noeud Z_t^i, alors ce noeud est dit *persistant*. Au sein d'un même pas de temps, les arcs peuvent être décrits de façon arbitraire tant que le graphe reste acyclique dirigé. D'une certaine manière, les arcs entre deux pas de temps sont vus comme la persistance d'un phénomène au cours du temps, alors que les arcs au sein d'un même pas de temps sont vus comme un effet causal immédiat.

La sémantique du réseau bayésien dynamique peut être définie en *déroulant* le 2TBN sur une longueur de T pas de temps. La distribution de probabilités jointes est alors :

$$P(Z_{1:T}) = \prod_{t=1}^{T} \prod_{i=1}^{N} P(Z_t^i|pa(Z_t^i))$$

Ceci revient à représenter un réseau bayésien dynamique *déroulé* comme un réseau bayésien

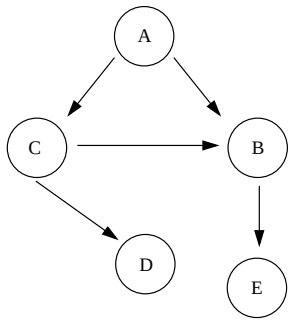


FIG. 4.2 – Un pas de temps pour un réseau bayésien dynamique

sous forme de chaîne dans lequel les variables sont aussi des réseaux bayésiens statiques. Par exemple, la figure 4.2 représente un pas de temps, et le réseau bayésien dynamique final de la figure 4.4. La figure 4.3 représente le 2TBN correspondant au modèle de base. La distribution de probabilités associée à chaque noeud du réseau de la figure 4.4 (la chaîne) est factorisée au sein d'un réseau bayésien *statique* : celui de la figure 4.2. Cependant, les dépendances temporelles entre chaque distribution de probabilités peuvent être mieux précisées. Ainsi, au lieu d'utiliser

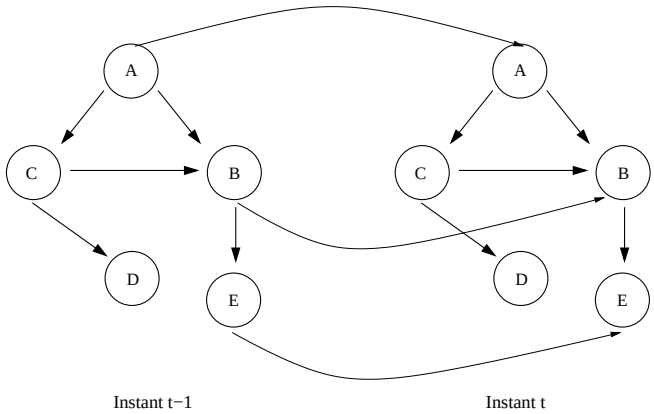


FIG. 4.3 – Un 2TBN où les liens causaux entre les pas temps relient les variables A, B et E à l'instant $t-1$ à leur homologue à l'instant t

deux réseaux, on en crée un seul dans lequel les arcs entre deux noeuds de pas de temps différents représentent les dépendances temporelles entre chaque pas de temps. L'intérêt est donc d'avoir un seul formalisme pour représenter à la fois un réseau statique et un réseau dynamique. La figure 4.5

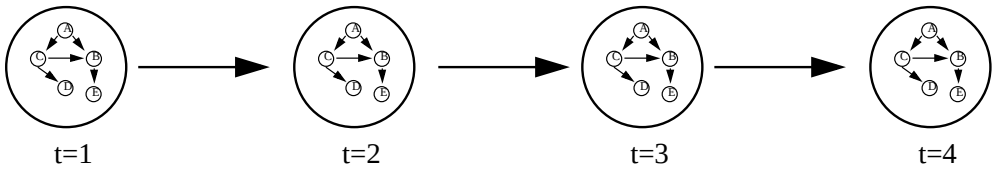


FIG. 4.4 – Un réseau bayésien contenant des réseaux bayésiens identiques dans chaque noeud : ici la relation entre les différents pas de temps est représentée par une *grosse* flèche, ce qui veut dire que chaque pas de temps influence directement le pas de temps suivant. Pour que le modèle soit complet, il faut expliciter les relations entre les pas de temps variables après variables, en utilisant le modèle fourni dans le 2TBN.

représente un réseau bayésien dynamique réel dans lequel il existe un phénomène de persistance uniquement entre les variables A_t, B_t et E_t. Ceci est beaucoup plus précis que le réseau précédent et permet ainsi de factoriser les dépendances temporelles de la même façon que sont factorisées les dépendances causales au sein d'un même pas de temps.

4.3 Des HMM aux réseaux bayésiens dynamiques

La principale différence entre les HMM et les réseaux bayésiens dynamiques est que dans un RBD les états cachés sont représentés sous forme distribuée par un ensemble de variables aléatoires $(X_t^1, \ldots, X_t^N)$. Ainsi, dans un HMM, l'espace d'états consiste en une seule variable aléatoire X_t. La figure 4.6 montre un HMM représenté dans sa forme graphique avec un réseau bayésien dynamique. Les noeuds en gris représentent les noeuds observés et les noeuds en blanc sont les noeuds cachés.

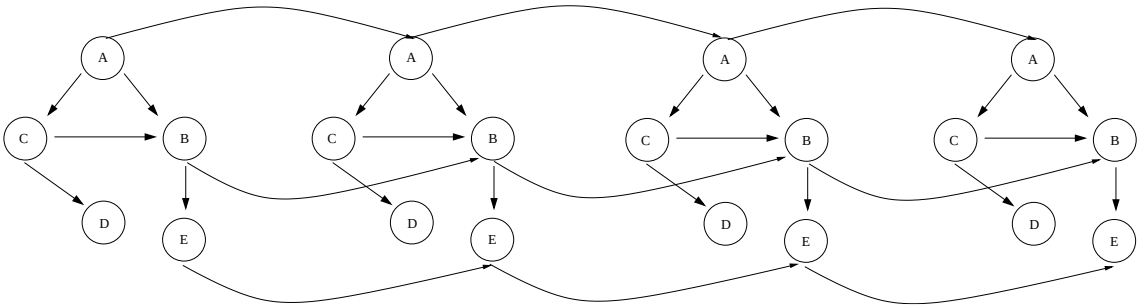


FIG. 4.5 – Un réseau bayésien dynamique *déroulé* sur 4 pas de temps

Le graphe représente les hypothèses d'indépendances conditionnelles suivantes :

– $X_{t+1} \perp X_{t-1}|X_t$ qui est la propriété de Markov,
– $Y_t \perp Y_{t'}|X_t$ pour tout $t \neq t'$. Cette hypothèse peut être relaxée dans le cas des HMM auto-régressifs (AR-HMM).

Dans la figure 4.6, suivant les notations utilisées dans la littérature sur les HMM, le noeud X_1 représente l'état initial π, avec $\pi(i) = P(X_1 = i)$. La matrice de transition A est représentée par les tables de probabilités entre les noeuds X_t et X_{t+1}, avec $A(i,j) = P(X_t = j|X_{t-1} = i)$. Enfin la matrice d'observation se retrouve dans les tables de probabilités entre les noeuds X_t et Y_t avec $B(i,j) = P(Y_t = j|X_t = i)$.

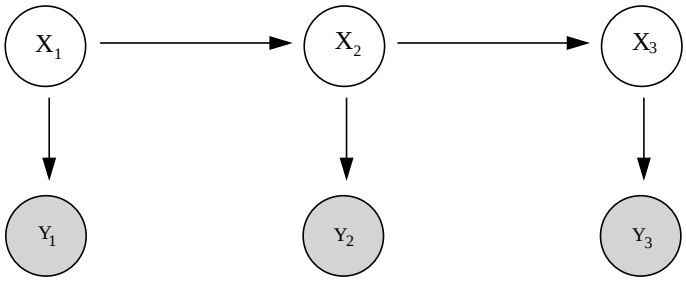


FIG. 4.6 – Un HMM représenté comme une instance de RBD *déroulé* sur 3 pas de temps

Ainsi, la spécification d'un HMM sous forme de réseau bayésien dynamique se fait simplement par la donnée des tables de probabilités pour $P(X_1)$, $P(X_t|X_{t-1})$ et $P(Y_t|X_t)$. Si l'on suppose que le modèle est invariant au cours du temps (matrice de transition et d'observations fixées au cours du temps) alors les données de $P(X_1)$, $P(X_2|X_1)$ et $P(Y_1|X_1)$ suffisent.

L'intérêt majeur des réseaux bayésiens dynamiques par rapport aux HMM est qu'il est très simple de créer des variantes aux HMM simplement en donnant une autre structure plus ou moins complexe au RBD. Le formalisme et les algorithmes restent les mêmes [Smyth et al., 1996]. Si l'on change les tables de distributions de probabilités (tables discrètes) par des distributions continues (par exemple des gaussiennes), alors il devient aussi possible de représenter des modèles basés sur les filtres de Kalman [Murphy, 2002]. Il est aussi possible de combiner ces modèles simplement en raccrochant divers RBD entre eux et fournir ainsi des modèles plus complexes.

4.4 Utilisation et inférence dans les réseaux bayésiens dynamiques

Comme les réseaux bayésiens dynamiques sont essentiellement utilisés pour déterminer des états cachés sachant une séquence d'observations, il est naturel que le problème de l'inférence dans un RBD soit d'estimer $P(X_t^i|y_{1:\mathcal{T}})$ où $\mathcal{T}$est la longueur de la séquence d'observations. Cependant, comme un réseau bayésien dynamique est avant tout un réseau bayésien, il est aussi possible, sachant la valeur d'un ensemble arbitraire de variables, d'inférer la distribution de l'ensemble des autres variables du réseau. Comme on l'a vu dans la section 3.3, si $\mathcal{T} = t$ alors il s'agit de filtrage, si $\mathcal{T} > t$ alors il s'agit de lissage, et si $\mathcal{T} < t$ alors il s'agit de prédiction.

Enfin, les réseaux bayésiens dynamiques sont particulièrement intéressants car il s'agit d'un modèle efficace permettant de représenter le temps dans un processus stochastique tout en ayant à notre disposition des algorithmes suffisamment efficaces. Néanmoins, d'autres représentations du temps ont été proposées dans le cadre des réseaux bayésiens, mais ils n'offrent pas la même puissance d'expression que les RBD, et posent parfois des problèmes incalculables autant pour l'inférence que pour l'apprentissage (voire les deux à la fois). [Aliferis and Cooper, 1995], [Dagum and Galper, 1995] et [Jr. and Young, 1999] proposent divers modèles de représentation du temps dans les réseaux bayésiens.

Je vais maintenant présenter un modèle à partir de réseaux bayésiens dont le but est la fusion d'informations hétérogènes, dans le cadre de la télémédecine. Ce modèle permet la détection de tendances dans l'état physiologique d'un patient. Il s'agit de la reformulation du problème de diagnostic médical du système Diatelic^{TM}v2 dans le cadre des réseaux bayésiens dynamiques.

5 Le système Diatelic^{TM}v3

Dans le chapitre 1, j'ai présenté le projet Diatelic^{TM}et la problématique qui y est associée :

– évaluer l'état d'hydratation d'un patient, sachant que cet état est une information cachée par rapport aux capteurs dont on dispose,
– mettre à jour un diagnostic en fusionnant les informations dont on dispose, même si elles sont incertaines, bruitées ou partiellement manquantes (par rapport à la quantité totale d'informations dont on devrait normalement disposer).

Voici maintenant un résumé qui servira de base au modèle à base de réseaux bayésiens dynamiques qui sera présenté par la suite.

Le problème du projet Diatelic^{TM}est de détecter des changements dans l'état d'hydratation d'un patient sous dialyse péritonéale à domicile. Chaque patient suivant cette procédure médicale de substitution des fonctions rénales doit fournir chaque jour diverses valeurs physiologiques. Les expériences menées sur le précédent système Diatelic^{TM}v2 ont montré que le poids et la tension sont les valeurs les plus intéressantes. La tension orthostatique (différence entre les tensions systoliques et diastolique) est une information à prendre en compte car, d'après les études médicales faites sur la dialyse péritonéale, il apparaît qu'elle est liée à l'évolution du poids du patient. Enfin, si le taux d'hydratation est l'information cachée du problème posé par Diatelic^{TM}, il existe une deuxième information nécessaire à la prise de décision qui est le poids idéal du patient. Ce poids idéal est le poids que le patient doit chercher à atteindre. Cependant, ce poids idéal est laissé à l'appréciation du médecin. Ce dernier peut le changer si il estime que le poids idéal actuel

ne correspond plus au patient. Divers critères entrent en jeu : physiologie du patient, tendance à prendre rapidement du poids (ou à en perdre rapidement), critères psychologiques, etc... Si le médecin est apte à déterminer un poids idéal sur une longue échéance, il est apparu que le choix d'un poids idéal au jour le jour est un problème beaucoup plus difficile. C'est pourquoi, le système Diatelic^{TM}doit aussi tenter d'estimer si le poids idéal choisi est toujours le bon. Dans le cas contraire, le médecin pourra éventuellement le modifier.

Pour la tension, la valeur moyenne est une moyenne mobile faite sur les n derniers jours d'observations du patient. Cette méthode simple permet en outre d'insister plus sur les changements brusques de tension et ainsi renforcer le fait qu'un changement brusque de tension est souvent précurseur d'une anomalie ou d'un incident. L'inconvénient est qu'une chute ou une augmentation de la tension très progressive (supérieure à 15 jours) passera inaperçue. Pour le poids, la valeur moyenne est bien évidemment le poids idéal. Ce poids idéal étant fixe au cours du temps (sauf changement volontaire de la part du médecin), les chutes ou prises de poids subites apparaissent clairement, mais une chute ou une augmentation du poids sur une longue période apparaîtra aussi. En ce qui concerne la tension, il n'existe pas de tension idéale car la tension dépend autant de l'activité physique du patient, que de son état mental ou encore de la période de l'année. De plus, le poids peut aussi influencer plus ou moins directement la tension du patient.

Parmi les choix importants, on doit noter celui des valeurs à attribuer aux variables. Le système Diatelic^{TM}(v2 et v3) travaillant sur la découverte de tendances, il est apparu évident de raisonner non pas sur les valeurs absolues des capteurs, mais sur leurs variations par rapport à une valeur moyenne. Ainsi, lorsque l'état du patient se dégrade, on souhaitera voir apparaître cette dégradation au fur et à mesure. Un taux d'hydratation anormal ne survient pas subitement du jour au lendemain, mais peut prendre plusieurs jours pour se déclarer. Cependant, avant d'atteindre un état critique, l'état se dégrade progressivement. La possibilité de voir l'évolution de l'état du patient et de distinguer les périodes où son état commence à se dégrader est une des exigences du projet Diatelic^{TM}. C'est une des raisons pour lesquelles il est nécessaire d'utiliser autant un historique des observations, qu'un historique des diagnostics pour calculer l'état d'hydratation courant du patient. Cette approche permet donc de prendre en compte l'évolution de l'état du patient dans le diagnostic courant.

5.1 Architecture du système Diatelic^{TM}v3

5.1.1 Présentation de l'architecture

L'architecture du système Diatelic^{TM}v3 peut être décliné selon deux modèles : l'architecture logicielle complète et l'architecture de la partie dite intelligente. L'architecture logicielle complète s'intéresse à la modélisation et l'implémentation d'un système de télémédecine complet basé sur des technologies telles que les bases de données relationnelles, la représentation de données en XML ou encore les services de communication de type Internet. On trouvera dans [Thomesse et al., 2001] et dans [Thomesse et al., 2002] un exposé plus complet sur la conception de tels systèmes de télémédecine.

L'architecture du sous-système intelligent est synthétisé dans la figure 4.7. Le point intéressant dans cette figure est l'utilisation d'opérateurs flous pour transformer les données issues des cap-

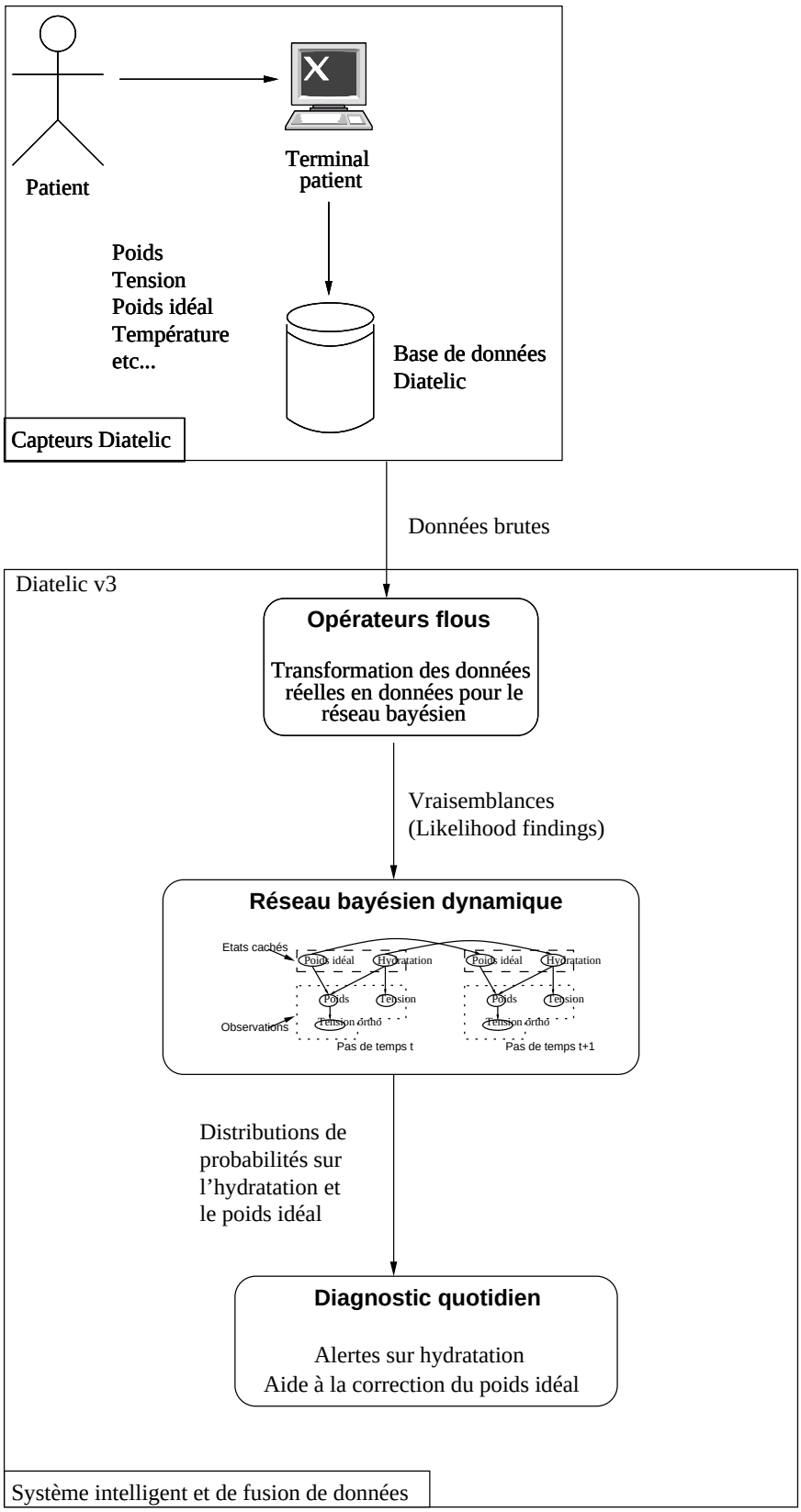


FIG. 4.7 – Architecture du système intelligent Diatelic^{TM}v3

teurs, en données dans un format unifié et acceptable directement par le réseau bayésien comme vraisemblances (connues aussi sous le nom de *likelihood findings*).

L'architecture ne présente pas de difficultés particulières. Il s'agit d'un système en couche particulièrement classique où les capteurs sont situés dans la couche la plus basse. La deuxième couche consiste à transformer les données brutes en données utilisables par le système de fusion. On notera en particulier que cette deuxième couche opère déjà une première fusion de données particulièrement simple, puisqu'elle fusionne le poids et le poids idéal ainsi que la tension et la tension moyenne (calculée sur 15 jours), en deux informations qui sont la variation du poids et la variation de la tension.

La troisième couche procède à la fusion proprement dite, puisqu'elle contient le réseau bayésien dynamique qui va donc fusionner, non seulement les différents types de données, issues des différents capteurs, mais aussi un historique des données et des diagnostics au cours du temps. Cette couche fait donc une fusion de données sensiblement plus complexe que dans la couche précédente.

La quatrième couche s'intéresse, quant à elle, à l'obtention d'un diagnostic simple, c'est-à-dire utilisable par un opérateur humain (le médecin en l'occurrence).

Une session de travail classique se déroule de la façon suivante : le patient se connecte au système en utilisant un navigateur web et en donnant nom et mot de passe. Il aura auparavant mesuré son poids et les différentes tension systolique et diastolique en position debout et couchée. Il saisie les données sur la page web présentée par le système Diatelic^{TM}pour les transmettre au serveur. A ce moment le serveur interroge le système expert en lui soumettant les nouvelles données. Le système expert calcule le diagnostic du jour et transmet le résultat au médecin et au patient. En cas de problème, une alarme est déclenchée en envoyant un message particulier au médecin contenant les informations nécessaires. Le médecin pourra dans ce cas intervenir en appelant le patient et en lui indiquant la nouvelle thérapie à suivre. Les expériences qui seront présentées par la suite utilisent des données issues du système 2 que je viens de présenter. Cependant, mon implémentation du système 3 n'est pas encore intégrée à l'expérimentation médicale faite sur le système 3. En effet, cela demanderait une modification du protocole expérimental pour intégrer les nouveautés apportées par le système 3 par rapport au système 2 (séparation du diagnostic et de l'aide à la décision sur le poids idéal, etc...). Ceci est une opération lourde à mettre en place et longue à réaliser. Une expérimentation médicale tel que Diatelic^{TM}v3 nécessite environ 2 années de test pour pouvoir être validée (cas de Diatelic^{TM}v2).

5.1.2 Implémentation

Le système implémenté au cours de cette thèse représente la partie Diatelic^{TM}v3, telle que présentée dans la figure 4.7. Le logiciel a été implémenté en C et C++ sous Linux et se compose de plusieurs parties.

La première partie sert à traiter le corpus et implémente les opérateurs flous. Il s'agit des mêmes opérateurs que le système v2 reprogrammés en langage C pour pouvoir être intégrés plus facilement dans le moteur bayésien écrit en C++ (la version 2 du système a été faite en Java). Ces opérateurs flous ont ensuite été intégrés dans un programme filtrant le corpus initial (données sur le poids, la tension, etc...) afin de fournir une entrée utilisable par le réseau bayésien dynamique.

La transformation effectuée durant cette phase sera amplement détaillée dans la section 5.4. Pour des raisons de simplicité et de robustesse du code, les différents programmes de cette partie ont été reliés avec des scripts écrits en C-shell.

La deuxième partie est le programme de diagnostic proprement dit. Ce programme utilise les résultats fournis par la première partie et calcule, chaque fois que de nouvelles observations sont disponibles, une mise à jour de l'état d'hydratation du patient. Il fournit en sortie l'ensemble des paramètres nécessaires (ceci est paramétrable) au diagnostic. Les courbes présentées dans la suite de ce chapitre sont issues directement de ce module de diagnostic. Ce module peut fournir une estimation brute de l'état d'hydratation du patient et aussi générer des alarmes pour le médecin.

Le programme de diagnostic a été écrit en C++ et utilise une bibliothèque de représentation et d'inférence dans les réseaux bayésiens. J'ai entièrement programmé cette bibliothèque en C++ : il s'agit d'un package générique servant à la représentation de n'importe quel réseau bayésien à variables discrètes et il implémente en outre la totalité de l'algorithme JLO (moralisation, triangulation, construction de l'arbre de jonction et propagation). Il peut donc traiter toute structure de réseau bayésien. Le respect de la norme ISO-C++ permet à cette bibliothèque d'être portable sur toute plate-forme. La bibliothèque possède en outre des objets de représentation de graphe et peut ainsi être étendue facilement à d'autres formes de réseaux bayésiens, comme ceux à variables continues gaussiennes.

La bibliothèque représente donc un réseau bayésien avec un objet `D_BNet` contenant un graphe et un ensemble de distributions de probabilités locales. Il hérite d'un objet `BN_Graph` représentant un graphe acyclique dirigé. Cet objet contient les primitives nécessaires à l'ajout de variables, la connexion des variables entre elles (pour former un graphe), la construction d'un arbre de jonction (moralisation, triangulation, algorithme de Kruskal).

L'objet `D_BNet` contient les primitives permettant d'insérer une évidence et une vraisemblance (*likelihood findings*) et de propager ces informations à l'ensemble du réseau. Les résultats sont récupérés grâce aux primitives permettant d'extraire une distribution de probabilités conditionnelles ou directement la distribution marginalisée de la variable d'intérêt. Sur un réseau bayesien contenant entre 50 et 100 noeuds (cas du système Diatelic^{TM}v3), la propagation prend environ 1ms sur une machine de bureau d'une puissance de 300 à 400 MIPS (millions d'opérations par seconde). Le temps de propagation est indépendant du nombre d'évidences insérées dans le réseau.

5.2 Le modèle de base

5.2.1 Définitions des variables du réseau

Pour utiliser le formalisme des réseaux bayésiens, nous avons besoin dans un premier temps d'exhiber les informations utiles à la représentation de notre problème. Ces informations sont représentées par des variables aléatoires à valeurs discrètes. Après une analyse du problème et en nous basant sur l'étude faite pour le système Diatelic^{TM}v2, nous avons déduit que les informations les plus pertinentes à traiter sont les suivantes :

– le poids,
– le poids idéal,
– la tension,
– la tension orthostatique,

– la moyenne de la tension sur n jours,
– le taux d'hydratation.

Cependant, ces informations ne sont pas immédiatement représentables par des variables aléatoires à valeurs discrètes. Il est donc nécessaire de procéder à une transformation préliminaire de ces informations. Comme nous souhaitons déterminer des tendances dans l'évolution de l'état du patient, il est donc nécessaire de raisonner sur les variations des grandeurs que l'on manipule plutôt que sur les valeurs absolues.

Il n'y a pas de modèle idéal d'un patient, cependant, le médecin souhaite que certaines grandeurs physiologiques restent dans des valeurs raisonnables. Ainsi le poids idéal est une valeur raisonnable pour le patient. Donc nous pouvons à partir de cela, raisonner sur les variations du poids par rapport au poids idéal. De même pour la tension, nous devons vérifier si le patient a une *bonne* tension en la comparant à sa tension moyenne calculée sur les n jours précédents. Classiquement n est égal à 15 jours, estimation donnée par le médecin néphrologue.

Nous ne connaissons pas la valeur acceptable pour l'hydratation. En effet, c'est ce que nous allons essayer de déduire en fusionnant les données physiologiques mesurées par le patient. Cependant, si des valeurs acceptables existent pour le taux d'hydratation et pour le poids idéal, alors on peut aisément imaginer que ces deux grandeurs puissent aussi être parfois trop grandes ou parfois trop basses par rapport à la valeur idéale.

A partir de cette réflexion, il apparaît clairement que les grandeurs que nous allons manipuler peuvent être représentées par des variables aléatoires pouvant prendre trois valeurs possibles : *trop haut*, *normal* ou *trop bas*. Pour clarifier le modèle, nous utiliserons les mêmes valeurs pour toutes nos variables aléatoires. Cependant, le formalisme des réseaux bayésiens autorise l'utilisation de variables aléatoires ayant chacune un nombre arbitraire de valeurs discrètes possibles.

Enfin, il faut noter que la valeur moyenne de la tension n'est pas nécessaire au raisonnement. Elle ne sert qu'à estimer si la tension est trop haute, normale ou trop basse. Par contre, il faut que le poids idéal soit représenté dans le modèle pour pouvoir par la suite l'évaluer et aider le médecin à décider si le poids idéal courant est bien adapté au patient qui est suivi. Ainsi, le problème peut être modélisé avec un réseau bayésien comprenant 5 variables aléatoires à trois valeurs discrètes chacune :

– le poids : c'est une variable d'observation,
– la tension et la tension orthostatique : ce sont deux variables d'observation,
– le taux d'hydratation : c'est une valeur cachée,
– le poids idéal : il sert à estimer la variation du poids du patient, mais comme on souhaite aider le médecin à déterminer si ce poids idéal est correct (aide à la décision thérapeutique), on n'insérera aucune observation dans cette variable : elle est donc considérée comme variable cachée. Ainsi, il sera possible d'estimer si le poids idéal est correct ou non et le cas échéant de prévenir le médecin.

Revenons à présent sur les deux dernières variables. L'intérêt de modéliser le taux d'hydratation de cette façon est que l'on peut immédiatement renseigner l'utilisateur du système (le médecin) sur la qualité du taux d'hydratation du patient : il est trop haut, ou trop bas ou normal. En fait, la valeur réelle du taux d'hydratation est quasiment impossible à déterminer de cette façon là. Nous n'avons pas assez d'informations sur le patient. Mais, on peut quand même estimer si ce taux

d'hydratation est anormal ou non et à quel point. Cette information est ensuite suffisante pour que le médecin puisse prendre une décision et éventuellement intervenir dans le traitement du patient.

Le poids idéal est connu, puisque c'est le médecin qui le fixe. Cependant, dans certaines situations, il se peut que le poids idéal ne soit pas correctement réglé. Le médecin souhaite donc savoir quand cette situation arrive. Dans ce cas, il pourra prendre la décision de modifier le poids idéal du patient. Par exemple, supposons que le patient soit dans un état de santé tout à fait satisfaisant, mais que le système déclare que le taux d'hydratation est trop élevé depuis une semaine. Dans ce cas, on est en droit de penser que le poids idéal est réglé trop bas. En effet, un taux d'hydratation trop élevé est souvent à l'origine d'une surcharge pondérale. Hors si le poids idéal est trop bas, la variable représentant le poids du patient prendra toujours la valeur *trop haut*. Et si le poids est trop haut, la probabilité que le taux d'hydratation soit trop haut augmente de même.

5.2.2 Modélisation de la structure du réseau

A présent, nous allons nous intéresser aux relations d'influence causale directe existant entre les variables. Cette étape permet de définir la structure du réseau bayésien utilisé.

La première observation est que le taux d'hydratation est un phénomène ayant une influence directe sur le poids et la tension. En effet, le poids total d'un être humain peut être décomposé en trois parties essentielles : la masse graisseuse, l'eau et les composantes solides (chair, muscle, os). Or si la quantité d'eau augmente, alors le poids augmente irrévocablement. De plus, si la quantité d'eau augmente, la tension artérielle est immédiatement modifiée car la pression appliquée à l'extérieur des veines est aussi modifié. Donc pour garder un certain équilibre, le coeur est obligé de modifier son effort. Ainsi, il est nécessaire de déclarer une influence directe de la variable d'hydratation sur les variables de poids et de tension.

Le poids idéal a de façon triviale une influence directe sur l'estimation de la variabilité du poids. Si le poids idéal est trop haut, on aura toujours tendance à estimer que le patient a trop perdu de poids, et réciproquement.

Enfin, bien que l'on ne sache pas actuellement si le taux d'hydratation a une influence directe sur la tension orthostatique, il est cependant évident que le poids du patient a une influence sur cette tension. En effet, une modification du poids (dû à une modification du taux d'hydratation ou dû à d'autres phénomènes) entraînera de toute façon une modification des tensions systoliques et diastoliques et par conséquent une modification de la tension orthostatique.

Ces informations, issues de l'expertise médicale, permettent de déduire directement le modèle graphique présenté en figure 4.8. Il relate simplement une modélisation simple des relations de cause à effet existants entre les variables qui nous paraissent pertinentes pour décrire le problème Diatelic^{TM}. Ce modèle ne prend pas en compte l'évolution du patient. Il est donc impossible de traiter une séquence d'observations. C'est pourquoi il va être étendu à un réseau bayésien dynamique, qui permet la modélisation d'un processus évoluant au cours du temps.

5.3 Modèle dynamique

Le modèle de base nous a permis de modéliser la connaissance médicale. Cependant, ce modèle ne permet pas la prise en compte de séquences d'observations : si un ensemble d'observations est

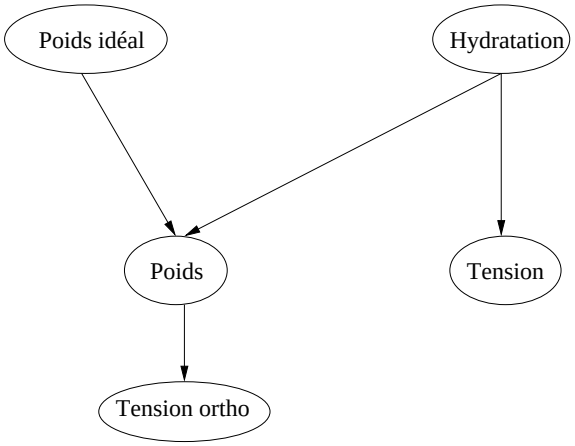


FIG. 4.8 – Modèle de base du système Diatelic^{TM}v3. Ce modèle est statique et ne peut pas prendre en compte l'évolution du patient.

inséré dans ce réseau bayésien, l'ordre dans lequel elles sont insérés n'a aucune importance. Or, dans le cas d'un raisonnement dans le temps, l'ordre dans lequel les observations sont utilisées est primordial : si la séquence d'observations est $\{o_1, o_2, o_3\}$, avec le modèle de base, l'ordre des observations n'a aucune importance et donnera le même résultat après propagation. Il est donc nécessaire d'étendre ce modèle de réseau bayésien à un modèle intégrant la notion de temps, en l'occurrence un réseau bayésien dynamique.

Dans cette section, le modèle de base est utilisé pour créer un modèle dynamique et de faire apparaître des tendances dans l'évolution de l'état de santé du patient. Nous avons vu, dans la section 4.2 que pour définir un réseau bayésien dynamique, il faut définir une paire $D = (B_1, B_\rightarrow)$ dans laquelle B_1 est le modèle initial et $B_\rightarrow$ est le modèle d'évolution. De plus, $B_\rightarrow$ est composé de trois ensembles de variables X_t, Y_t et U_t pour respectivement les états cachés, les observations et les variables de commandes. Ici, seul X_t et Y_t sont à définir, puisque le système ne prend en compte aucune action extérieure (commande de l'utilisateur par exemple).

Si nous appelons B_b le modèle de base, alors le modèle initial est $B_1 = B_b$. En effet, la distribution initiale des probabilités peut être représentée à partir du modèle de base directement. Il est à noter que les paramètres du modèle de base ont été définis à partir d'une expertise humaine (en l'occurrence les médecins-néphrologues de l'ALTIR à Nancy).

Le modèle d'évolution $B_\rightarrow$ est défini par le graphe acyclique dirigé de la figure 4.9 dans lequel les noeuds représentant le poids idéal et le taux d'hydratation ont été reliés d'un pas de temps à l'autre. Puisque nous faisons une hypothèse markovienne d'ordre un, il n'est pas utile de définir plus de deux pas de temps dans le modèle. Nous noterons cependant que dans un RBD, il n'est pas obligatoire de relier deux noeuds de même type d'un pas de temps à l'autre. Dans le modèle Diatelic^{TM}v3, si cela avait été juste et nécessaire, nous aurions aussi le droit d'écrire que le poids idéal à l'instant $t-1$ influe directement sur la tension orthostatique à l'instant t. Ceci ne reflète pas la réalité de Diatelic^{TM}, mais le formalisme des RBD permet d'écrire cela.

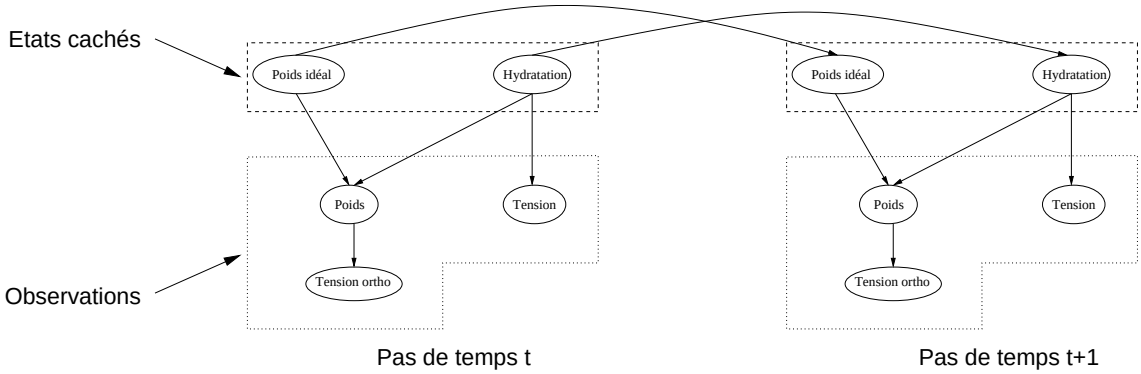


FIG. 4.9 – Modèle d'évolution du système Diatelic^{TM}v3. Les noeuds représentant les *états cachés* sont reliés d'un pas de temps à l'autre pour modéliser la relation markovienne d'ordre un qui existe entre eux.

Dans le modèle Diatelic^{TM}v3, le taux d'hydratation d'un jour a, de façon triviale, une influence directe sur le taux d'hydratation du lendemain. Pour pouvoir étudier l'évolution du poids idéal, il est aussi nécessaire que le poids idéal d'un jour influe sur le poids idéal du lendemain.

Enfin, nous faisons une hypothèse d'indépendance sur les variables d'observation. Bien que ceci ne reflète pas forcément la réalité (par exemple le poids du patient d'un jour influe sur son poids le lendemain), cette hypothèse permet de grandement simplifier le modèle. Donc dans ce modèle dynamique, les observations à l'instant t sont dépendantes uniquement de l'état du patient au même instant.

Dans la figure 4.9, les variables d'observations sont encadrées en pointillés et les variables cachées sont encadrées par des tirets. Cependant cette distinction est une vue de l'esprit. En effet, si l'on se réfère à la théorie sur les réseaux bayésiens présentée dans le chapitre 3, il n'y a aucune distinction possible entre les variables et elles sont toutes utilisées de la même façon lors de la phase de propagation des observations. Ainsi, si l'on souhaite simuler le comportement d'un patient dans, par exemple, un état d'hyperhydratation, il est possible d'instancier la variable $hydratation$ à la valeur $trop\,haut$ et d'utiliser un algorithme d'inférence (tel que JLO) pour tenter de prédire l'évolution du poids ou de la tension du patient.

La détermination d'une tendance se fait en utilisant, ou plutôt en fusionnant au sein d'un même modèle, un historique des observations et des diagnostics. Les réseaux bayésiens se prêtent naturellement à ce type de problème, car il est possible de *dérouler* le réseau sur une période arbitraire d'observations et d'utiliser ainsi un historique des observations. Les relations de dépendances indirectes entre les différentes parties du réseau permettront ainsi de propager l'influence des observations passées sur l'état courant du patient. Dans le cas du réseau Diatelic^{TM}v3, étant donné que les variables relatives à l'hydratation et au poids sec ne sont jamais instanciées, un changement sur l'une aura de l'influence tout au long de la chaîne et influencera donc l'hydratation ou le poids idéal à l'autre extrémité du réseau bayésien.

Avant de présenter des résultats significatifs du système Diatelic^{TM}v3, je vais présenter son architecture complète et en particulier l'utilisation d'opérateurs flous pour le passage de données numériques sur un espace continu à des données plus simples utilisables directement dans le réseau bayésien dynamique.

5.4 Les opérateurs flous

Dans le chapitre 2, une étude préliminaire du projet Diatelic^{TM}selon l'approche de la fusion de données, nous a permis d'identifier un certain nombre de sources de données hétérogènes. En effet, les trois grandeurs qui sont intéressantes sont le poids, la tension et la tension différentielle.

Le modèle présenté ici utilise, pour des raisons de vitesse de calcul et de simplicité de la représentation, des variables à trois états : $\{trop\,haut, normal, trop\,bas\}$. L'ensemble des variables du réseau bayésien possède ces trois états. Ce choix a été fait uniquement dans le but de simplifier et d'unifier le modèle.

Pour expliquer la sémantique des variables, je vais prendre un exemple. La variable $poids$ possède trois états. Si le poids idéal du patient est 70 kg et que le poids mesuré est 75 kg, alors il est évident que le patient est particulièrement au-dessus de son poids idéal. Donc il faut alors que la probabilité *a priori* $P(poids = trop\,haut)$ soit grande. Si au contraire le poids du patient est de 70,1 kg, alors la probabilité *a priori* $P(poids = normal)$ doit être nettement plus grande que $P(poids = trop\,haut)$ et $P(poids = trop\,bas)$. Comme $70,1 > 70$, on aura malgré tout $P(poids = trop\,haut) > P(poids = trop\,bas)$.

Ceci est valable pour un patient stéréotypé. Cependant, les patients ont chacun une physiologie différente, et certains prennent du poids très rapidement, mais sans conséquence grave, alors que pour d'autres, la prise de poids, même minime, peut s'avérer plus dangereuse. De plus, pour une personne pesant 100 kg, un kilogramme supplémentaire n'aura que peu d'incidence, alors que pour une personne pesant 50 kg, 1 kg de plus est une situation à prendre sérieusement en compte.

Donc pour passer d'une mesure absolue (le poids du patient) à une mesure relative prenant en compte la physiologie du patient, il est nécessaire d'utiliser trois opérateurs flous, un pour chaque valeur de la variable aléatoire. Ces opérateurs flous donnent les vraisemblances suivantes :

$$\begin{aligned} p_1 &= P(Poids > Poids\,ideal) \\ p_2 &= P(Poids = Poids\,ideal) \\ p_3 &= P(Poids < Poids\,ideal) \end{aligned}$$

Les opérateurs flous sont utilisés de la même manière avec la tension et la tension différentielle.

Pour créer les opérateurs flous et leur donner une phase de transition suffisamment douce d'un état à un autre, on utilise une approximation d'une fonction sigmoïdale pour les valeurs $trop\,haut$ et $trop\,bas$, et une fonction en cloche pour la valeur $normal$ qui est le complémentaire des deux autres fonction [Jeanpierre, 2002]. Ceci nous permet d'avoir trois valeurs de vraisemblance sommant à un, ce qui est une condition nécessaire pour être qualifiées de *likelihood findings* et être utilisables dans un réseau bayésien.

Afin de ne pas trop prendre en compte les variations très faibles et le bruit issu des capteurs (poids, tension, etc...), un facteur de tolérance a été associé au calcul des vraisemblances. Il a été déterminé par les médecins néphrologues : il est de 1,5 kg pour le poids et de 1,2 pour la tension. Ceci nous permet donc de définir les fonctions suivantes.

La déviation du capteur par rapport à la valeur normale (poids idéal, tension moyenne, etc...) est définie par :

$$dev(x) = \frac{x - base}{tolerance}$$

où x est la mesure issue du capteur, $base$ est la valeur normale (poids idéal, tension moyenne, etc...) et tolérance est le facteur de tolérance associé au capteur. Cette fonction nous permet d'obtenir une valeur normalisée et qui est ensuite utilisée pour calculer la vraisemblance de la mesure issue du capteur grâce aux fonctions suivantes :

$$
\begin{array}{rcl}
x = dev(y) \quad avec \quad y\, la\, valeur\, issue\, du\, capteur & & \\
sigmo^{-}(x) & = & \left\{ \begin{array}{llll} si & x < -1 & alors & 1 \\ si & x > 0 & alors & 0 \\ sinon & x^2(3+2x) & & \end{array} \right. \\
sigmo^{+}(x) & = & \left\{ \begin{array}{llll} si & x < 0 & alors & 0 \\ si & x > 1 & alors & 1 \\ sinon & x^2(3-2x) & & \end{array} \right. \\
cloche(x) & = & 1 - (sigmo^{-}(x) + sigmo^{+}(x))
\end{array}
$$

où x est la valeur obtenue en résultat de la fonction dev décrite précédemment.

La figure 4.10 montre la fonction floue $sigmo^{-}(x)$. L'utilisation de conditions dans la fonction permet d'obtenir des valeurs cohérentes avec leur utilisation dans un réseau bayésien.

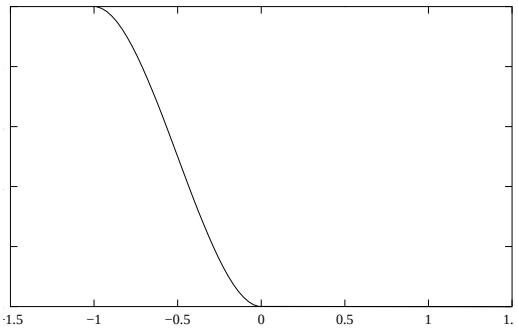


FIG. 4.10 – Opérateur flou $sigmo^{-}(x)$ utilisé pour estimer la valeur $trop\, bas$

La figure 4.11 montre la fonction floue $sigmo^{+}(x)$. Cette fonction permet de modéliser la valeur $trop\, haut$. Ici aussi il est fait usage de conditionnelle.

Enfin, la figure 4.12 donne la fonction floue intermédiaire $cloche(x)$ qui représente la valeur $normal$.

La figure 4.13 donne un aperçu des trois opérateurs flous et montre la complémentarité de ces trois opérateurs : ils couvrent l'ensemble des valeurs possibles. On notera que ces fonctions peuvent éventuellement être adaptées à la spécificité du patient. En effet, en modifiant les paramètres associés au calcul des opérateurs flous, on modifie la réponse du capteur et, par conséquent l'information insérée dans le RBD. Le principal avantage de cette approche est de pouvoir adapter la réponse des capteurs aux spécificités de chaque patient. Les paramètres modélisant les paramètres spécifiques d'un patient sont appelés *profil patient.*

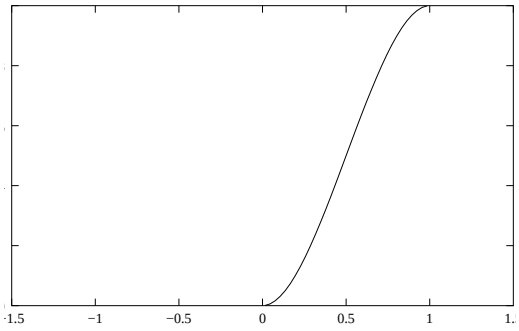


FIG. 4.11 – Opérateur flou $sigmo^{+}(x)$ utilisé pour estimer la valeur $trop\,haut$

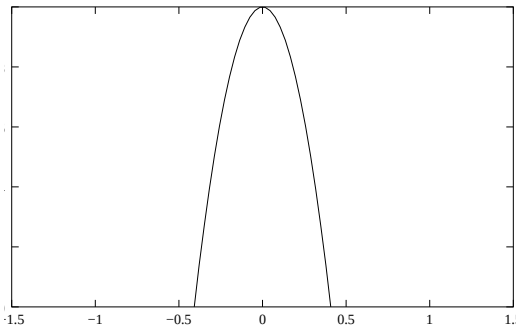


FIG. 4.12 – Opérateur flou $cloche(x)$ utilisé pour estimer la valeur $normal$

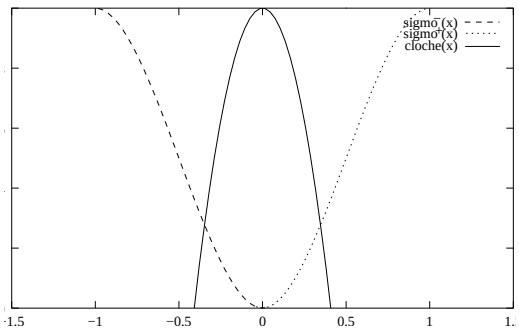


Figure 4.13: Les trois opérateurs flous. Cette représentation graphique montre la complémentarité de ces trois opérateurs permettant de couvrir complètement l'ensemble des valeurs possibles.

Nous verrons par la suite qu'il est aussi possible d'adapter les paramètres du réseau bayésien dynamique de manière à mieux refléter les spécificités d'un patient. Ceci nous a permis en particulier de n'utiliser qu'un seul jeu d'opérateurs flous et de concentrer l'ensemble des informations concernant le patient dans les paramètres du réseau bayésien. Le profil patient est donc défini par les tables de probabilités associées aux variables du réseau et non plus par des paramètres associés aux opérateurs flous, comme c'était le cas dans Diatelic^{TM} v2. Cette approche n'a pour seul intérêt que d'uniformiser le modèle. Une approche plus complète consistera à paramétriser les opérateurs flous et à utiliser en même temps un jeu de paramètres pour le RBD pour chaque patient. Néanmoins cette dernière approche est d'une plus grande complexité et nécessite de redéfinir les moyens d'intéractions dont dispose le médecin pour agir sur le système expert. Il serait en effet inconcevable de le laisser manipuler en même temps les paramètres des opérateurs flous (ce qui est simple en soit) et les paramètres du RBD (ce qui est particulièrement long et délicat).

5.5 Utilisation du réseau bayésien dynamique dans Diatelic^{TM}v3

5.5.1 Introduction

Cette section va présenter les résultats d'expériences significatives sur le comportement du réseau bayésien dynamique fait pour Diatelic^{TM}v3. Ces expériences ont pour but d'illustrer l'approche et de présenter les principales caractéristiques du fonctionnement d'un tel modèle dans Diatelic^{TM}. J'ai choisi des exemples significatifs. On trouvera en annexe A, les résultats sur l'ensemble des patients du système Diatelic^{TM}v3.

Le système doit donc permettre de :

1. estimer l'état d'hydratation du patient,
2. permettre de détecter les tendances d'aggravation de l'état du patient en utilisant un historique des diagnostics et des observations.

Les expériences ont été faites avec et sans fusion des données issues des différents capteurs. Cette approche permet de mettre en évidence la nécessité d'utiliser l'ensemble des informations disponibles, mais aussi que le raisonnement initial consistant souvent à n'utiliser que l'évolution du poids du patient conduit à ne pas détecter un certain nombre de situations graves.

Les résultats obtenus par le système sont encourageants et montrent qu'il est possible d'obtenir des tendances significatives dans l'évolution du patient. Cependant, pour contrôler l'exactitude des diagnostics fournis par le système, il va être nécessaire de lancer une nouvelle campagne de test à l'aide d'un médecin néphrologue. En effet, les données utilisées ici sont celles provenant du système 2. Pour montrer que le système 3 permet d'obtenir des diagnostics significatifs et détecte correctement les aggravations de l'état du patient, il est nécessaire de confronter ce système à de nouveaux patients. Ce type de validation est particulièrement difficile, car elle nécessite la mise en place d'une expérimentation médicale nouvelle, la disponibilité d'un ensemble significatif de patients (une moitié participant à l'expérience et l'autre moitié n'y participant mais servant de référence) et d'un ou plusieurs médecins néphrologues assurant le suivi de l'expérimentation et sa validation médicale. Une prochaine étape dans l'élaboration complète du système 3 serait donc de réunir, à l'instar du système 2, un ensemble de patients et de les équiper avec le nouveau système

(en parallèle avec l'ancien). Ce type d'expérimentation se fait sur une longue période : 2 ans pour le système 2. Une période d'un an est généralement considérée comme un minimum.

Une telle expérimentation sera cependant nécessaire pour valider pleinement le système et permettra d'obtenir des données supplémentaires primordiales : le retour du médecin. Ce retour est tout simplement l'acceptation (ou non) du diagnostic fait par le système au jour le jour. Ainsi, à partir de l'avis quotidien du médecin sur l'ensemble des patients, il sera possible d'évaluer la pertinence des diagnostics faits par le système et de modifier le système pour que celui-ci fournisse un diagnostic de plus en plus précis.

Cependant, à travers cette expérimentation, nous cherchons à obtenir un système qui sache correctement utiliser les données de poids et de tension au cours du temps, afin de fournir une bonne estimation de l'état d'hydratation. En effet, suivant les patients, les phénomènes observés ne sont pas les mêmes. Certains ont un poids qui varie vite, pour d'autres c'est la tension. Certains ont toujours une surcharge pondérale, mais sans gravité, d'autres ne peuvent tolérer le moindre écart de poids. Dans certains cas, des modifications importantes du poids pourraient laisser à penser que le patient est entré dans un état d'hyper- ou de deshydratation. Cependant, si la tension reste normale, alors il est nécessaire de la prendre en compte, de manière à modérer le diagnostic. Il en est de même avec la tension. De plus, un poids ou une tension anormale un jour doivent être confirmés les jours suivants de manière à pouvoir en déduire que le patient est dans un état grave d'hydratation. Il se peut que son poids varie brutalement un jour, mais redevienne normal le lendemain. Dans ce cas, le système ne doit pas déclarer que le patient est dans un état grave. La prise en compte d'un historique est donc une solution pour que le système puisse déterminer correctement l'état du patient.

Dans la suite de cette section, je vais présenter les détails du corpus de données et des expériences sur plusieurs exemples significatifs. L'ensemble des résultats est reporté en annexe A.

5.5.2 Corpus de données

Le corpus de données utilisé pour les expérimentations porte sur un ensemble de 15 patients observés sur une période d'expérimentation allant de 99 à 500 jours, selon les patients. Les données utilisées ont été recueillies pendant la campagne d'expérimentation du système 2 de Diatelic^{TM}, basé sur l'utilisation d'un POMDP.

Le but du système n'est pas de proposer directement une thérapie, mais de suivre l'évolution de l' état de santé. Le système est donc clairement un outil d'aide à la décision, et non pas un outil thérapeutique. Il aide le médecin à décider de ce que sera la meilleure thérapie au jour le jour. Il agit donc comme un capteur intelligent capable de percevoir une information que le médecin ne peut absolument pas consulter au jour le jour (le taux d'hydratation).

La décision thérapeutique appartient pleinement au médecin. Bien qu'il utilise le diagnostic fait par le système, il reste le seul responsable et le seul à pouvoir décider de l'évolution de la thérapie. Comme la dialyse est un traitement palliatif, le but n'est pas de guérir le patient mais de le garder dans un état de santé acceptable jusqu'à, par exemple, une greffe de rein.

Il serait faux de considérer que le système 2 ou 3 serve à maintenir le patient dans un état d'hydratation normal constant, et ceci pour plusieurs raisons :

– lorsqu'un problème d'hydratation est résolu, un autre peut survenir après, dû par exemple à un repas plus important que la moyenne, à une absorption plus importante de boissons (soirée, fête, etc...). Ainsi, il est difficile de garder le patient dans un état stable, car toutes les informations ne sont pas disponibles. Pour cela, il faudrait enregistrer une quantité d'informations considérable sur la vie quotidienne du patient.
– le médecin, en fonction des patients, peut choisir de laisser volontairement un patient dans un état d'hydratation anormal. Ceci dépend de la physiologie, et parfois de la psychologie du patient.

Il est donc normal que les données issues du système 2 contiennent des anomalies d'hydratation. Le but de Diatelic^{TM}est de suivre l'état du patient, non pas de le soigner. Il aide à la décision thérapeutique, il ne fait pas d'action thérapeutique.

Lors de l'expérimentation du système 2, la validation du médecin a été faite quotidiennement. Le médecin a suivi l'ensemble des patients et n'est intervenu sur les paramètres du système 2 que lorsqu'il estimait que le système donnait un diagnostic vraiment faux. A ce moment, il a modifié les paramètres du profil patient (dans le système 2, il s'agit essentiellement des paramètres des opérateurs flous).

Pour réaliser une expérimentation médicale complète avec le système 3, il serait alors nécessaire de saisir chaque jour l'avis du médecin sur la précision et la validité du diagnostic donné par le système à base de réseaux bayésiens. Cet avis, sous forme par exemple d'une information qualitative, pourrait éventuellement être utilisé pour adapter en continu les paramètres du réseau bayésien et ainsi fournir chaque jour un diagnostic plus précis et plus certain. Ceci constituera alors un corpus efficace pour éventuellement apprendre de nouveaux modèles de diagnostic avec les réseaux bayésiens.

La durée d'expérimentation pour chaque patient est variable et l'arrêt du recueil des données pour un patient peut correspondre à de multiples raisons. Le choix est laissé à l'initiative du médecin néphrologue. Les courbes qui seront présentées par la suite montrent des variations particulièrement brutales dans le poids et la tension des patients. Ceci est typique d'un patient sous dialyse et ne correspond en rien aux variations observées sur un patient ayant des reins pleinement fonctionnels.

Nous noterons enfin que le système 2 a permis, après 2 ans d'expérimentations, d'obtenir des résultats significatifs au niveau de l'amélioration de l'état de santé du patient et de sa sécurité. Parmi les résultats intéressants, on notera une baisse significative des coûts moyens d'hospitalisation, et pour les patients, une tension moyenne en général plus basse. Il est clair que ce type d'approche de la télémédecine est donc particulièrement profitable aux patients mais aussi à la collectivité, car elle permet une meilleure gestion des dépenses de santé liées à cette pathologie.

5.5.3 Déroulement d'une expérience

Le réseau bayésien mis en oeuvre dans le système Diatelic^{TM}v3 est utilisé sur une période de n jours d'observations. La plupart du temps, on choisira d'utiliser un $n = 15$. En effet, ceci est une durée médicalement correcte pour prendre en compte un historique d'observations. Plusieurs expériences ont été réalisées avec un historique inférieur à 15. Entre 10 et 15 jours, les différences entre les résultats sont faibles. En deçà de 10 jours et jusqu'à 3 jours d'historique, on note une dégradation de la précision du diagnostic. Pour une historique très petit (3 ou 4 jours), le système

reste en permanence dans un état de grande incertitude : les probabilités que le patient soit dans un état hyperhydraté, normal ou déshydraté oscillent en permanence autour de 30%. Il devient donc assez difficile de prendre une décision puisque le système *n'arrive pas à se décider*. La longueur de l'historique est donc importante.

Le réseau est donc *déroulé* sur n jours et les observations sont insérées dans le réseau sous forme *d'évidences vraisemblables* (likelihood findings) et ensuite propagées à l'ensemble des variables du réseau jusqu'à atteindre un nouvel état d'équilibre. Ceci est possible, car avant propagation, le réseau était dans un état d'équilibre. L'insertion de nouvelles informations (les évidences) brise cet équilibre et grâce à l'algorithme JLO présenté dans le chapitre 3 un nouvel état d'équilibre est atteint et permet ainsi d'obtenir la probabilité de l'ensemble des variables sachant les observations. Le dernier pas de temps correspond aussi au dernier jour d'observation et par conséquent au jour courant. Ainsi, l'état d'hydratation du patient sera exprimé dans la variable $Hydratation$ du dernier pas de temps (le plus à droite dans le graphe du réseau bayésien dynamique) du réseau bayésien. Cette forme de réseau bayésien dynamique est inspiré d'un modèle utilisé en reconnaissance de la parole sur des séquences sonores de longueurs fixes comme dans [Smyth et al., 1996],[Zweig and Russell, 1997], [Murphy, 1999], [Daoudi et al., 2000] ou encore [Murphy, 2002].

La figure 4.14 montre un réseau Diatelic^{TM}où $n = 3$. On notera en particulier que le premier pas de temps ne contient pas d'observation et est utilisé comme état initial du système modélisé. Il s'agit de l'instant 0. Les observations sont insérées dans les pas de temps 1, 2 et 3. Comme il s'agit d'un processus de monitoring, le dernier pas de temps reçoit aussi des observations sous forme d'évidence. L'état du patient correspond, dans cet exemple, à la variable $Hydratation3$. Les pas de temps 1 et 2 reçoivent donc les observations des ante-pénultième et avant-dernier jour.

Pour traiter une séquence d'observations complète d'un patient donné, le réseau est ré-initialisé avec les paramètres de base (profil patient typique) et une nouvelle séquence d'observation est introduite. Ainsi, si l'expérience dure *N* jours, alors l'algorithme suivant sera utilisé pour calculer le diagnostic complet :

Algorithm 1 Diagnostic d'un patient sur N jours

Pour $i = n$ **à** N **faire**
début
Insérer les observations de $i - n$ à i (où i correspond au jour courant)
Propager en utilisant l'algorithme JLO
Marginaliser la variable $Hydratationn$ pour obtenir $P(Hydrationn|o_{i-n} \dots o_i)$ (où $o_{i-n} \dots o_i$ sont les observations)
fin

Chaque jour, un diagnostic est donc extrait. Si le taux d'hydratation dépasse un seuil défini à l'avance, alors une alerte est déclenchée pour prévenir le médecin d'une aggravation de l'état d'hydratation du patient. Deux valeurs sont observées : le taux d'hydratation est trop bas, et le taux d'hydratation est trop haut. Le même travail est effectué sur le poids idéal, de manière à aider le médecin dans le choix d'un nouveau poids idéal si celui-ci s'avérait inefficace pour le patient (poids idéal trop haut ou trop bas).

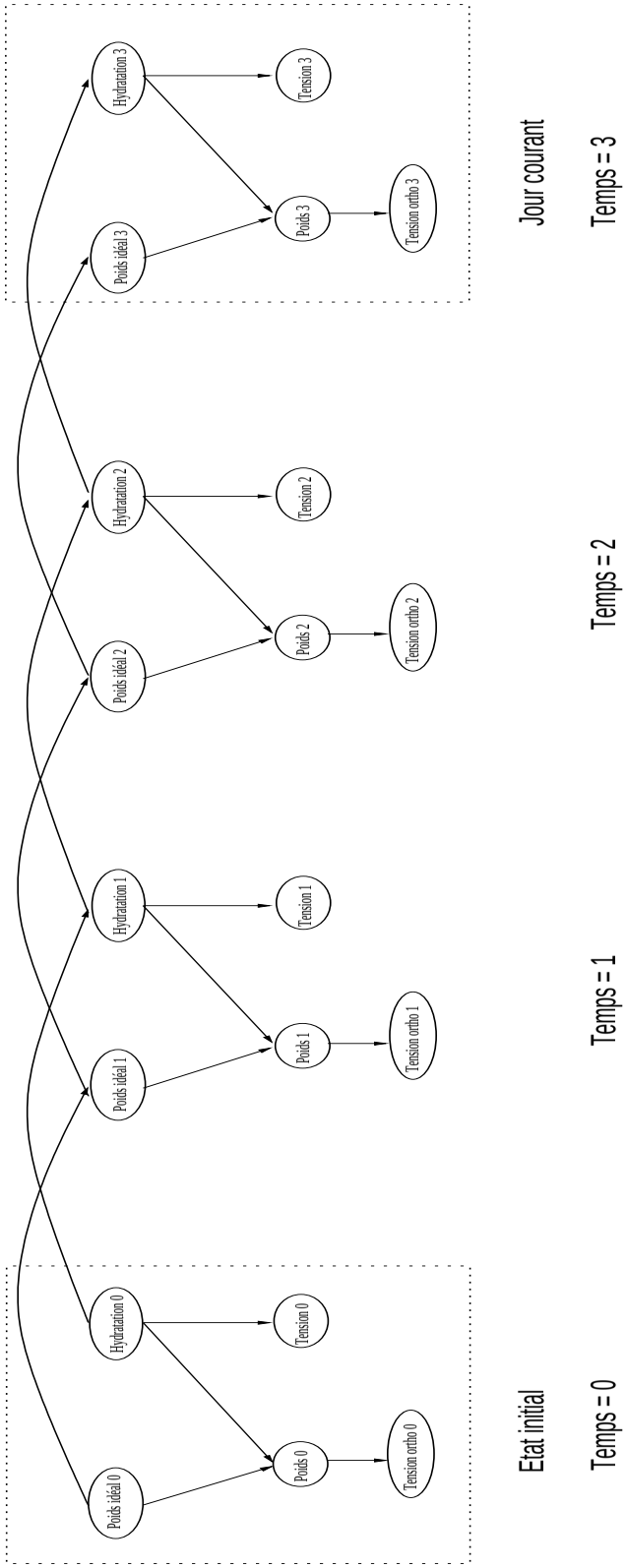


FIG. 4.14 – Réseau bayésien Diatelic^{TM}v3 déroulé sur 3 jours

On notera enfin que les paramètres du réseau bayésien dynamique ont été déterminés à partir de l'expertise médicale. En effet, étant donné la difficulté *de connaître la vérité* dans le cadre du monitoring médical, il aurait été impossible d'apprendre automatiquement les paramètres du réseau. Cependant, l'expertise médicale et les paramètres utilisés dans le POMDP du système 2, nous ont permis de déduire des paramètres appropriés pour le RBD.

5.6 Expérimentations

Pour montrer les capacités et le comportement d'un réseau bayésien dynamique en fusion de données, je vais maintenant présenter une série d'expérimentations simples mais significatives faites sur des patients réels. Ces patients subissent une dialyse péritonéale. Ils sont tous localisés dans la région Lorraine et dépendent du service de traitement des dialysés de l'ALTIR[3] au Centre Hospitalier Universitaire de Nancy-Brabois. Ces patients font donc partie de l'expérimentation en grandeur réelle du système Diatelic^{TM}v2.

5.6.1 Diagnostic simple

Cette première expérience est basée sur un patient type qui présente au cours du temps un certain nombre de problèmes d'hydratation. Ce patient a été observé pendant une période de 110 jours. La figure 4.15 montre l'évolution du patient n°1010. Les pics de poids sont souvent synonymes d'hyperhydratation. Les chutes de poids indiquent plutôt un état de déshydratation. Ici, le poids idéal du patient n'a pas été modifié par le médecin pendant toute la durée de l'expérience. Les pics de poids sont souvent indicateurs d'un état d'hyperhydratation et les chutes de poids, d'un état de déshydratation. La fusion avec la tension et avec l'historique des observations et des diagnostics partiels permet de confirmer ou d'infirmer l'état d'hydratation réel du patient.

La figure 4.16 représente l'évolution de la tension de ce même patient. La courbe en pointillée représente la moyenne mobile de cette tension calculée typiquement sur les n derniers jours d'observations. Les chutes et les pics de tension représentent aussi de précieuses indications sur l'état d'hydratation du patient. D'une façon empirique, il apparaît clairement que l'utilisation de la tension seule n'est pas suffisante pour détecter correctement les désordres d'hydratation. En effet, la variabilité de la tension est grande pour un patient sous dialyse, et cette information ne peut être qu'un complément au poids. C'est pourquoi il est nécessaire de fusionner cette information aux autres informations (poids, historiques des observations et des diagnostics, etc...). Dans le cas contraire, l'utilisation du poids seul ou de la tension seule provoquerait un nombre d'alarmes particulièrement important et sans aucune pertinence vis-à-vis de la pathologie.

Le diagnostic est effectué sur l'ensemble des valeurs et le résultat est observé sur la figure 4.17. La courbe en trait plein représente l'évolution de l'état d'hyperhydratation. La courbe en pointillé représente l'évolution de l'état de déshydratation. L'état normal n'est pas tracé sur cette courbe pour des raisons de clarté, mais il représente le complémentaire des deux courbes puisque la somme des probabilités pour ces trois états est toujours égale à un. La figure 4.18 représente la même évolution du taux d'hydratation du patient n°1010 mais après un filtrage simple par seuil. Le seuil d'alerte a été fixé à 55% de taux d'hyper ou de déshydratation. En deçà de ce taux, le patient est considéré

[3] Association Lorraine de Traitement de l'Insuffi sance Rénale

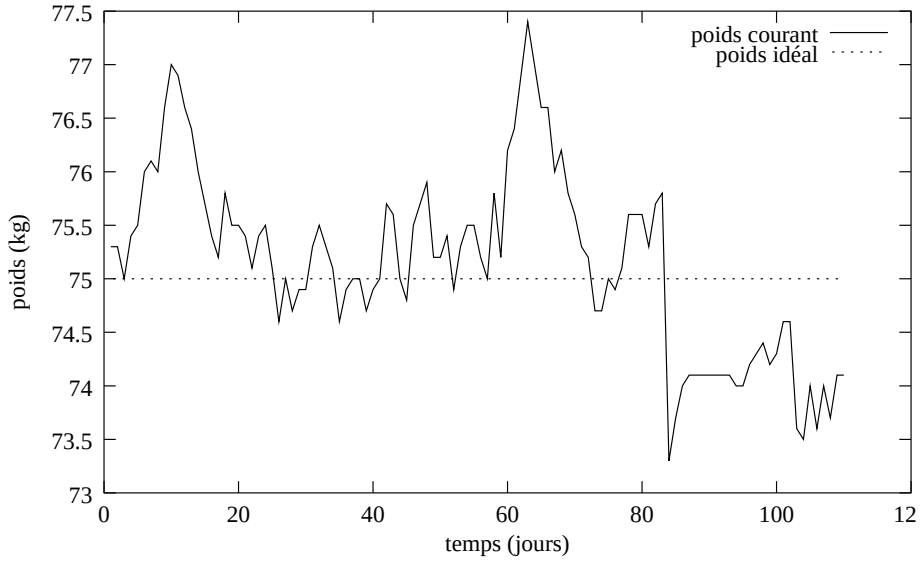


FIG. 4.15 – Évolution du poids d'un patient.

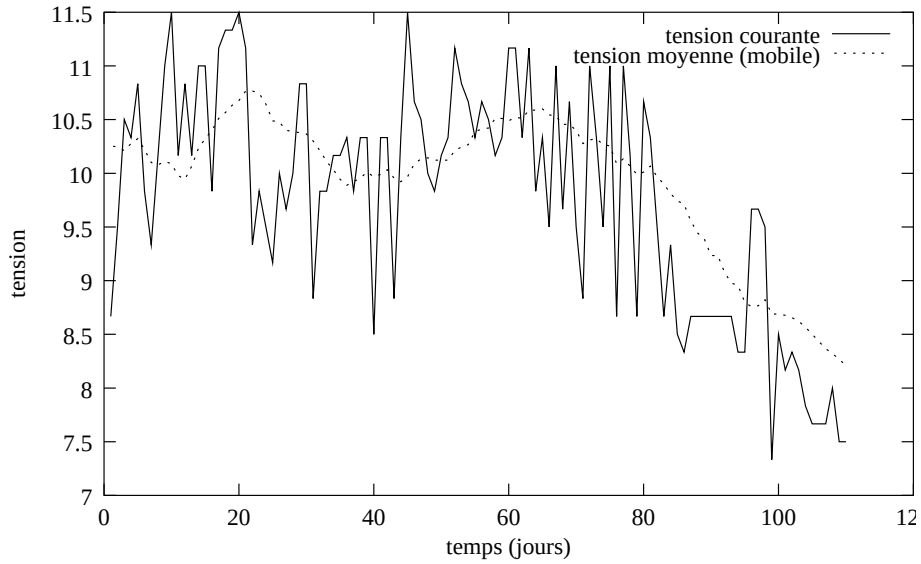


FIG. 4.16 – Évolution de la tension d'un patient. La courbe en pointillés représente la moyenne mobile de la tension calculée sur n jours d'observations.

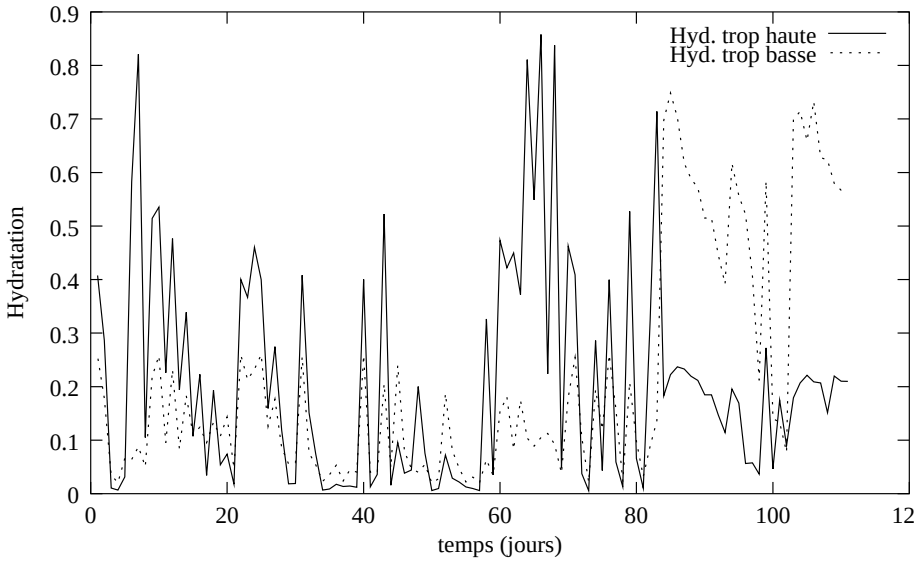


FIG. 4.17 – Évolution de l'état d'hydratation du patient n°1010

comme étant dans un état normal. La valeur 1 représente une hyperhydratation et la valeur -1 une déshydratation. Ceci nous permet de voir apparaître plus clairement les périodes d'aggravation de l'état du patient. Ce seuil de 55% est arbitraire et permet de ne prendre en compte que les cas où l'état d'hydratation anormal est bien avéré. La somme des probabilités que le patient soit dans un état d'hydratation anormale inverse ou dans l'état normal est donc de 45%. La courbe de la figure 4.18 est directement issue des courbes de la figure 4.17.

5.6.2 Deuxième diagnostic

Ce patient n°1011 est intéressant pour illustrer un problème courant dans ce type d'application en télémédecine où le patient est responsable de la saisie et de l'envoi des données. En effet, les figures 4.19 (poids) et 4.20 (tension) montrent que ce patient n'a pas saisi de données durant une longue période. L'expérience a duré 184 jours pour ce patient, mais durant 53 jours, il n'a pas saisi de données. Dans ce cas, le système n'a d'autre choix que d'estimer que le patient est dans un état normal, ce qui, quelquepart, est rassurant sur la bonne tenue du système. En effet, les données recueillies sont normalisées par le système 2 pendant 2 jours, mais après, comme le patient ne rentre plus de données, nous n'avons plus d'informations pour faire le diagnostic. On considère par défaut qu'il est normal : les paramètres du RBD décrivent en effet un patient normal. Bien qu'un réseau bayésien dynamique puisse prédire l'état futur d'un phénomène, on assiste ici à une situation où la quantité de données manquantes est trop importante pour que le réseau puisse estimer *a priori* l'état correct du patient. Cependant, lorsque le patient recommence à saisir ses données, le système est apte à fournir un diagnostic avec un temps de latence particulièrement réduit. En effet, les figures 4.21 et 4.22 montrent l'évolution de l'état d'hydratation du patient et l'on voit clairement le phénomène induit par le manque de données. Au bout de quelques jours,

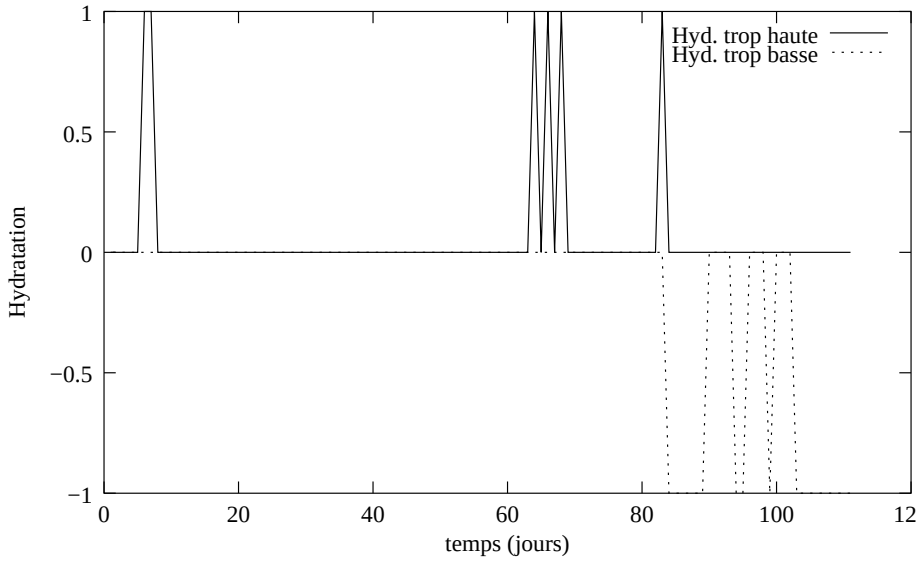


FIG. 4.18 – Évolution de l'état d'hydratation du patient n°1010. Ici, seules les situations estimées comme réellement grave par le nouveau système Diatelic^{TM}v3 ont été reportées.

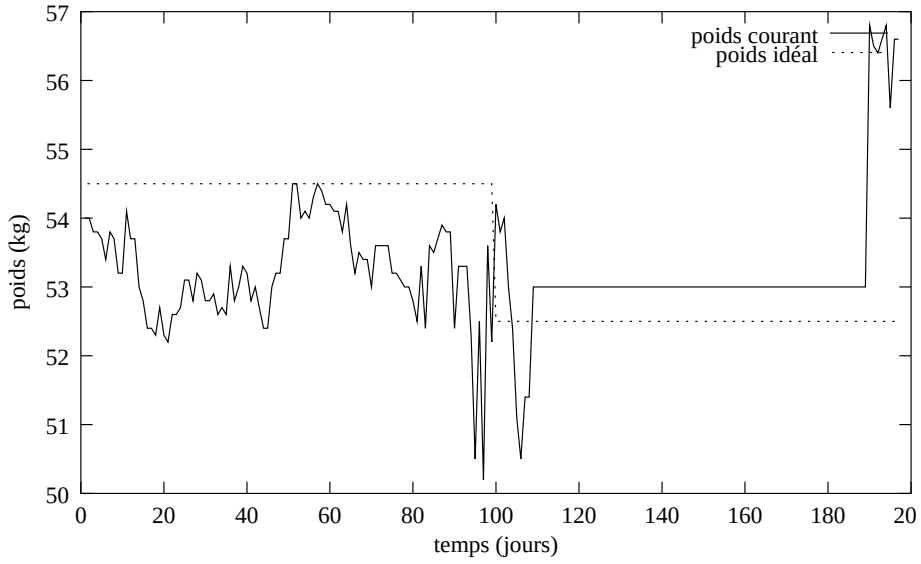


FIG. 4.19 – Évolution du poids du patient n°1011

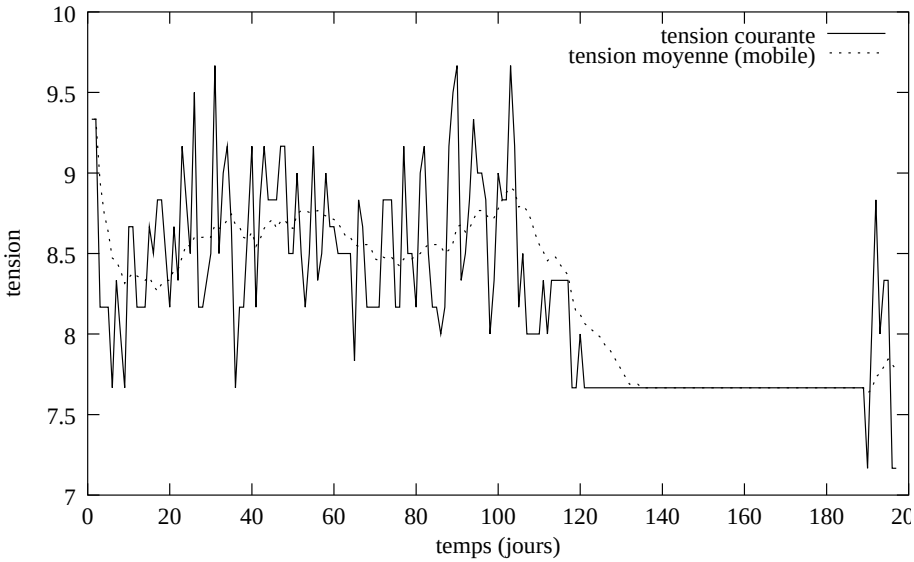


FIG. 4.20 – Évolution de la tension du patient n°1011

le système n'est plus en mesure de donner un résultat fiable. Et inversement, lorsque de nouvelles données sont disponibles, le système est capable de fournir très rapidement un diagnostic fiable.

5.6.3 Intérêt de la fusion dans Diatelic^{TM}v3

Les deux premiers exemples que je viens de présenter illustrent le fonctionnement normal du système Diatelic^{TM}v3 à base de réseaux bayésiens dynamiques. Dans des situations où les données manquent ou sont imparfaites (erreur ou oubli de saisie du patient), le système fournira un diagnostic en utilisant les données dont il dispose. Si une valeur est particulièrement fausse, elle ne sera pas confirmée par les autres valeurs et son influence sur le diagnostic final sera amoindrie. Si une valeur manque, le système donnera un diagnostic inutilisable, mais il préviendra en plus le médecin que le patient n'a pas saisi de données.

La valeur manquante ne sera pas insérée dans le réseau. Le corpus utilisé dans ces expériences, issus du système Diatelic^{TM}v2, a cependant une caractéristique particulière : lorsqu'un patient n'a pas saisit ses données quotidiennes, le système 2 fait une moyenne des jours précédents et *saisit* lui-même les données. Ceci est fait pour une période de 2 jours, en espérant qu'il s'agit d'un oubli passager du patient. Si le patient ne saisit toujours par de données, alors le médecin est prévenu et le système 2 arrête de faire un diagnostic (qui à ce moment n'a forcément plus aucun sens).

Avec un RBD, la situation est différente. Si le patient ne saisit pas de données, il n'est pas nécessaire de créer des données artificielles. Il est possible d'inférer sans avoir d'observation. Bien sûr, à l'instar du système 2, cette situation n'est supportable que durant quelques jours. Au-delà, le système ne peut plus raisonnablement donner de diagnostic. Dans le cas du RBD, cette période peut se prolonger théoriquement jusqu'à 14 jours (en considérant un historique d'une taille $n = 15$). Mais je pense qu'au delà de 3 à 4 jours, le diagnostic fait sans observation devient aberrant et donc

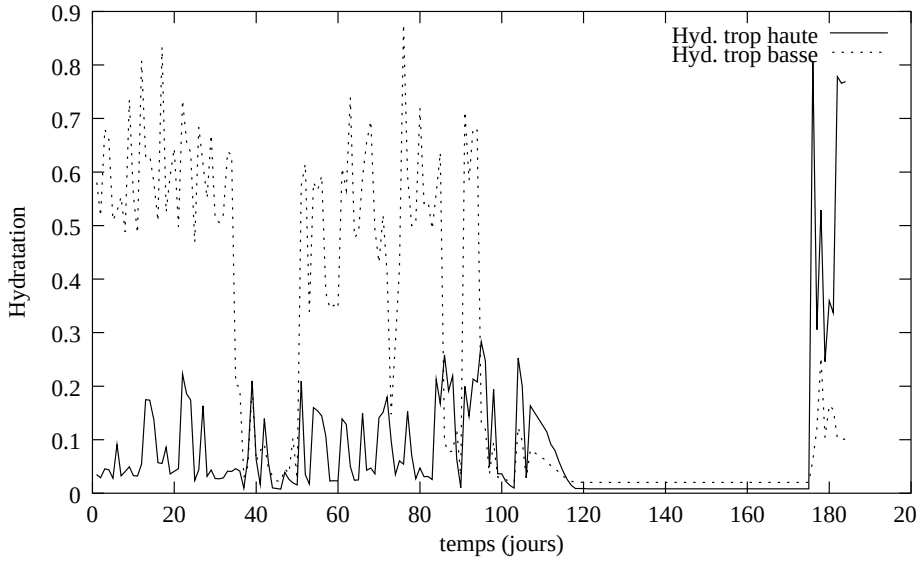


FIG. 4.21 – Évolution de l'état d'hydratation du patient n°1011.

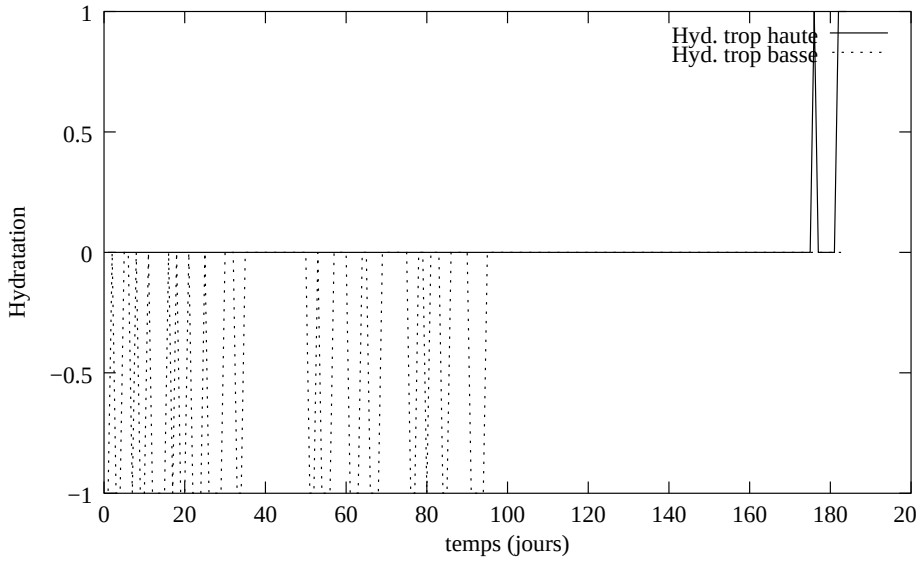


FIG. 4.22 – Évolution de l'état d'hydratation du patient n°1011. Ici, seules les situations définies comme réellement grave par le nouveau système Diatelic^{TM}v3 ont été reportées.

inutilisable. Néanmoins, l'approche à base de RBD est *robuste* vis-à-vis de manque de données très ponctuelles.

Ce dernier exemple est très important, car il montre la nécessité de fusionner l'ensemble des données dont on dispose pour avoir un diagnostic fiable. Dans cette expérimentation, un seul type de données a été utilisé. Dans le premier cas, le poids seul est utilisé, dans le deuxième cas, la tension seule est utilisée. Cette expérience est intéressante à plusieurs titres :

– elle a fait apparaître une grande instabilité dans le diagnostic, lorsque les données ne sont pas fusionnées,
– elle démontre une plus grande incertitude sur le diagnostic. En effet, le système n'arrive plus à fournir de réponse avec un fort taux de confiance (c'est-à-dire une probabilité élevée),
– certains phénomènes tels qu'une amélioration suivie d'une rechute disparaissent et ne sont pas correctement reporté par le système.

Patients	Poids seul	Tension seule
1004	62%	71%
1006	95%	92%
1007	81%	90%
1008	38%	81%
1010	82%	80%
1011	81%	76%
1012	37%	86%
1014	77%	89%
1015	48%	86%
1016	75%	90%
1017	83%	87%
1019	67%	76%
1020	61%	85%
1023	44%	63%
1024	40%	61%

TAB. 4.1 – Pourcentage de bons diagnostics quand un seul type de données à la fois est utilisé

En général, si le diagnostic fourni par le système utilisant l'ensemble des données est considéré comme juste (donc en supposant que la fusion des données donne 100% de bons résultats), en général on observe une nette diminution de la précision du diagnostic lorsqu'un seul type d'information est utilisé. Le tableau 4.1 récapitule quelques résultats. Le diagnostic fourni par le système a été filtré avec la fonction à simple seuil utilisée dans les sections précédentes pour générer les alertes. Le corpus a été ensuite traité en utilisant toutes les données disponibles, puis seulement le poids, et enfin seulement la tension. Les valeurs indiquées dans le tableau 4.1 correspondent au pourcentage d'alertes correctes données par le système lorsque seul le poids ou la tension sont utilisés. En résumé, l'utilisation du poids seul permet au système de fournir en moyenne 64.7% de bonnes réponses. Si seule la tension est utilisée, alors on obtient $80,8\%$ de bonnes réponses. On considère ici que le taux de 100% est atteint lorsque l'ensemble des données sont utilisées. Les

taux de bonnes réponses montrent donc que tout n'est pas découvert lorsque l'on prive le système de données. La précision et la complétude du système est donc amoindrie dans ce cas.

Ces résultats confirment la nécessité de fusionner les données et montrent clairement l'insuffisance du système lorsque les données sont manquantes.

Pour terminer cette section, voici les résultats obtenus sur le patient n°1010 en utilisant l'ensemble des informations, puis seulement le poids ou la tension. La figure 4.23 montre l'évolution de l'état d'hyperhydratation en n'utilisant que le poids comme information. La figure 4.24 montre des résultats similaires, mais au cours de cette expérience, seule la tension a été utilisée. Ce qui est remarquable dans ces deux expériences est le fait que le diagnostic obtenu avec des informations partielles reflètent à peu près bien le diagnostic réalisé avec l'ensemble des informations. Cependant, pour ce patient, comme pour l'ensemble des patients du corpus, on assiste à un phénomène d'incertitude. En effet, le système fournit des probabilités beaucoup plus faibles lorsque l'on n'utilise pas l'ensemble des données disponibles. Ceci traduit le fait que l'utilisation de toutes les données permet bien sûr de renforcer et de confirmer un diagnostic. Sans cela, le système reste relativement incertain sur l'état réel du patient.

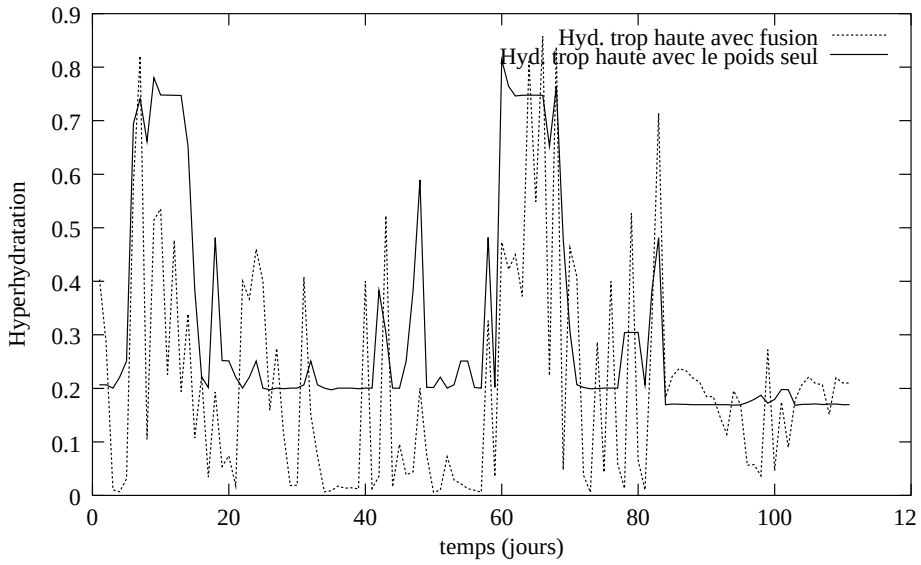


FIG. 4.23 – Évolution de l'état d'hyperhydratation du patient n°1010 avec et sans fusion des données hétérogènes. Ici seul le poids est utilisé.

5.6.4 Synthèse des résultats des expérimentations

L'utilisation de données partielles ne permet pas d'obtenir un diagnostic sûr et laisse le système dans un état d'incertitude permanent. De plus, l'utilisation de données partielles ne permet pas la découverte de l'ensemble des cas d'aggravation de l'état d'hydratation du patient. Ainsi, un diagnostic utilisant seulement le poids ou seulement la tension risque de compromettre la robustesse du système. Il est donc nécessaire de fusionner l'ensemble des données.

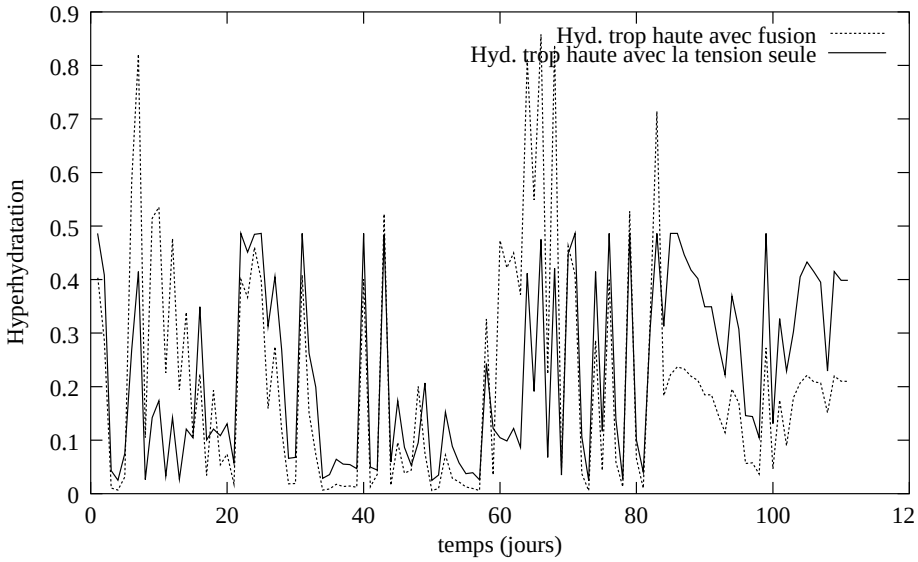


FIG. 4.24 – Évolution de l'état d'hyperhydratation du patient n°1010 avec et sans fusion des données hétérogènes. Ici seule la tension est utilisée.

Les expérimentations ont néanmoins montré un bon comportement du système lorsqu'on utilise l'ensemble des données : prise en compte de l'évolution du patient, et nécessité d'utiliser le poids et la tension en même temps en modérant les variations trop abruptes d'un capteur quand l'autre capteur a un comportement plus stable.

Dans le cas du poids seul, 64% des situations graves sont découvertes. Dans le cas de la tension seule, 80% des situations graves sont exhibées par le système. Il apparaît donc que ce modèle à base de réseaux bayésiens dynamiques est suffisamment robuste pour ce type de diagnostic médical.

6 Conclusion et perspectives

Ce chapitre a présenté le problème de la modélisation des systèmes dynamiques par des méthodes stochastiques, en mettant particulièrement l'accent sur le formalisme des réseaux bayésiens dynamiques. L'approche qui y est décrite démontre l'utilisation conjointe d'opérateurs flous et d'un réseau bayésien afin de progressivement et uniformément intégrer l'ensemble des données hétérogènes et incertaines que l'on recueille au cours du temps. La nécessité de fusionner l'ensemble des données a été exhibée et il est apparu que le manque de données emmène le système à fournir des résultats plus incertains que lorsqu'on utilise l'ensemble des données disponibles.

Il sera cependant nécessaire pour valider ce modèle de mettre en place une expérimentation médicale complète et de confronter le système à base de réseaux bayésiens à une situation réelle. Ceci devrait permettre en particulier de mesurer l'impact de cette approche sur la santé des patients et de proposer une aide à la décision pour réguler l'état du patient. En effet, il est possible dans ce

modèle d'estimer les valeurs des capteurs sachant le diagnostic que l'on voudrait obtenir, donc de faire le raisonnement inverse de celui que j'ai présenté dans ce chapitre.

Les perspectives envisageables à partir de ces travaux sont multiples :

– comment adapter les paramètres du réseau au cours du temps en ayant un référentiel seulement qualitatif, comme un conseil ou un souhait exprimé par le médecin, dans le cas du problème Diatelic^{TM} ?
– comment adapter la structure du réseau (le graphe) de manière efficace, sans avoir à recalculer l'ensemble des diagnostics ou des modèles intermédiaires à chaque fois que de nouvelles données sont disponibles, tout en assurant la validité du modèle, surtout dans des applications critiques telles que le diagnostic médical ?

A l'heure actuelle, il existe des travaux portant sur l'adaptation au cours du temps d'un modèle à base de RBD. Cependant, aucun résultat significatif n'est apparu, dû à la grande complexité du problème. Dans le cas du problème posé par Diatelic^{TM}, un autre problème s'ajoute encore : il n'existe pas de modèle de référence pour confronter les résultats fournis par le système (pas de corpus de diagnostic déjà validé par un médecin). Ce problème est essentiellement dû au fait qu'il est particulièrement difficile d'estimer l'état d'hydratation d'un patient de façon quotidienne. Les méthodes sont simplement empiriques ou nécessitent au contraire un appareillage très lourd, incompatible avec les contraintes de la télémédecine. Le système de télémédecine ne doit pas (trop) déranger le patient dans sa vie quotidienne.

L'adaptation dynamique de modèles à partir d'informations symboliques semble donc être une voie possible pour l'amélioration de la qualité de tels systèmes de diagnostic. Le médecin pourra en effet donner un avis non quantitatif sur le diagnostic donné par le système. Mais cet avis qualitatif servira, dans ce contexte, à ré-estimer les paramètres du RBD de manière à fournir, à partir du lendemain, un diagnostic plus correct et plus en relation avec la façon de penser du médecin.

Conclusion et perspectives

Le travail présenté dans cette thèse se situe dans le cadre de la fusion de données appliquée à la télémédecine. Il s'est appuyé sur l'utilisation de modèles à base de réseaux bayésiens dynamiques pour la résolution du problème du *monitoring* et du diagnostic en continu.

Les travaux présentés ont eu pour but de proposer des solutions au problème de la fusion de données dans des environnements dynamiques et incertains afin de réaliser un diagnostic et sa mise à jour au cours du temps. Une application dans le cadre du domaine médical en télémédecine a été conçue et a servi de base expérimentale aux modèles présentés dans cette thèse. De plus, une étude de la télémédecine a été présentée car elle a servi de cadre applicatif à l'ensemble de la thèse.

Le diagnostic médical, tel que nous l'avons perçu, a été transformé en un problème de monitoring de l'état physiologique d'un patient et a été résolu en utilisant plusieurs sources de données hétérogènes, incertaines et bruitées que nous avons fusionnées au sein d'un réseau bayésien dynamique. Ce type de modèle a montré son efficacité et a permis, dans notre cas, de pouvoir effectuer un diagnostic et une mise à jour de ce diagnostic, chaque fois que de nouvelles observations ont été disponibles.

Parmi les résultats les plus intéressants, on notera en particulier la proposition d'une approche originale de la fusion de données dans les systèmes dynamiques. Cette approche propose une classification des sources de données et des capteurs utilisés dans un processus de fusion de données. Elle s'intéresse de plus à la notion de gain qualifié dans un processus de fusion de données. Cette approche a été publiée dans la conférence internationale Fusion, en juillet 2002. La notion de gain qualifié apporte une sémantique originale à la notion plus générale d'utilité de la fusion de données.

Le cadre générique proposé a alors été appliqué à la modélisation du système Diatelic^{TM}, dans sa version 3. Le système a ensuite été expérimenté sur un corpus de patients subissant une dialyse péritonéale à domicile. Les résultats montrent que le système est capable de fournir une estimation quotidienne et de former des tendances sur l'évolution de l'état du patient. Ceci permet en partie de détecter les débuts d'aggravation de l'état du patient avant qu'il ne soit réellement dans une situation grave. En outre, les expériences mettent en avant la nécessité de fusionner un maximum d'informations, afin de fournir un diagnostic le plus précis possible avec la plus grande certitude possible. Enfin, il est clairement apparu qu'une expérimentation à plus grande échelle, dans une situation réelle, sera nécessaire afin de valider le modèle proposé. Une telle expérimentation doit avoir pour but la validation du modèle, mais aussi permettre de mesurer l'impact d'un tel système sur les patients. Ceci ouvre en particulier la perspective de la recommandation d'actions, où les ca-

pacités de prédiction des réseaux bayésiens dynamiques devraient fournir au médecin la meilleure action thérapeutique pour réguler l'état de santé du patient.

Parmi les perspectives les plus intéressantes, on notera le problème de l'adaptation en continu des paramètres du réseau. Cette adaptation passerait pas l'utilisation d'informations qualitatives multiples, fournies en général par le médecin, pour adapter les paramètres du réseau dynamique et permettre ainsi la production d'un diagnostic plus fiable. Ce type de problème n'a actuellement pas encore de solution et reste largement ouvert. La définition d'une ontologie de qualification d'un diagnostic serait donc nécessaire, et permettrait ainsi d'aider à inférer les modifications des paramètres nécessaires afin de mieux cibler les spécificités de chaque patient. Le modèle proposé actuellement utilise des paramètres issus de l'expertise médicale. Ce modèle, bien que fonctionnant correctement, ne peut à long terme donner des résultats satisfaisants, car les connaissances médicales ainsi codées sont génériques et donc mal adaptées aux cas particuliers.

La deuxième perspective de ce travail concerne la mise en place de modèles dynamiques et hiérarchiques dans lesquels plusieurs échelles de temps seront considérées en même temps. Ceci permettrait de raisonner à plusieurs niveaux d'abstraction : évolution sur le jour, tendance de la semaine, comportement du patient sur le mois, etc... Le but de ce type de modèle serait d'étudier à long terme la tolérance et l'acception de la thérapie par le patient et d'évaluer l'évolution à long terme de la thérapie. Même si le problème soulevé par Diatelic^{TM}ne nécessite pas forcément un modèle d'une telle complexité, ce type d'approche permettrait de traiter d'autres problèmes issus de la télémédecine, comme la surveillance de personnes âgées à domicile ou encore la surveillance de patients souffrant d'une maladie cardiaque. Dans ce type de problèmes, les échelles de temps à prendre en considération varient d'une fraction de seconde à plusieurs heures, voire plusieurs jours dans le cas des personnes âgées. Il est donc nécessaire de pouvoir raisonner sur plusieurs échelles de temps en même temps et d'utiliser au maximum les connaissances de chaque échelle de temps pour améliorer les performances et la qualité du raisonnement effectué sur les autres échelles de temps. Ce type de modèle conduira vraisemblablement à la proposition de réseaux bayésiens dynamiques hiérarchisés.

Annexe A

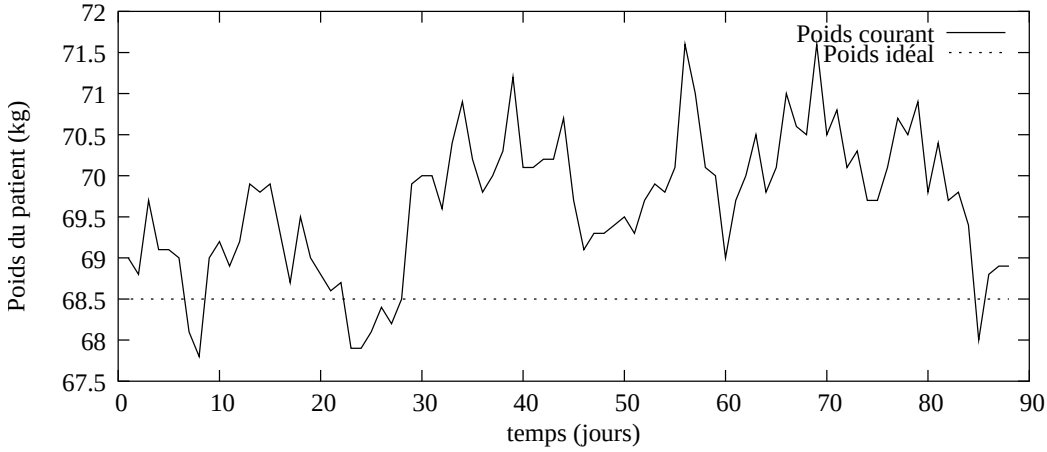
Diagnostics sur l’ensemble du corpus

Cette annexe présente l’ensemble des diagnostics effectués sur l’ensemble des patients du corpus. Pour chaque patient, quatre courbes sont présentées :

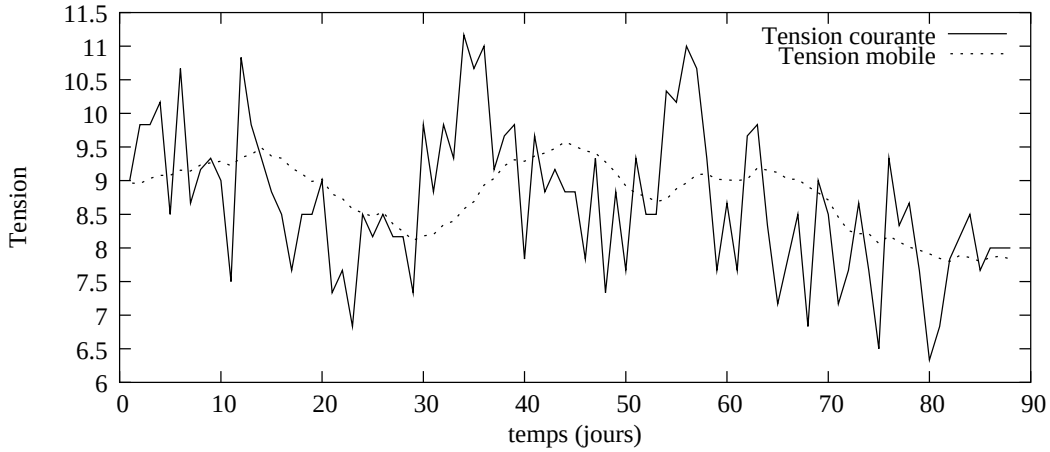
– l’évolution du poids du patient avec son poids idéal en pointillés,
– l’évolution de la tension du patient avec sa tension moyenne calculée sur 15 jours en pointillés,
– l’évolution de l’estimation de son taux d’hydratation avec en trait plein l’estimation que le taux est trop élevé et en pointillé l’estimation que le taux est trop bas.
– l’occurence des états de deshydratation (ligne en pointillés en bas) et d’hyperhydratation (ligne en trait plein en haut).

Patient n°1004

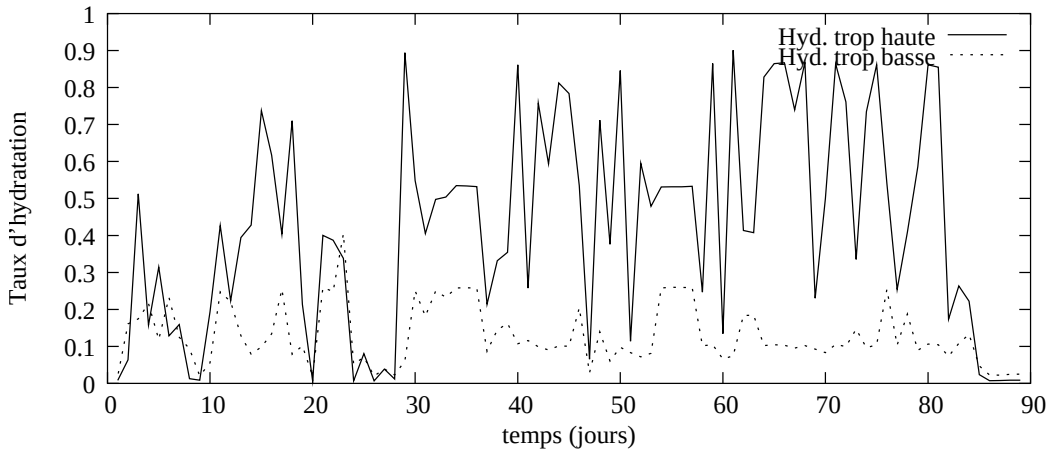
Poids du patient



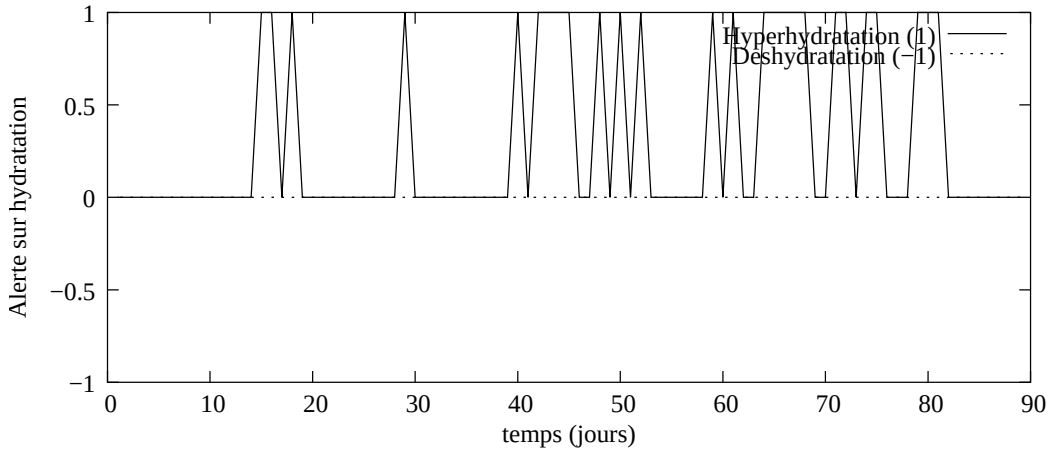
Tension du patient



Taux d’hydratation du patient



Alerte sur hydratation

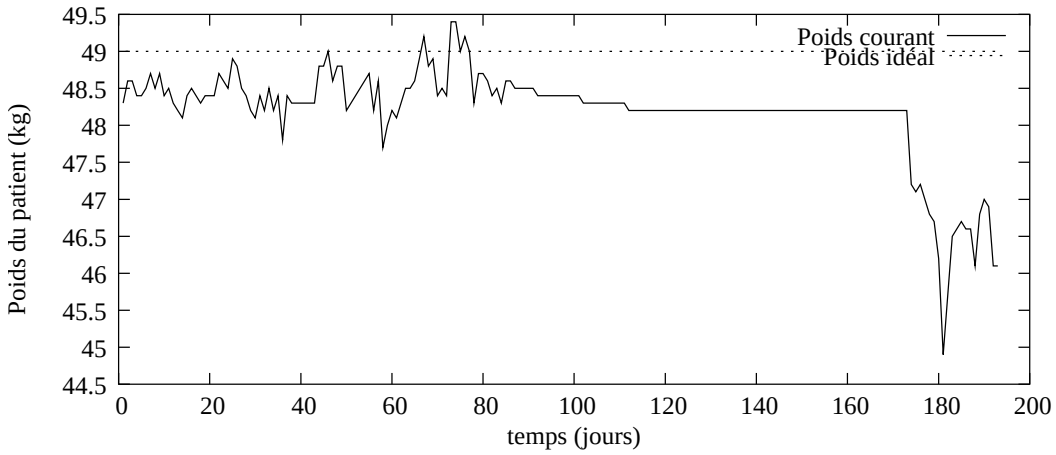


Note

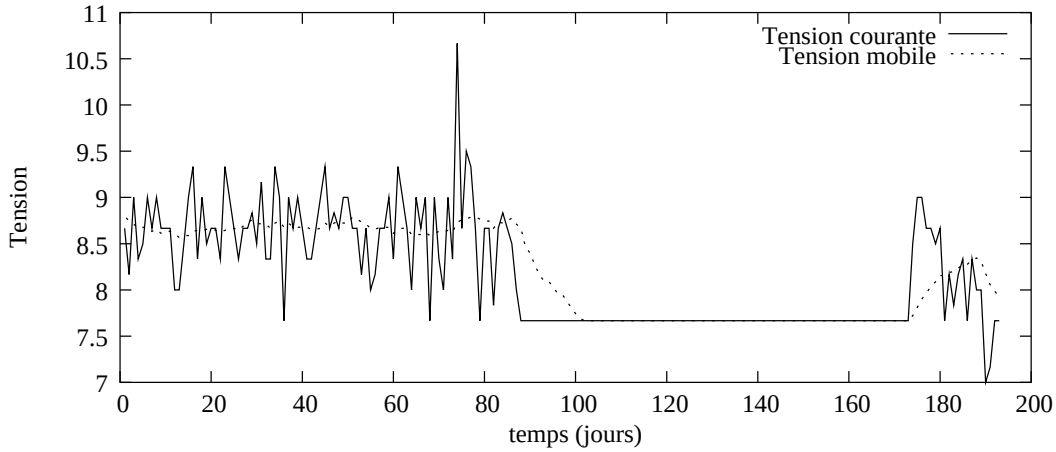
Ce patient a, pendant les trois quart de l’expérience, un poids trop élevé par rapport à son poids sec. Il s’agit d’un patient pour lequel le médecin a choisi de ne pas ré-évaluer le poids idéal pour forcer le patient à maigrir. Cette situation entraîne le système à déclarer pendant toute cette durée le patient comme étant hyperhydraté. Or la courbe de la tension présente une certaine régularité autour de la tension moyenne qui ne varie que peu. En conséquence le système hésite sur l’état réel du patient et propose une solution avec une grande variabilité, ce qui corrobore les pics de la courbe d’hyperhydratation (trait plein).

Patient n°1006

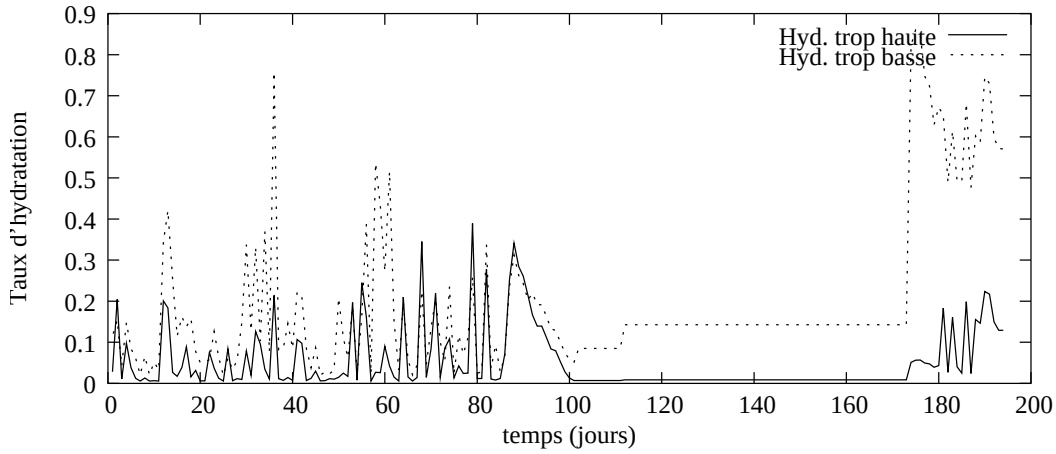
Poids du patient



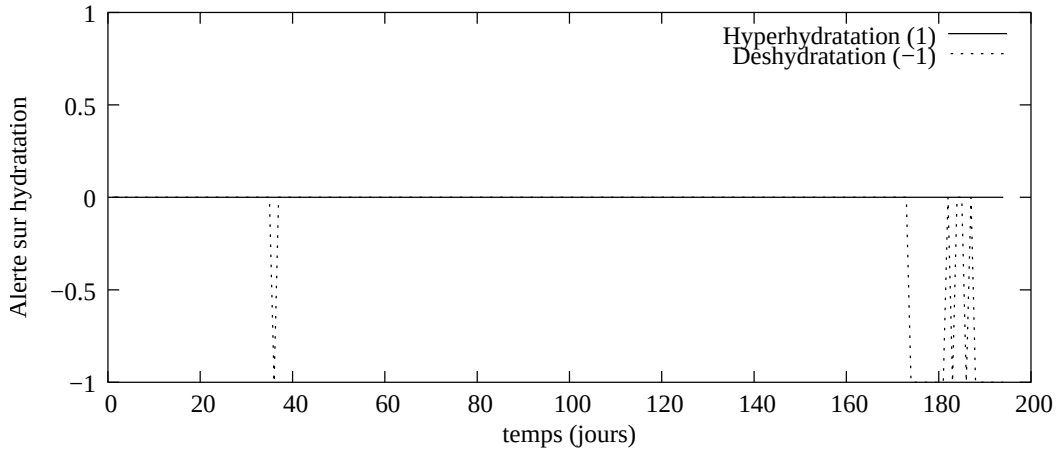
Tension du patient



Taux d’hydratation du patient



Alerte sur hydratation

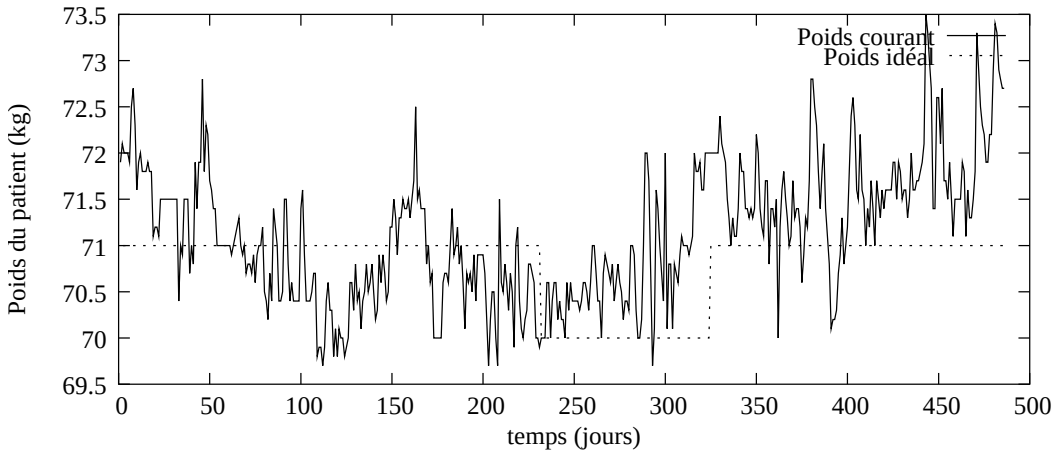


Note

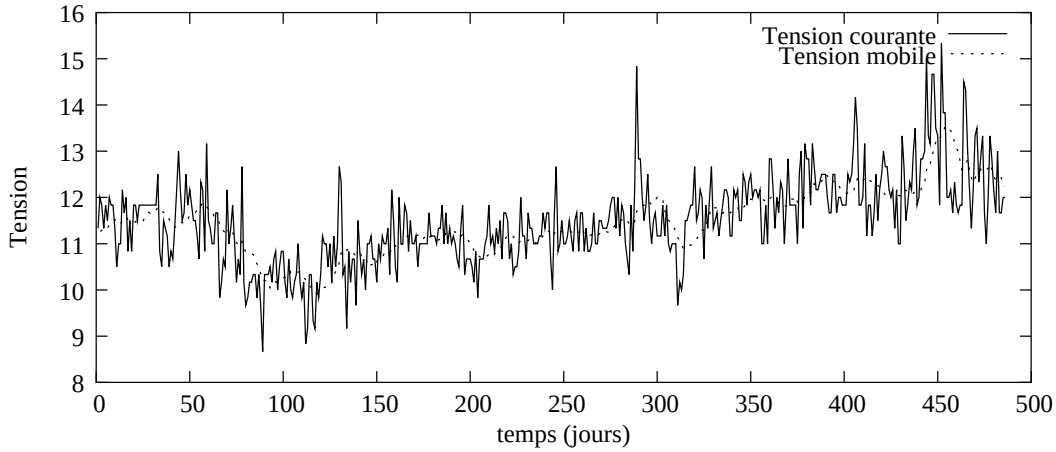
Ce patient a oublié de saisir les données pendant une assez longue période. Le système 2, par défaut, prolonge la dernière valeur durant 2 jours. Dans ce type de situation, le système à base de réseaux bayésiens dynamiques continue à fonctionner, mais donne un résultat qui est forcément faux par rapport à la réalité : aucune donnée n’est disponible, donc le système ne peut pas fournir de diagnostic. Le médecin sera en outre prévenu de cette situation.

Patient n°1007

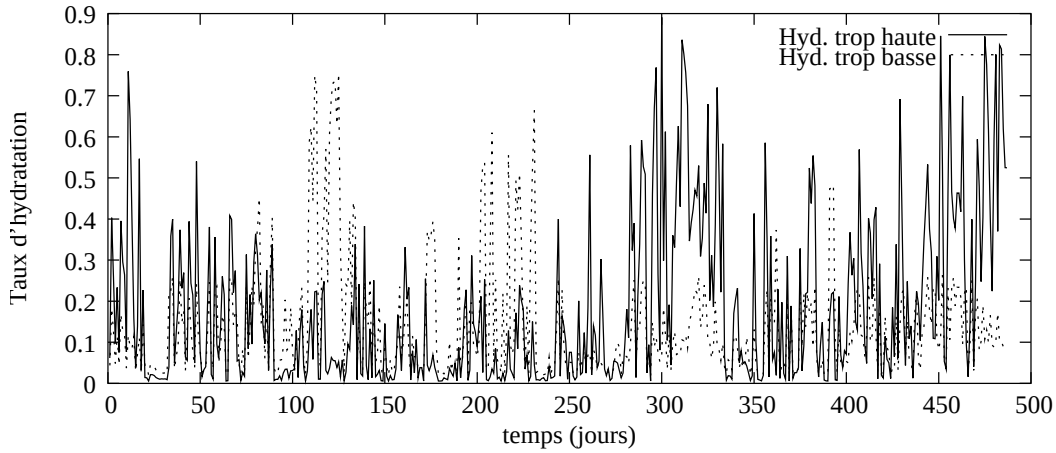
Poids du patient



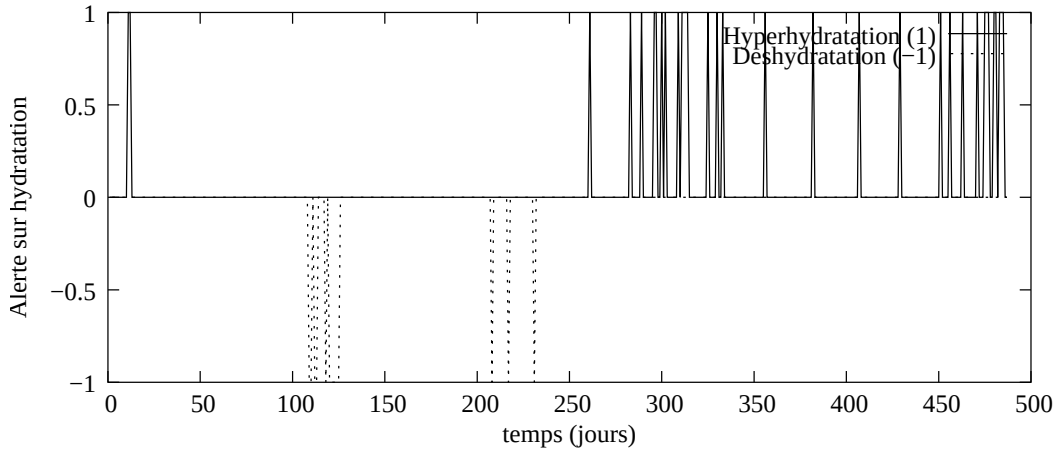
Tension du patient



Taux d’hydratation du patient



Alerte sur hydratation

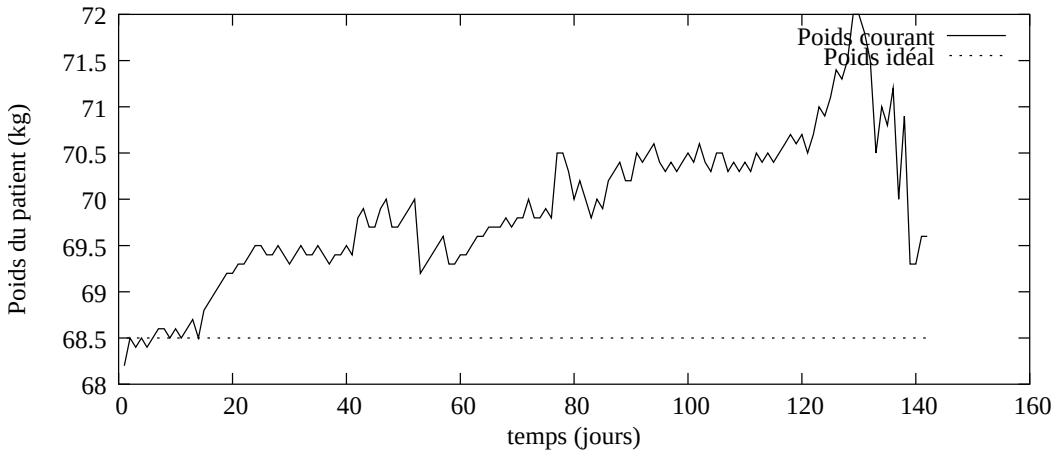


Note

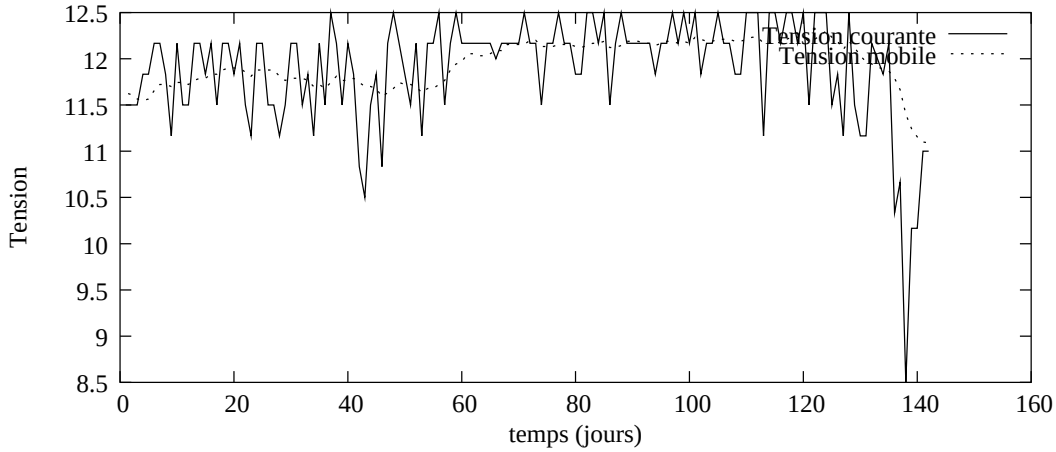
Il s’agit d’une des plus longues expériences. Ce patient est dans un état d’hyperhydratation répété durant la deuxième partie de l’expérience. Ceci est essentiellement dû à un poids trop élevé durant cette période. L’hésitation du système tient au fait que le patient a une tension qui reste dans un intervalle acceptable.

Patient n°1008

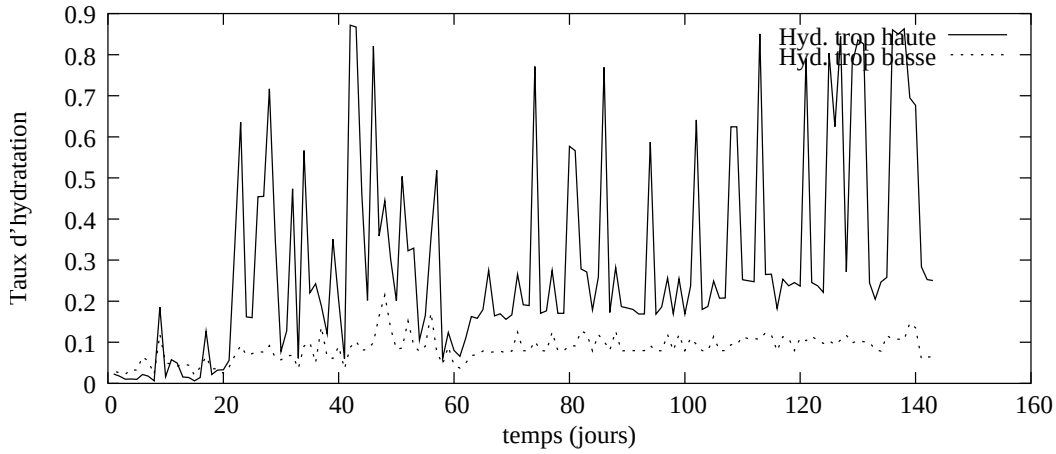
Poids du patient



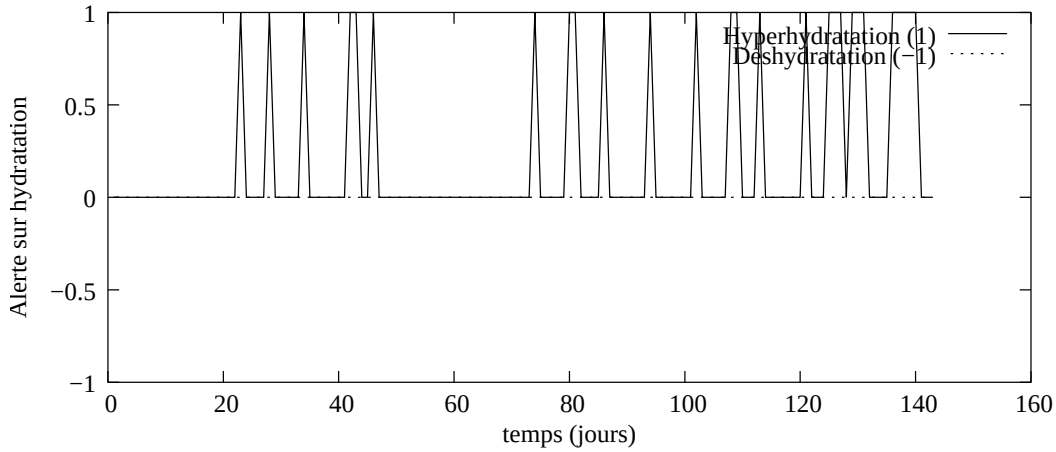
Tension du patient



Taux d’hydratation du patient



Alerte sur hydratation

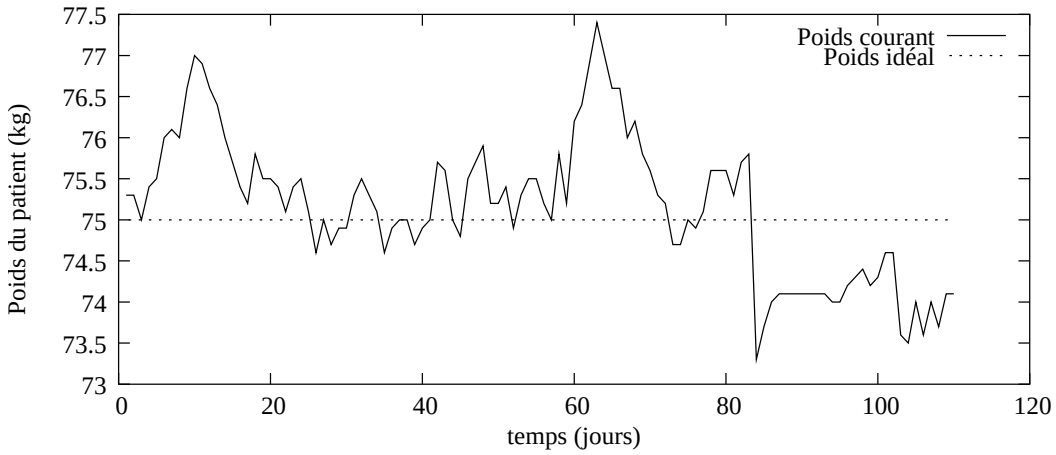


Note

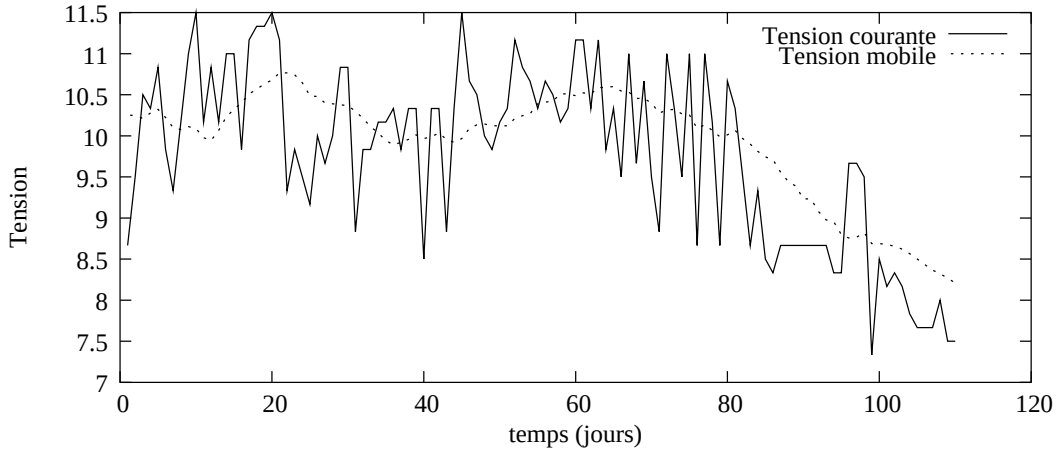
A la fin de l’expérience, ce patient a subi une chute de tension importante, qui explique en partie le fait que la période d’hyperhydratation soit un peu plus longue que d’habitude. Cette chute de tension est due à une chute de poids, qui reste néanmoins largement supérieure à la valeur conseillée par le médecin.

Patient n°1010

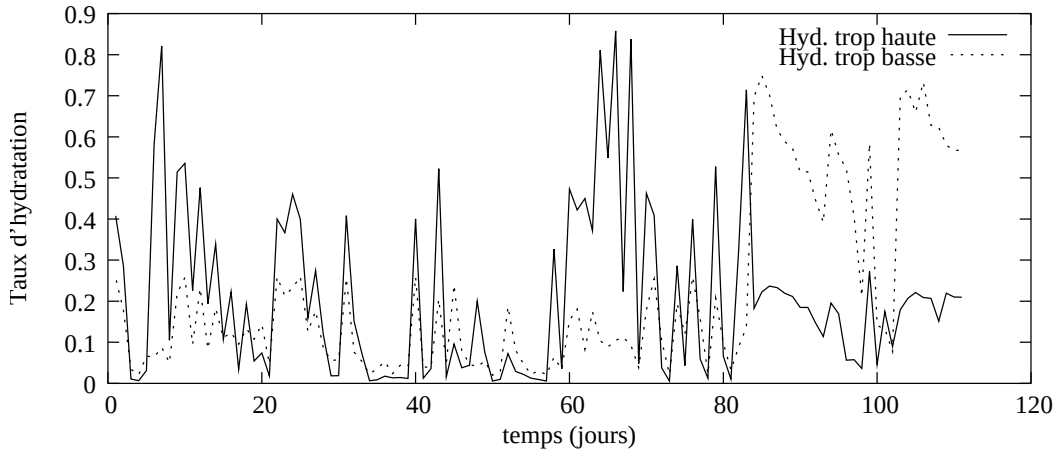
Poids du patient



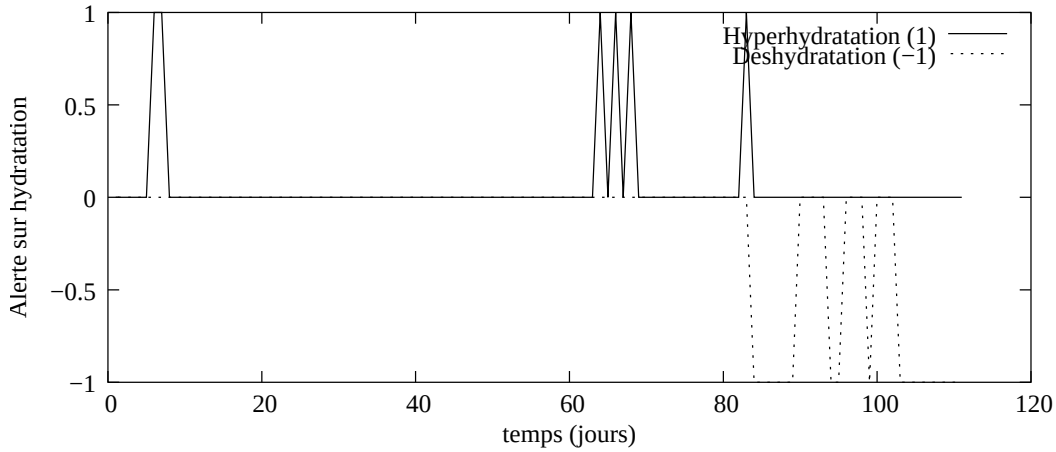
Tension du patient



Taux d’hydratation du patient



Alerte sur hydratation

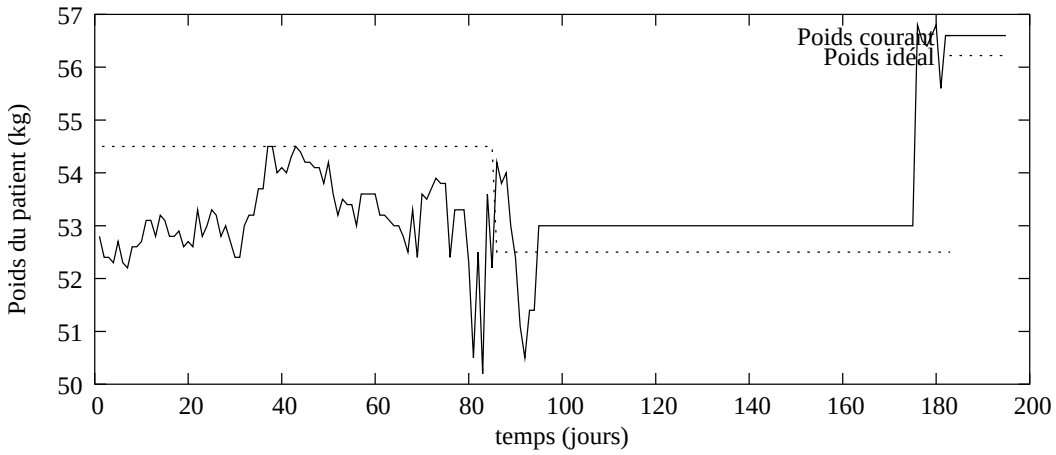


Note

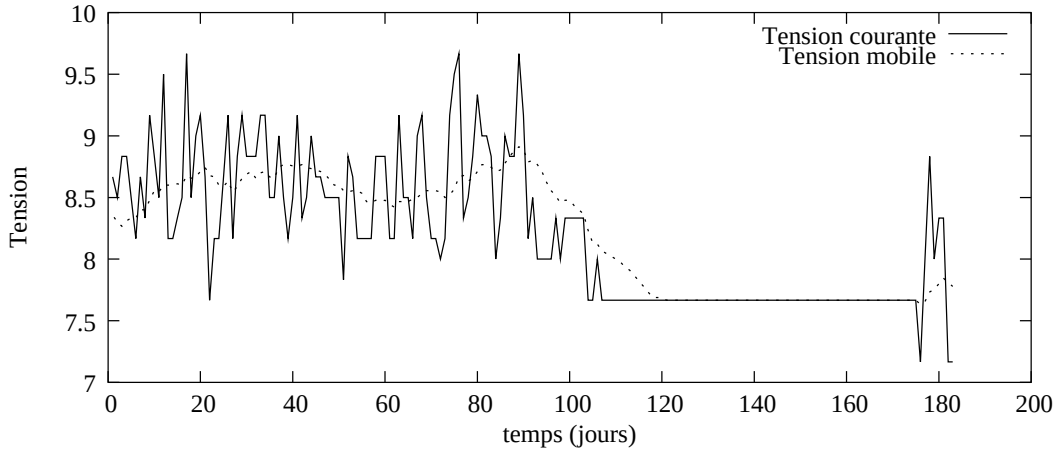
La fin de l’expérience montre un cas typique de patient hyperhydraté passant dans un état de deshydratation. Il a perdu énormément de poids : environ 2,5 kg en l’espace de deux jours.

Patient n°1011

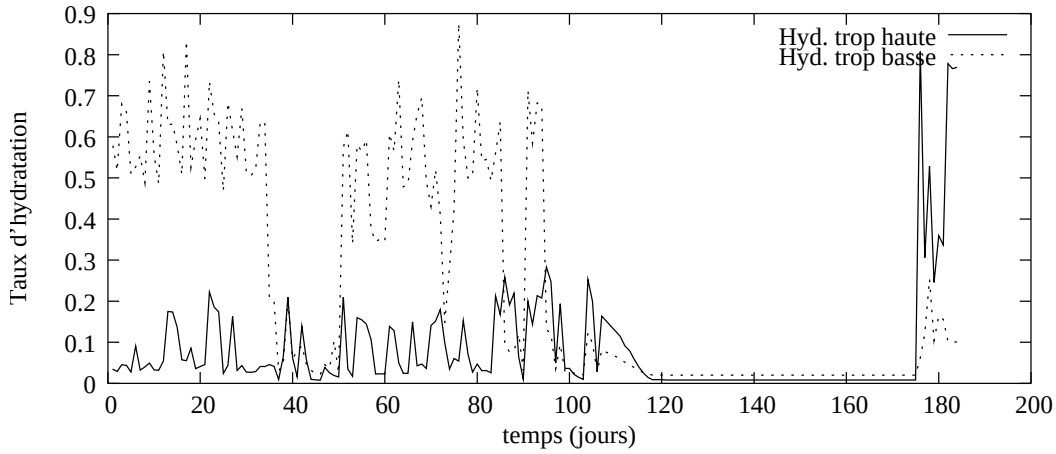
Poids du patient



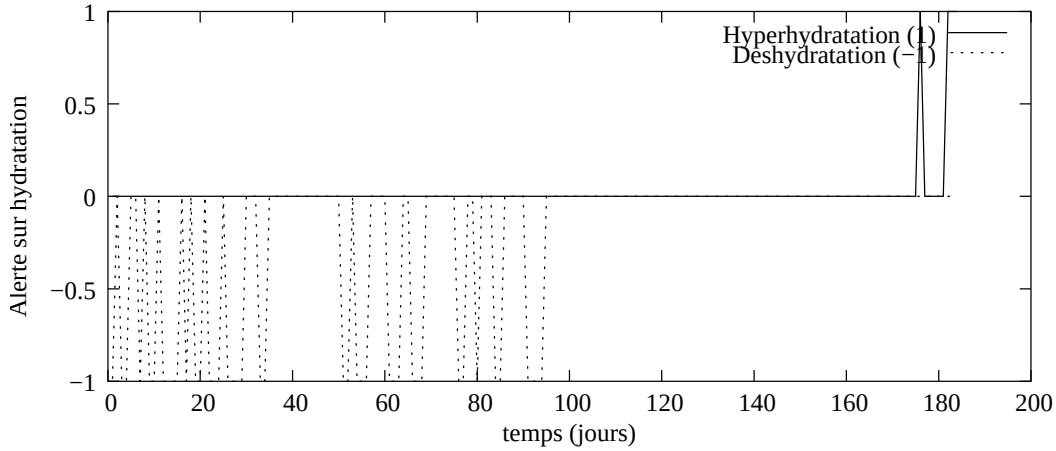
Tension du patient



Taux d’hydratation du patient



Alerte sur hydratation

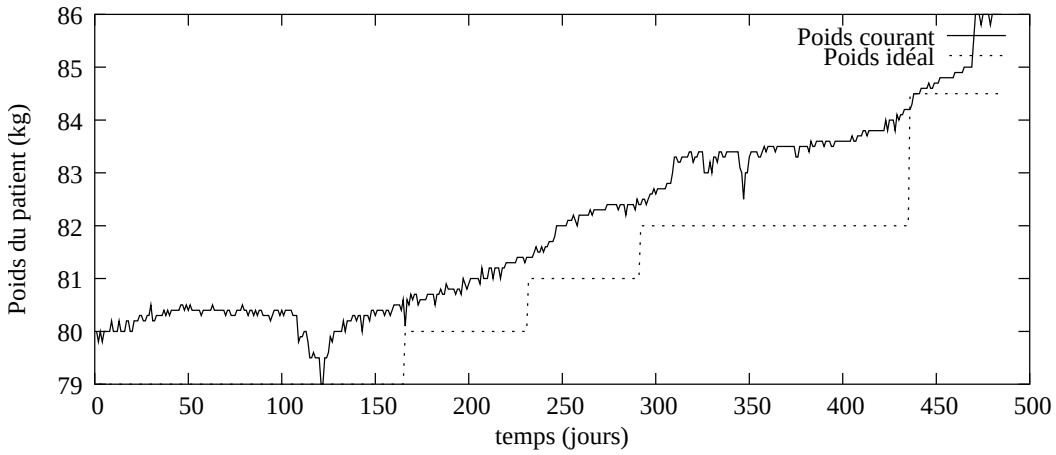


Note

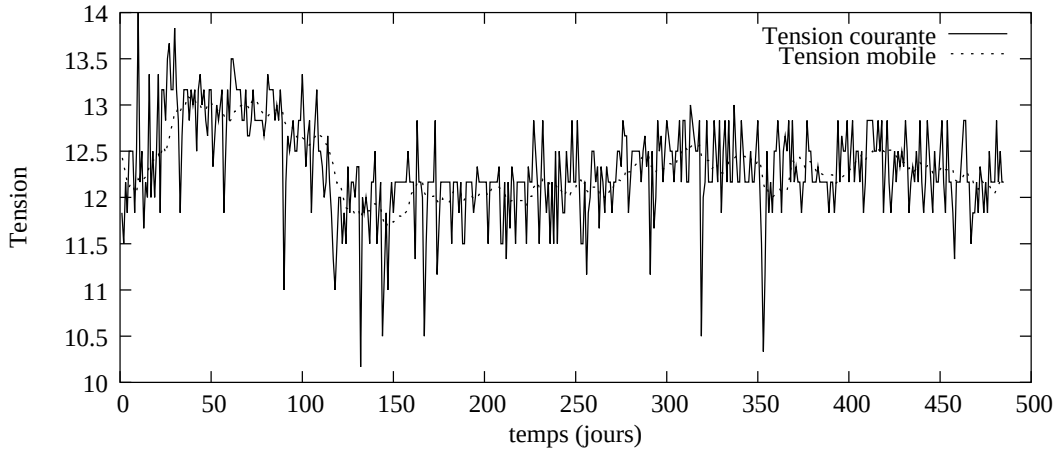
Le début de l’expérience montre que le système reporte bien la période où le patient a regagné du poids : son état d’hydratation est redevenu temporairement normal (du 36ème au 50ème jours).

Patient n°1012

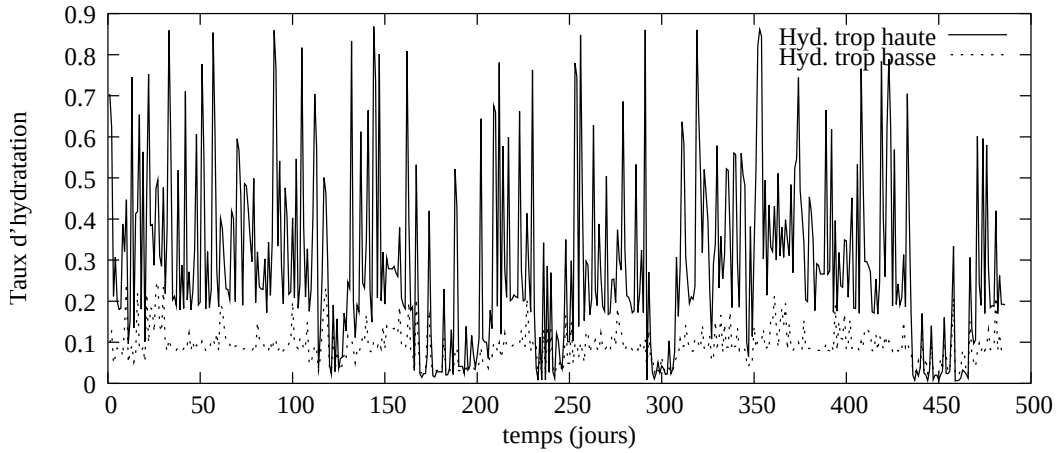
Poids du patient



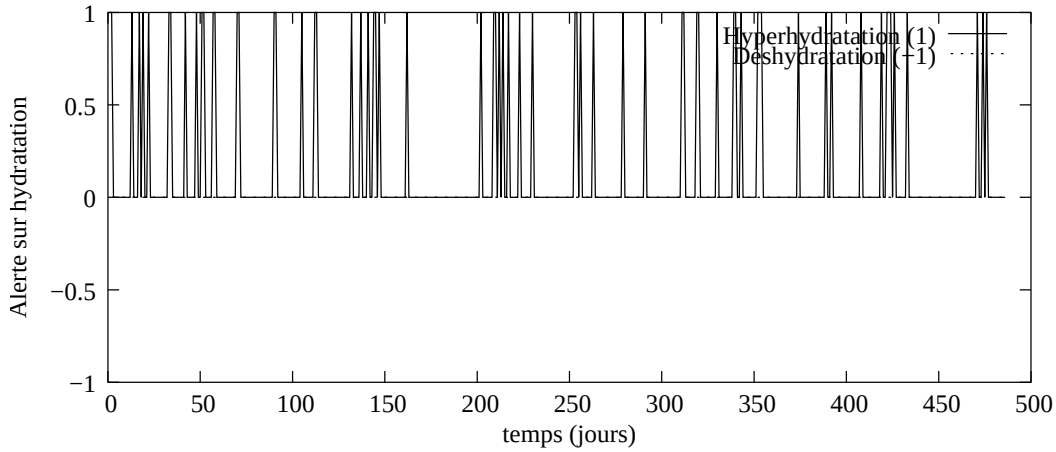
Tension du patient



Taux d’hydratation du patient



Alerte sur hydratation

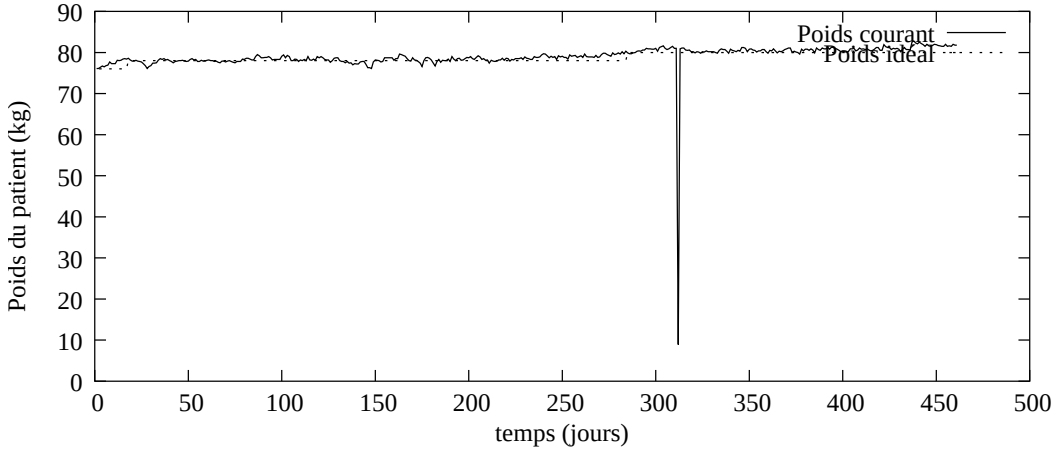


Note

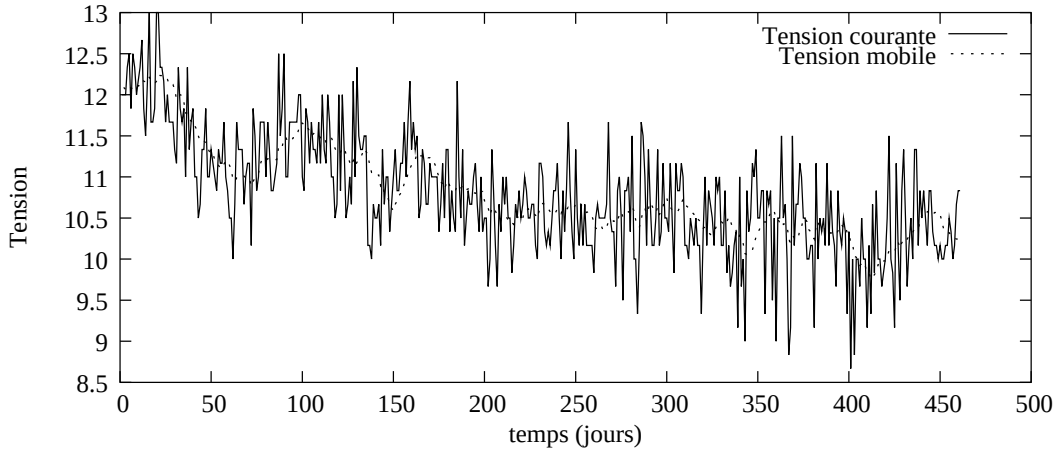
Il s’agit d’une longue expérience (486 jours) qui est intéressante car le médecin a changé le poids idéal du patient a de nombreuses reprises, en l’augmentant sans cesse. Ce cas arrive lorsque le patient est trop maigre et doit absolument grossir. Dans ce cas, le médecin règle le poids idéal à une valeur trop basse pour forcer le patient à grossir.

Patient n°1014

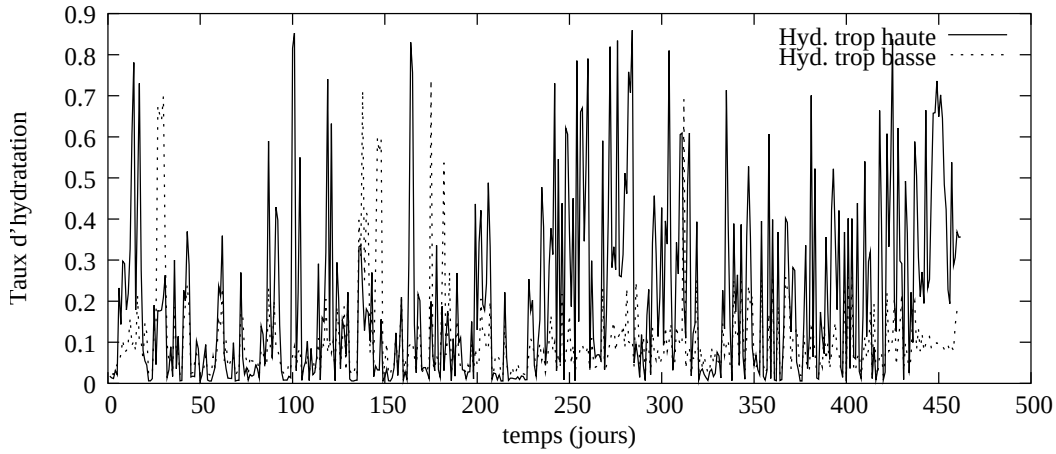
Poids du patient



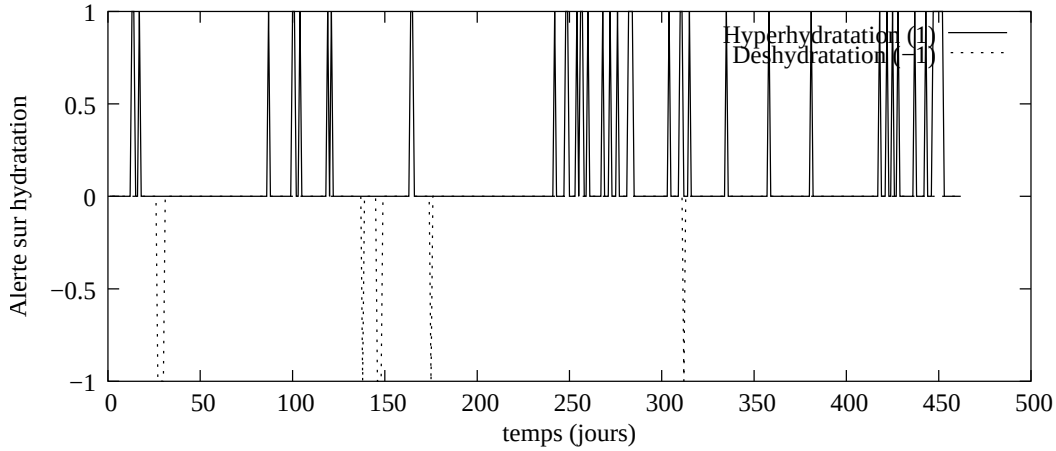
Tension du patient



Taux d’hydratation du patient



Alerte sur hydratation

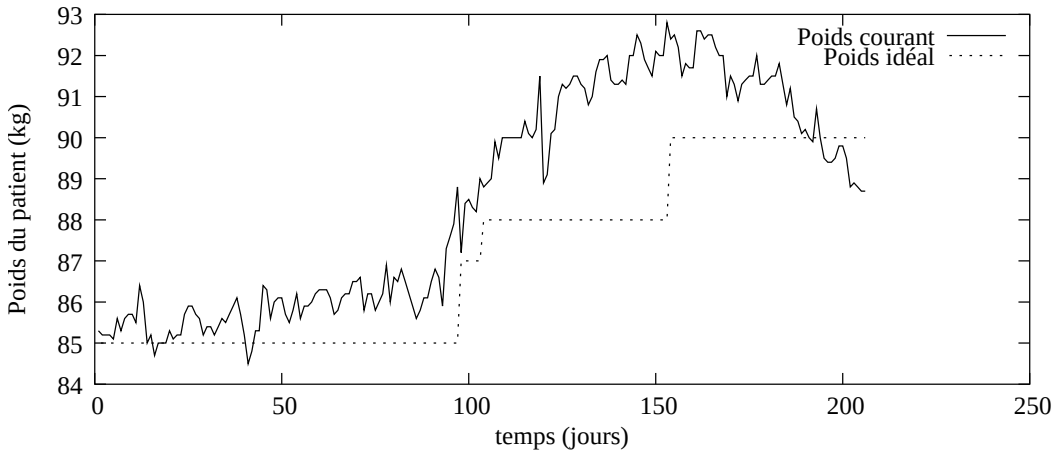


Note

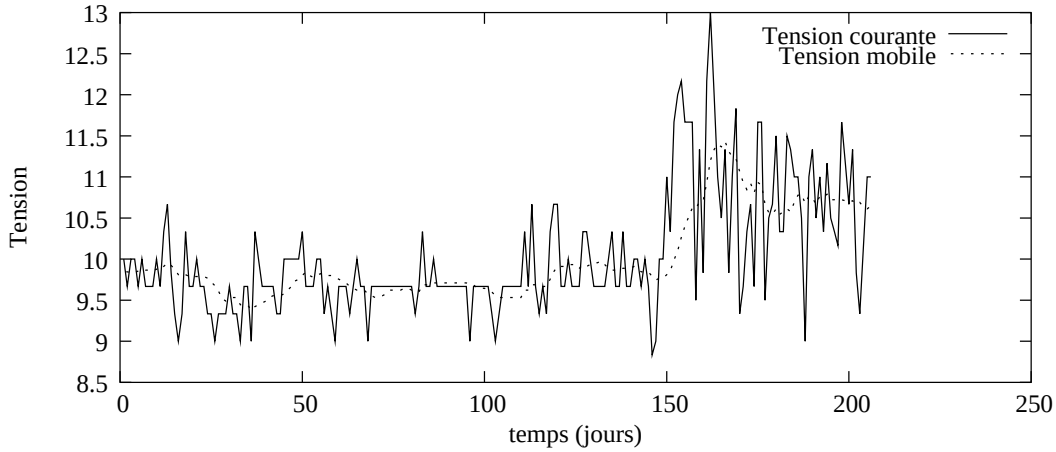
Il y a eu une erreur de saisie de poids de la part du patient. Néanmoins, cette erreur n’a pas eu d’influence notable sur le résultat car elle n’est pas corroboré par les données des jours précédents et suivants.

Patient n°1015

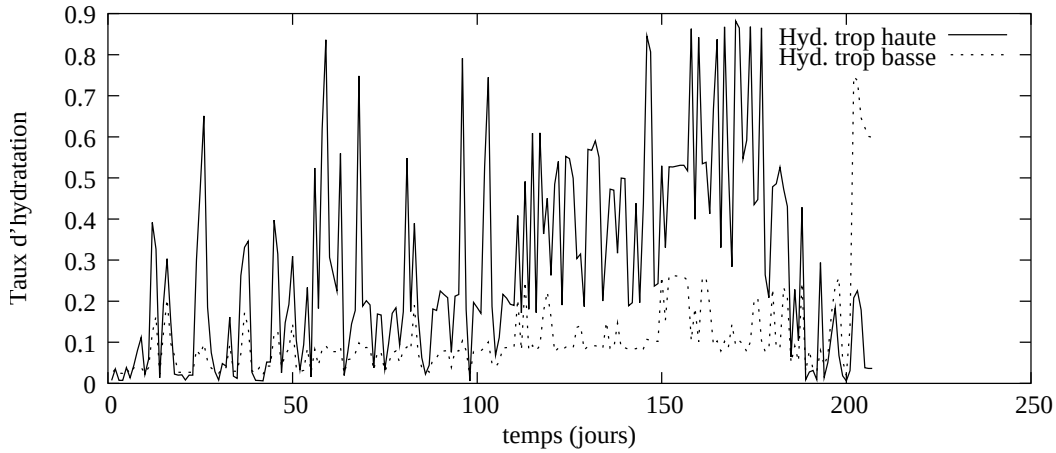
Poids du patient



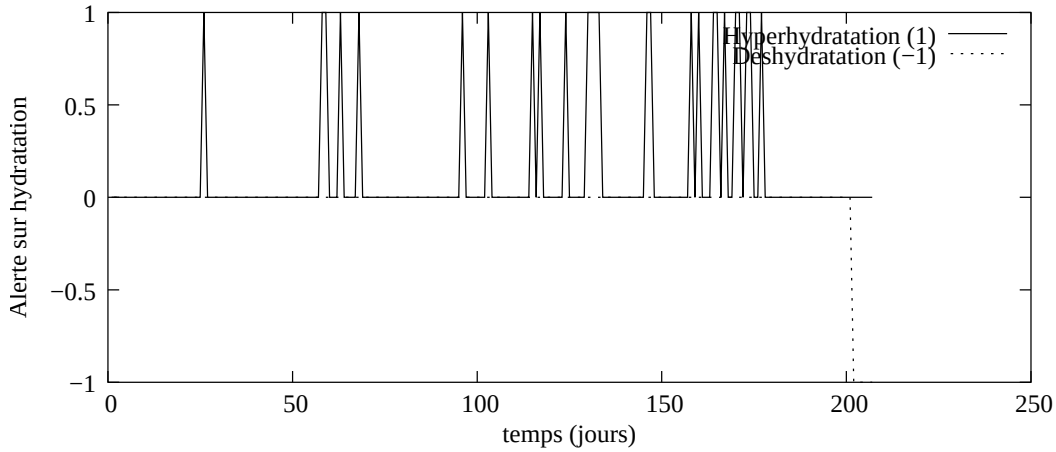
Tension du patient



Taux d’hydratation du patient



Alerte sur hydratation

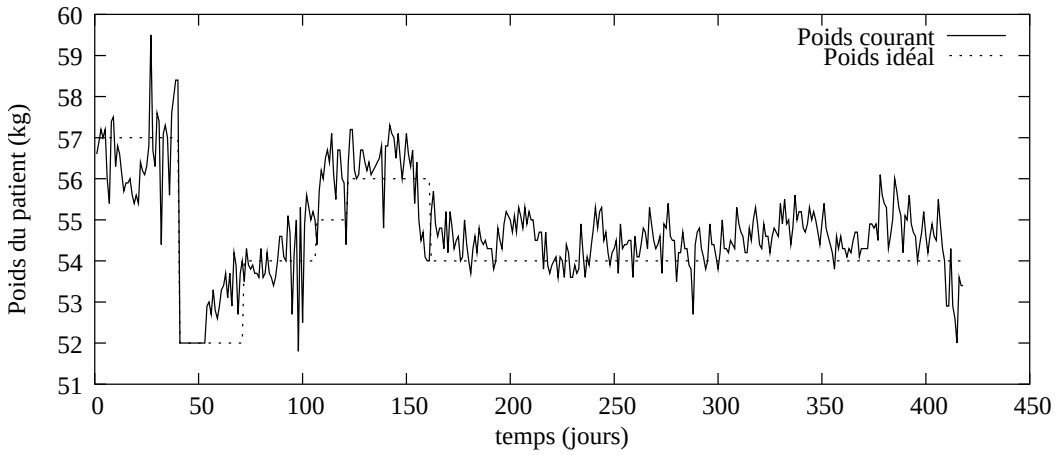


Note

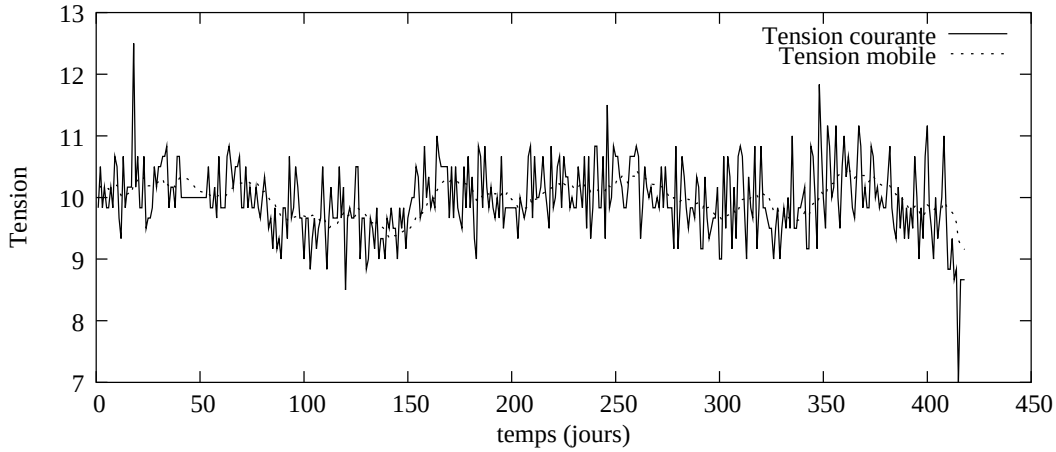
Là encore, le médecin a progressivement monté le poids du patient pour l’obliger à grossir. Cependant, à la fin de l’expérience, le patient perd subitement du poids, ce qui entraîne un état de deshydratation immédiat, reporté par le système.

Patient n°1016

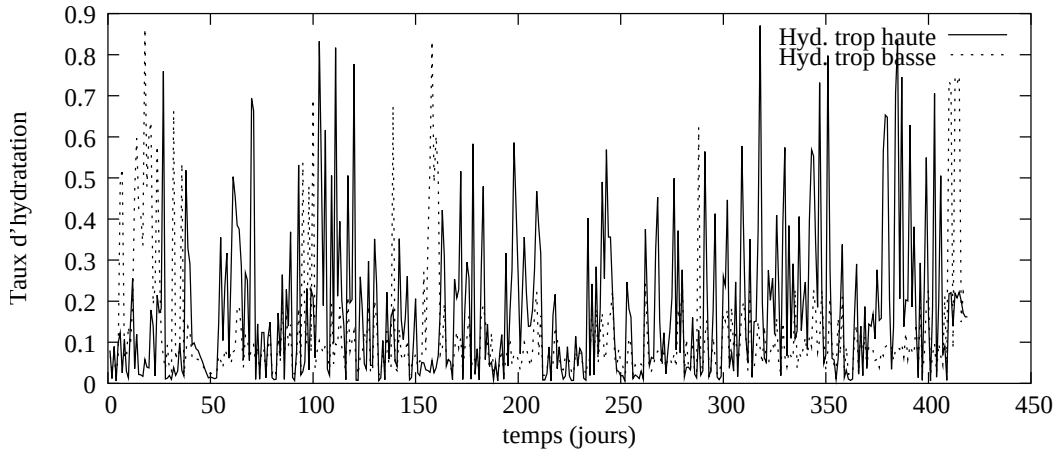
Poids du patient



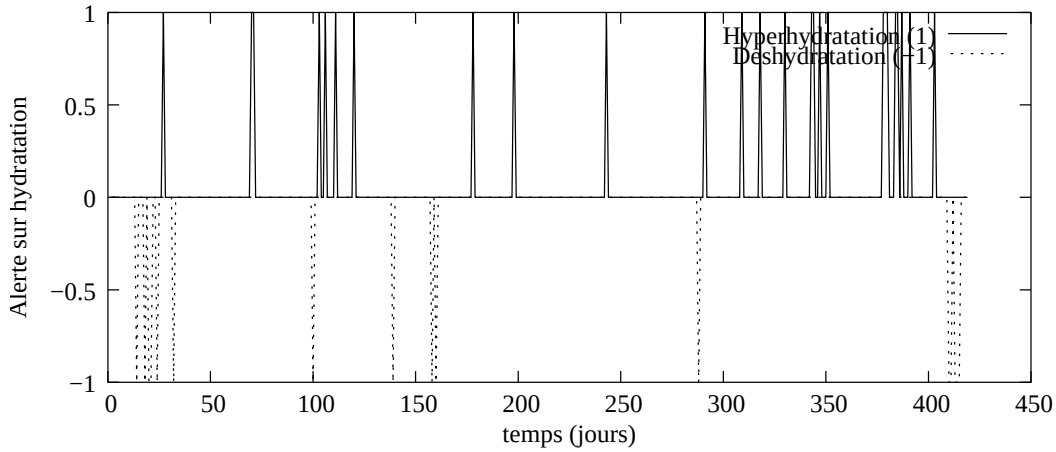
Tension du patient



Taux d’hydratation du patient



Alerte sur hydratation

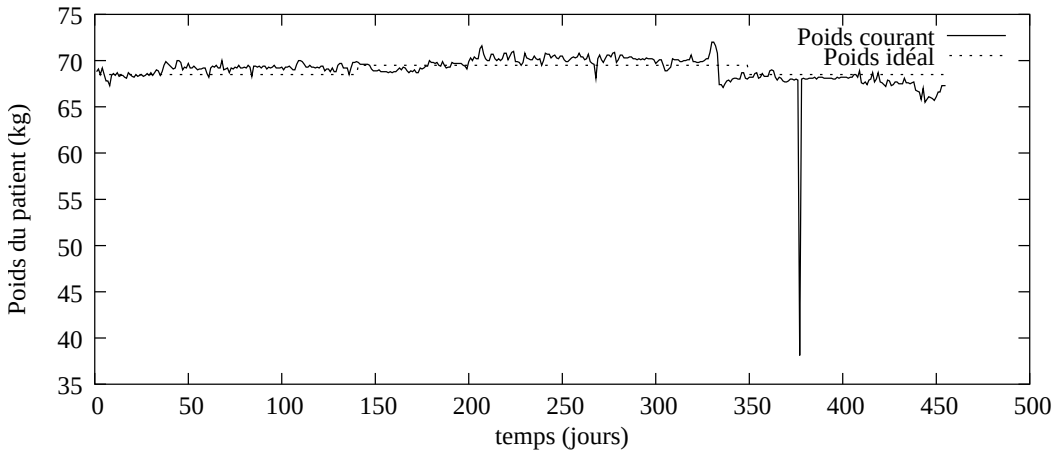


Note

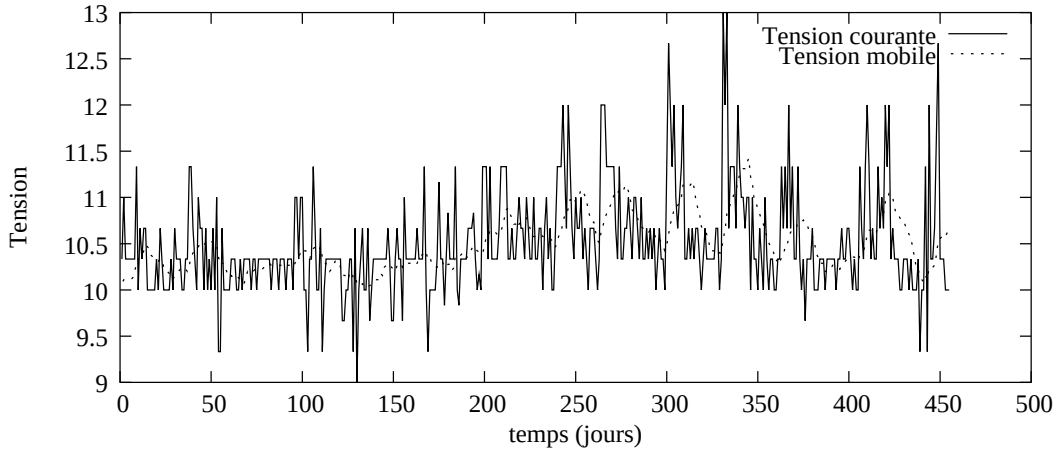
Il s’agit d’un cas typique où la tension intervient autant que le poids dans le diagnostic, car aucun des deux paramètres ne fait d’écart notable. Mais les petits écarts de tension et de poids influent conjointement sur l’état d’hydratation du patient.

Patient n°1017

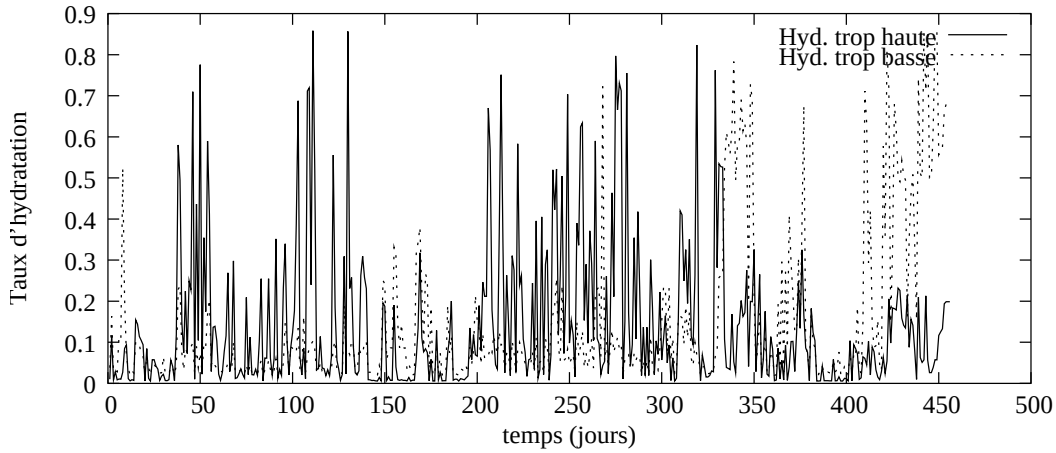
Poids du patient



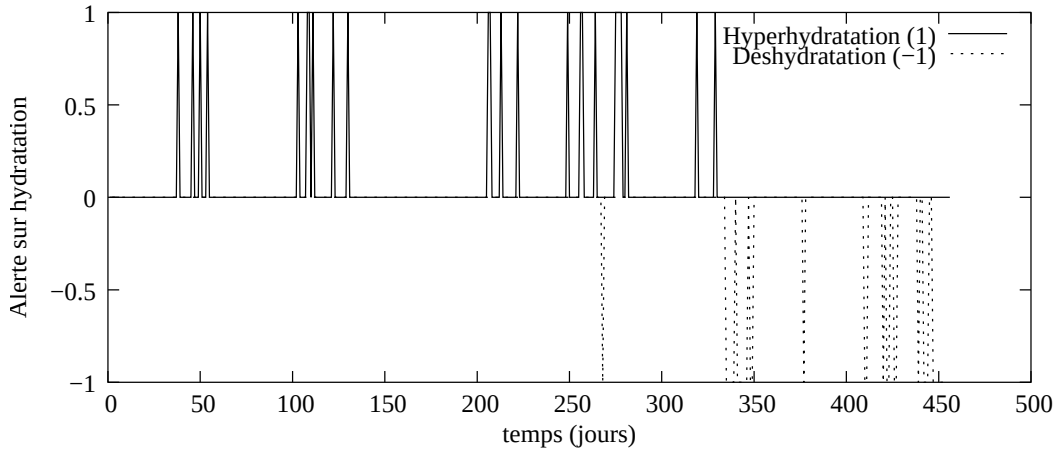
Tension du patient



Taux d’hydratation du patient



Alerte sur hydratation

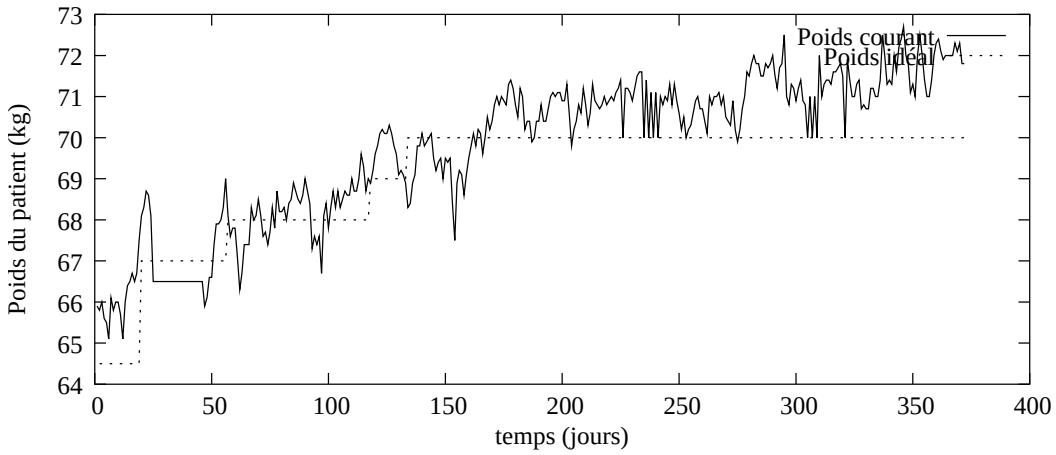


Note

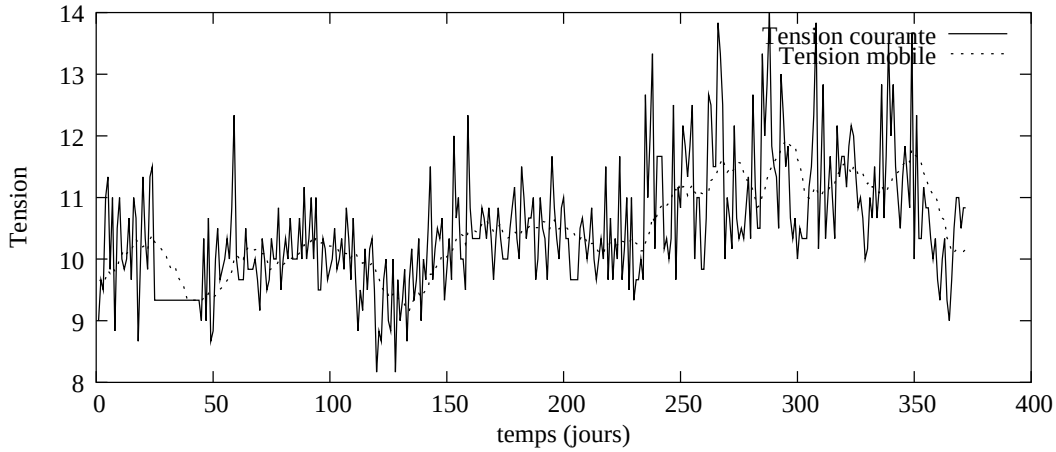
Ici encore, l’erreur de saisie n’a pas eu d’influence notable sur le diagnostic délivré par le système.

Patient n°1019

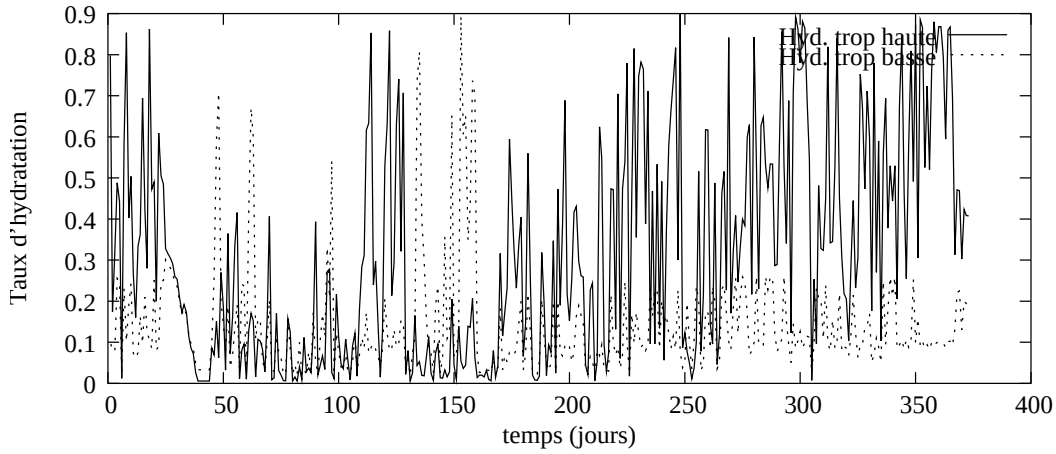
Poids du patient



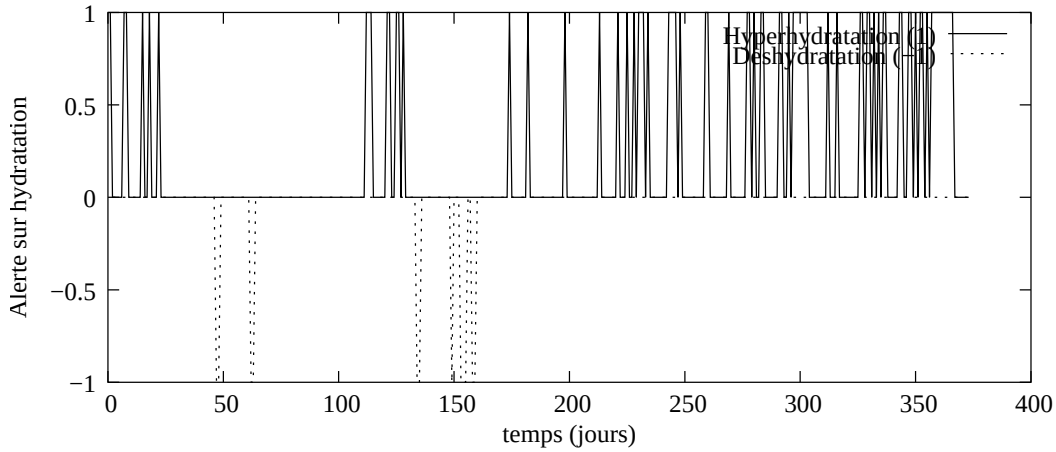
Tension du patient



Taux d’hydratation du patient



Alerte sur hydratation

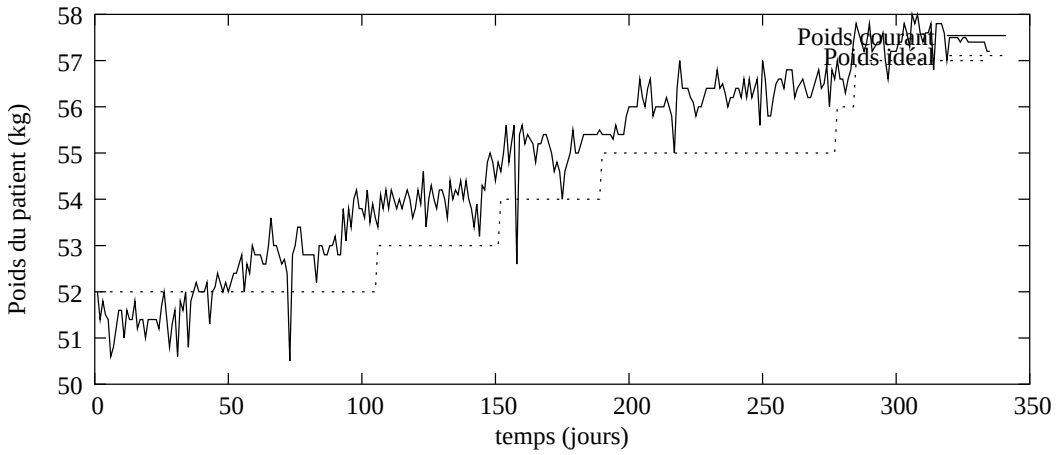


Note

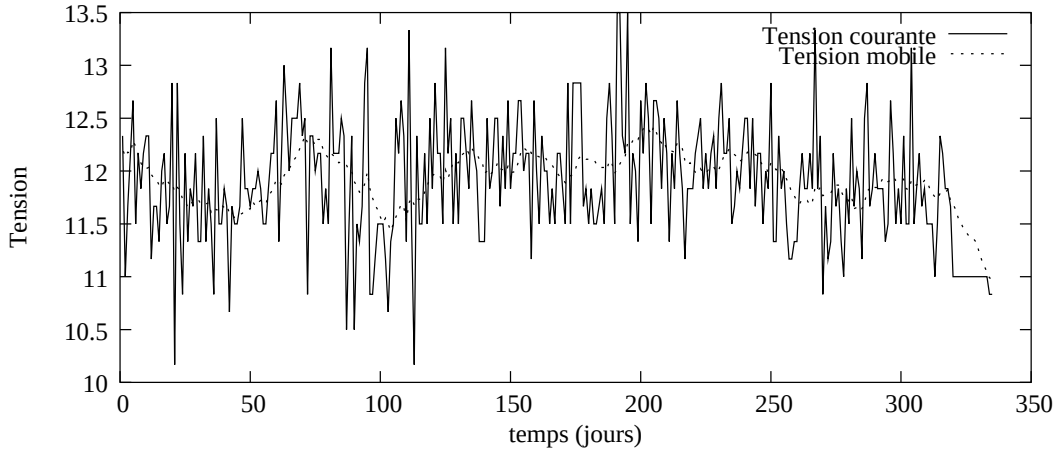
Après une première période où le médecin a régulièrement ajusté le poids idéal, le patient a continué à grossir. Ceci a eu pour conséquence une assez longue période d’hyperhydratation.

Patient n°1020

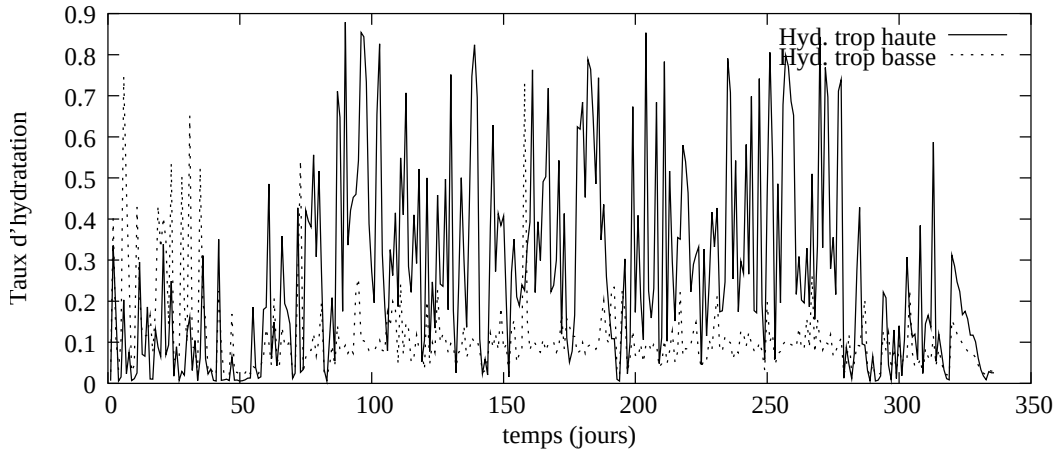
Poids du patient



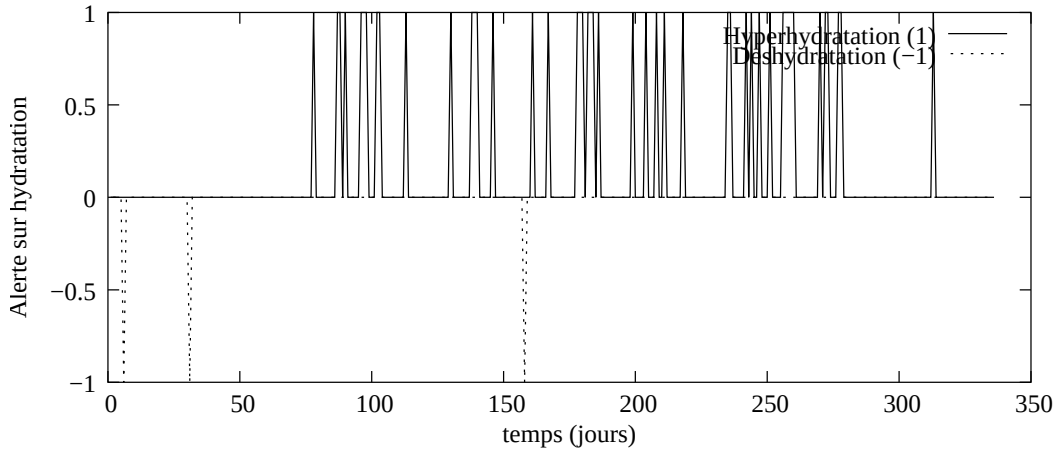
Tension du patient



Taux d’hydratation du patient



Alerte sur hydratation

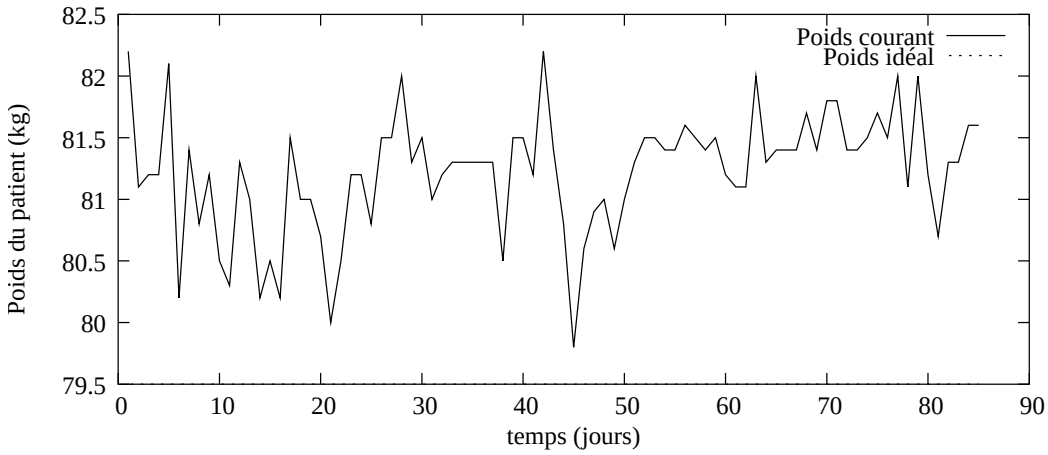


Note

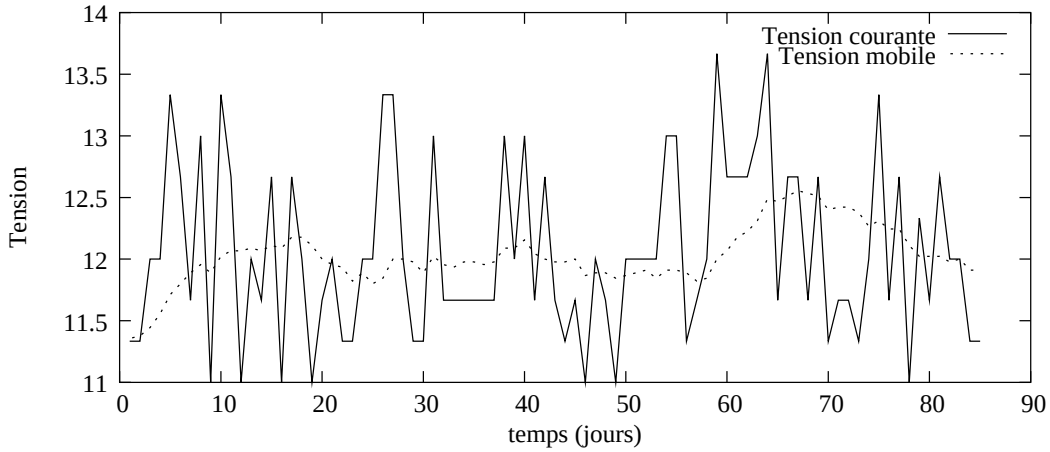
Le patient prend du poids régulièrement et le médecin ajuste le poids idéal à la hausse. De plus les variations importantes de la tension provoquent, chez ce patient, une répétition des cas d’hyperhydratation. Le cas de deshydratation au milieu de la courbe est dû à une chute importante de poids durant quelques jours.

Patient n°1023

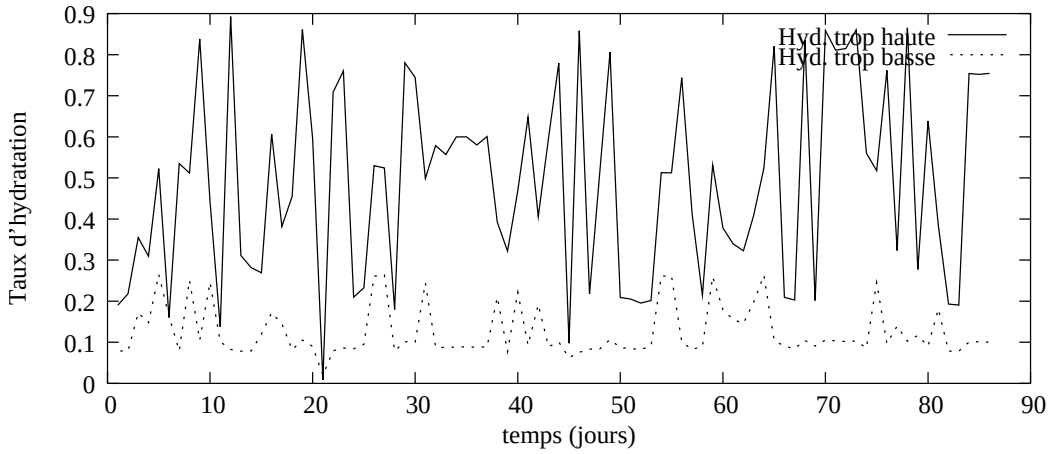
Poids du patient



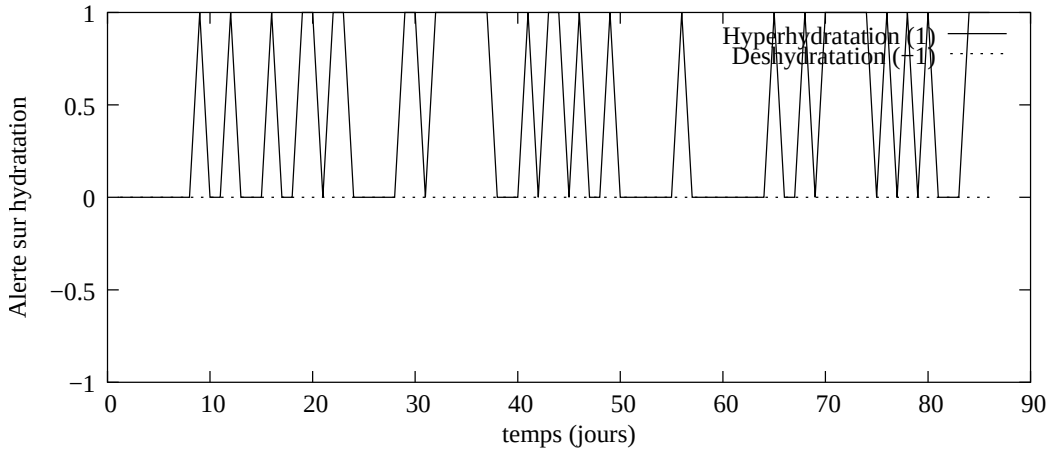
Tension du patient



Taux d’hydratation du patient



Alerte sur hydratation

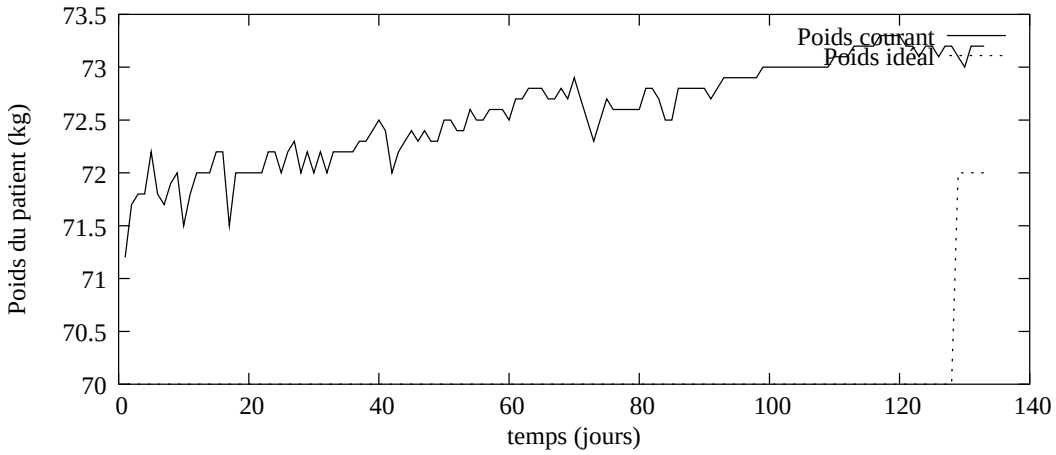


Note

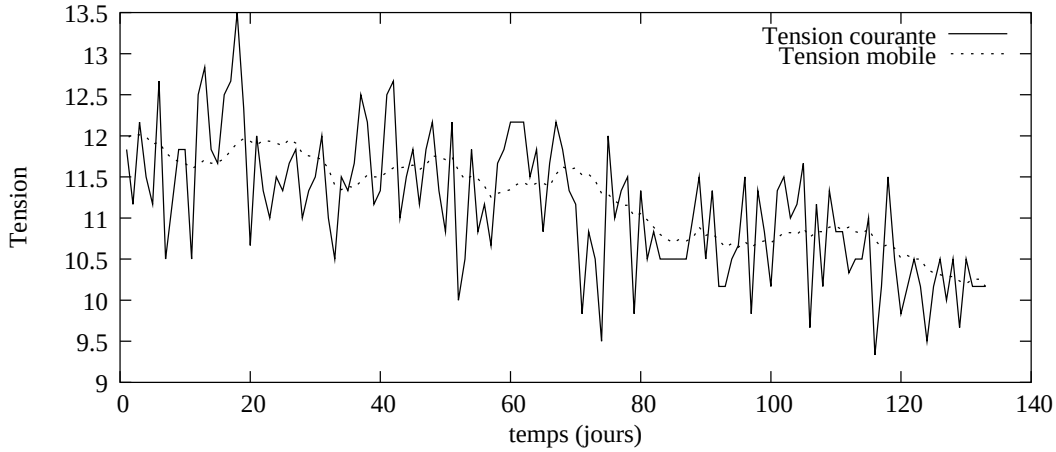
Il s’agit typiquement d’un patient ayant un poids trop élevé en permanence. Ceci explique la répétition des périodes d’hyperhydratation.

Patient n°1024

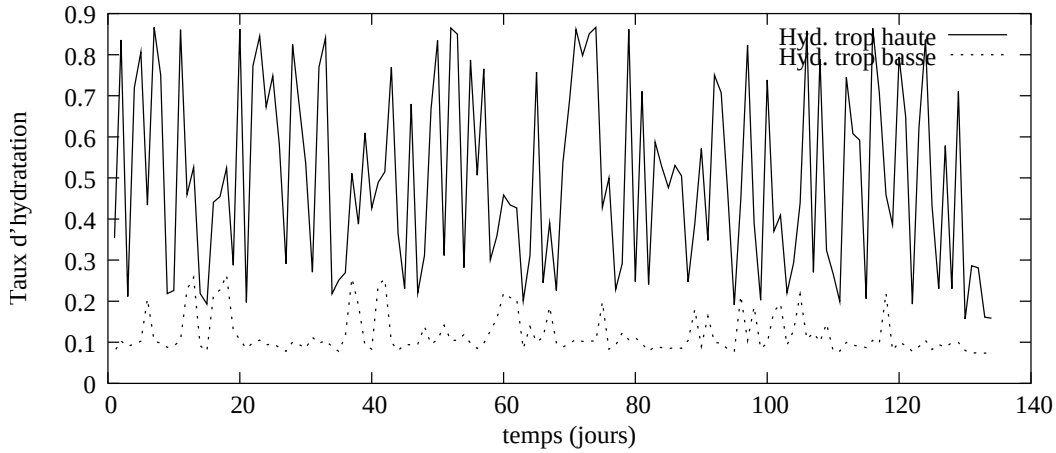
Poids du patient



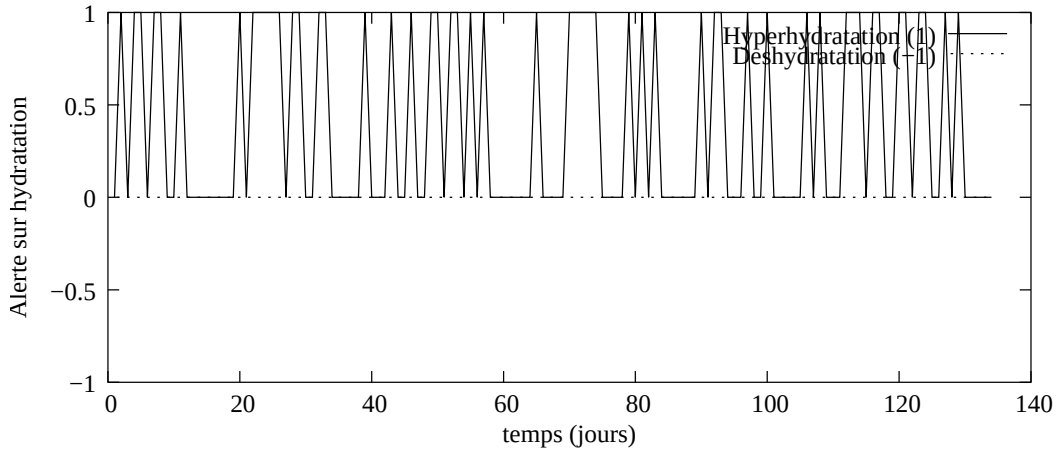
Tension du patient



Taux d’hydratation du patient



Alerte sur hydratation



Note

Même remarque que précédemment. Vers la fin de l’expérience, le médecin a néanmoins modifié fortement le poids idéal du patient. Il se peut justement que son estimation ait été fausse durant l’expérience. Il a donc décidé de procéder à une forte modification du poids idéal.

Annexe B

Publications personnelles

Conférences Internationales avec comité de lecture

– D. Bellot, A. Boyer, and F. Charpillet. A new definition of qualified gain in a data fusion process : application to telemedicine. In *FUSION 2002*, Annapolis, Maryland, USA, 2002.

D. Bellot, A. Boyer, and F. Charpillet. Designing smart agent based telemedicine systems using dynamic bayesian networks : an application to kidney disease people. In *HealthCom 2002*, Nancy, France, 2002.

J.P. Thomesse, D. Bellot, A. Boyer, E. Campo, M. Chan, F. Charpillet, J. Fayn, C. Leschi, N. Noury, V. Rialle, L. Romary, P. Rubel, N. Selmaoui, F. Steenkeste, and G. Virone. Integrated Information Technologies for patients remote follow-up and homecare. In *HealthCom 2001*, 2001.

Conférence nationale

– J.P Thomesse, D. Bellot, A. Boyer, E. Campo, M. Chan, F. Charpillet, J. Fayn, C. Leschi, N. Noury, V. Rialle, L. Romary, P. Rubel, N. Selmaoui, F. Steenkeste, and G. Virone. TIIS-SAD : Technologies de l'Information Intégrées aux Services des Soins A Domicile. In *AIM 2001 "Télémédecine et eSanté"*, http ://www.biomath.jussieu.fr/aim2001/, 2001.

Rapports

– D. Bellot, A. Boyer, and F. Charpillet. Vers une approche formelle de la fusion de données en intelligence artificielle : application en télémédecine. Technical Report A02-R-021, LORIA - INRIA Lorraine, 2002.

J.P. Thomesse, D. Bellot, A. Boyer, E. Campo, M. Chan, F. Charpillet, J. Fayn, C. Leschi, N. Noury, V. Rialle, L. Romary, P. Rubel, N. Selmaoui, F. Steenkeste, and G. Virone. Tiissad, technologies de l'information intégrées aux services des soins a domicile - compte rendu de fin

de recherche d'opération financée par le ministère de l'éducation nationale de la recherche et de la technologie. Technical report, LORIA, 2002.

D. Bellot. Notion de causalité et de dépendances dans les théories d'action. Mémoire, 1997. DEA d'Informatique de Marseille.

Exposés à des séminaires

– Fusion de données en télémédecine. Séminaire TIISSAD, Toulouse - Mars 2000

Réseaux bayésiens : principes et inférence avec l'algorithme JLO Séminaire TIISSAD, Strasbourd - Juin 2000

Bibliographie

[Abidi et al., 1992] Abidi, Mongi, A., Gonzalez, Rafael, C., and Elfes, A. (1992). *Data fusion in robotics and machine intelligence*, chapter 3, pages 137–163. Academic Press.

[Abidi and Gonzalez, 1992] Abidi, M. and Gonzalez, R. (1992). *Data Fusion in Robotics and Machine Intelligence*. Academic Press.

[Aliferis and Cooper, 1995] Aliferis, C. and Cooper, G. (1995). A structurally and temporally extended bayesian belief network model : Definitions, properties and modeling techniques. In *UAI*.

[Arnborg et al., 2000] Arnborg, S., Brynielson, J., Artman, H., and Wallenius, K. (2000). Information awareness in command and control : Precision, quality, utility. In *Fusion 2000*.

[Ayari, 1996] Ayari, I. (1996). *Fusion multi-capteurs dans un cadre multi-agents : application à un robot mobile*. PhD thesis, Université Henri Poincaré - Nancy I.

[Bahl et al., 1974] Bahl, L., Cocke, J., Jelinek, F., and Raviv, J. (1974). Optimal decoding of linear codes for minimizing symbol error rate. *IEEE Transactions on Information Theory*, 20 :284–287.

[Baldi and Brunak, 1998] Baldi, P. and Brunak, S. (1998). *Bioinformatics, the Machine Learning Approach*. MIT Press.

[Bar-Shalom and Li, 1995] Bar-Shalom, Y. and Li, X.-R. (1995). *Multitarget-Multisensor Tracking : Principles and Techniques*. Number 3rd printing. YBS.

[Barret, 1990] Barret, I. (1990). *Synthèse d'algorithmes de poursuite multi-radars d'avions civils manouevrants*. PhD thesis, Ecole Nationale supérieure de l'aeronautique et de l'espace.

[Bayes, 1763] Bayes, T. (1763). A essay toward solving a problem in the doctrine of chance. *Philosophical Transactions of the Royal Society*, 53 :370,418.

[Becker and Naïm, 1999] Becker, A. and Naïm, P. (1999). *Les réseaux bayésiens*. Eyrolles, eyrolles edition.

[Bellot et al., 2002a] Bellot, D., Boyer, A., and Charpillet, F. (2002a). Designing smart agent based telemedicine systems using dynamic bayesian networks : an application to kidney disease people. In *Proc. HealtCom 2002*, pages 90–97, Nancy, France.

[Bellot et al., 2002b] Bellot, D., Boyer, A., and Charpillet, F. (2002b). A new definition of qualified gain in a data fusion process : application to telemedicine. In *Fusion 2002*, Annapolis, Maryland, USA.

[Berkson, 1946] Berkson, J. (1946). Limitations of the application of fourfold table analysis to hospital data. *Biometrics Bulletin*, 2 :47–53.

[Beschta et al., 1993] Beschta, A., Dressler, O., Freitag, H., Montag, M., and Struss, P. (1993). Dpnet-a second generation expert system for localizing faults in power transmission networks. In *Proceedings International Conference on Fault Diagnosis (Tooldiag-93)*, pages 1019–1027, Toulouse, France.

[Bigi, 2000] Bigi, B. (2000). *Contribution à la modélisation du langage pour des applications de recherche documentaire et de traitement de la parole*. PhD thesis, Université d'Avignon et des Pays du Vaucluse, Avignon, France.

[Bloch and Maître, 1994] Bloch, I. and Maître, H. (1994). Fusion de données en traitement d'images : modèles d'information et décisions. *Traitement du Signal*, 11(6) :435–446.

[Brogi et al., 1988] Brogi, A., Filipi, R., Gaspari, M., and Turini, F. (1988). An expert system for data fusion based on blackboard architecture. In *Proc. 8th Int. Workshop Expert Systems and their Applications*, pages 147–165, Avignon, France.

[Buchanan and Shortliffe, 1984] Buchanan, B. and Shortliffe, E. (1984). *Rule-based Expert Systems : the MYCIN Experiments of the Stanford Heuristic Programming Project*. Addison-Wesley, MA.

[Cain and Turner, 1995] Cain, N. M. and Turner, H. (1995). A causal theory of ramifications and qualifications. In *IJCAI'95*, pages 1978–1984.

[Chandrasekaran, 1988] Chandrasekaran, B. (1988). Generic tasks as building blocks for knowledge-based systems : the diagnosis and routine examples. *The Knowledge Engineering Review*, 3 :183–210.

[Chanliau et al., 2001] Chanliau, J., Charpillet, F., Durand, P., Hervy, R., Pierrel, J., Romary, L., and Thomesse, J. (2001). Un système de télésurveillance de malades à domicile. Brevet français n° 00 00903.

[Chaodhury et al., 2002] Chaodhury, T., Pentland, A., Regh, J., and Pavlovic, V. (2002). Boosted learning in dynamic bayesian networks for multimodal detection.

[Cheng and Bell, 1997] Cheng, J. and Bell, D. (1997). Learning bayesian networks from data : an efficient approach based on information theory. In *Proceeding of the sixth ACM International Conference on Information and Knowledge Management*.

[Cooper, 1990] Cooper, G. (1990). Computational complexity of probabilistic inference using bayesian belief networks. *Artificial Intelligence*, 42 :393–405.

[Cooper and Herskovitz, 1992] Cooper, G. and Herskovitz, E. (1992). A bayesian method for the induction of probabilistic networks from data. *Machine Learning*, 9 :309–347.

[Cormen et al., 1990] Cormen, T., Leiserson, C., and Rivest, R. (1990). *Introduction à l'algorithmique*.

[Cowell et al., 1999] Cowell, R., Dawid, A., Lauritzen, S., and Spiegelhalter, D. (1999). *Probabilistic Networks and Expert Systems*. ISBN : 0-387-98767-3.

[Cox and Wermuth, 1996] Cox, D. and Wermuth, N. (1996). *Multivariate Dependancies - Models, Analysis and Interpretation*. Chapman Hall, London.

[Crowley and Demazeau, 1993] Crowley, J. and Demazeau, Y. (1993). Principles and techniques pour sensor data fusion. *Signal processing*, 32 :5–27.

[Dague, 1994] Dague, P. (1994). Model-based diagnosis of analog electronic circuits. 11 :439–492.

[Dagum and Galper, 1995] Dagum, P. and Galper, A. (1995). Time-seris prediction using belief network models. *Intl. Journal of Human-Computer Studies*, 42 :617–632.

[Daoudi et al., 2000] Daoudi, K., Fohr, D., and Antoine, C. (2000). A new approach for multi-band speech recognition based on probabilistic graphical models. In *ICSLP'2000*, Beijing, China.

[Dawid, 1992] Dawid, A. (1992). Applications of a general propagation algorithm for probabilistic expert systems. *Statistics and Computing*, 2 :25–36.

[Dechter, 1996a] Dechter, R. (1996a). Bucket elimination : a unifying framework for probabilistic inference. In Horvitz, E. and Jensen, F., editors, *Proceedings of the 12th Annual Conference on Uncertainty in Artificial Intelligence*, pages 211–219, San Francisco, California. Morgan Kaufman.

[Dechter, 1996b] Dechter, R. (1996b). Topological parameters for time-space tradeoff. In Horvitz, E. and Jensen, F., editors, *Proceedings of the 12th Conference on Uncertainty in Artificial Intelligence*, pages 220–227, San Francisco. Morgan Kaufman.

[Dechter, 1998] Dechter, R. (1998). *Bucket elimination : a unifying framework for probabilistic inference*. MIT Press.

[Delone and McLean, 1992] Delone, W. and McLean, E. (1992). Information systems success : the quest for the dependant variable. *Information Systems Research*, 3 :60–95.

[Dempster, 1967] Dempster, A. (1967). Upper and lower probabilities induced by a multivalued mapping. *Ann. Math. Stat.*, pages 325–339.

[Downing, 1993] Downing, K. (1993). Physiological applications of consistency-based diagnosis. *Artificial Intelligence in Medicine*, 5 :9–30.

[Durand and Kessler, 1998] Durand, P.-Y. and Kessler, M. (1998). *La dialyse péritonéanle automatisée*.

[Elfes, 1989] Elfes, A. (1989). Using occupancy grids for mobile robot perception and navigation. *IEEE Computer*, 6 :46–57.

[Fine et al., 1998] Fine, S., Singer, Y., and Tishby, N. (1998). The hierarchical hidden markov model : Analysis and applications. *Machine Learning*, pages 32–41.

[Frey, 1998] Frey, B. (1998). *Graphical Models for Machine Learning and Digital Communications*. Cambridge, Massachusetts.

[Friedman, 1998] Friedman, N. (1998). The Bayesian structural EM algorithm. In Kaufmann, M., editor, *Proc. Fourteenth Conference on Uncertainty in Artificial Intelligence (UAI '98)*, pages 129–138, San Francisco, CA.

[Gales, 1999] Gales, M. (1999). Semi-tied covariance matrices for hidden markov models. *IEEE Transaction on Speech and Audio Processing*, 7 :272–281.

[Gebhardt and Kruse, 1998] Gebhardt, J. and Kruse, R. (1998). Information Source Modelling for Consistent Data Fusion. In Hamid R. Arabnia and Dongping (Daniel) Zhu, editors, *Proceedings of the International Conference on Multisource-Multisensor Information Fusion - Fusion'98*, volume I, pages 27–34, Las Vegas, Nevada, USA. CSREA Press.

[Geiger et al., 1988] Geiger, D., Verma, T., and Pearl, J. (1988). Identifying independance in bayesian networks. *Networks*, 20 :507–534.

[Hall and Llinas, 1997] Hall, D. and Llinas, J. (1997). An introduction to multisensor data fusion. In IEEE, editor, *Proceedings of the IEEE*, volume 85, pages 6–23.

[Hamilton, 1994] Hamilton, J. (1994). *Time Series Analysis*.

[Haton et al., 1998] Haton, J., Charpillet, F., and Haton, M. (1998). Numeric/symbolic approaches to data and information fusion. In *Proceedings of the International Conference on Multisource-Multisensor Information Fusion - Fusion'98*, volume II, pages 888–895, Las Vegas, Nevada, USA. CSREA Press.

[Heckerman and et al., 1995] Heckerman, D. and et al., D. G. (1995). Learning bayesian networks : the combination of knowledge and statistical data. *Machine Learning*, 20 :197–243.

[Hertz and Krogh, 1991] Hertz, J. and Krogh, A. (1991). Palmer : Introduction to the theory of neural computation.

[Isham, 1981] Isham, V. (1981). An introduction to spatial point processes and markov random fields. *International Statistical Review*, 49 :21–43.

[Jaakkola and Jordan, 1999] Jaakkola, T. and Jordan, M. I. (1999). Variational probabilistic inference and the QMR-DT network. *Journal of Artificial Intelligence Research*, 10 :291–322.

[Jeanpierre, 2002] Jeanpierre, L. (2002). *Apprentissage et adaptation pour la modélisation stochastique de systèmes dynamiques réels*. PhD thesis, Université Henri Poincaré - Nancy I, Nancy, France.

[Jeanpierre and Charpillet, 2002] Jeanpierre, L. and Charpillet, F. (2002). Hidden markov models for medical diagnosis. In *Proc. HealtCom 2002*, pages 98–102, Nancy, France.

[Jelinek, 1997] Jelinek, F. (1997). *Statistical methods for speech recognition*. MIT Press.

[Jensen, 1996] Jensen, F. (1996). *An Introduction to Bayesian Networks*. UCL Press.

[Jensen et al., 1990] Jensen, F., Lauritzen, S., and Olesen, K. (1990). Bayesian updating in recursive graphical models by local computations. *Computational Statistical Quaterly*, 4 :269–282.

[Jordan, 1999] Jordan, M. I., editor (1999). *Learning in Graphical Models*. MIT Press.

[Jordan et al., 1999] Jordan, M. I., Ghahramani, Z., Jaakkola, T., and Saul, L. K. (1999). An introduction to variational methods for graphical models. *Machine Learning*, 37(2) :183–233.

[Jr. and Young, 1999] Jr., E. S. and Young, J. (1999). Probabilistic temporal networks : A unified framework for reasoning with time and uncertainty. *Intl. Journal of Approximate Reasoning*, 20 :191–216.

[Kask et al., 2001] Kask, K., Dechter, R., Larrosa, J., and Cozman, F. (2001). Bucket-elimination for automated reasoning. Technical Report R92, UC Irvine ICS.

[Kiiveri et al., 1984] Kiiveri, H., Speed, T., and Carlin, J. (1984). Recursive causal models. *Journal of the Australian Mathematical Society*, 36 :30–52.

[Kim and Pearl, 1983] Kim, J. and Pearl, J. (1983). A computational model for combined causal and diagnostic reasoning in inference systems. In *Proceedings IJCAI-83*, pages 190–193, Karlsruhe, Germany.

[Kjaerulff, 1992] Kjaerulff, U. (1992). Optimal decomposition of probabilistic networks by simulated annealing. *Statistics and Computing*, 2 :7–17.

[Kjaerulff, 1998] Kjaerulff, U. (1998). *Nested junction trees*, pages 51–74. Kluwer Academic Publishers, Dordrecht, The Netherlands.

[Koenig and Simmons, 1996] Koenig, S. and Simmons, R. (1996). Unsupervised learning of probabilistic models for robot navigation. In *Proceedings of the IEEE International Conference on Robotics and Automation*.

[Koller and Pfeffer, 1997] Koller, D. and Pfeffer, A. (1997). Object-oriented bayesian networks. In *Proceedings of the Thirteenth Conference on Uncertainty in Artificial Intelligence (UAI-97)*, pages 302–313.

[Kong, 1986] Kong, A. (1986). *Multivariate Belief Functions and Graphical Models*. PhD thesis, Department of Statistics, Harvard University, Massachussetts.

[Kschischang et al., 2001] Kschischang, F., Frey, B., and Loeliger, H. (2001). Factor graphs and the sum-product algorithm. *IEEE Transactions on Information Theory*.

[Lauritzen, 1982] Lauritzen, S. (1982). *Lectures on Contengency Tables*. University of Aalborg Press, Aalborg, Denmark, 2 edition.

[Lauritzen, 1988] Lauritzen, S. (1988). Local computation with probabilities on graphical structures and their application to expert systems. *Journal of the Royal Statistical Society*, 50 :157.

[Lauritzen, 1996] Lauritzen, S. (1996). *Graphical Models*. Clarendon, Oxford, UK.

[Lauritzen et al., 1990] Lauritzen, S., Dawid, A., Larsen, B., and Leimer, H. (1990). Independance properties of directed markov fields. *Networks*, 20 :491–505.

[Lauritzen and Wermuth, 1989] Lauritzen, S. and Wermuth, N. (1989). Graphical models for associations between variables, some of which are qualitative and some quantitative. *Annals of Statictics*, 17 :31–57.

[Lin, 1996] Lin, F. (1996). Embracing causality in specifying the indeterminate effects of actions. In *AAAI'96*, pages 670–676.

[Ling and Rudd, 1988] Ling, X. and Rudd, W. (1988). Combining opinions from several experts. *Applied Artificial Intelligence*, 3 :439–452.

[Llinas and Antony, 1993] Llinas, J. and Antony, R. (1993). Blackboard Concepts for Data Fusion Applications. *Int. Journal of Pattern Recognition and Artificial Intelligence*, 7(2) :285–308.

[Lucas, 1998] Lucas, P. (1998). Analysis of notions of diagnosis. *Artificial Intelligence*, 105 :295–343.

[Luo and Kay, 1989] Luo, R. and Kay, M. (1989). Multisensor integration and fusion in intelligent systems. *IEEE Trans. on Systems, Man, and Cybernetics*, 19(5) :901–931.

[Madsen and Jensen, 1998] Madsen, A. and Jensen, F. (1998). Lazy propagation in junction trees. In Cooper, G. and Moral, S., editors, *Proceedings of the 14th Annual Conference on Uncertainty in Artificial Intelligence*, pages 362–369, San Francisco, California. Morgan Kaufman.

[Martin and Moravec, 1996] Martin, M. and Moravec, H. (1996). Robot evidence grids. Technical Report CMU-RI-TR-96-06, Carnegie Mellon University, The Robotics Institute Carnegie Mellon University Pittsburgh, Pennsylvania 15213.

[Maybeck, 1990] Maybeck, P. (1990). The kalman filter : An introduction to concepts. pages 194–204, New York, NY, USA. Springer-Verlag.

[McMichael et al., 1996] McMichael, D., Halgamuge, S., Hamlyn, G., Karan, M., and Okello, N. (1996). *Multisensor Data Fusion*. Lecture Notes.

[Meek et al., 2002] Meek, C., Chickering, D., and Heckerman, D. (2002). Autoregressive tree models for time-series analysis. In *Proceedings of the Second International SIAM Conference on Data Mining*, pages 229–244, Arlington, VA.

[Moravec, 1987] Moravec, H. (1987). Sensor fusion in certainty grids for mobile robots. pages 253–276.

[Murphy, 1999] Murphy, K. (1999). Filtering, smoothing and the junction tree algorithm. Technical report, University of Berkeley.

[Murphy, 2002] Murphy, K. (2002). *Dynamic Bayesian Networks : Representation, Inference and Learning*. PhD thesis, University of California, Berkeley.

[Murphy and Paskin, 2001] Murphy, K. and Paskin, M. (2001). Linear time inference in hierarchical hmms. In *Proceedings of the NIPS'01 Conference*.

[Naumann, 1998] Naumann, F. (1998). Data fusion and data quality. In *In Proc. of the New Techniques and Technologies for Statistics Seminar (NTTS)*, Sorrento, Italy.

[Olmsted, 1983] Olmsted, S. (1983). *On Representing and Solving Decision Problems*. PhD thesis, Department of Engineering-Economic Systems, Standford University, Stanford, California.

[Organization, 1997] Organization, W. H. (1997). Technical report, http ://www.who.int.

[Parsing, 1999] Parsing, S. (1999). Computational linguistics. pages 573–605.

[Pearl, 1982] Pearl, J. (1982). Reverand bayes on inference engines : A distributed hierarchical approach. In *Proceedings of the AAAI National Conference on AI*, pages 133–136, Pittsburgh.

[Pearl, 1988] Pearl, J. (1988). *Probabilistic reasoning in intelligent systems : Networks of plausible inference*. Morgan Kaufman Publishers, Inc., San Mateo,CA, 2nd edition.

[Pearl, 2001] Pearl, J. (2001). *Causality - Models, reasoning and inference*. Cambrige University Press.

[Poole, 1994] Poole, D. (1994). Representing diagnosis knowledge. *Ann. Math. Artificial Intelligence*, 11 :33–50.

[Poole et al., 1987] Poole, D., Goebel, R., and Aleliunas, R. (1987). *Theoris : a logical reasoning system for defaults and diagnosis*. Springer, Berlin.

[Rabiner, 1989] Rabiner, L. (1989). A tutorial on hidden Markov models and selected applications in speech recognition. In *Proceedings of the IEEE*, volume 77, pages 257–285.

[Rao, 1991] Rao, B. (1991). *Data Fusion Methods in Decentralized Sensing Systems*. PhD thesis, Dept. of Engineering Sciences, Oxford University.

[Shachter et al., 1994] Shachter, R., Andersen, S., and Szolovits, P. (1994). *Global conditionning for probabilistic inference in belief networks*, pages 514–524. Morgan Kaufman, San Francisco.

[Shafer, 1976] Shafer, G. (1976). *A mathematical theory of evidence*. Princeton University Press, Princeton, N.J.

[Shenoy, 1997] Shenoy, P. (1997). Binary join trees for computing marginals in the shenoy-shafer architecture. *International Journal of Approximate Reasoning*, 17 :239–263.

[Smets, 1988] Smets, P. (1988). *Belief functions. Non-Standard Logics for Automated Reasoning*. Academic Press, San Diego.

[Smyth et al., 1996] Smyth, P., Heckerman, D., and Jordan, M. (1996). Probabilistic Independance Networks for Hidden Markov Probability Models. Technical Report MSR-TR-96-03, Microsoft Research.

[Thomesse et al., 2002] Thomesse, J., Bellot, D., Boyer, A., Campo, E., Chan, M., Charpillet, F., Esteve, D., Fayn, J., Leschi, C., Noury, N., Rialle, V., Romary, L., Rubel, P., and Steenkeste, F. (2002). TIISSAD, Technologies de l'Information Intégrées aux Services de Soins à Domicile. Technical Report Décision d'aide 99 B0611 à 616, LORIA. Compte rendu de fin de recherche.

[Thomesse et al., 2001] Thomesse, J., Bellot, D., Boyer, A., Campo, E., Chan, M., Charpillet, F., Fayn, J., Leschi, C., Noury, N., Rialle, V., Romary, L., Rubel, P., Selmaoui, N., Steenkeste, F., and Virone, G. (2001). Integrated Information Technologies for patients remote follow-up and homecare. In *HealthCom 2001*.

[Thrun and Bücken, 1996] Thrun, S. and Bücken, A. (1996). Integrating grid-based and topological maps for mobile robot navigation. In AAAI, editor, *Proceedings of the Thirteenth National Conference on Artificial Intelligence*, Portland, Oregon. AAAI, AAAI.

[Thrun et al., 1998] Thrun, S., Gutmann, J., Fox, D., Burgard, W., and Kuipers, B. (1998). Integrating topological and metric maps for mobile robot navigation : A statistical approach. In *Proceedings of AAAI-98*. AAAI. http ://www.cs.cmu.edu/ thrun/papers/full.html.

[Verma and Pearl, 1988] Verma, T. and Pearl, J. (1988). Causal networks : Semantics and expressiveness. In *Proceedings of the 4th Workshop on Uncertainty in Artificial Intelligence*, pages 352–359, Mountain View, CA.

[Viterbi, 1967] Viterbi, A. (1967). Error bounds for convolutional codes and an asymptotically optimum decoding algorithm. *IEEE Trans. on Information Theory*, pages 260–269.

[Waltz and Llinas, 1990] Waltz, E. and Llinas, J. (1990). *Multisensor Data Fusion*. Boston-London.

[Wand and Wang, 1996] Wand, Y. and Wang, R. (1996). Anchoring data quality dimensions in ontological foundations. *Communications of the ACM*.

[Wang et al., 1996] Wang, R., Strong, D., and Guarascio, L. (1996). Beyond accuracy : What data quality means to data consumers.

[Wermuth and Lauritzebn, 1983] Wermuth, N. and Lauritzebn, S. (1983). Graphical and recursive models for contingency tables. *Biometrika*, 70 :37–52.

[Wu et al., 2001] Wu, X., Lucas, P., Kerr, S., and Dijkhuizen, R. (2001). Learning bayesian-network topologies in realistic medical domains. pages 302–308.

[Xu et al., 1992] Xu, L., Krzyzak, A., and Suen, C. (1992). Methods of combining multiple classifiers and their application to handwriting recognition. *IEEE Trans. on Systems, Man, and Cybernetics*, 22(3) :418–435.

[Yamauchi, 1995] Yamauchi, B. (1995). *Exploration and spatial learning in dynamic environments*. PhD thesis, Case Western Reserve University, Department of Computer Engineering and Science.

[Yannakakis, 1981] Yannakakis, M. (1981). Computing the minimum fill-in is np-complete. *SIAM Journal on Algebraic and Discrete Methods*, 2 :77–79.

[Zilberstein, 1995] Zilberstein, S. (1995). Models of Bounded Rationality : a concept paper. In *AAAI Fall Symposium on Rational Agency*, Cambridge, Massachusetts.

[Zweig and Russell, 1997] Zweig, G. and Russell, S. (1997). Compositional modeling with DPNs. Technical Report CSD-97-970.

Résumé

Cette thèse présente une nouvelle approche de la fusion de données et applique ces notions à la modélisation et au diagnostic probabiliste dans le cadre de la télémédecine. Notre contribution se situe au niveau de la définition d'une notion de gain dans un processus de fusion de données, ainsi que de l'application des réseaux bayésiens dynamiques au diagnostic en télémédecine. Le but final est de réguler, à distance, l'état physiologique d'un patient.

Une première étude du domaine de la fusion de données a permis d'exhiber les concepts de base de la fusion de données : processus de fusion de données et notion de gain qualifié. Elle a aussi permi de structurer et de typer les sources de données et les résultats d'un processus de fusion. Cette approche a servi de cadre général à la seconde partie de la thèse qui a portée sur la modélisation et le diagnostic médical dans le cadre d'une application de télémédecine. Il s'agit typiquement d'un problème où interviennent plusieurs sources de données incertaines et hétérogènes. Le projet Diatelic^{TM}d'assistance à domicile de personnes souffrant d'insuffisance rénale, vise à monitorer l'état d'hydratation d'un patient subissant une dialyse péritonéale. Les données physiologiques recueillies quotidiennement auprès du patient sont incertaines, hétérogènes et bruitées.

Les réseaux bayésiens dynamiques permettent de modéliser des dépendances causales, typiques de la connaissance médicale, mais aussi de gérer efficacement le problème de l'incertitude à travers le formalisme probabiliste. Le modèle à base de réseaux bayésiens permet une fusion efficace : notre but est de maximiser le gain en certitude, c'est-à-dire de détecter avec la plus grande confiance possible l'état d'hydratation du patient à partir des informations fournies par les capteurs. Ce travail théorique a donné lieu à l'implémentation d'un moteur d'inférence bayésienne, permettant d'expérimenter nos modèles. Une première version du système Diatelic^{TM}(v3) a été réalisée.

Le processus de fusion que nous modélisons permet une prise de décision plus efficace par le médecin car il indique avec précision l'état physiologique du patient. On peut ainsi réguler son état de santé et éviter les aggravations médicales. Ce travail s'est ouvert à d'autres problématiques : adaptation en-ligne de modèles, quantification du gain et prise en compte de multiples échelles de temps dans un réseau bayésien dynamique.

Mot-clés : fusion de données, réseaux bayésiens dynamiques, télémédecine

Abstract

This thesis presents a new approach of data fusion and applies these notions to the modelisation of probabilistic diagnosis in telemedicine. Our contribution is a new definition of qualified gain in a data fusion process, and an application of dynamic bayesian network to telemedicine diagnosis. The final goal is to remotely regulate the physiological state of a patient.

A first study of the field has shown basic data fusion concepts : data fusion process and qualified gain. Structures and types of data sources and results has emerged during this study too. This approach forms a general framework for the second part of this thesis : modelizing and medical diagnosis for a telemedicine application. This is a typical problem where multiple, uncertain and heterogeneous data sources are needed. The Diatelic^{TM}project's goal is to assist kidney disease people at home by monitoring their hydration rate. Physiological data are uncertain, noisy and heterogeneous.

Dynamic bayesian networks are used to modelize causal dependancies, which are typical of medical knowledge. They can deal with uncertainty by using the strong probabilistic formalism. A bayesian network model allows us to do efficient data fusion, in particular, by maximizing the certainty gain, i.e. by detecting hydration rate problems with the highest confidence. For this purpose, we are implemented a bayesian engine, to deal with our experiments. The Diatelic^{TM}(v3) has been implemented with it.

The physician is able to take the right decision using our data fusion process, because he/she has a precision estimation of the hydration rate of the patient. Health state of the patient could be regulated through the use of this system. New problems have arise during this PhD thesis work : on-line models adaptation, quantifying the data fusion gain and dealing with multiple time-scale bayesian networks.

Keyworks : data fusion, dynamic bayesian networks, telemedicine