THÈSE

présentée à

L'Institut National Polytechnique de Grenoble - I.N.P.G.
Laboratoire LEIBNIZ - IMAG

pour obtenir le titre de

DOCTEUR

Spécialité:

Informatique
(Arrêté ministériel du 30 mars 1992)

INSS : UN SYSTEME HYBRIDE NEURO-SYMBOLIQUE POUR L'APPRENTISSAGE AUTOMATIQUE CONSTRUCTIF

par

Fernando Santos Osório

Soutenue le 3 février 1998 devant le jury composé de:

MM.	Philippe JORRAND	Directeur
	Alain GRUMBACH	Rapporteur
	Christian PELLEGRINI	Rapporteur, Président
	Bernard AMY	Examinateur (co-directeur)
	Frédéric ALEXANDRE	Examinateur
	Vincent RIALLE	Examinateur

A Denise et Felipe,

mes deux compagnons de toutes les heures

embarqués dans cette aventure avec moi.

Remerciements:

Je tiens tout d'abord à remercier Bernard Amy, qui m'a accepté dans son équipe, a consacré beaucoup de son temps à mon travail, et a fait de nombreuses remarques, toujours intéressantes, sur cette thèse. Bernard a toujours su diriger mes recherches dans la bonne direction; je dois beaucoup à lui pour ses nombreuses contributions à la réussite de cette thèse.

Je tiens à remercier aussi Alain Grumbach, Professeur à l'ENST de Paris, et à Christian Pellegrini, Professeur à l'Université de Genève, pour avoir accepté de juger ce travail.

Je tiens également à remercier Vincent Rialle, Maître de Conférences à l'Université Joseph Fourrier de Grenoble, et à Fréderic Alexandre, Chercheur à l'INRIA Lorraine, qui ont gentiment accepté de participer à mon jury de thèse.

Merci beaucoup à Philippe Jorrand, Directeur du Laboratoire Leibniz - IMAG/INPG, qui a accepté de m'accueillir dans ce laboratoire, et aussi, qui a accepté de passer toutes ces années à mon côté, comme mon directeur de thèse.

Je remercie à la CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nivel Superior) et plus particulièrement au programme de coopération Franco-Brésilien CAPES-COFECUB, de m'avoir accordé un financement pour la réalisation de cette thèse. Le financement ainsi que le support administratif fourni par ces deux organismes a été une pièce fondamentale pour permettre mon séjour en France et la réalisation de cette thèse. Merci à Brigitte Plateau et Claudio Geyer, coordinateurs du projet de coopération CAPES-COFECUB entre l'UFRGS et l'INPG, de m'avoir permis de participer de son projet de coopération. Je tiens aussi à remercier particulièrement à Philippe Navaux de l'UFRGS.

Je remercie aussi à l'UNISINOS (Universidade do Vale do Rio dos Sinos, Brésil), pour m'avoir accordé un détachement de mes fonctions d'enseignant de 1993 à 1998, et principalement, pour son soutien constant qui m'a permis de mener mes travaux de doctorat jusqu'à la fin.

Je tiens aussi à remercier :

Les membres des laboratoires LIFIA et Leibniz qui m'ont beaucoup aidé dans la réalisation de cette thèse. Plus particulièrement, je remercie aux 'anciens de l'équipe' Arnaud Giacometti, Bruno Orsier, Abderrahim Labbi, Nicolas Szilas, Maria Malek, Françoise Fessant et Emmanuel Duriez pour leurs inestimables contributions sur le plan scientifique et aussi pour leur amitié. Je remercie aussi à tous ceux qui ont travaillé a mon côté dans le développement et l'utilisation du système INSS (Loïc Decloedt, Pierre-Yves Blond, Gerardo Reyes et Benjamin Deuker), ainsi qu'à Vincent Danel et Vincent Rialle pour leurs contributions dans l'application médicale. Je remercie particulièrement à Henri Saya, notre ingénieur système, à Jaime Sichman et à Remis Balaniuk, qui m'ont beaucoup

aidé à "m'insérer" dans ce laboratoire d'un point de vue technique et aussi humain. Je remercie ~~beaucop~~ beaucoup ;^) aux *héros* qui ont fait la révision de ce document, les critiques sur son contenu, et la correction de mes innombrables fautes commises dans la langue de Molière: Nico, Eric Gauthier, Daniel Memmi, Sophie Rouzier et Bernard.

Tous les membres de l'équipe RESEAUX dirigée par Bernard Amy au laboratoire Leibniz pour leur accueil très chaleureux et leur disponibilité constante, avec qui j'ai eu des rapports aussi divers qu'enrichissants. *Une thèse est principalement le résultat d'un vrai travail d'équipe!*

Merci aussi à tous mes amis des autres groupes de recherche de Grenoble : l'équipe LAPLACE (en particulier à Olivier Lebeltel pour sa inestimable aide avec le petit Khepera), les équipes MAGMA et PRIMA, les amis des laboratoires TIMA, TIMC et LMC, aux membres de l'association InCognito, et tant d'autres compagnons de recherche et d'amitié.

Merci à Scott Fahlman, à Geoffrey Towell, à Olivier Michel, et à tous les autres chercheurs qui ont mis leurs outils, articles et travaux accessibles à travers l'Internet, et merci à tous les membres du projet Européen MIX. Vos contributions ont beaucoup apporté au développement de cette thèse.

Et bien sûr, merci à toute la communauté brésilienne de Grenoble et à tous mes amis avec qui j'ai garde des souvenirs très agréables des bonnes moments vécus ensemble. Nos échanges de messages (et des "flames") à travers la "penduick-fm" resteront à jamais enregistrés dans la mémoire non-volatile des nos âmes... "SaluTCHE et Obrigado!".

Parmi la communauté brésilienne de Grenoble, je garderai un souvenir très particulier de tous ceux qui ont été toujours prêts à m'encourager dans ce long parcours : l'unique et inestimable João Paulo Kitajima, Silvio Nabeta, Ricardo Duarte et Silvana, Alvaro Guarda et Vera, Gilson Braviano et Teresinha, Paulo Fernandes et Marilia, José Celso et Néia, Gilberto Marchioro (Giba) et Camila, Alexandre Ribeiro et Helena.

Je remercie aussi à tous ceux et celles qui sont là-bas de l'autre côté du monde, au Brésil, et qui ont donné leur soutien à la réalisation de cette thèse. Merci particulièrement à mes beaux-parents Egydio et Eneida, à mes grands-parents Luiz Carlos et Maria, et à ma mère Beatriz, je n'aurais pas pu arriver jusqu'ici sans vous.

Finalement, merci, merci et encore merci à mon épouse qui a toujours su me soutenir (dans tous les sens de ce mot, et principalement dans le sans brésilien - "suportar, agüentar") et me donner les forces nécessaires pour arriver jusqu'à la fin de cette thèse. Je n'aurais jamais écrit chacune des 300 pages qui composent cette thèse sans son inestimable aide et amour. DeniSe, ma vraie DS à moi, tu es for me... formidable!

Merci à Denise et Felipe, sources de mon inspiration.

Merci à toute ma famille et plus particulièrement à mes parents Beatriz et Luiz Carlos.

Résumé:

Plusieurs méthodes ont été développées par l'Intelligence Artificielle pour reproduire certains aspects de l'intelligence humaine. Ces méthodes permettent de simuler les processus de raisonnement en s'appuyant sur les connaissances de base disponibles. Chaque méthode comporte des points forts, mais aussi des limitations. La réalisation de systèmes hybrides est une démarche courante qui permet de combiner les points forts de chaque approche, et d'obtenir ainsi des performances plus élevées ou un champ d'application plus large. Un autre aspect très important du développement des systèmes hybrides intelligents est leur capacité d'acquérir de nouvelles connaissances à partir de plusieurs sources différentes et de les faire évoluer.

Dans cette thèse, nous avons développée des recherches sur les systèmes hybrides neuro-symboliques, et en particulier sur l'acquisition incrémentale de connaissances à partir de connaissances théoriques (règles symboliques) et de connaissances empiriques (exemples). Un nouveau système hybride, nommé système INSS - *Incremental Neuro-Symbolic System,* a été étudié et réalisé. Ce système permet le transfert de connaissances déclaratives (règles symboliques) d'un module symbolique vers un module connexionniste (réseau de neurones artificiel - RNA) à travers un convertisseur de règles en réseau. Les connaissances du réseau ainsi obtenu sont affinées par un processus d'apprentissage à partir d'exemples. Ce raffinement se fait soit par ajout de nouvelles connaissances, soit par correction des incohérences, grâce à l'utilisation d'un réseau constructif de type Cascade-Correlation. Une méthode d'extraction incrémentale de règles a été intégrée au système INSS, ainsi que des algorithmes de validation des connaissances qui ont permis de mieux coupler les modules connexionniste et symbolique. Le système d'apprentissage automatique INSS a été conçu pour l'acquisition constructive (incrémentale) de connaissances. Le système a été testé sur plusieurs applications, en utilisant des problèmes académiques et des problèmes réels (diagnostic médical, modélisation cognitive et contrôle d'un robot autonome). Les résultats montrent que le système INSS a des performances supérieures et de nombreux avantages par rapport aux autres systèmes hybrides du même type.

Abstract:

Various Artificial Intelligence methods have been developed to reproduce intelligent human behaviour. These methods allow to reproduce some human reasoning process using the available knowledge. Each method has its advantages, but also some drawbacks. Hybrid systems combine different approaches in order to take advantage of their respective strengths. These hybrid intelligent systems also present the ability to acquire new knowledge from different sources and so to improve their application performance.

This thesis presents our research in the field of hybrid neuro-symbolic systems, and in particular the study of machine learning tools used for constructive knowledge acquisition. We are interested in the automatic acquisition of theoretical knowledge (rules) and empirical knowledge (examples). We present a new hybrid system we implemented: INSS - Incremental Neuro-Symbolic System. This system allows knowledge transfer from the symbolic module to the connectionist module (Artificial Neural Network - ANN), through symbolic rule compilation into an ANN. We can refine the initial ANN knowledge through neural learning using a set of examples. The incremental ANN learning method used, the Cascade-Correlation algorithm, allows us to change or to add new knowledge to the network. Then, the system can also extract modified (or new) symbolic rules from the ANN and validate them. INSS is a hybrid machine learning system that implements a constructive knowledge acquisition method. We conclude by showing the results we obtained with this system in different application domains: ANN artificial problems (The Monk's Problems), computer aided medical diagnosis (Toxic Comas), a cognitive modelling task (The Balance Scale Problem) and autonomous robot control. The results we obtained show the improved performance of INSS and its advantages over others hybrid neuro-symbolic systems.

TABLE DES MATIERES

1. INTRODUCTION 3

2. APPRENTISSAGE AUTOMATIQUE .....

2.1 LES SYSTEMES EXPERTS ET L'ACQUISITION DE CONNAISSANCES 6
2.1.1 Les Systèmes Experts 8
2.1.2 La Représentation des Connaissances 12
2.1.2.1 Approches Connexionnistes 15
2.1.2.2 Approches Symboliques 17
2.1.2.2.a Description de Cas Pratiques 17
2.1.2.2.b Règles de Production et Formules Logiques 18
2.1.2.2.c Réseaux Sémantiques 20
2.1.2.2.d Objets Structurés et Frames 20
2.1.2.2.e Méta-Connaissances 21
2.1.3 L'Acquisition de Connaissances 22
2.1.3.1 Méthodes d'Apprentissage Empirique 24
2.1.3.1.a Apprentissage par Analogie 24
2.1.3.1.b Apprentissage par Induction 25
2.1.3.2 Méthodes d'Apprentissage fondées sur l'Explication 25
2.1.3.3 Considérations Finales sur l'Acquisition de Connaissances 26
2.2 APPRENTISSAGE AUTOMATIQUE - APPROCHES SYMBOLIQUES 27
2.2.1 Arbres de Décision 27
2.2.1.1 L'Apprentissage dans les Arbres de Décision 27
2.2.1.2 Les Systèmes fondés sur les Arbres de Décision 31
2.2.1.3 Discussion sur les Arbres de Décision 34
2.2.2 Les Algorithmes Génétiques 35
2.2.2.1 Principes de Base 35
2.2.2.2 Induction de Règles à partir d'Exemples 36
2.2.2.3 Discussion sur les Algorithmes Génétiques 37
2.2.3 Le Raisonnement Fondé sur des Cas - CBR 38
2.2.3.1 Principes de Base 38
2.2.3.2 Discussion sur le Raisonnement Fondé sur des Cas 39
2.3 APPRENTISSAGE AUTOMATIQUE - APPROCHES CONNEXIONNISTES 40
2.3.1 Classification et Propriétés 43
2.3.1.1 Apprentissage Connexionniste 43
2.3.1.2 Type des Unités 47
2.3.1.3 Type d'Architecture du Réseau 49
2.3.1.4 Type d'Application du Réseau 53
2.3.2 Le modèle du Perceptron Multi-Couches 54
2.3.3 Discussion sur les Réseaux Connexionnistes 61
2.4 VERS LES SOLUTIONS HYBRIDES 65
2.4.1 Systèmes et Méthodes Hybrides d'Apprentissage Automatique 66
2.4.2 Vers les Systèmes Hybrides Neuro-Symboliques 70

3. SYSTEMES HYBRIDES NEURO-SYMBOLIQUES 72

3.1 GENERALITES 72
3.2 TYPES D'INTEGRATION ET CLASSIFICATION DES SHNS 74
3.2.1 Types d'Intégration 75
3.2.2 Degré de Couplage 77
3.2.3 Mode d'Intégration 78
3.2.4 Transferts de Connaissances 79
3.2.5 Type de Représentation des Connaissances 81
3.2.6 Type de Codage des Connaissances 82
3.2.7 Type de Méthodes de Raisonnement 83
3.2.8 Mode d'Acquisition de Connaissances 83
3.2.9 Classification des SHNS 84
3.2.10 Discussion sur les SHNS 89
3.3 REALISATIONS DE SYSTEMES HYBRIDES 95
3.3.1 Système SYNHESYS 95
3.3.1.1 Module Symbolique du Système SYNHESYS 96
3.3.1.2 Module Connexionniste du Système SYNHESYS 97

3.3.1.3 Transfert de Connaissances dans le Système SYNHESYS ... 100
3.3.1.4 Coopération entre modules dans le Système SYNHESYS ... 102
3.3.1.5 Discussion sur le modèle de SHNS proposé par SYNHESYS ... 105
3.3.2 Réseaux KBANN ... 107
3.3.2.1 Compilation de Règles dans KBANN ... 108
3.3.2.2 L'Apprentissage dans les Réseaux KBANN ... 116
3.3.2.3 Extraction de Règles des Réseaux KBANN ... 118
3.3.2.3.a Algorithme SUBSET d'extraction de règles ... 120
3.3.2.3.b Algorithme NofM d'extraction de règles ... 122
3.3.2.4 Discussion sur les Réseaux KBANN ... 127
3.3.3 Systèmes Connexionnistes de Raffinement de Connaissances ... 130
3.4 CONCLUSION SUR L'ETUDE DES SHNS ... 134

4. SYSTEME INSS ... 135

4.1 MODULE SYMBOLIQUE ... 139
4.2 INSERTION DE CONNAISSANCES A PRIORI ... 141
4.2.1 Compilation de Règles d'Ordre 0^+ ... 144
4.2.2 Discussion sur la Compilation de Règles ... 155
4.3 MODULE CONNEXIONNISTE ... 156
4.3.1 Méthode d'Apprentissage Cascade-Correlation ... 157
4.3.2 Avantages de l'Utilisation de l'Algorithme Cascade-Correlation ... 161
4.3.3 L'Algorithme Cascade-Correlation dans NeuSim ... 162
4.3.4 Le Module NeuSim en tant que simulateur de réseaux ... 164
4.3.5 Discussion sur le Module Connexionniste ... 169
4.4 EXTRACTION DE CONNAISSANCES ... 170
4.4.1 Sélection des Unités pour l'Extraction de Règles ... 172
4.4.2 Sélection des Connexions pour l'Extraction de Règles ... 175
4.4.3 Application de l'algorithme SUBSET dans INSS ... 177
4.4.4 Discussion sur le Module d'Extraction de Règles ... 179
4.5 VERIFICATION ET VALIDATION DE CONNAISSANCES ... 181
4.5.1 Discussion sur le Module de Vérification et Validation de Connaissances ... 185
4.6 CONSIDERATIONS FINALES SUR LE SYSTEME INSS ET L'ACQUISITION CONSTRUCTIVE DE CONNAISSANCES ... 187

5. APPLICATIONS ... 191

5.1 LES PROBLEMES DU MOINE (THE MONK'S PROBLEMS) ... 192
5.1.1 Description du Problème ... 192
5.1.2 Les Résultats Expérimentaux obtenus avec INSS ... 194
5.1.3 Discussion Finale sur l'étude des Problèmes du Moine ... 198
5.2 DIAGNOSTIC MEDICAL - COMAS TOXIQUES ... 199
5.2.1 Description du Problème ... 199
5.2.2 Description des données disponibles sur le problème ... 200
5.2.3 Les Résultats Expérimentaux obtenus avec INSS ... 203
5.2.4 Discussion Finale sur l'Application Médicale ... 211
5.3 PROBLEME DE LA BALANCE ... 212
5.3.1 Description du Problème ... 213
5.3.2 Les Résultats Expérimentaux Obtenus avec INSS ... 216
5.3.3 Discussion Finale sur le Problème de la Balance ... 222
5.4 ROBOTIQUE AUTONOME ... 223
5.4.1 Description de l'Environnement de Tests Utilisé ... 225
5.4.1.1 Le Robot Khepera ... 225
5.4.1.2 Le Simulateur de Khepera ... 227
5.4.2 Les Résultats Expérimentaux Obtenus avec INSS ... 230
5.4.3 Discussion Finale sur l'Application en Robotique Autonome ... 241
5.5 CONSIDERATIONS FINALES ... 242

6. CONCLUSIONS ET PERSPECTIVES ... 244

7. REFERENCES BIBLIOGRAPHIQUES ... 247

7.1 BIBLIOGRAPHIE ... 272

8. ANNEXE A ... 283

9. ANNEXE B ... 285

10. GLOSSAIRE ET ABREVIATIONS 299

TABLE DES FIGURES

FIGURE 1 - L'ARCHITECTURE D'UN SYSTÈME EXPERT **9**
FIGURE 2 - ACQUISITION DE CONNAISSANCES....11
FIGURE 3 - DIFFÉRENTES FAÇONS DE REPRÉSENTER LES CONNAISSANCES SUR LA FONCTION LOGIQUE XOR....13
FIGURE 4 - LES RÉSEAUX CONNEXIONNISTES....16
FIGURE 5 - *FRAME* DÉCRIVANT LE CONCEPT DE VIN....21
FIGURE 6 - EXEMPLE D'ARBRE DE DÉCISION SIMPLE....28
FIGURE 7 - EXEMPLE D'ARBRE DE DÉCISION COMPLEXE....30
FIGURE 8- EXEMPLE DE NEURONE BIOLOGIQUE....41
FIGURE 9- EXEMPLE DE NEURONE FORMEL....42
FIGURE 10- APPRENTISSAGE : ERREUR DANS L'ENSEMBLE D'APPRENTISSAGE ET DE TEST....45
FIGURE 11 - PROTOTYPES D'UN RÉSEAUX À DEUX ENTRÉES....48
FIGURE 12 - SÉPARATION DES CLASSES PAR UN PERCEPTRON À DEUX ENTRÉES....49
FIGURE 13 - ARCHITECTURES D'INTERCONNEXION DES UNITÉS D'UN RÉSEAU....51
FIGURE 14- SIGMOÏDE ASYMÉTRIQUE....54
FIGURE 15 - DESCENTE DE GRADIENT SUR UNE SURFACE D'ERREUR ET MINIMA LOCAUX....56
FIGURE 16 - DÉTAIL DU PROCESSUS DE RÉTRO-PROPAGATION....61
FIGURE 17 - INTÉGRATION NEURO-SYMBOLIQUE....75
FIGURE 18 - MODES D'INTÉGRATION NEURO-SYMBOLIQUE....78
FIGURE 19 - ARCHITECTURE DU SYSTÈME SYNHESYS....96
FIGURE 20 - EXEMPLE TYPIQUE D'UNE BASE DE RÈGLES SYMBOLIQUES DE SYNHESYS....97
FIGURE 21 - PROTOTYPES DANS UN RÉSEAUX À HYPER-RECTANGLES AVEC DEUX ENTRÉES....98
FIGURE 22 - ARCHITECTURE MULTICOUCHES DU RÉSEAU À PROTOTYPES UTILISÉ DANS SYNHESYS....99
FIGURE 23 - L'ACQUISITION INTERACTIVE DE CONNAISSANCES FAIT PAR SYNHESYS....104
FIGURE 24 - SCHÉMA DE L'INTÉGRATION NEURO-SYMBOLIQUE UTILISÉ DANS LES RÉSEAUX KBANN....108
FIGURE 25 - SYNTAXE ET EXEMPLE DE RÈGLE SYMBOLIQUE UTILISÉS DANS KBANN....110
FIGURE 26 - CORRESPONDANCE ENTRE LA STRUCTURE D'UNE BASE DE RÈGLES ET D'UN RÉSEAU DE NEURONES....111
FIGURE 27 - OBTENTION DES POIDS ET DES SEUILS DES UNITÉS D'UN RÉSEAU KBANN....113
FIGURE 28 - PROCESSUS DE TRANSFORMATION DE RÈGLES DANS UN RÉSEAU KBANN....114
FIGURE 29 - EXEMPLE DE CODAGE D'ATTRIBUTS NUMÉRIQUES DANS KBANN....115
FIGURE 30 - EXEMPLE D'APPLICATION DE L'ALGORITHME SUBSET....122
FIGURE 31 - EXEMPLE D'APPLICATION DE L'ALGORITHME NOFM....126
FIGURE 32 - SCHÉMA GLOBAL DU SYSTÈME HYBRIDE INSS....136
FIGURE 33 - L'ACQUISITION CONSTRUCTIVE DE CONNAISSANCES DANS LE SYSTÈME INSS....137
FIGURE 34 - LES CONNAISSANCES DANS LE SYSTÈME INSS....138
FIGURE 35 - EXEMPLE DE PROGRAMME EN CLIPS (FONCTION XOR)....140
FIGURE 36 - RÈGLES SYMBOLIQUES UTILISÉES PAR NEUCOMP....141
FIGURE 37 - APERÇU DE LA SYNTAXE UTILISÉE PAR NEUCOMP....142
FIGURE 38 - EXEMPLE DE FICHIER DE RÈGLES SYMBOLIQUES (STRUCTURES SYNTAXIQUES)....143
FIGURE 39 - FICHIER DE CONFIGURATION DU COMPILATEUR NEUCOMP....144
FIGURE 40 - FONCTION GREATER_THAN (TRAIT, VALEUR)....145
FIGURE 41 - FONCTION GREATER_THAN (TRAIT, TRAIT)....146
FIGURE 42 - FONCTION LESS_THAN (TRAIT, VALEUR)....147
FIGURE 43 - FONCTION LESS_THAN (TRAIT, TRAIT)....148
FIGURE 44 - FONCTION EQUAL (TRAIT, VALEUR)....149
FIGURE 45 - FONCTION EQUAL (TRAIT, VALEUR)....151
FIGURE 46 - FONCTION EQUAL (TRAIT, TRAIT)....152
FIGURE 47 - SORTIE DE LA FONCTION GREATER_THAN PAR RAPPORT À LA VALEUR DE LA SENSIBILITÉ....153
FIGURE 48 - SORTIE DES FONCTIONS GREATER_THAN ET IN_RANGE PAR RAPPORT À LA VALEUR DE LA CONFIANCE....154
FIGURE 49 - SORTIE DE LA FONCTION IN_RANGE PAR RAPPORT À LA VALEUR DE WF....154
FIGURE 50 - ARCHITECTURE DES RÉSEAUX CASCOR (ÉTAT INITIAL ET APRÈS L'AJOUT DE DEUX UNITÉS)....158

FIGURE 51 - L'ALGORITHME CASCOR ET L'ÉVOLUTION DE LA STRUCTURE D'UN RÉSEAU DANS INSS....163
FIGURE 52 - FICHIERS UTILISÉS PAR NEUSIM....165
FIGURE 53 - EXEMPLE D'UN FICHIER DE CONFIGURATION DE NEUSIM....166
FIGURE 54 - SÉLECTION DES UNITÉS BASÉE SUR LA POURCENTAGE D'ACTIVATION....174
FIGURE 55 - SÉLECTION DES CONNEXIONS LES PLUS FAIBLES....176
FIGURE 56 - PROCESSUS DE VÉRIFICATION ET DE VALIDATION DE CONNAISSANCES....183
FIGURE 57 - DESCRIPTION DU PROBLÈME DU MOINE: "MONK'S PROBLEM 1"....193
FIGURE 58 - MONK1 : TAUX DE GÉNÉRALISATIONS DU MS, MC ET SYSTÈME HYBRIDE....196
FIGURE 59 - PROBLÈME MONK1 : RÈGLES EXTRAITES....197
FIGURE 60 - RÈGLES OBTENUES À PARTIR D'UN RÉSEAU (APPRENTISSAGE DE LA BASE COMA21-ADT)....207
FIGURE 61 - LE SYSTÈME EUROTOX-INSS, PROJET MIX....209
FIGURE 62 - FORMULAIRE ÉLECTRONIQUE POUR LA CONSULTATION DU SYSTÈME EUROTOX-INSS....210
FIGURE 63 - LES SOUS-PROBLÈMES DE LA BALANCE ET LES TAUX DE RÉUSSITE (EN %)....214
FIGURE 64 - EXEMPLE DE L'ÉVOLUTION DU TAUX DE BONNES RÉPONSES....220
FIGURE 65 - LE ROBOT KHEPERA....225
FIGURE 66 - EMPLACEMENT DES CAPTEURS SUR KHEPERA....226
FIGURE 67 - INTERFACE GRAPHIQUE DU SIMULATEUR DE KHEPERA....227
FIGURE 68 - EXEMPLE D'UN FICHIER D'APPRENTISSAGE UTILISÉ PAR NEUSIM....229
FIGURE 69 - SITUATIONS TYPIQUES D'ÉVITEMENT D'OBSTACLE....230
FIGURE 70 - EVITEMENT D'OBSTACLES : TRAJECTOIRE DU ROBOT APRÈS L'APPRENTISSAGE....231
FIGURE 71 - ENSEMBLE DE RÈGLES SYMBOLIQUES POUR L'ÉVITEMENT D'OBSTACLES....233
FIGURE 72 - EVOLUTION DU TAUX DE RÉUSSITE PENDANT L'APPRENTISSAGE (RÈGLES + EXEMPLES)....235
FIGURE 73 - EVITEMENT D'OBSTACLES : UTILISATION D'UNE BASE DE RÈGLES....237
FIGURE 74 - APPRENTISSAGE D'UN COMPORTEMENT DE SUIVI DU CONTOUR D'UN MUR....238
FIGURE 75 - TRAJECTOIRES DU ROBOT DANS LES ENVIRONNEMENTS DE TEST (SUIVI DU CONTOUR)....239
FIGURE 76 - ENVIRONNEMENT DE TEST AVEC LE ROBOT RÉEL....240
FIGURE 77 - TRAJECTOIRE DU ROBOT DANS UNE TÂCHE DE SUIVI DE CONTOUR DES MURS....241

1. INTRODUCTION

De nos jours, l'informatique à pris une place importante dans notre société. Les systèmes informatiques deviennent de plus en plus complexes, comme les tâches qui leur sont confiées. Les systèmes d'Intelligence Artificielle (I.A.) ont été créés avec l'espoir de pouvoir aider l'homme dans les tâches les plus complexes, celles qui requièrent l'utilisation de l'intelligence[1] ou la réalisation de procédés dits intelligents. Les résultats obtenus jusqu'à présent sont encore assez limités par rapport à notre conception d'un vrai système intelligent. A quelques rares exceptions près, l'homme reste toujours supérieur à la machine pour la réalisation de tâches complexes.

Plusieurs domaines de recherche en I.A. ont pour but la reproduction de certains aspects de l'intelligence humaine, et plusieurs méthodes ont été développées à cette fin. Ces méthodes nous permettent de simuler les processus de raisonnement humain (par exemple : l'inférence, l'analogie et la déduction) en s'appuyant sur les connaissances de base disponibles. Chaque méthode comporte des points forts, mais aussi des limitations. La réalisation de systèmes hybrides est une démarche courante qui nous permet de combiner les points forts de chaque approche, et d'obtenir ainsi des performances plus élevées ou un champ d'application plus large. Un autre aspect très important du développement des systèmes intelligents est leur capacité d'acquérir de nouvelles connaissances (parfois à partir de plusieurs sources différentes) et de les faire évoluer.

Dans cette thèse, nous avons développé des recherches sur les système hybrides neuro-symboliques (SHNS), et en particulier sur l'acquisition incrémentale de connaissances à partir de connaissances théoriques (règles symboliques) et de connaissances empiriques (exemples). Un nouveau système hybride, nommé système INSS - *Incremental Neuro-Symbolic System,* a été étudié et réalisé. Ce système permet le transfert de connaissances déclaratives (règles symboliques) d'un module symbolique vers un module connexionniste (réseau de neurones artificiel - RNA) à travers un convertisseur de règles en réseaux. Les connaissances du réseau

[1] La définition de "l'intelligence", ainsi que les tentatives de la quantifier, demeurent des sujets controversés. Nous considérons l'homme comme notre modèle d'intelligence, et pour nous, les systèmes dits "intelligents" sont ceux capables de reproduire certaines caractéristiques propres au raisonnement humain.

ainsi obtenu sont affinées par un processus d'apprentissage à partir d'exemples. Ce raffinement se fait soit par ajout de nouvelles connaissances, soit par correction des inconsistances, grâce à l'utilisation d'un réseau constructif. Une méthode d'extraction incrémentale de règles a été intégrée au système INSS, ainsi que des algorithmes de validation des connaissances qui ont permis de mieux coupler les modules connexionniste et symbolique. INSS a été conçu de façon à créer un système d'apprentissage automatique pour l'acquisition constructive (incrémentale) de connaissances. Le système a été testé sur plusieurs applications, en utilisant des problèmes académiques et des problèmes réels (diagnostic médical et contrôle d'un robot autonome).

Ce mémoire est organisé de la manière suivante :

- Le Chapitre 2 est consacré à une description des systèmes intelligents (systèmes experts) et des méthodes d'acquisition et de représentation des connaissances. Nous présentons plus particulièrement les systèmes d'apprentissage automatique, que nous avons divisés en deux grands groupes : les approches symboliques d'apprentissage et les approches connexionnistes d'apprentissage. Nous détaillons le mode de fonctionnement des principales méthodes d'apprentissage qui sont liées directement aux études développées dans le cadre de cette thèse. A l'issue de ce chapitre nous présentons une liste de solutions hybrides utilisées dans le domaine de l'apprentissage automatique, et nous expliquons notre choix des systèmes hybrides neuro-symboliques.

- Le Chapitre 3 est consacré à un état de l'art critique sur les systèmes hybrides neuro-symboliques. Nous présentons une liste d'exemples de ce type de systèmes, en les classifiant par rapport à leurs principales propriétés. Nous avons choisi d'étudier plus en détails deux d'entre eux en raison de leurs propriétés et de leurs caractéristiques : le système SYNHESYS et le système KBANN. Nous décrivons leur implémentation, car ces deux systèmes ont beaucoup influencé le développement du système INSS. Nous présentons aussi une introduction aux systèmes connexionnistes de raffinement de connaissances, une classe de systèmes dont les systèmes SYNHESYS et KBANN sont des exemples typiques.

- La Chapitre 4 concerne la description du système hybride que nous avons développé, le système INSS. Nous présentons l'organisation générale du système INSS et ensuite nous détaillons le mode de fonctionnement de chacun des modules : le module d'inférence symbolique, le module d'insertion de règles symboliques dans un réseau, le module de simulation du réseau connexionniste, le module d'extraction de règles à partir des réseaux et le module de validation de connaissances. Puis nous concluons par une analyse des propriétés de notre système.

- Le Chapitre 5 est consacré à la présentation des résultats pratiques obtenus avec l'utilisation du système INSS. Ce chapitre est divisé en quatre sous-sections, qui décrivent les principales applications que nous avons développée avec l'aide d'INSS : les problèmes du Moine (problème académique d'apprentissage), le système d'aide au diagnostic des comas toxiques (système expert d'aide au diagnostic médical développé en coopération avec le CHU de Grenoble), l'étude du problème de la balance (un problème de modélisation du développement cognitif chez l'enfant), et une application de contrôle d'un robot autonome (en utilisant d'abord un simulateur puis le vrai robot). Les résultats obtenus dans les différents tests réalisés montrent bien les propriétés et les qualités de notre système.

Les chapitres 2 et 3 constituent un état de l'art respectivement dans le domaine de l'apprentissage automatique et dans celui des systèmes hybrides neuro-symboliques. C'est pourquoi nous conseillons aux lecteurs qui ont des connaissances bien fondées sur l'apprentissage automatique et les réseaux de neurones de passer directement au chapitre 3. Par contre, nous pensons que la lecture du chapitre 3 (et en particulier de la section 3.3) est très importante pour la bonne compréhension des sujets qui seront abordés par la suite. De plus, nous avons ajouté à la fin de ce mémoire un glossaire et une liste des abréviations utilisées, qui pourront aider les lecteurs à mieux parcourir cet ouvrage.