

## **Syllabus des UE du parcours Intelligence Artificielle**

UE AP	Apprentissage
UE LMAI	Logiques Modales et Applications à l'Interaction
UE LTR	Logique et Types de Raisonnement
UE MAB	Modèles et Algorithmes pour la Bioinformatique
UE MPOSC	Méthodes de Preuve, d'Optimisation et de Satisfaction de Contraintes
UE PLD	Planification et Décision
UE RTI	Représentation et Traitement de l'Incertitude
UE TALN	Traitement Automatique de la Langue Naturelle écrite

## Titre : APprentissage (AP)

Mots-clés : apprentissage par exploration, apprentissage par optimisation, hypothèses probablement approximativement correctes

**Objectifs du cours :** L'apprentissage est un champ de l'intelligence artificielle où l'on se propose de concevoir des machines dont les performances s'améliorent avec l'expérience. En d'autres termes, une telle machine simule le comportement évolutif d'un être humain. On tente dans ce cours de donner un aperçu des méthodes classiques d'apprentissage, et de présenter le problème de l'évaluation de la complexité des différentes méthodes.

Le cours présente d'abord le cadre général de l'apprentissage de concepts en mode supervisé: on apprend à décider si des objets représentés par  $n$  attributs sont oui ou non des instances d'une certaine classe dont on ne connaît pas la définition mais dont on connaît des exemples et des contre-exemples. On introduit ensuite quelques représentations courantes pour les concepts à apprendre, et les algorithmes classiques permettant d'apprendre ces représentations :

1- Les arbres de décision sont des représentations compactes d'ensembles de règles de la forme :

SI Attribut1 = a ET ... ET Attributn = z ALORS oui/non.

L'algorithme classique d'apprentissage d'arbres de décision est ID3.

2- Les programmes logiques sont des ensembles de règles de la forme:

SI propriété-1(Attribut1,Attribut2,Attribut3) = vrai

ET ... ET propriété-i(Attribut1,Attribut2,Attribut3) = vrai

ALORS oui/non

Un algorithme classique d'apprentissage de programmes logiques procède par couverture séquentielle.

3- Les séparatrices linéaires sont utilisées dans le cas d'attributs numériques, elles correspondent à des règles de la forme :

SI  $\text{coef1} \times \text{Attribut1} + \dots + \text{coefn} \times \text{Attributn} > 0$  ALORS oui/non.

Les fonctions à noyaux permettent de représenter et apprendre des séparatrices non linéaires en passant par des séparatrices linéaires dans un espace de dimension supérieure.

4- Les réseaux de neurones permettent de représenter des séparatrices (linéaires ou non) par une grille de fonctions linéaires élémentaires, appelées <<neurones>>.

Les coefficients de ces neurones sont appris par une méthode de descente en gradient.

Le cours aborde ensuite l'apprentissage non supervisé. On présente finalement une théorie computationnelle de l'apprentissage: on donne des outils probabilistes pour définir quand une règle de décision apprise est <<Probablement Approximativement Correcte (PAC)>>, et évaluer le temps ou le nombre d'exemples nécessaires pour apprendre une règle ayant cette propriété.

Le module est composé de 9 cours :

1. Apprentissage de concepts
2. Apprentissage d'arbres de décision
3. Apprentissage de règles et programmation logique inductive
4. Apprentissage bayésien
5. Réseaux de neurones
6. Séparateurs à vaste marge
7. Classification non supervisée
8. Découverte automatique d'associations
9. Théorie computationnelle de l'apprentissage

### *Bibliographie:*

- Machine Learning. Tom Mitchell. McGraw Hill, 1997.

- Apprentissage artificiel - Concepts et algorithmes. Antoine Cornuéjols, Laurent Miclet et Yves Kodratoff. Eyrolles, 2002.

# **Titre : Logiques Modales et Applications à l'Interaction (LMAI)**

Mots-clés : Logiques modales, Logiques épistémique, dynamique et déontique, Formalisation

**Objectifs du cours :** Ce cours présente la sémantique des mondes possibles en tant qu'outil de formalisation de problèmes, en focalisant sur la formalisation de l'interaction. Il couvre les principaux systèmes de logique modale (présentation sémantique et axiomatique) et introduit des concepts permettant d'analyser les agents et leurs interactions (états mentaux, actions, communication, obligations).

**Introduction :** Histoire et motivation en IA

## **Systèmes monomodaux et multimodaux**

1. Généralités sur la sémantique
2. Système K
  - Langage et Axiomatique
  - Sémantique des mondes possibles / Satisfaction, validité, vérité
  - Caractérisation
3. Les axiomes D, T, B, 4, 5
  - Lecture des axiomes, axiomatique
  - Sémantique du 2ème ordre
  - Sémantique du 1er ordre
4. Déduction automatique et décidabilité : la méthode des tableaux
  - Règles classiques / pour K / pour D, T, 4
  - Adéquation et complétude
  - Terminaison et test de boucle / décidabilité
  - Stratégies complètes et terminantes ; complexité
5. Systèmes multimodaux : multi-K et axiomes d'interaction
  - Axiomatique
  - Sémantique du 1er ordre

## **Application à la formalisation de l'interaction**

1. Croyance et connaissance
  - Omniscience
  - Introspection
  - Connaissance versus croyance
  - Connaissance commune, connaissance distribuée
2. Action
  - Pré-conditions et effets
  - Persistance (frame problem)
  - Causalité
  - Actes de communication, logique des annonces publiques
3. Intention, état mental
  - intention = choix + engagement
  - logiques BDI (belief-desire-intention)
4. Norme, réglementation, obligation, permission
  - Mondes idéaux
  - Violation, paradoxe de Ross
  - Obligations secondaires (contrary-to-duties)

### *Bibliographie*

- Modal Logic: an introduction. Brian Chellas. Cambridge University Press, 1988.
- A companion to modal logic. G.E. Hughes, M.J. Cresswell. Methuen, 1968.
- Reasoning about Knowledge. Fagin, J. Halpern, Y. Moses, M. Vardi. MIT Press, 2004.
- Dynamic Logic. David Harel, Dexter Kozen et Jerzy Tiuryn. MIT Press, 2000.

## **Titre : Logique et Types de Raisonnements (LTR)**

Mots-clés :Logique des prédicats, déduction, abduction

**Objectifs du cours** : présentation d'un langage logique (la logique des prédicats) et des méthodes de raisonnement formel associées suivant le type du raisonnement recherché.

### **1. Logique des prédicats**

- syntaxe et sémantique
- théorie des modèles
- théorie de la preuve
- méta-propriétés de la logique des prédicats
- théories du premier ordre

### **2. Extensions de la logique des prédicats suivant le type de raisonnement recherché**

- typologie des différents types de raisonnements
- raisonnement déductif avec application au raisonnement non-monotone, à la logique des défauts
- raisonnement abductif avec application aux ATMS, à l'argumentation

### *Bibliographie :*

Mise à niveau : <http://www.irit.fr/recherches/LILAC/Pers/Herzig/C/>

Pour la partie 1 :

- Éléments de logique contemporaine avec exercices et corrigés. François Lepage. Presses Universitaires de Montréal.
- An Introduction to Mathematical Logic and Type Theory: To Truth Through Proof. P.B. Andrews, Applied Logic Series , Vol. 27, Springer.

Pour la partie 2 :

- Consequence finding algorithms. P. Marquis. In Handbook of defeasible reasoning and uncertainty management systems (Gabbay & Smets Edts), Vol5. Kluwer Academic Publisher.
- Nonmonotonic reasoning: an overview. G. Brewka, J.Dix, K. Konolige, CLSI Publications, Stanford University, 1997.

## **Titre : Modèles et Algorithmes pour la Bioinformatique (MAB)**

Mots-clés : bioinformatique, biologie, évolution, algorithmique, complexité, modèles stochastiques, optimisation.

**Objectifs du cours :** Le cours forme une introduction à la bioinformatique pour les informaticiens. Après une rapide présentation de quelques fondements de la biologie moderne (fonctionnement cellulaire, génome, principales molécules supports de l'information biologique) indispensable à la compréhension des problèmes, différents problèmes fondamentaux de la bioinformatique seront présentés en montrant comment les techniques issues de l'intelligence artificielle et de l'informatique en général permettent de les modéliser, de les résoudre et produire ainsi de l'information biologique.

### **Contenu du cours:**

Introduction à la biologie : Le génome et son expression : la réplication, la transcription et la traduction. Méthodes à haut débit d'acquisition des données.

Modélisation et résolution de problèmes d'analyse de séquence, génomes et données biologiques. Les problèmes abordés pourront comprendre:

- Evolution des génomes (I): évolution moléculaire. Problèmes de comparaison de séquence, alignement: modélisation et résolution par programmation dynamique. Reconstruction phylogénétique.

- Evolution des génomes (II): évolution chromosomique (ordre des gènes). Calcul de distances de remaniements entre génomes.

- Structure des gènes de protéines: modèles HMM (chaînes de Markov cachées) et à base de graphes, prédiction de gènes (algorithme de Viterbi, plus court chemin).

- Structure des gènes d'ARN: modèles déterministes et stochastiques des structures. Recherche des structures dans les génomes (réseaux de contraintes et grammaires stochastiques).

- Analyse de données de typage dans les pedigres : modélisation par réseau bayésien, recherche d'explication de probabilité maximum. Haplotyping.

De nouveaux problèmes pourront éventuellement remplacer les problèmes ci-dessus selon l'évolution du domaine.

### *Bibliographie:*

- Biological Sequence Analysis. Probabilistic models of proteins and nucleic acids. R. Durbin, S. Eddy, A. Krogh, G. Mitchison.
- Introduction to Computational Molecular Biology, Setubal/Meidanis, PWS Publishing Company.
- Bioinformatics, A practical guide to the analysis of genes and proteins. A. Baxevanis & B Ouellette. Wiley and Sons.
- Bioinformatics, Sequence and Genome Analysis, David. W. Mount. Cold Spring Harbor Lab. Press.

## **Titre : Méthodes de Preuve, d'Optimisation et de Satisfaction de Contraintes (MPOSC)**

Mots clés : Programmation par Contraintes, Optimisation, Logique Propositionnelle, Recherche Complète, Recherche Heuristique, Théorie de la Complexité

**Objectifs du cours :** Ce cours représente une introduction à la modélisation et à la résolution de problèmes de recherche et d'optimisation combinatoire s'appuyant principalement sur des cadres de représentations issus de l'intelligence artificielle (réseaux de contraintes (CSP), logique propositionnelle (SAT)) et les grandes classes d'algorithmes associés (résolution/élimination de variable, cohérence locale, recherche arborescente et locale). Ces techniques sont utilisées dans de nombreux domaines (vision, planification, diagnostic, ordonnancement, bioinformatique, déduction automatique, gestion de ressources...).

Le cours comprend une partie fondamentale consacrée à la théorie de la complexité (classes polynomiales, problèmes NP-complets, preuves de NP-complétude) et une partie consacrée à la présentation et à la compréhension des principaux algorithmes et techniques de résolution formant une boîte à outils pour le résolution de problèmes combinatoires en intelligence artificielle.

### **Contenu détaillé du cours :**

*Algorithmes énumératifs complets :* Principe de base et illustration sur un problème (voyageur de commerce) . Rappel sur le problème SAT : la procédure de Davis et Putnam Définition du cadre des CSP et exemples. Algorithme de backtrack.

*Production de conséquences :* Introduction : pourquoi la recherche de conséquence (inférence, synthèse d'information, propagation). Illustration sur SAT : résolution, recherche d'impliqués premiers. Amélioration d'algorithmes énumératifs par recherche de conséquence (propagation): illustration sur les CSP (backtrack intelligent), sur SAT.

*Recherche incomplète :* Algorithmes énumératifs partiels et algorithmes de recherche locale (e.g. descente, recuit simulé, recherche taboue)

*Notions de complexité computationnelle :* Notion de complexité des problèmes. Réductions entre problèmes, complétude pour une classe. Illustration : quelques problèmes polynomiaux (2-SAT, Horn-SAT, ) et NP-complets (SAT, 3-SAT, consistance d'un CSP, etc...).

*Travaux Pratiques :* expérimentations de solvers SAT et CSP.

*Prérequis :* De bonnes bases en logique propositionnelle et en algorithmique générale.

### *Bibliographie :*

- Computational Complexity. C.M. Papadimitriou. Addison, 1994.
- Handbook of Constraint Programming. Elsevier, 2006.
- Intelligence artificielle et Informatique théorique. Alliot, Schiex, Brisset, Garcia. Editions Cépaduès, 2002.

## **Titre : PPlanification et Décision (PLD)**

Mots-clés : Planification, Décision, Incertitude

**Objectifs du cours :** La planification, domaine essentiel de l'IA, cherche à concevoir des systèmes capables de générer automatiquement, par une procédure formalisée, un plan-solution. Ce dernier est une collection organisée de descriptions d'opérations destinée à guider l'action d'agents exécuteurs (systèmes robotiques ou humains) qui doivent agir dans un monde particulier (et souvent incertain) pour atteindre un but préalablement défini. Ce cours présente les aspects fondamentaux de ce domaine.

### **Introduction générale**

1. Qu'est ce qu'un problème de décision/planification? Pourquoi planifier? Le mirage de la planification.
2. Typologie :
  - des actions, des tâches, et leur représentation (frame problem, ramification...),
  - des plans,
  - des problèmes.
3. La planification modélisée comme :
  - réutilisation de plans,
  - expertise,
  - réaction,
  - synthèse de plans.
4. Quelques applications de la planification.

### **Algorithmes pour la synthèse de plans d'actions dans le cadre classique**

1. Le cadre classique, les problèmes posés.
2. Le langage STRIPS et ses extensions (ADL, PDDL...).
3. Algorithmique de la synthèse de plans :
  - a. la recherche dans les espaces d'états (HSP, HSPR, FF, SHAPER...),
  - b. la recherche dans les espaces de plans (Tweak, Pweak, UCPOP...),
  - c. les méthodes GRAPHPLAN, LCGP...
  - d. les méthodes SATPLAN, CSP-PLAN
4. Extension du cadre pour la prise en compte de ressources et du temps

### **Extension du cadre classique : décision et planification en présence d'incertitude**

1. Du cadre classique à la prise en compte de l'incertitude.
2. La théorie de la décision (individuelle) :
  - a. modélisation des préférences,
  - b. décision individuelle dans le risque et dans l'incertain,
  - c. attitude d'un agent face au risque.
3. Planification dans l'incertain et processus décisionnels markoviens :
  - a. définitions et typologie,
  - b. MDP totalement observables et méthodes de résolution,
  - c. MDP partiellement observables (POMDP), méthodes de résolution et d'approximation.

### *Prérequis :*

Algorithmes de recherche dans les espaces d'états (largeur, profondeur, Hill Climbing, A\*...),  
Logique propositionnelle, SATisfiabilité, logique des prédicats  
Notions de problèmes de satisfaction de contraintes (CSP).

### *Bibliographie :*

- Algorithmique de la planification en IA. P. Régnier. Editions Cépaduès, 2004.
- Automated Planning. M. Ghallab, D. Nau, P. Traverso. Editions Morgan Kaufmann, Elsevier, 2004.

## **Titre : Représentation et Traitement de l'Incertain (RTI)**

Mots-clés : Incertitude, Plausibilité, Inférence, Révision

**Objectifs du cours :** Ce cours traite de la représentation et du traitement de l'incertitude dans les systèmes à base de connaissances.

On décrit dans ce cours quelques théories numériques et ordinales de l'incertitude, comme la théorie des probabilités, la théorie des possibilités et la théorie des fonctions de croyance.

On aborde également des problèmes importants tels que :

- représenter l'incertitude sur les informations (approches numériques ou ordinales)
- inférer (déduire/calculer) les informations les plus plausibles
- réviser (ajouter une nouvelle information)
- fusionner (des informations incertaines)

**I Typologie des imperfections des informations :** Incomplétude, inconsistance, incertitude, gradualité, indiscernabilité

### **II Quelques approches numériques à la représentation de l'incertain**

1. Rappels sur la théorie des probabilités
2. Introduction aux réseaux bayésiens
3. Théorie des possibilités
4. Théorie des fonctions de croyance
5. Probabilités imprécises

*Prérequis :* La logique classique, la théorie des ensembles naïve.

*Bibliographie :*

- Concepts et méthodes pour l'aide à la décision, Outils de modélisation (Vol. 1, Chapitre 3). D. Bouyssou, D. Dubois, M. Pirlot, H. Prade. Editions Hermès, 2006.
- An introduction to Bayesian Networks. F. Jensen. UCL Press, 1996.

## **Titre : Traitement Automatique de la Langue Naturelle écrite (TALN)**

Mots clés : lexique, grammaire, sémantique

**Objectifs du cours :** La communication Homme-Machine en langage naturel, sous une grande diversité de formes (langue simplifiée, schémas, langue standard, interrogation, les modes textuels, en situation de dialogue, dans le WEB, dans les autoroutes de l'information et le multi-media, etc.) se trouve au coeur de nombreuses préoccupations (interrogation de services, dialogue homme-machine, traductions, aide à la rédaction technique, systèmes tutoriels). Elle fait appel à des outils conceptuels ou techniques largement répandus (ontologies, terminologies, systèmes de traits,...).

Ce cours propose une analyse des fondements théoriques du traitement automatique du langage naturel écrit. Il propose les éléments fondamentaux linguistiques, syntaxiques et sémantiques, cognitifs et conceptuels, de représentation des connaissances et de traitement informatique.

### **Organisation du cours :**

1. Objectifs de la communication en langue naturelle :  
les enjeux, les exigences, familles d'applications  
le langage et son environnement

2. Rappel d'éléments d'informatique théorique :  
grammaires, automates, systèmes de traits, stratégies d'analyse et de génération, unification, lambda-calcul

3. Typologie des applications incluant des éléments de langue :  
familles d'applications et leurs principales caractéristiques.

4. Fondements linguistiques :

Lexique : morphologie, sous-catégorisation, rôles thématiques, relations lexicales

fondements de la syntaxe, techniques d'analyse syntaxique et de construction de représentations.

fondements de la sémantique : notions de sens, principe de compositionnalité, formalismes sémantiques

5. Méthodes d'analyse et d'évaluation :

Méthodes d'analyse de problèmes (par ex. sur corpus), dépouillement de données langagières, annotations.

Evaluation des coûts et des besoins.

Techniques d'évaluation des ressources, protocoles pour l'évaluation des productions (analyse, génération).

### *Bibliographie :*

Mathematical Methods in Linguistics. B. Partee et ali. Kluwer academic, 1993.