

Département des sciences exactes

Spécialité : Informatique

Polycopié de cours destiné aux :
Étudiants de la 5^{ème} année PES informatique

TECHNIQUES AVANCEES

Proposé par :

Dr. BENKADDOUR Fatima Zohra, Maitre de conférences « A ».

Année universitaire 2025-2026

À mes étudiantes et étudiants,
qui transforment chaque séance d'enseignement
en un espace d'échange, de réflexion et de progression
scientifique partagée.

Table des matières

Avant-propos	12
I Bio Systèmes	18
1 Méta-heuristiques et Systèmes Bio-inspirés	19
1.1 Introduction	19
1.2 Méta-heuristiques	20
1.2.1 Définition	20
1.2.2 Caractéristiques	21
1.3 Systèmes Bio-inspirés	21
1.3.1 Historique	22
1.4 Conclusion	23
2 Algorithmes génétiques (AG)	24
2.1 Introduction	24
2.2 Caractéristiques	25
2.3 Principe général	25
2.4 Codage des individus	26
2.4.1 Codage binaire	27
2.4.2 Codage réel	27
2.4.3 Codage de Gray	28
2.5 Fonction d'adaptation (Fitness)	28
2.6 Opérateurs génétiques	28
2.7 Fonctionnement d'un Algorithme Génétique (AG) standard	30
2.8 Exemple graphique et application d'un AG au TSP	31
2.9 Domaine d'application	33
2.10 Conclusion	34

3	Algorithme de la colonie des fourmis	35
3.1	Introduction	35
3.2	Historique	36
3.3	Colonies de fourmis et l'informatique	36
3.3.1	Modèle de comportement des fourmis éclaïreuses	36
3.3.2	Exemple illustratif : choix du chemin le plus court	37
3.4	Limites des algorithmes de colonies de fourmis	37
3.5	Description générale	38
3.6	Algorithme de base de l'Ant System (AS)	38
3.6.1	Choix de la ville suivante	38
3.6.2	Dépôt de phéromones	39
3.6.3	Évaporation des phéromones	39
3.6.4	Synthèse du processus de fonctionnement	40
3.7	Conclusion	40
4	Réseaux de Neurones (RN)	41
4.1	Introduction	41
4.2	Neurones biologiques	42
4.2.1	Structure et fonctionnement	42
4.2.2	Cheminement de l'information	42
4.3	Neurones artificiels	43
4.3.1	Structure d'un neurone artificiel	43
4.4	Réseaux de neurones	44
4.4.1	Réseaux monocouches (Perceptron)	44
4.4.2	Réseaux multicouches	44
4.5	Apprentissage des réseaux de neurones	45
4.6	Applications des RNs	45
4.7	Limites des RNs	46
4.8	Conclusion	46
II	Extraction et Gestion de la connaissance	47
5	Extraction et Gestion des connaissances	48
5.1	Introduction	49

5.2	Extraction de la connaissance	49
5.3	Data Mining	49
5.3.1	Définitions	50
5.3.2	Raisons du développement	50
5.3.3	Objets du Data Mining	51
5.3.4	Processus d'Extraction de Connaissances à partir des Données (ECD) . .	51
5.4	Textmining	52
5.4.1	Importance du Text Mining	53
5.4.2	Techniques de base du Text Mining	53
5.4.3	Fonctionnement du Text Mining	54
5.4.4	Processus d'Extraction de Connaissances à partir de texte)	54
5.4.5	Domaines d'application du Text Mining et du Data Mining	55
5.5	Gestion de la connaissance (Knowledge management	56
5.5.1	Les formes de la connaissance	56
5.5.2	Niveaux organisationnels de création des connaissances	58
5.5.3	Dimensions du KM	58
5.6	Processus de gestion des connaissances	58
5.6.1	Repérer les connaissances	59
5.6.2	Capitaliser les connaissances	59
5.6.3	Transférer les connaissances	59
5.6.4	Actualiser les connaissances	59
5.7	Conclusion	60
6	Systèmes d'Information Géographique (SIG)	61
6.1	Introduction	62
6.2	Historique des SIG	62
6.3	Définition d'un SIG	62
6.4	Composants d'un SIG	63
6.4.1	Le matériel	63
6.4.2	Les logiciels	63
6.4.3	Les données	63
6.4.4	Les méthodes	64
6.4.5	Les utilisateurs	64
6.5	Types de données géographiques	64

6.5.1	Données vectorielles	64
6.5.2	Données raster	65
6.6	Fonctionnalités d'un SIG	65
6.7	Domaines d'application	66
6.8	Conclusion	67
7	Systèmes d'Aide à la Décision (SAD)	68
7.1	Introduction	68
7.2	Évolution de l'aide à la décision	68
7.2.1	Définitions de base de la décision	69
7.2.2	Processus décisionnel	71
7.2.3	Processus de décision collaborative	73
7.3	Le travail collaboratif	76
7.4	Intelligence collective	78
7.5	Domaines d'application	79
7.6	Conclusion	79
III	Interface Humain-Machine	80
8	Interfaces Homme-Machine	81
8.1	Introduction	81
8.2	Définitions de bases	82
8.3	Objectifs d'une interface homme-machine	82
8.4	Typologie d'une interface homme-machine	82
8.5	Composants d'une interface homme-machine	83
8.5.1	L'utilisateur	83
8.5.2	La tâche	83
8.6	Techniques et styles d'interaction	84
8.6.1	Modes d'interaction	84
8.6.2	Styles d'interaction	84
8.7	Grands modèles d'architecture d'interface	85
8.7.1	Le modèle Seeheim	85
8.8	Interfaces Web	85
8.8.1	Ergonomie des interfaces Web	86

8.9 Conclusion	87
IV Examens Types et Exercices	88
9 Examen Type 1	89
10 Corrigé Type 1	92
11 Examen Type 2	95
12 Corrigé type 2	97
13 Travail Dirigé	99
Références	101

Table des figures

2.1	Exemple du codage binaire	27
2.2	Exemple de croisement	29
2.3	Roulette wheel	30
2.4	Exemple du problème du voyageur de commerce	32
3.1	Exemple du comportement de l'algorithme colonie de fourmis	37
4.1	Shéma d'un neurone biologique	42
4.2	Système Nerveux Central (SNC)	43
4.3	Processus de cheminement de l'information	43
4.4	Composition d'un neurone artificiel	43
4.5	Illustration d'un Perceptron	44
4.6	Shéma montrant un réseau de neurone multicouche	45
5.1	Processus d'Extraction de Connaissances (ECD)	51
5.2	Connaissances explicites et connaissances tacites	58
6.1	Les composants d'un SIG	63
6.2	Les types de modèles géographiques	64
6.3	Les Les trois volets d'un SIG.	66
7.1	Processus décisionnel de Simon.	71
7.2	Cycle de la décision Courbon.	72
7.3	Processus décisionnel de Lavergne	72
7.4	Modèle DTL (Decision Time Line)	72
7.5	Processus décisionnel de Cauvin.	73
7.6	Modèle du processus de conception collaborative.	73
7.7	Modèle du processus de conception collaborative.	73
7.8	Modèle du processus de conception collaborative.	74

7.9	Modèle du processus de prise de décision collective.	74
7.10	Processus du processus de prise de décision collaborative basé sur la connaissance.	75
7.11	Cycle de collaboration.	77
8.1	Composants d'une interface homme-machine	83
8.2	Les modes d'interaction	84
8.3	Le modèle Seeheim	85

Liste des tableaux

1.1	Comparaison des méta-heuristiques à solution unique et à population	22
5.1	Exemples d'applications du Text Mining et du Data Mining	55
6.1	Évolution historique des systèmes d'information géographique	62
7.1	Définitions de la décision	70

Résumé

Ce polycopié du cours **Techniques Avancées** offre une présentation complète des méthodes et outils modernes pour l'optimisation, l'apprentissage et la prise de décision dans des contextes complexes. Il couvre les méta-heuristiques et systèmes bio-inspirés, tels que les algorithmes génétiques, les colonies de fourmis et les réseaux de neurones, ainsi que les techniques d'extraction et de gestion de la connaissance, incluant le data mining, le text mining, et l'exploitation des systèmes d'information géographique (SIG) et des systèmes d'aide à la décision (SAD).

Le polycopié aborde également les interfaces homme-machine (IHM) et l'ergonomie, en insistant sur leur rôle dans l'efficacité de l'interaction avec les systèmes et la qualité de la prise de décision. Chaque chapitre combine fondements théoriques, illustrations pratiques, favorisant la compréhension des concepts avancés, le développement des compétences analytiques et la capacité à mettre en œuvre et évaluer des solutions adaptées aux problèmes réels. La partie finale propose des examens types, leurs corrigés et des exercices, permettant aux étudiants de consolider leurs acquis et de préparer efficacement l'évaluation.

Mots clés : Techniques avancées, heuristiques, méta-heuristiques, optimisation, algorithmes génétiques, colonies de fourmis, réseaux de neurones, intelligence artificielle, SIG, SAD, extraction et gestion de la connaissance, interfaces homme-machine, ergonomie, exercices, examens.

Avant-propos

Ce document résulte d'un cours dispensé aux étudiants de cinquième année (5^e année) des parcours PES, ainsi qu'aux étudiants de Master Informatique. Il s'inscrit dans le cadre de l'enseignement des **Techniques avancées**, en mettant l'accent sur des domaines à forte valeur scientifique et applicative.

Ce polycopié a pour objectif de fournir aux étudiants les fondements théoriques et pratiques nécessaires à la compréhension et à la maîtrise de concepts avancés en informatique, tout en établissant des liens entre les approches algorithmiques, les systèmes intelligents et l'interaction homme-machine.

Il s'adresse principalement :

- aux étudiants de la 5^e année des parcours PES des ENS d'Algérie ;
- aux étudiants des cycles universitaires en informatique ;
- aux apprenants souhaitant approfondir leurs connaissances dans les systèmes intelligents ;
- à toute personne désireuse d'acquérir une vision structurée des techniques avancées en informatique.

Le document est organisé en **quatre grandes parties** :

- **Première partie** : *Bio-systèmes informatiques*

Elle introduit les approches informatiques inspirées du vivant, en abordant les principes de l'informatique bio-inspirée, les algorithmes génétiques, les algorithmes de colonies de fourmis, les réseaux de neurones et leurs principales applications.

- **Deuxième partie** : *Extraction et gestion des connaissances*

Cette partie traite du passage des données à l'information, puis à la connaissance. Elle couvre les techniques de data mining, d'apprentissage automatique, la gestion des connaissances, les bases de connaissances et les systèmes d'aide à la décision.

- **Troisième partie** : *Interfaces Humain-Machine (IHM)*

Elle présente les principes fondamentaux de l'ergonomie, de l'utilisabilité et de l'accessibilité, ainsi que les interfaces avancées.

— **Quatrième partie :** *Examens Types et Exercices*

Cette partie fournit des exercices corrigés et des examens types afin de permettre aux étudiants de tester leurs connaissances et de se préparer efficacement aux évaluations. Elle inclut des problèmes pratiques, des études de cas, ainsi que des questions théoriques couvrant l'ensemble des notions abordées dans les trois premières parties.

Ce support pédagogique est conçu pour accompagner les cours magistraux, les travaux dirigés, les travaux pratiques et les projets. Il constitue également un outil de révision pour la préparation aux examens.

Vos remarques, corrections d'erreurs ou suggestions d'amélioration sont accueillies avec gratitude.

Fiche matière

Public cible

Ce cours s'adresse principalement aux :

- Étudiants de cinquième année PES informatique, département des sciences exactes, ENS d'Oran – Ammour Ahmed.
- Étudiants en informatique, département d'informatique, faculté des sciences exactes et appliquées.

Prérequis

Pour suivre ce cours, l'étudiant doit posséder :

- Des notions générales en informatique à savoir : structures de données et l'algorithmique.
- Des bases en mathématiques appliquées, incluant l'algèbre linéaire, les probabilités et les statistiques.
- Une compréhension générale des systèmes d'information et de la gestion des données.

Objectifs du cours

À l'issue de ce cours, l'étudiant sera capable de :

- **Comprendre** les fondements théoriques des heuristiques, méta-heuristiques et systèmes bio-inspirés appliqués aux problèmes d'optimisation.
- **Expliquer** le fonctionnement des algorithmes génétiques, des colonies de fourmis et des réseaux de neurones, ainsi que leurs domaines d'application.
- **Appliquer** les méthodes d'intelligence artificielle et les techniques d'optimisation pour résoudre des problèmes concrets et complexes.

- **Exploiter** les techniques d'extraction et de gestion de la connaissance (data mining, text mining, SIG, SAD) pour analyser et structurer les données.
- **Analyser** et **évaluer** la performance, la robustesse et la pertinence des différentes approches selon les caractéristiques des problèmes et les critères de qualité.
- **Concevoir** et **adapter** des heuristiques, des solutions bio-inspirées, des systèmes d'aide à la décision et des interfaces homme-machine pour répondre aux besoins spécifiques des utilisateurs.
- **Intégrer** les connaissances acquises pour résoudre de manière autonome des problèmes complexes et participer à la conception de projets ou d'études appliquées.
- **Préparer** et **réaliser** des exercices pratiques, études de cas et examens, en consolidant les acquis théoriques et méthodologiques.

Fonctionnement du cours

Tout le matériel du cours (polycopié, sujets et exposés) peut être transmis par mail. Pour toute question ou problème, l'étudiant peut contacter l'enseignant à :

- `benkaddour.fatima@ens-oran.dz`
- `benkaddourfatima@gmail.com`

Organisation de l'enseignement

L'enseignement est structuré en cours magistraux et travaux dirigés (TD) :

- **Cours magistral** : 1h30 par semaine.
- **Travaux dirigés (TD)** : 1h30 par semaine.

Coefficient

Cette matière est de coefficient **3**.

Modalités d'évaluation

L'évaluation des connaissances et compétences acquises repose sur :

- Contrôle continu.

- Exposé individuel ou en groupe sur un sujet pratique ou théorique.
- Examen final.

Compétences développées

À l'issue de ce cours, l'étudiant sera capable de :

- Résoudre des problèmes d'optimisation et de décision en utilisant des heuristiques, méta-heuristiques et systèmes bio-inspirés.
- Développer et appliquer des algorithmes génétiques, des colonies de fourmis et des réseaux de neurones à des cas réels.
- Exploiter et gérer efficacement les connaissances issues de grands ensembles de données grâce au data mining, text mining, SAD et SIG.
- Concevoir des interfaces homme-machine ergonomiques et intuitives, adaptées aux besoins des utilisateurs.
- Évaluer et comparer différentes méthodes et outils en termes de performance, de complexité et de pertinence pour un problème donné.
- Mettre en œuvre une démarche scientifique rigoureuse dans la conception, l'adaptation et l'implémentation de solutions techniques.

Carte mentale du cours



Première Partie : BIOSYSTEMES

« Cette partie explore les systèmes bio-inspirés, ou bio-informatique, en présentant des approches telles que les algorithmes génétiques, les colonies de fourmis et les réseaux de neurones, qui s'inspirent des mécanismes naturels pour résoudre des problèmes complexes d'optimisation et d'apprentissage ».

1

Méta-heuristiques et Systèmes Bio-inspirés

Table des matières

1.1	Introduction	19
1.2	Méta-heuristiques	20
1.2.1	Définition	20
1.2.2	Caractéristiques	21
1.3	Systèmes Bio-inspirés	21
1.3.1	Historique	22
1.4	Conclusion	23

1.1 Introduction

Les problèmes d'optimisation complexes dépassent souvent les capacités des méthodes exactes classiques. Les méta-heuristiques offrent alors des stratégies générales et flexibles permettant d'obtenir des solutions de bonne qualité en un temps raisonnable, en équilibrant exploration et exploitation. Parmi elles, les approches bio-inspirées s'appuient sur des mécanismes observés dans la nature, tels que l'évolution, l'intelligence collective ou l'auto-organisation. Ces méthodes imitent le comportement de systèmes biologiques capables de s'adapter à des environnements complexes. Ce chapitre présente un aperçu des méta-heuristiques et des systèmes bio-inspirés, développés pour résoudre des problèmes d'optimisation complexes lorsque les méthodes exactes deviennent inefficaces.

1.2 Méta-heuristiques

Les méthodes de résolution exacte, telles que les algorithmes classiques d'optimisation, permettent de déterminer des solutions dont l'optimalité est garantie. Cependant, dans de nombreux problèmes complexes ou de grande taille, le coût computationnel associé à ces méthodes peut devenir prohibitif.

Dans ce contexte, il est souvent préférable de rechercher des solutions de qualité satisfaisante, même si leur optimalité ne peut être assurée, afin de réduire significativement le temps de calcul. C'est dans cette perspective que s'insèrent les méta-heuristiques, un ensemble de méthodes d'optimisation généralistes et adaptables à différents types de problèmes.

Ces techniques permettent d'explorer efficacement l'espace des solutions, mais présentent l'inconvénient de ne fournir aucune garantie sur la qualité des solutions obtenues. L'objectif de ce chapitre est de présenter les principes fondamentaux des méta-heuristiques, d'en expliquer les mécanismes généraux, et d'illustrer leur application à travers des exemples concrets d'optimisation.

1.2.1 Définition

Le terme **méta-heuristique** provient des mots grecs *meta* (au-delà) et *heuriskein* (trouver). Il désigne un ensemble de techniques d'optimisation qui dépassent les heuristiques classiques en proposant une approche générique et adaptable à différents types de problèmes.

Il n'existe pas de consensus strict sur la définition exacte des heuristiques et des méta-heuristiques. Toutefois, on peut retenir les distinctions suivantes :

- Une **heuristique** est une méthode de résolution spécialisée pour un problème donné. Elle ne garantit pas que la solution obtenue est optimale, mais vise à fournir une solution de qualité raisonnable dans un temps réduit.
- Une **méta-heuristique** est une heuristique générique, c'est-à-dire une méthode flexible qui peut être adaptée à différents problèmes. Elle fournit un cadre méthodologique permettant d'explorer l'espace des solutions de manière efficace, sans toutefois garantir l'optimalité finale.

Ainsi, les méta-heuristiques combinent la généralité et l'adaptabilité, offrant des solutions robustes pour des problèmes complexes où les méthodes exactes sont impraticables.

1.2.2 Caractéristiques

Les méta-heuristiques constituent des stratégies de recherche conçues pour guider efficacement la recherche de solutions dans un espace complexe. Leur objectif principal est d'explorer cet espace de manière structurée afin de déterminer des points proches de l'optimum, tout en conservant un coût computationnel raisonnable.

Ces techniques peuvent varier d'une simple procédure de recherche locale à des processus plus sophistiqués intégrant des mécanismes d'apprentissage ou d'adaptation. En général, les méta-heuristiques sont **non déterministes** et ne garantissent pas l'optimalité des solutions obtenues. Elles peuvent être décrites de manière abstraite, indépendamment d'un problème spécifique, ce qui en fait des outils génériques et flexibles.

Les méta-heuristiques peuvent intégrer des heuristiques spécifiques au problème étudié, mais ces dernières sont toujours contrôlées par une stratégie de niveau supérieur, assurant une recherche guidée et cohérente. De plus, elles peuvent exploiter l'expérience accumulée au cours de la recherche pour améliorer la qualité et l'efficacité des itérations suivantes.

En résumé, les caractéristiques fondamentales des méta-heuristiques sont les suivantes :

- capacité à guider la recherche vers des solutions de qualité proche de l'optimum ;
- généralité et adaptabilité à différents types de problèmes ;
- nature non déterministe, sans garantie d'optimalité ;
- possibilité d'intégrer des heuristiques spécifiques contrôlées par une stratégie globale ;
- exploitation de l'expérience acquise lors du processus de recherche.

Les méta-heuristiques sont généralement classées en deux grandes catégories, en fonction de la manière dont elles explorent l'espace des solutions (voir la Table 1.1).

1.3 Systèmes Bio-inspirés

L'informatique bio-inspirée, également appelée informatique naturo-inspirée, s'intéresse à l'exploitation des systèmes naturels comme source d'inspiration pour la conception et le développement de systèmes informatiques. Cette approche repose sur l'observation de phénomènes naturels complexes, capables de produire des comportements efficaces, robustes et adaptatifs, souvent sans contrôle centralisé.

Parmi les principales sources d'inspiration figurent les systèmes biologiques, tels que le cerveau humain, la génétique des populations, le système immunitaire, les écosystèmes, ainsi que

TABLE 1.1 – Comparaison des méta-heuristiques à solution unique et à population

Caractéristique	Solution unique	Population
Principe	Manipule un seul point à chaque itération	Manipule simultanément une population de points à chaque itération
Exploration de l'espace	Locale	Globale, grâce à la diversité et à l'interaction entre solutions
Exemples	Recherche Tabou, Recuit Simulé	Algorithmes Evolutionnaires, Algorithmes de Colonies de Fourmis
Avantages	Simple à implémenter et efficace sur des problèmes locaux	Exploration plus large, meilleure adaptation aux problèmes complexes
Inconvénients	Risque de rester bloqué dans un optimum local	Plus complexe à gérer et coût computationnel plus élevé

les colonies d'insectes sociaux. Ces systèmes présentent des propriétés remarquables, notamment l'auto-organisation, l'adaptation à l'environnement et la capacité à résoudre collectivement des problèmes complexes.

La modélisation de ces mécanismes naturels a conduit au développement de nombreuses méthodes et algorithmes en informatique, en particulier dans les domaines de l'optimisation, de l'apprentissage et de la recherche de solutions approchées. Ces approches bio-inspirées constituent aujourd'hui un cadre fondamental pour l'étude et la mise en œuvre des méta-heuristiques, qui s'appuient sur des principes généraux issus de la nature pour explorer efficacement de vastes espaces de solutions.

1.3.1 Historique

En 1860, Charles Darwin publie son ouvrage fondateur intitulé *L'origine des espèces au moyen de la sélection naturelle, ou la lutte pour l'existence dans la nature*. Dans cet ouvrage, Darwin propose une théorie révolutionnaire visant à expliquer l'évolution des espèces à travers le mécanisme de la sélection naturelle. Cette théorie repose notamment sur quatre principes fondamentaux :

- la loi de croissance et de reproduction ;
- la loi de l'hérédité, étroitement liée au processus de reproduction ;
- la loi de la variabilité, résultant des conditions d'existence ;
- la loi de multiplication des espèces, conduisant à la lutte pour l'existence et ayant pour conséquence la sélection naturelle.

Ces principes ont constitué une source majeure d’inspiration pour de nombreux travaux ultérieurs en informatique et en intelligence artificielle. Dans les années 1960, John Holland s’intéresse aux systèmes évolutifs et pose les bases théoriques des mécanismes d’adaptation artificielle. En 1975, il introduit le premier modèle formel des algorithmes génétiques, connu sous le nom *Canonical Genetic Algorithm (CGA)*, dans son ouvrage *Adaptation in Natural and Artificial Systems*.

Par la suite, en 1989, David E. Goldberg publie un ouvrage de vulgarisation des algorithmes génétiques qui contribue largement à leur diffusion et à leur adoption dans la communauté scientifique. Goldberg enrichit la théorie des algorithmes génétiques en établissant notamment les analogies suivantes :

- un individu est lié à son environnement par son code génétique (ADN) ;
- une solution est liée à un problème par un indice de qualité, communément appelé fonction d’évaluation.

Ces travaux ont permis de structurer le domaine des algorithmes génétiques et d’en faire un pilier fondamental des approches d’optimisation bio-inspirées.

1.4 Conclusion

Les méta-heuristiques bio-inspirées constituent une approche efficace pour traiter des problèmes d’optimisation complexes. Elles s’appuient sur des mécanismes naturels afin d’obtenir des solutions robustes et adaptatives. Le chapitre suivant présente l’algorithme génétique, en détaillant ses principes fondamentaux et son fonctionnement.

2

Algorithmes génétiques (AG)

Table des matières

2.1	Introduction	24
2.2	Caractéristiques	25
2.3	Principe général	25
2.4	Codage des individus	26
2.4.1	Codage binaire	27
2.4.2	Codage réel	27
2.4.3	Codage de Gray	28
2.5	Fonction d'adaptation (Fitness)	28
2.6	Opérateurs génétiques	28
2.7	Fonctionnement d'un Algorithme Génétique (AG) standard	30
2.8	Exemple graphique et application d'un AG au TSP	31
2.9	Domaine d'application	33
2.10	Conclusion	34

2.1 Introduction

Les algorithmes génétiques (AG) sont des méta-heuristiques inspirées des mécanismes de la sélection naturelle et de la génétique des populations. Ils sont particulièrement efficaces pour résoudre des problèmes complexes où les méthodes exactes ou analytiques sont difficiles à appliquer, notamment lorsque l'espace de recherche est vaste, discontinu ou non linéaire. Leur force réside dans leur capacité à explorer simultanément de nombreuses solutions et à produire

des résultats de qualité en un temps de calcul raisonnable.

Ce chapitre a pour objectif d'introduire les concepts fondamentaux des algorithmes génétiques, de présenter leurs principaux mécanismes (codage, population, sélection, croisement et mutation) et de montrer comment ces techniques bio-inspirées peuvent être appliquées à l'optimisation, à l'apprentissage automatique et à la modélisation de systèmes complexes.

2.2 Caractéristiques

Les algorithmes génétiques (AG) sont des méthodes de recherche inspirées de la sélection naturelle et de la génétique. Ils combinent la *survie des plus aptes* avec des échanges aléatoires structurés, comme le croisement et la mutation, pour explorer efficacement l'espace des solutions.

Les AG sont des méthodes approchées : ils ne garantissent pas l'optimalité de la meilleure solution obtenue en un temps fini, mais offrent une grande robustesse pour les problèmes complexes.

Le fonctionnement repose sur l'évolution d'une population d'individus (solutions candidates), génération après génération, en appliquant des mécanismes d'hérédité, de mutation et de sélection. La population initiale est créée aléatoirement et se transforme progressivement vers des solutions de meilleure qualité.

2.3 Principe général

Un algorithme génétique (AG) a pour objectif de rechercher le ou les extrema d'une fonction définie sur un espace de solutions. Pour le mettre en œuvre, cinq éléments essentiels doivent être définis :

1. **Codage des individus** : chaque point de l'espace de recherche est représenté par une structure de données appelée *chromosome*. Le codage est généralement établi après une modélisation mathématique du problème et conditionne fortement le succès de l'algorithme.
2. **Population initiale** : il s'agit de générer un ensemble d'individus diversifiés servant de base pour les générations futures. Le choix de la population initiale influence la rapidité de convergence vers l'optimum global.
3. **Fonction à optimiser** : également appelée fonction *fitness* ou fonction d'évaluation,

elle attribue à chaque chromosome une valeur représentant la qualité de la solution. Un chromosome est considéré meilleur lorsque sa valeur de fitness est élevée.

4. **Opérateurs génétiques** : ils permettent de diversifier la population et d'explorer l'espace de recherche.
 - *Croisement (crossover)* : combine les gènes de deux ou plusieurs chromosomes pour créer de nouveaux individus (enfants).
 - *Mutation* : modifie aléatoirement certains gènes pour garantir l'exploration de nouvelles zones de l'espace d'états.
5. **Paramètres de dimensionnement** : taille de la population, nombre de générations ou critère d'arrêt, probabilités d'application des opérateurs de croisement et de mutation.

Forme générale d'un algorithme génétique

Un algorithme génétique peut être décrit de manière générique par le schéma suivant :

Algorithm 1 Algorithme génétique générique

- 1: **Initialiser** la population initiale P
 - 2: **Évaluer** la population P
 - 3: **while** critère de convergence non atteint **do**
 - 4: Sélectionner les parents dans P pour former P'
 - 5: Appliquer l'opérateur de **croisement** sur P'
 - 6: Appliquer l'opérateur de **mutation** sur P'
 - 7: Remplacer les anciens individus de P par les descendants de P'
 - 8: **Évaluer** la population P
 - 9: **end while**
 - 10: Retourner la meilleure solution trouvée
-

Critères de convergence

Le critère de convergence peut être défini de différentes manières, par exemple :

- Atteinte d'un niveau minimum de fitness souhaité pour la population.
- Dépassement d'un temps de calcul maximum.
- Combinaison de plusieurs critères, tels que la fitness et le temps.

2.4 Codage des individus

Chaque paramètre d'une solution est assimilé à un gène, et l'ensemble des paramètres constituant une solution forme un chromosome, également appelé génotype. Les différents éléments

d'un chromosome sont ainsi des gènes, chacun étant repérable par sa position spécifique sur le chromosome, appelée locus. Les différentes valeurs que peut prendre un gène sont désignées sous le terme d'allèles. Un individu est alors représenté par un ensemble de chromosomes, décrivant une solution candidate au problème étudié. Enfin, une population correspond à un ensemble d'individus, constituant l'espace de recherche exploré par l'algorithme génétique (voir la Figure 2.1).

2.4.1 Codage binaire

Le codage binaire constitue l'une des méthodes de représentation les plus courantes pour modéliser les solutions potentielles d'un problème dans les algorithmes génétiques, en utilisant une structure inspirée des mécanismes de la génétique biologique.

- Chaque paramètre de la solution est un **gène**.
- Un **chromosome** est un ensemble de gènes.
- La position d'un gène est le **locus**.
- Les valeurs possibles d'un gène sont les **allèles**.
- Un **individu** est représenté par un ou plusieurs chromosomes.
- Une **population** est un ensemble d'individus.

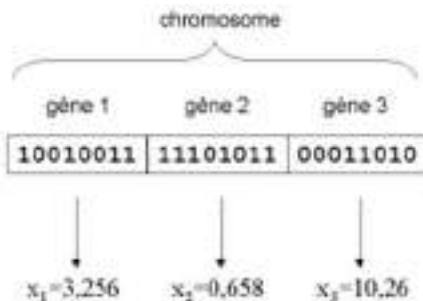


FIGURE 2.1 – Exemple du codage binaire

2.4.2 Codage réel

Le codage réel constitue une autre méthode de représentation utilisée dans les algorithmes génétiques, particulièrement adaptée aux problèmes d'optimisation où les variables prennent des valeurs continues.

- Les paramètres sont directement représentés par des nombres réels.

- Utile pour les problèmes d’optimisation continue.
- Les opérateurs génétiques doivent être adaptés aux valeurs réelles.

2.4.3 Codage de Gray

Le codage de Gray est une technique de représentation utilisée dans les algorithmes génétiques afin d’améliorer certaines limitations du codage binaire classique lors de l’exploration de l’espace de recherche.

- Corrige une limitation du codage binaire : les voisins en Hamming peuvent être éloignés dans l’espace de recherche.
- Deux éléments consécutifs diffèrent par un seul bit.
- Améliore la convergence de l’algorithme.

2.5 Fonction d’adaptation (Fitness)

Dans les algorithmes génétiques, la fonction d’adaptation (ou fitness function) joue un rôle central dans l’évaluation des solutions candidates et dans le processus d’évolution de la population.

- Mesure la qualité ou l’utilité d’un individu pour le problème.
- Détermine la probabilité de sélection pour la reproduction.
- Une bonne définition influence la convergence.
- Peut être maximisée ou minimisée selon le problème.

2.6 Opérateurs génétiques

Dans les algorithmes génétiques, les opérateurs génétiques constituent les mécanismes fondamentaux permettant de faire évoluer la population de solutions d’une génération à une autre. Inspirés des processus biologiques de l’évolution, ces opérateurs assurent l’exploration de l’espace de recherche et l’amélioration progressive de la qualité des solutions.

Sélection

L’opérateur de sélection constitue une étape essentielle du processus évolutif, car il permet de choisir les individus qui participeront à la reproduction et à la génération suivante.

- Choisit les individus qui serviront de parents pour la nouvelle population P' .
- Probabilité de survie proportionnelle à la fitness.
- Méthodes principales :
 1. **Roulette Wheel** : chance de sélection proportionnelle à la fitness.
 2. **Élitiste** : sélection des meilleurs individus.
 3. **Tournoi** : tirage de deux individus, le plus adapté est choisi avec probabilité p .
 4. **Universelle stochastique** : peu utilisée, faible variance.

Croisement (Crossover)

L'opérateur de croisement (crossover) est un mécanisme de reproduction qui permet de combiner l'information génétique de deux individus afin de produire de nouvelles solutions.

- Appliqué sur la population P' après sélection.
- Formation de $n/4$ couples pour générer $n/2$ descendants.
- Les chromosomes sont recombinaés pour produire des enfants.
- Appliqué avec une probabilité élevée pour explorer l'espace de solutions (voir exemple montré à la Figure 2.2).

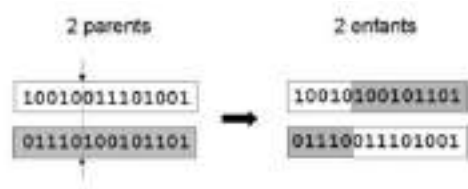


FIGURE 2.2 – Exemple de croisement

Mutation

L'opérateur de mutation introduit des modifications aléatoires dans les chromosomes des individus afin de maintenir la diversité génétique au sein de la population et d'éviter une convergence prématurée vers des solutions sous-optimales.

- Inversion aléatoire d'un ou plusieurs bits d'un gène.
- Appliquée avec faible probabilité.
- Introduit de la diversité et évite la stagnation dans des optima locaux.
- Rôle secondaire par rapport au croisement.

2.7 Fonctionnement d'un Algorithme Génétique (AG) standard

Le fonctionnement d'un AG standard repose sur un cycle itératif composé de plusieurs étapes clés :

Initialisation

La population initiale est générée de manière aléatoire ou semi-aléatoire. Le choix de cette population conditionne fortement la rapidité et l'efficacité de l'algorithme.

Évaluation

Chaque individu de la population est évalué à l'aide d'une *fonction d'adaptation (fitness)*, qui attribue une valeur numérique représentant la qualité de la solution codée par le chromosome.

Sélection

La sélection consiste à choisir les individus qui seront autorisés à se reproduire. La probabilité de sélection dépend de leur valeur de fitness : plus un individu est adapté, plus il a de chances d'être choisi comme parent. Les méthodes classiques sont : roulette wheel, tournoi, sélection élitiste, ou universelle stochastique tel montré à la Figure 2.3.



FIGURE 2.3 – Roulette wheel

Reproduction

Les individus sélectionnés subissent les ****opérateurs génétiques**** :

- **Croisement (Crossover)** : deux parents sont combinés pour produire un ou plusieurs descendants, selon une probabilité de croisement.
- **Mutation** : modification aléatoire de certains gènes des descendants, selon une faible probabilité, pour maintenir la diversité génétique.

La reproduction crée une nouvelle population de descendants qui remplace tout ou partie de la population courante.

Retour et itération

Après l'application des opérateurs, la nouvelle population devient la population courante, et les étapes d'évaluation, sélection et reproduction sont répétées.

Critère d'arrêt

Le processus se répète tant que l'une des conditions d'arrêt n'est pas vérifiée :

- Nombre maximal de générations atteint.
- La population n'évolue plus significativement.
- La valeur de fitness d'un individu dépasse un seuil prédéfini.

Lorsque l'une de ces conditions est remplie, l'algorithme retourne la meilleure solution trouvée.

2.8 Exemple graphique et application d'un AG au TSP

Le problème du voyageur de commerce (PVC) consiste à déterminer un trajet de coût minimal permettant à un voyageur de visiter un ensemble de n villes, chacune une seule fois, avant de revenir à la ville de départ. L'objectif principal est la minimisation du coût total du parcours, lequel peut représenter soit la distance totale parcourue, soit le temps global de déplacement. Pour ce faire, on considère un ensemble de n villes ainsi que les distances (ou durées) séparant chaque paire de villes. Ce problème peut être modélisé de manière formelle à l'aide d'un graphe pondéré, où les sommets correspondent aux villes et les arêtes aux liaisons entre celles-ci, pondérées par les distances associées. Une représentation équivalente consiste à utiliser une matrice de distances décrivant le graphe, dans laquelle les villes sont généralement indexées de 0 à $n - 1$, et chaque élément de la matrice indique le coût du déplacement entre deux villes données.

On peut également considérer des graphes incomplets, dans lesquels certains arcs (ou chemins) entre villes n'existent pas. Dans ce cas, la matrice de représentation du graphe tient compte de ces absences en attribuant aux arcs inexistantes une valeur de distance strictement négative, généralement fixée à -1 , afin de les distinguer clairement des distances valides. Par ailleurs, les arcs reliant une ville à elle-même ne sont pas pris en considération, puisque les déplacements réflexifs n'ont aucun sens dans le cadre du problème du voyageur de commerce. Cette modélisation permet de représenter des situations réalistes où toutes les villes ne sont pas nécessairement directement accessibles entre elles (voir la Figure 2.4).

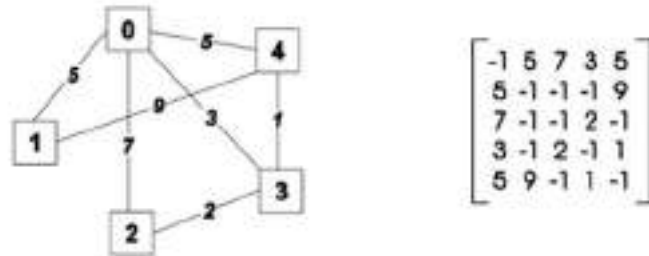


FIGURE 2.4 – Exemple du problème du voyageur de commerce

Application d'un AG au TSP

L'application d'un algorithme génétique (AG) au problème du voyageur de commerce consiste à utiliser des mécanismes inspirés de l'évolution biologique afin de rechercher une tournée optimale ou quasi optimale minimisant la distance totale parcourue entre un ensemble de villes. Ce problème, connu en optimisation combinatoire sous le nom de Travelling Salesman Problem (TSP), est largement utilisé comme cas d'étude pour l'évaluation des métaheuristiques.

1. **Codage** : chaque solution (tournée) est représentée par un vecteur indiquant l'ordre de visite des villes, par exemple : $[0, 2, 4, 3, 1]$.
2. **Population initiale** : générer aléatoirement plusieurs tournées distinctes.
3. **Évaluation (fitness)** : calculer la distance totale parcourue pour chaque tournée.
4. **Sélection** : choisir les tournées les plus courtes pour servir de parents.
5. **Croisement (Crossover)** : combiner deux parents pour créer de nouvelles tournées (exemple : ordre partiel, cycle crossover).
6. **Mutation** : échanger deux villes dans la tournée pour maintenir la diversité.
7. **Remplacement** : former une nouvelle population à partir des descendants et, éventuellement, des meilleurs parents.

8. **Itération** : répéter jusqu'à atteindre le critère d'arrêt (nombre de générations ou distance minimale atteinte).

Remarques pédagogiques

- Chaque permutation de villes constitue un individu dans l'AG.
- Le croisement et la mutation permettent d'explorer l'espace des solutions sans avoir à examiner toutes les permutations possibles.
- La fitness guide l'algorithme vers les tournées les plus courtes.

Formulation formelle du Problème du Voyageur de Commerce (TSP)

Soit un ensemble de n villes $\{1, 2, \dots, n\}$ et une matrice des coûts de déplacement

$$C = (c_{ij}) \quad \text{pour } 1 \leq i, j \leq n,$$

où c_{ij} représente le coût (distance, temps ou autre) pour aller de la ville i à la ville j .

Objectif

Le TSP consiste à trouver une **permutation** π des villes qui minimise le coût total d'une tournée :

$$\text{Coût}(\pi) = \sum_{k=1}^n c_{\pi(k)\pi(k+1)},$$

avec $\pi(n+1) = \pi(1)$ pour revenir à la ville de départ.

Autrement dit, il s'agit de trouver un **circuit de longueur minimale** dans un graphe valué complet :

- **Graphe orienté** : chaque arc a une direction.
- **Graphe non orienté** : les arcs n'ont pas de direction.

Classification

- **TSP symétrique** : si $c_{ij} = c_{ji}$ pour tout $i, j \in \{1, \dots, n\}$.
- **TSP asymétrique** : si $c_{ij} \neq c_{ji}$ pour au moins un couple (i, j) .

Remarque

Le coût d'une tournée est simplement la somme des coûts de tous les arcs appartenant à cette tournée. La difficulté du problème réside dans le fait que le nombre de permutations possibles est $n!$, ce qui rend le TSP NP-difficile pour des instances de grande taille.

2.9 Domaine d'application

Les principaux domaines d'utilisation des algorithmes génétiques sont les suivants.

Optimisation

Les algorithmes génétiques sont largement utilisés pour la résolution de problèmes d'optimisation, notamment lorsque l'espace de recherche est de grande dimension ou fortement contraint. Ils permettent de rechercher des solutions quasi optimales sans hypothèses fortes sur la forme de la fonction objectif.

Apprentissage automatique et intelligence artificielle

Dans le domaine de l'apprentissage automatique, les algorithmes génétiques interviennent comme outils d'optimisation et d'adaptation. Ils sont utilisés pour améliorer les performances des modèles et pour ajuster leurs paramètres ou leur structure.

Programmation automatique

Les algorithmes génétiques sont à l'origine de la programmation automatique, qui vise à faire évoluer des programmes ou des règles de décision capables de résoudre un problème donné sans intervention humaine directe.

Modélisation des systèmes naturels et sociaux

Les algorithmes génétiques sont également employés pour la modélisation et la simulation de systèmes complexes issus du monde réel. Leur caractère évolutif et adaptatif les rend adaptés à l'étude de phénomènes dynamiques et collectifs.

2.10 Conclusion

En conclusion, l'algorithme génétique illustre la puissance des mécanismes évolutionnaires pour l'exploration et l'optimisation d'espaces de recherche complexes. Sa flexibilité et sa robustesse en font une méthode largement utilisée dans de nombreux domaines. Le chapitre suivant est consacré à l'algorithme de colonie de fourmis, qui exploite quant à lui les principes d'intelligence collective et de communication indirecte.

3

Algorithme de la colonie des fourmis

Table des matières

3.1	Introduction	35
3.2	Historique	36
3.3	Colonies de fourmis et l'informatique	36
3.3.1	Modèle de comportement des fourmis éclaireuses	36
3.3.2	Exemple illustratif : choix du chemin le plus court	37
3.4	Limites des algorithmes de colonies de fourmis	37
3.5	Description générale	38
3.6	Algorithme de base de l'Ant System (AS)	38
3.6.1	Choix de la ville suivante	38
3.6.2	Dépôt de phéromones	39
3.6.3	Évaporation des phéromones	39
3.6.4	Synthèse du processus de fonctionnement	40
3.7	Conclusion	40

3.1 Introduction

L'algorithme de la colonie de fourmis est une méta-heuristique bio-inspirée basée sur les principes de l'intelligence collective. Il s'inspire du comportement des fourmis réelles, qui utilisent les phéromones pour découvrir des chemins optimaux. Cette méthode repose sur la coopération entre agents simples et un équilibre entre exploration et exploitation. Ce chapitre présente ses principes fondamentaux et ses principales applications en optimisation.

3.2 Historique

L'Ant System (AS) a été introduit par Dorigo, Colorni et Maniezzo en 1991. Il est le premier algorithme de fourmis à apparaître et le prototype de plusieurs autres méthodes. L'AS a été créé au départ pour la résolution du problème du voyageur de commerce. Le comportement des fourmis est un comportement collectif. Chaque fourmi a pour priorité le bien être de la communauté. Chaque individu de la colonie est à priori indépendant et n'est pas supervisé d'une manière ou d'une autre. Ce concept est appelé Hétérarchie (s'opposant à la Hiérarchie). Chaque individu est aidé par la communauté dans son évolution et en retour il aide au bon fonctionnement de celle-ci. La colonie est donc auto-contrôlée par le biais de mécanismes relativement simples à étudier.

3.3 Colonies de fourmis et l'informatique

En observant une colonie de fourmis à la recherche de nourriture dans les environs du nid, on s'aperçoit qu'elle résout des problèmes tels que celui de la recherche du plus court chemin. Les fourmis résolvent des problèmes complexes par des mécanismes assez simples à modéliser. Leur technique repose sur la pose de marqueurs chimiques, les phéromones, déposés sur les trajets parcourus. Cela peut paraître surprenant au premier abord mais un chemin plus court reçoit plus de phéromones qu'un chemin plus long. Les fourmis suivantes ayant tendance à emprunter les chemins les plus marqués, une solution va progressivement se dessiner. Il est ainsi assez simple de simuler leur comportement par des algorithmes.

3.3.1 Modèle de comportement des fourmis éclaireuses

Un modèle expliquant ce comportement repose sur le rôle des fourmis appelées **éclaireuses**. Celles-ci parcourent l'environnement autour de la colonie de manière plus ou moins aléatoire. Lorsqu'une éclaireuse découvre une source de nourriture, elle revient au nid en suivant le chemin le plus direct possible, en laissant sur son parcours une piste de **phéromones**. Ces phéromones, attractives, incitent les autres fourmis passant à proximité à suivre la piste, de façon plus ou moins directe. En revenant au nid, ces mêmes fourmis renforcent la piste de phéromones, augmentant ainsi son attractivité et guidant progressivement l'ensemble de la colonie vers la source de nourriture.

Si plusieurs chemins permettent d'atteindre la même source, le chemin le plus court sera

parcouru par un plus grand nombre de fourmis dans le même intervalle de temps. La piste courte sera donc de plus en plus renforcée et deviendra de plus en plus attractive, tandis que la piste longue finira par disparaître, les phéromones étant volatiles. À terme, l'ensemble de la colonie aura naturellement déterminé et « choisi » le chemin le plus court, optimisant ainsi collectivement le déplacement vers la source de nourriture.

3.3.2 Exemple illustratif : choix du chemin le plus court

Pour illustrer le mécanisme, considérons une situation simple : la première fourmi trouve la source de nourriture F en empruntant un chemin quelconque a , puis revient au nid N en déposant sur son parcours une piste de **phéromones** b .

Les autres fourmis, initialement indifférentes, peuvent emprunter l'un des quatre chemins possibles. Cependant, le renforcement progressif de la piste de phéromones rend le chemin le plus court de plus en plus attractif. Les fourmis finissent ainsi par privilégier ce chemin, tandis que les portions plus longues des autres chemins voient leurs phéromones disparaître, guidant collectivement la colonie vers la solution optimale (voir la Figure 3.1).

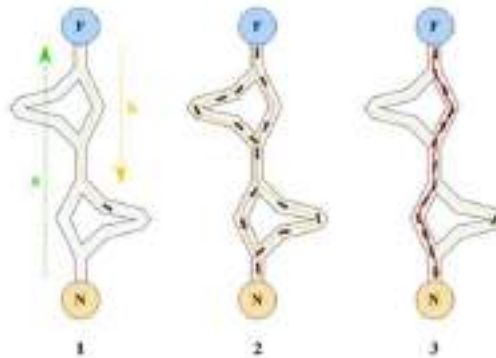


FIGURE 3.1 – Exemple du comportement de l'algorithme colonie de fourmis

3.4 Limites des algorithmes de colonies de fourmis

L'algorithme de colonies de fourmis a été initialement conçu pour produire des solutions quasi-optimales au problème du voyageur de commerce, et plus généralement pour les problèmes d'optimisation combinatoire.

Cependant, certaines limites doivent être prises en compte. Un risque majeur survient lorsqu'un chemin non optimal est fortement marqué par des phéromones. Les fourmis à proximité

de ce chemin seront incitées à l'emprunter, renforçant encore sa piste et pouvant entraîner la colonie dans un minimum local.

Pour limiter ce phénomène, l'algorithme utilise un mécanisme d'**évaporation des phéromones** qui réduit progressivement la concentration sur tous les chemins. Ce processus favorise l'exploration de nouvelles solutions et augmente les chances de découvrir le chemin optimal. Le taux d'évaporation constitue ainsi l'un des paramètres principaux de l'algorithme, déterminant l'équilibre entre exploration et exploitation.

3.5 Description générale

L'algorithme général de colonies de fourmis repose sur un ensemble de fourmis, chacune parcourant un trajet possible entre les villes. Le fonctionnement de chaque fourmi obéit aux règles suivantes :

- Chaque ville ne peut être visitée qu'une seule fois par une fourmi.
- Plus une ville est éloignée, moins elle a de chances d'être choisie, ce critère étant appelé **visibilité**.
- Plus l'intensité de la phéromone sur l'arête reliant deux villes est grande, plus le trajet a de chances d'être sélectionné.
- Une fois le trajet terminé, la fourmi dépose des phéromones sur les arêtes parcourues, la quantité étant proportionnelle à la qualité du parcours (par exemple, un trajet plus court reçoit davantage de phéromones).
- À chaque itération, les phéromones s'évaporent, diminuant l'attractivité des mauvaises solutions et favorisant l'exploration de nouvelles routes.

Ainsi, au fur et à mesure des itérations, les arêtes du meilleur chemin sont progressivement renforcées par un dépôt répété de phéromones, tandis que les arêtes correspondant à des solutions moins bonnes perdent leur attractivité à cause de l'évaporation. Ce mécanisme conduit progressivement la colonie vers des solutions optimales ou quasi-optimales.

3.6 Algorithme de base de l'Ant System (AS)

L'Ant System (AS) fonctionne itérativement sur un graphe représentant le problème à résoudre. À chaque itération t ($1 \leq t \leq t_{\max}$), chaque fourmi k ($k = 1, \dots, m$) construit un trajet complet de $n = |N|$ étapes, où $|N|$ est le nombre total de villes.

3.6.1 Choix de la ville suivante

Lorsque la fourmi k se trouve sur la ville i , le choix de la ville j à visiter dépend de trois paramètres :

1. La liste des villes déjà visitées, définissant l'ensemble des mouvements possibles $\mathcal{J}_i^k$;
2. La visibilité $\eta_{ij} = 1/d_{ij}$, inverse de la distance entre les villes i et j , qui favorise les déplacements vers les villes proches ;
3. L'intensité de la phéromone $\tau_{ij}(t)$ sur l'arête (i, j) , représentant l'attractivité accumulée des trajets parcourus par les fourmis précédentes.

La probabilité pour que la fourmi k choisisse la ville $j \in \mathcal{J}_i^k$ est donnée par la **règle aléatoire de transition proportionnelle** :

$$p_{ij}^k(t) = \frac{[\tau_{ij}(t)]^\alpha \cdot [\eta_{ij}]^\beta}{\sum_{l \in \mathcal{J}_i^k} [\tau_{il}(t)]^\alpha \cdot [\eta_{il}]^\beta} \quad (3.1)$$

où α et β contrôlent respectivement l'importance relative de l'intensité de la piste et de la visibilité.

- $\alpha = 0$: seules les distances sont prises en compte, la fourmi choisit toujours la ville la plus proche.
- $\beta = 0$: seules les phéromones influencent le choix, la fourmi suit les pistes existantes.

3.6.2 Dépôt de phéromones

Après avoir terminé son trajet $T^k(t)$, la fourmi dépose une quantité de phéromone $\Delta\tau_{ij}^k(t)$ sur chaque arête parcourue, proportionnelle à la qualité de la solution :

$$\Delta\tau_{ij}^k(t) = \begin{cases} \frac{Q}{L_k(t)} & \text{si la fourmi } k \text{ a utilisé l'arête } (i, j) \text{ dans son trajet } T^k(t) \\ 0 & \text{sinon} \end{cases} \quad (3.2)$$

où $L_k(t)$ est la longueur de la tournée de la fourmi k et Q est un paramètre fixé.

3.6.3 Évaporation des phéromones

Pour éviter que le système ne reste piégé dans des solutions sous-optimales, l'additivité des phéromones est contrebalancée par une évaporation à chaque itération. La mise à jour des

phéromones sur chaque arête est donc :

$$\tau_{ij}(t+1) = (1 - \rho) \cdot \tau_{ij}(t) + \sum_{k=1}^m \Delta\tau_{ij}^k(t) \quad (3.3)$$

où $\rho \in (0, 1]$ est le taux d'évaporation et $\tau_0 \geq 0$ est la quantité initiale de phéromones sur toutes les arêtes.

3.6.4 Synthèse du processus de fonctionnement

L'algorithme Ant System repose sur un compromis entre deux mécanismes complémentaires :

- **Intensification** : renforcement des arêtes correspondant aux meilleurs trajets ;
- **Diversification** : exploration de nouveaux chemins grâce à la probabilité stochastique et à l'évaporation des phéromones.

Au fil des itérations, les fourmis convergent vers des solutions optimales ou quasi-optimales.

Algorithm 2 Algorithme de base de l'Ant System (AS)

```

1: Initialiser  $\tau_{ij}(0) = \tau_0$  pour toutes les arêtes  $(i, j)$ 
2: for  $t = 1$  to  $t_{\max}$  do
3:   for  $k = 1$  to  $m$  do
4:     Construire un trajet complet  $T^k(t)$ 
5:     while toutes les villes ne sont pas visitées do
6:       Calculer  $p_{ij}^k(t)$  pour chaque ville  $j$ 
7:       Sélectionner la ville suivante selon  $p_{ij}^k(t)$ 
8:     end while
9:     Calculer  $L_k(t)$ 
10:    Déposer la phéromone  $\Delta\tau_{ij}^k(t)$ 
11:  end for
12:  Mettre à jour les phéromones :  $\tau_{ij}(t+1) = (1 - \rho)\tau_{ij}(t) + \sum_{k=1}^m \Delta\tau_{ij}^k(t)$ 
13: end for
14: Retourner le meilleur trajet

```

3.7 Conclusion

En conclusion, l'algorithme de la colonie de fourmis montre comment des interactions simples entre agents peuvent produire des solutions globales efficaces. Sa capacité d'adaptation et de coopération en fait un outil puissant pour l'optimisation combinatoire. Le chapitre suivant s'intéresse aux réseaux de neurones, qui s'inspirent du fonctionnement du cerveau pour traiter et apprendre à partir de données complexes.

4

Réseaux de Neurones (RN)

Table des matières

4.1	Introduction	41
4.2	Neurones biologiques	42
4.2.1	Structure et fonctionnement	42
4.2.2	Cheminement de l'information	42
4.3	Neurones artificiels	43
4.3.1	Structure d'un neurone artificiel	43
4.4	Réseaux de neurones	44
4.4.1	Réseaux monocouches (Perceptron)	44
4.4.2	Réseaux multicouches	44
4.5	Apprentissage des réseaux de neurones	45
4.6	Applications des RNs	45
4.7	Limites des RNs	46
4.8	Conclusion	46

4.1 Introduction

Les réseaux de neurones sont des modèles mathématiques inspirés du cerveau humain, utilisés pour apprendre des relations complexes à partir de données. Ils permettent de faire des prédictions continues (régression) ou discrètes (classification). Ce chapitre présente leurs principes, architectures et méthodes d'apprentissage.

4.2 Neurones biologiques

Le neurone est l'élément fondamental du cerveau humain, qui en compte environ 10^{11} , chacun connecté à d'autres par environ 10^{14} synapses. Ces cellules communiquent par impulsions électriques transmises le long des axones et au niveau des synapses (voir la Figure 4.1).

La cellule nerveuse élémentaire (neurone) est composée de trois régions distinctes :

- **entrées** : les dendrites ou arbre d'axones qui reçoit l'information d'autres neurones à travers d'axones ;
- **le corps cellulaire (soma)** : considéré comme l'unité de traitement ;
- **Sortie** : L'axone qui transporte l'impulsion du neurone aux voisins.

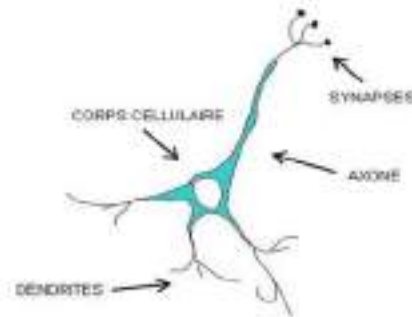


FIGURE 4.1 – Shéma d'un neurone biologique

4.2.1 Structure et fonctionnement

La cellule nerveuse est constituée de trois parties :

- Les dendrites, qui reçoivent les signaux d'autres neurones ;
- Le soma (corps cellulaire), qui traite les informations ;
- L'axone, qui transmet le signal aux neurones voisins.

Au niveau des synapses, l'information est transmise sous forme d'impulsions électriques. Chaque synapse possède un poids représentant son influence, qui peut être excitateur (+) ou inhibiteur (-). Le neurone s'active lorsque la somme des impulsions excitatrices dépasse un seuil θ , générant ainsi une impulsion électrique transmise via l'axone.

4.2.2 Cheminement de l'information

Les informations sensorielles sont captées par des récepteurs spécifiques et converties en impulsions électriques, qui sont ensuite traitées par le système nerveux central (SNC) pour

déclencher des actions ou réguler le fonctionnement des organes (voir les Figures 4.2, 4.3).

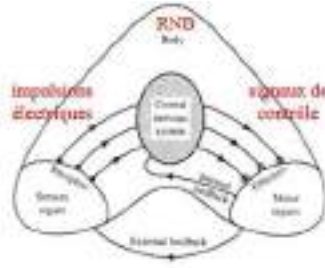


FIGURE 4.2 – Système Nerveux Central (SNC)

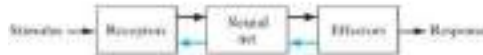


FIGURE 4.3 – Processus de cheminement de l'information

4.3 Neurones artificiels

Les réseaux de neurones artificiels (RNA) sont des systèmes adaptatifs inspirés du cerveau biologique. Ils apprennent à partir d'exemples en ajustant leurs poids synaptiques pour atteindre un objectif donné. Plusieurs modes d'apprentissage existent, notamment supervisé et par renforcement (voir la Figure 4.4).

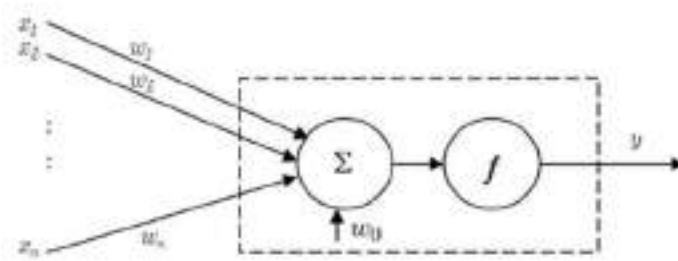


FIGURE 4.4 – Composition d'un neurone artificiel

4.3.1 Structure d'un neurone artificiel

Un neurone artificiel reçoit un vecteur d'entrées x_i ($i = 1, \dots, N$), pondérées par des coefficients w_i . La sortie du neurone est calculée comme :

$$S = \sum_{i=1}^N w_i x_i - w_0, \quad y = f(S)$$

où w_0 est le biais et f la fonction d'activation, qui ajuste la sortie selon le problème.

Les éléments clés d'un neurone artificiel sont :

- Les connexions pondérées ;
- L'additionneur calculant la somme pondérée ;
- La fonction d'activation.

Fonctions d'activation courantes :

Heaviside :

$$f(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases}$$

Sigmoïde :

$$f(x) = \frac{1}{1 + e^{-x}}$$

4.4 Réseaux de neurones

Les réseaux de neurones sont constitués d'unités de traitement interconnectées, appelées neurones artificiels, organisées selon différentes architectures. La complexité et la capacité d'apprentissage d'un réseau dépendent de sa structure, qu'il s'agisse d'une seule couche de neurones ou de plusieurs couches imbriquées.

4.4.1 Réseaux monocouches (Perceptron)

Un perceptron est un réseau à une seule couche de neurones. Chaque neurone reçoit toutes les entrées x_i et produit une sortie (voir la Figure 4.5) :

$$y_j = f\left(\sum_{i=1}^N w_{ij}x_i - w_{0j}\right), \quad j = 1, \dots, P$$

où P est le nombre de neurones de sortie (ex. le nombre de classes).

4.4.2 Réseaux multicouches

Dans les réseaux multicouches, les neurones sont organisés en plusieurs couches. La sortie d'un neurone k de la couche l est :

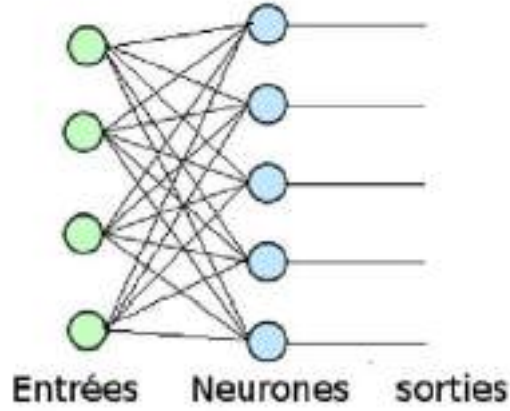


FIGURE 4.5 – Illustration d'un Perceptron

$$y_k^{(l)} = f \left(\sum_j w_{kj}^{(l)} y_j^{(l-1)} - w_{0k}^{(l)} \right)$$

Les couches intermédiaires sont appelées couches cachées et permettent de modéliser des fonctions non linéaires complexes (voir la Figure 4.6).

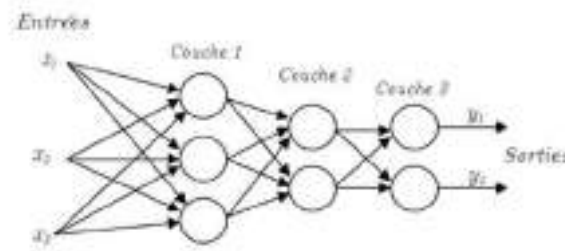


FIGURE 4.6 – Schéma montrant un réseau de neurone multicouche

4.5 Apprentissage des réseaux de neurones

L'apprentissage consiste à ajuster les poids w_i pour minimiser l'erreur entre la sortie produite y_i et la sortie attendue d_i :

$$E = \frac{1}{2} \sum_{i=1}^N (d_i - y_i)^2$$

Dans les réseaux multicouches, la correction des poids se fait par rétropropagation (back-propagation) :

$$\Delta w_{kj}^{(l)} = \eta \delta_k^{(l)} y_j^{(l-1)}$$

où η est le taux d'apprentissage et $\delta_k^{(l)}$ l'erreur locale du neurone k à la couche l .

4.6 Applications des RNs

Les réseaux de neurones trouvent des applications dans :

- La reconnaissance d'écriture manuscrite et vocale ;
- La reconnaissance faciale et la prévision météo ;
- La détection de maladies via l'imagerie médicale ;
- La finance, le traitement du langage naturel et la découverte de médicaments.

4.7 Limites des RNs

Parmi les limites principales des algorithmes réseaux de neurones qu'on peut citer :

- Les réseaux sont des *boîtes noires*, difficiles à interpréter ;
- L'ordre des exemples d'apprentissage peut influencer les résultats ;
- Sensibilité aux minima locaux et risques de divergence.

4.8 Conclusion

En conclusion, les réseaux de neurones constituent des outils puissants pour le traitement de signaux complexes et la prédiction. Leur efficacité dépend du choix de la structure, des fonctions d'activation et des méthodes d'apprentissage, tout en restant limitée par des questions d'interprétabilité et de convergence. Ces limites soulignent l'importance de disposer de méthodes complémentaires pour organiser et exploiter les informations. Le chapitre suivant aborde l'extraction et la gestion de la connaissance, qui permet de structurer, analyser et valoriser efficacement les données issues de grands ensembles d'information.

Deuxième Partie :

EXTRACTION ET GESTION DE LA CONNAISSANCE

« Cette partie présente l'extraction et la gestion des connaissances, en mettant l'accent sur les méthodes permettant de collecter, organiser et exploiter les informations. Elle couvre le data mining, le text mining, la gestion des connaissances, ainsi que l'intégration avec les SIG et les SAD pour transformer les données en connaissances utiles à la décision ».

5

Extraction et Gestion des connaissances

Table des matières

5.1	Introduction	49
5.2	Extraction de la connaissance	49
5.3	Data Mining	49
5.3.1	Définitions	50
5.3.2	Raisons du développement	50
5.3.3	Objetsifs du Data Mining	51
5.3.4	Processus d'Extraction de Connaissances à partir des Données (ECD) . .	51
5.4	Textmining	52
5.4.1	Importance du Text Mining	53
5.4.2	Techniques de base du Text Mining	53
5.4.3	Fonctionnement du Text Mining	54
5.4.4	Processus d'Extraction de Connaissances à partir de texte)	54
5.4.5	Domaines d'application du Text Mining et du Data Mining	55
5.5	Gestion de la connaissance (Knowledge management	56
5.5.1	Les formes de la connaissance	56
5.5.2	Niveaux organisationnels de création des connaissances	58
5.5.3	Dimensions du KM	58
5.6	Processus de gestion des connaissances	58
5.6.1	Repérer les connaissances	59

5.6.2	Capitaliser les connaissances	59
5.6.3	Transférer les connaissances	59
5.6.4	Actualiser les connaissances	59
5.7	Conclusion	60

5.1 Introduction

L'extraction et la gestion des connaissances constituent des enjeux majeurs pour les organisations modernes confrontées à la croissance rapide des volumes de données et d'informations. La capacité à transformer des données brutes en connaissances exploitables est devenue un facteur déterminant pour la prise de décision, l'innovation et l'amélioration des performances organisationnelles.

Dans ce chapitre, nous aborderons les notions de *data mining* et de *text mining*, qui représentent les principales approches permettant d'extraire des connaissances à partir des données structurées et non structurées. Nous nous intéresserons ensuite à la gestion des connaissances, en présentant les processus, les méthodes et les outils visant à capitaliser, organiser, partager et valoriser les connaissances au sein des organisations.

5.2 Extraction de la connaissance

L'extraction de connaissances vise à transformer de grands volumes de données brutes en informations pertinentes et exploitables. Elle s'appuie principalement sur des techniques de *Data Mining* pour l'analyse des données structurées, et de *Text Mining* pour l'exploitation des données textuelles non structurées, permettant ainsi de découvrir des modèles, des tendances et des relations utiles à la prise de décision.

5.3 Data Mining

Le « data mining » est l'art d'extraire des connaissances à partir de données. Les données peuvent être stockées dans des entrepôts (« data warehouse »), dans des bases de données distribuées ou sur Internet (« web mining »). Le « data mining » ne se limite pas au traitement des données structurées sous forme de tables numériques ; il offre des moyens pour aborder les corpus en langage naturel (« text mining »), les images (« image mining »), le son (« sound

mining ») ou la vidéo et dans ce cas, on parle alors plus généralement de « multimedia mining ».

5.3.1 Définitions

1. Le data-mining est un processus de découverte de règle, relations, corrélations et/ou dépendances à travers une grande quantité de données, grâce à des méthodes statistiques, mathématiques et de reconnaissances de formes.
2. Le data-mining est un processus d'extractions automatique d'informations predictives à partir de grandes bases de données.

5.3.2 Raisons du développement

Le développement du *Data Mining* s'explique par la convergence de plusieurs facteurs technologiques, économiques et organisationnels.

Explosion des données

L'augmentation continue des volumes de données générées, souvent désignée par le terme *Big Data*, constitue l'une des principales raisons du développement du Data Mining. Les données proviennent de sources diverses telles que les réseaux sociaux, les plateformes collaboratives, les services de messagerie et les contenus multimédias. À titre d'exemple :

- Twitter génère environ 50 millions de tweets par jour, soit près de 7 téraoctets de données ;
- Facebook produit environ 10 téraoctets de données par jour ;
- YouTube voit plus de 50 heures de vidéos mises en ligne chaque minute ;
- environ 2,9 millions de courriels sont échangés chaque seconde.

Évolution de la puissance de calcul

L'augmentation de la puissance de calcul, notamment grâce à la loi de Moore, a fortement contribué au développement du Data Mining. Cette évolution est renforcée par l'émergence du calcul massivement distribué, qui permet de traiter efficacement de très grands volumes de données en parallèle.

Création de valeur ajoutée

Le Data Mining vise également la création de valeur ajoutée à partir des données. L'analyse approfondie des données permet de mieux comprendre les comportements des clients, d'optimiser les produits et les services, et de passer d'une logique centrée sur le produit à une approche orientée vers le client.

Extraction de connaissances

Enfin, le Data Mining constitue un ensemble de techniques essentielles pour l'extraction de connaissances à partir des *Big Data*. Il permet d'identifier des régularités, des tendances et des modèles pertinents, contribuant ainsi à la prise de décision et à l'amélioration des performances des organisations.

5.3.3 Objectifs du Data Mining

Data Mining utilise des outils issus de nombreuses spécialités différentes : statistiques, algorithmes génétiques, réseaux neuronaux, etc. C'est un champ de recherche multidisciplinaire qui s'est progressivement constitué au cours des dernières années. On peut regrouper les objectifs des méthodes de Data Mining en cinq grandes fonctions : classification, estimation, segmentation, prédiction, explication. Le choix de la méthode dépendra de la nature du problème posé et du type de données dont on dispose.

5.3.4 Processus d'Extraction de Connaissances à partir des Données (ECD)

Le processus d'extraction de connaissances à partir des données repose sur une succession d'étapes complémentaires visant à transformer des données brutes en connaissances exploitables. Ces étapes sont les suivantes (montrées à la Figure) :

- **Nettoyage des données** : cette phase consiste à éliminer le bruit et les données inutiles. Les données sont filtrées, triées, homogénéisées et corrigées. En effet, les données peuvent être incomplètes, contradictoires ou contenir des erreurs d'origine humaine ou informatique.
- **Intégration des données** : elle consiste à regrouper des données issues de sources multiples afin de les associer sous une forme unique et cohérente (consolidation). Cette in-

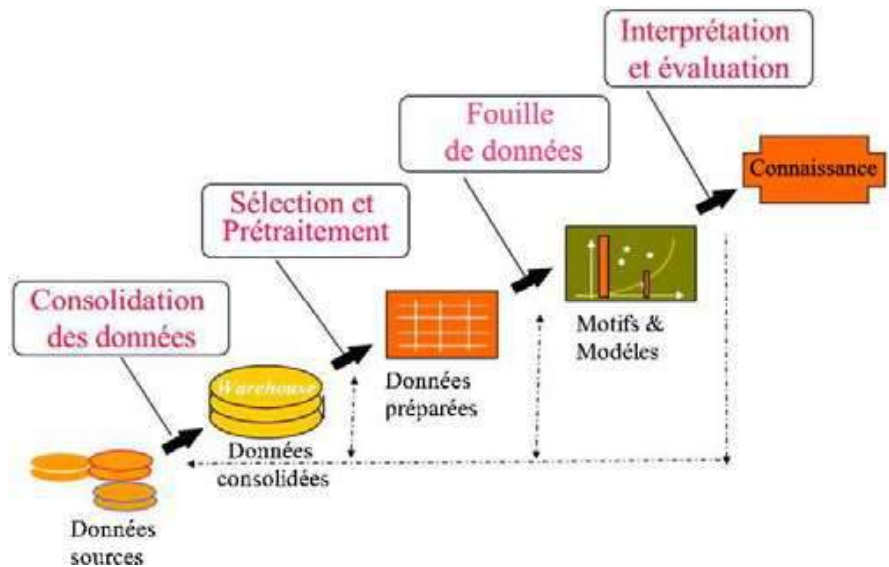


FIGURE 5.1 – Processus d'Extraction de Connaissances (ECD)

tégration s'effectue généralement dans le cadre d'une architecture d'entrepôt de données (*Data Warehouse*).

- **Sélection des données** : seules les données ayant un lien direct avec l'analyse demandée sont retenues et extraites de la base de données.
- **Transformation des données** : les données sélectionnées sont regroupées, normalisées et transformées dans un format adapté à la fouille de données. Cette étape constitue un prétraitement essentiel avant l'application des techniques de *Data Mining*.
- **Data Mining (fouille de données)** : il s'agit de l'étape centrale du processus. Elle consiste à appliquer des méthodes dites « intelligentes » afin d'extraire des éléments remarquables appelés *patterns*. Ces derniers correspondent à des structures de données inhabituelles, des corrélations imprévues, des écarts statistiques ou tout autre comportement sortant de l'ordinaire.
- **Évaluation et interprétation des patterns** : cette étape vise à identifier les patterns intéressants, c'est-à-dire ceux représentant une information pertinente. Leur intérêt est évalué à l'aide de règles objectives, fondées sur la structure des patterns et les mesures statistiques associées, ainsi que de règles subjectives basées sur les croyances et les connaissances des utilisateurs. Ces connaissances peuvent être stockées dans une base de connaissances.
- **Présentation de la connaissance** : des techniques de visualisation et de représentation sont utilisées afin de présenter clairement à l'utilisateur le savoir extrait des données. Ces représentations peuvent prendre la forme de tableaux, arbres, règles, graphiques, courbes,

matrices, cubes de données, etc.

5.4 Textmining

Le *text mining*, ou fouille de textes, est le processus d'extraction d'informations utiles et non triviales à partir de données textuelles non structurées. Cette discipline repose sur la combinaison de plusieurs domaines, notamment la statistique, l'apprentissage automatique (*Machine Learning*), la fouille de données (*Data Mining*) et la linguistique informatique.

5.4.1 Importance du Text Mining

L'importance du *text mining* peut être mise en évidence à travers les points suivants :

- **Gestion des données non structurées** : une part significative des données générées aujourd'hui est non structurée, en particulier les textes issus de sources variées telles que les réseaux sociaux, les courriers électroniques et les documents numériques ;
- **Extraction de connaissances** : le *text mining* permet d'extraire des connaissances pertinentes et de découvrir des structures, des régularités et des motifs (*patterns*) au sein de grandes collections de données textuelles ;
- **Aide à la prise de décision** : les informations extraites peuvent soutenir les entreprises et les organisations dans leurs processus décisionnels en s'appuyant sur des analyses objectives des données textuelles.

5.4.2 Techniques de base du Text Mining

Le *text mining* repose sur un ensemble de techniques fondamentales permettant d'extraire, d'analyser et d'interpréter l'information contenue dans des données textuelles non structurées. Parmi les principales techniques utilisées, on distingue les suivantes :

- **Prétraitement des données** : cette étape est essentielle et consiste à nettoyer les données textuelles afin de supprimer le bruit et les informations non pertinentes. Elle comprend généralement des opérations telles que la suppression des mots vides (*stop words*), la normalisation, la tokenisation, la lemmatisation ou la racinisation (*stemming*), ainsi que la préparation des données pour les phases d'analyse ultérieures.
- **Analyse de sentiments** : cette technique vise à identifier et à analyser l'attitude, l'opinion ou l'émotion exprimée dans un texte. Elle est couramment utilisée pour évaluer la

perception des utilisateurs à l'égard d'un produit, d'un service ou d'un événement, en classifiant les sentiments en catégories telles que positives, négatives ou neutres.

- **Classification de documents** : elle consiste à attribuer automatiquement un document textuel à une ou plusieurs catégories prédéfinies. Cette technique est largement employée dans des applications telles que le filtrage de courriers électroniques, l'indexation de documents, la catégorisation d'articles ou l'organisation de grandes collections textuelles.

5.4.3 Fonctionnement du Text Mining

Le fonctionnement du *Text Mining* est comparable à celui du *Data Mining*, à la différence qu'il se concentre principalement sur l'analyse de données non structurées ou partiellement structurées. Or, une part très importante des données disponibles aujourd'hui — estimée à environ 80 % — est de nature non structurée. Les techniques de Text Mining visent ainsi à faciliter le traitement, la structuration et l'exploitation de grands volumes de documents textuels.

Pour atteindre cet objectif, les données textuelles sont analysées, organisées et catégorisées à l'aide de méthodes quantitatives et qualitatives issues notamment du traitement automatique du langage naturel (*Natural Language Processing*) et du *Deep Learning*. Ces technologies permettent de transformer les textes bruts en représentations exploitables pour l'analyse et la fouille de connaissances.

5.4.4 Processus d'Extraction de Connaissances à partir de texte)

Le processus de Text Mining peut être décomposé en plusieurs étapes principales :

- **Collecte et prétraitement des données textuelles** : les textes sont collectés à partir de sources variées telles que les courriels, les documents, les contenus de sites Web ou les bases de données thématiques. Après la collecte, les données sont structurées, normalisées et nettoyées. Cette phase comprend la suppression des caractères spéciaux et des mots vides, la tokenisation des textes, ainsi que la réduction des mots à leur forme canonique par la racinisation (*stemming*) ou la lemmatisation, afin de préparer les données aux analyses ultérieures.
- **Préparation du texte** : cette étape consiste à identifier les mots-clés, expressions pertinentes, structures ou modèles communs dans l'ensemble de données préparé. Elle inclut également l'extraction de caractéristiques textuelles (telles que la fréquence des termes), l'étiquetage, le regroupement et la catégorisation des documents.

- **Analyse** : une fois les données prétraitées, différents modèles d'analyse sont appliqués afin de mettre en évidence des connaissances et des structures significatives. Parmi les techniques couramment utilisées figurent le clustering hiérarchique, la modélisation thématique, l'analyse de sentiments, l'extraction de mots-clés et le résumé automatique de textes. Ces méthodes permettent de faire émerger des entités, des relations et des modèles pertinents.
- **Interprétation et modélisation** : les résultats issus des phases d'analyse sont interprétés et traduits en modèles exploitables, en stratégies décisionnelles ou en prévisions. L'extraction d'informations, l'analyse des tendances et la reconnaissance de modèles permettent ainsi d'identifier des opportunités d'optimisation pour les produits et services, ainsi que de traiter efficacement de grandes quantités de données textuelles.

5.4.5 Domaines d'application du Text Mining et du Data Mining

Les logiciels de *Text Mining* et de *Data Mining* sont utilisés dans de nombreux secteurs à des fins commerciales, scientifiques ou de sécurité. Le tableau 5.1 présente quelques applications courantes du Text Mining.

5.5 Gestion de la connaissance (Knowledge management)

On appelle *Knowledge Management* (KM), ou *management des connaissances*, l'ensemble des méthodes, des processus et des outils permettant d'identifier, de capitaliser, d'organiser et de diffuser les connaissances au sein d'une organisation. L'objectif principal du Knowledge Management est de valoriser le capital de connaissances de l'entreprise afin d'améliorer la performance, la prise de décision et l'innovation.

5.5.1 Les formes de la connaissance

Nous présentons dans les sous-sections suivantes les différentes formes de la connaissance.

Différences sémantiques

- **Source** : Une source désigne l'origine humaine d'une information ou l'origine technique d'une donnée.
 - Une source primaire : présente des données ou informations fournies par un témoin humain ou un dispositif technique de première main.

TABLE 5.1 – Exemples d’applications du Text Mining et du Data Mining

Domaine	Description de l’application
Service client	Optimisation de l’expérience client à travers l’analyse des retours utilisateurs (chatbots, tickets d’assistance, enquêtes, réseaux sociaux). Le Text Mining permet d’identifier rapidement les problèmes, d’améliorer les services et d’accroître la fidélisation des clients.
Analyse du sentiment	Analyse des avis, commentaires et communications afin d’évaluer la perception du public vis-à-vis des marques, produits ou campagnes, et d’adapter les stratégies commerciales en conséquence.
Gestion des risques	Surveillance des variations de sentiment et détection de signaux faibles dans les rapports, études de marché ou documents financiers, permettant une meilleure compréhension des tendances et des risques.
Entretien et maintenance	Identification de modèles, de tendances ou de faiblesses dans les processus de maintenance industrielle, détection des causes de pannes, d’arrêts ou d’erreurs de production.
Santé	Analyse et catégorisation de la littérature médicale afin d’identifier rapidement des informations sur les symptômes, maladies et traitements, et de soutenir la recherche médicale.
Filtrage anti-spam	Détection de courriers indésirables, de logiciels malveillants et de tentatives de cyberattaques à partir de l’analyse de structures linguistiques et de motifs récurrents.
Screening de candidatures	Analyse structurée des dossiers de candidature pour identifier les profils correspondant aux compétences clés recherchées et faciliter le processus de recrutement.
Recherche d’information	Amélioration de la recherche et de l’extraction d’informations (<i>Information Retrieval</i>), notamment pour les moteurs de recherche, le référencement (SEO) et l’accès à l’information pertinente.

- Une source secondaire : décrit des sources primaires en y ajoutant des commentaires, analyses et critiques.
- Une source tertiaire : correspond à une sélection et une compilation de sources primaires et secondaires.
- **Donnée** : Une donnée est une description élémentaire d’une réalité. Il peut s’agir d’une observation ou d’une mesure effectuée selon un étalon de référence.

Exemple : 19/12/1991.

On distingue au moins deux catégories de données : les données brutes et les jeux de données (data).

- **Information** : Les informations correspondent à des données interprétées et porteuses de sens. Elles sont construites selon des règles et dans un contexte donné. Elles sont

subjectives, car liées à l'intention de l'émetteur de transmettre un message.

Lorsque les informations reposent sur des données falsifiées, on parle alors de falsification ou de désinformation.

Exemple : le lecteur sait désormais que la donnée 19121991 correspond à une date : le 19 décembre 1991.

- **Connaissance :** En gestion des connaissances, une connaissance correspond à l'appropriation et à l'interprétation des informations par les êtres humains. Elle peut être formalisée afin d'être utilisable à des fins opératoires. Elle est vérifiable et orientée vers l'action.

Le concept de connaissance mobilise des notions telles que le langage, la sémantique, les croyances et la conscience. Le passage de la donnée à l'information, puis à la connaissance, nécessite un traitement cognitif.

Exemple : En l'absence de contexte particulier, le lecteur peut interpréter la date 19/12/1991 comme une date de naissance. En revanche, la date 05/07/1962 peut être perçue comme une date historique selon ses connaissances.

- **Savoir-faire :** En gestion des connaissances comme en cognition industrielle, on distingue la donnée brute, l'information, la connaissance et les savoirs, qui mettent en perspective les connaissances ponctuelles sur le long terme. L'utilisation des connaissances dans un contexte d'activité constitue le savoir-faire.
- **Compétence :** Les compétences regroupent les attitudes et aptitudes mobilisables dans une situation concrète afin de réussir une action ou de prendre une décision.
- **Expertise :** L'expertise résulte de la conceptualisation des savoir-faire, des compétences et de l'expérience acquise dans différentes situations de travail. L'expert ne se limite pas à énoncer des vérités, mais mesure des écarts par rapport à des normes ou à un cahier des charges.
- **Communauté :** Une communauté est un groupe auto-organisé partageant un intérêt commun, une identité reconnue, et des connaissances en construction faisant l'objet d'une capitalisation.

Différences pratiques

La typologie des connaissances la plus citée est celle établie par Nonaka et Takeuchi (1995) qui classe les connaissances selon deux dimensions :

- **Connaissances tacites** : Les connaissances tacites relèvent des représentations mentales, profondément ancrées dans les individus et leur vécu. Elles sont difficiles à formaliser par écrit et se transmettent principalement par interaction directe, imitation et imprégnation, notamment entre experts et apprentis.
- **Connaissances explicites** : Ce sont des connaissances clairement formalisées sous forme de documents, de systèmes informatiques ou d'automatismes stockés dans la mémoire des machines, comme illustré à la Figure 5.2.

Le management des connaissances vise ainsi à faciliter la transformation des connaissances tacites en connaissances explicites, leur partage entre les individus, les groupes et l'organisation, et leur exploitation efficace dans les processus métier.

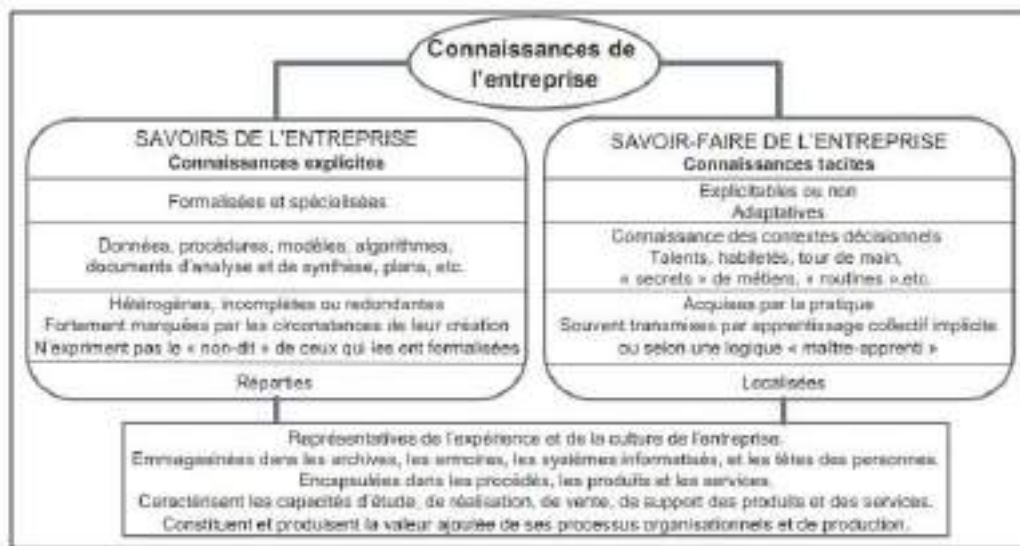


FIGURE 5.2 – Connaissances explicites et connaissances tacites

5.5.2 Niveaux organisationnels de création des connaissances

La création des connaissances ne se limite pas à l'individu. Elle s'opère à plusieurs niveaux organisationnels :

- **Individuel** : apprentissage personnel et expérience.
- **Collectif** : partage et interaction au sein des groupes de travail.
- **Organisationnel** : capitalisation et diffusion à l'échelle de l'entreprise.

Les interactions entre ces niveaux, combinées aux échanges entre connaissances tacites et explicites, constituent le fondement de la dynamique de création des connaissances.

5.5.3 Dimensions du KM

Selon (Grundstein et al., 2002) : Il propose de considérer la fonction de gestion des connaissances comme un prisme constitué de quatre dimensions complémentaires qui sont liées et qui interagissent entre elles,

5.6 Processus de gestion des connaissances

La démarche de gestion des connaissances repose sur un ensemble de processus complémentaires visant à assurer l'identification, la préservation, le partage et l'évolution des savoirs au sein de l'organisation. Ces processus sont généralement regroupés en quatre étapes fondamentales : *repérer*, *capitaliser*, *transférer* et *actualiser* les connaissances.

5.6.1 Repérer les connaissances

Le processus de repérage consiste à identifier les connaissances clés nécessaires au bon fonctionnement et au développement de l'entreprise. Il s'agit notamment de :

- repérer les experts et les détenteurs de savoirs critiques ;
- identifier les connaissances tacites et explicites stratégiques ;
- analyser les besoins en connaissances en fonction des objectifs de l'organisation.

5.6.2 Capitaliser les connaissances

La capitalisation vise à conserver les connaissances afin d'éviter leur perte, notamment lors du départ d'experts ou de changements organisationnels. Elle consiste à formaliser, structurer et enregistrer les connaissances clés dans des supports adaptés tels que des bases de données, des documents, des référentiels ou des systèmes d'information dédiés.

5.6.3 Transférer les connaissances

Le transfert des connaissances permet d'assurer leur diffusion et leur appropriation par les différents acteurs de l'organisation. Il peut s'effectuer selon deux modes principaux :

- **Transfert direct** : l'apprenant acquiert les connaissances par l'observation, l'échange et l'interaction avec l'expert (compagnonnage, tutorat).
- **Transfert indirect** : l'apprenant s'appuie sur des connaissances formalisées et retranscrites sous forme de documents, de procédures ou de supports numériques.

5.6.4 Actualiser les connaissances

Les connaissances évoluent rapidement et peuvent devenir obsolètes. Le processus d'actualisation consiste donc à :

- mettre à jour régulièrement les connaissances existantes ;
- intégrer de nouveaux savoirs issus de l'innovation, de la recherche ou de l'expérience ;
- assurer une veille technologique et organisationnelle continue.

L'actualisation garantit ainsi la pertinence et la durabilité du capital de connaissances de l'entreprise.

5.7 Conclusion

Dans ce chapitre, nous avons présenté les principes fondamentaux de l'extraction et de la gestion des connaissances. Les approches de *data mining* et de *text mining* ont été mises en évidence comme des moyens essentiels pour transformer des données structurées et non structurées en connaissances exploitables. Ces techniques permettent de révéler des régularités, des modèles et des informations pertinentes, contribuant ainsi à une meilleure compréhension des données.

Nous avons également abordé la gestion des connaissances. L'extraction et la gestion des connaissances apparaissent comme des composantes complémentaires et indispensables des systèmes d'information modernes.

Le chapitre suivant présente les Systèmes d'Informations Géographiques (SIG).

6

Systèmes d'Information Géographique (SIG)

Table des matières

6.1	Introduction	62
6.2	Historique des SIG	62
6.3	Définition d'un SIG	62
6.4	Composants d'un SIG	63
6.4.1	Le matériel	63
6.4.2	Les logiciels	63
6.4.3	Les données	63
6.4.4	Les méthodes	64
6.4.5	Les utilisateurs	64
6.5	Types de données géographiques	64
6.5.1	Données vectorielles	64
6.5.2	Données raster	65
6.6	Fonctionnalités d'un SIG	65
6.7	Domaines d'application	66
6.8	Conclusion	67

6.1 Introduction

Les systèmes d'information géographique (SIG) occupent aujourd'hui une place centrale dans de nombreux domaines tels que l'aménagement du territoire, l'environnement, les transports, la santé, l'agriculture ou encore la gestion des risques. Ils permettent de manipuler et d'analyser des données spatiales afin d'aider à la prise de décision et à la compréhension de phénomènes localisés dans l'espace. Ce chapitre aborde les principes fondamentaux des SIG, leurs composantes, les méthodes d'analyse et leurs principales applications pratiques.

6.2 Historique des SIG

L'évolution des systèmes d'information géographique s'est faite progressivement, en parallèle des avancées de l'informatique et des technologies de l'information (voir la Table 6.1).

Période	Évolution des SIG
Fin des années 1950 – milieu des années 1970	Début de l'informatique et apparition des premières cartographies automatiques.
Milieu des années 1970 – début des années 1980	Diffusion des outils de cartographie automatique et des SIG au sein des organismes d'État (armée, cadastre, services topographiques, etc.).
Depuis les années 1980	Croissance du marché des logiciels SIG, développement des applications sur PC et mise en réseau des bases de données géographiques.
Depuis les années 1990	Développement des SIG sur Internet et démocratisation de l'information géographique (cartographie en ligne, calcul d'itinéraires, GPS, applications embarquées).

TABLE 6.1 – Évolution historique des systèmes d'information géographique

6.3 Définition d'un SIG

Un système d'information géographique (SIG) est un système informatique permettant, à partir de données provenant de sources diverses, de collecter, stocker, organiser, gérer, analyser, croiser et représenter des informations localisées géographiquement.

Un SIG est ainsi un système d'information dédié au traitement d'objets spatiaux, caractérisés à la fois par leur position dans l'espace et par des attributs descriptifs. Il constitue un outil d'aide à la décision en offrant des fonctionnalités d'analyse spatiale et de visualisation cartographique.

6.4 Composants d'un SIG

Un SIG repose sur cinq composantes principales tel montré à la Figure 6.1 :



FIGURE 6.1 – Les composants d'un SIG

6.4.1 Le matériel

Il comprend les ordinateurs, serveurs, périphériques de stockage, dispositifs GPS et équipements réseaux nécessaires au fonctionnement du système.

6.4.2 Les logiciels

Les logiciels SIG permettent la gestion des données, l'analyse spatiale et la visualisation cartographique (exemples : QGIS, ArcGIS). Les logiciels assurent les six fonctions suivantes :

1. la saisie des informations géographiques sous forme numérique (*acquisition*) ;
2. la gestion des bases de données (*archivage*) ;
3. la manipulation et l'interrogation des données géographiques (*analyse*) ;
4. la mise en forme et la visualisation des données (*affichage*) ;
5. la représentation simplifiée du monde réel (*abstraction*) ;
6. l'aide à la prospective et à l'anticipation (*anticipation*).

6.4.3 Les données

Les données constituent l'élément central d'un SIG. Elles peuvent être :

- spatiales (géométrie) ;

- attributaires (descriptions) ;
- temporelles.

6.4.4 Les méthodes

Elles regroupent les procédures, modèles et règles d'analyse permettant l'exploitation cohérente des données géographiques. La mise en œuvre et l'exploitation d'un SIG ne peut s'envisager sans le respect de certaines règles et procédures propres à chaque organisation. Un SIG fait appel à une connaissance technique et à divers savoir-faire et donc divers métiers qui peuvent être effectués par une ou plusieurs personnes.

6.4.5 Les utilisateurs

Un SIG étant avant tout un outil, c'est son utilisation (et donc, son ou ses utilisateurs) qui permet d'en exploiter la quintessence. Les SIG s'adressent à une très grande communauté d'utilisateurs depuis ceux qui créent et maintiennent les systèmes, jusqu'aux personnes utilisant dans leur travail quotidien la dimension géographique. Avec l'avènement des SIG sur internet, la communauté des utilisateurs de SIG s'agrandit de façon importante chaque jour et il est raisonnable de penser qu'à brève échéance, nous serons tous à des niveaux différents des utilisateurs de SIG.

Les utilisateurs sont donc les acteurs qui conçoivent, alimentent, analysent et interprètent les résultats produits par le SIG.

6.5 Types de données géographiques

Les Systèmes d'Information Géographique exploitent deux différents types de modèles Géographiques tel montré à la Figure 6.2.

6.5.1 Données vectorielles

Les données vectorielles représentent les objets géographiques sous forme de :

- **Le point** : l'objet le plus simple, il peut représenter à grande échelle des arbres, des bornes d'incendie, des collecteurs d'ordures, etc. Mais à des échelles plus petites de type carte routière au 1/1000000ème, il représente une capitale régionale.

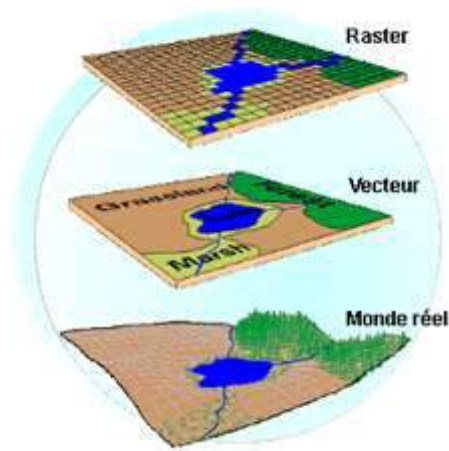


FIGURE 6.2 – Les types de modèles géographiques

- **La ligne** : la ligne représente les réseaux de communication, d'énergie, hydrographiques, d'assainissement, etc. Elle peut être fictive, en représentant l'axe d'une route, ou virtuelle en modélisant des flux d'information, d'argent.
- **La surface** : elle peut matérialiser une entité abstraite comme la surface d'une commune ou des entités ayant une existence géographique comme une forêt, un lac, une zone bâtie, etc.

6.5.2 Données raster

Les données raster sont constituées d'une grille de cellules (pixels), chacune associée à une valeur. Elles sont couramment utilisées pour les images satellitaires, les modèles numériques de terrain et les cartes thématiques.

La donnée raster ou maillée donne une information en chaque point du territoire à partir d'appareil photo ou de caméra aéroportée (avion, ballon, etc.) il est possible d'avoir de nombreux détails de la surface de la terre. Elle peut être scannée, numérique ou orthorectifiée (corrigée des déformations d'échelle :

- **Plan scanné ou carte scannée** : la carte scannée est un bon référentiel visuel car elle est souvent issue de carte papier destinée au grand public (plan cadastral, carte routière) ;
- **Image satellite** : image issue de capteurs embarqués dans des satellites d'observation placés sur des orbites. Elle doit subir plusieurs traitements radiométriques et géométriques avant d'être utilisable dans un SIG ;
- **Image satellite radar** : image enregistrée par des capteurs embarqués dans des satellites d'observation. Elle représente la réponse du sol à l'onde envoyée par le capteur.

6.6 Fonctionnalités d'un SIG

Un SIG offre plusieurs fonctionnalités essentielles :

- acquisition et intégration des données ;
- stockage et gestion des bases de données spatiales ;
- analyse spatiale (proximité, recouvrement, réseau) ;
- visualisation et cartographie thématique ;
- aide à la décision.

Un système d'information géographique (SIG) permet principalement de gérer, d'analyser et d'afficher des informations géographiques. Celles-ci sont représentées par une série de jeux de données géographiques qui modélisent la géographie à l'aide de structures de données génériques simples. De nombreux outils sont intégrés au SIG, afin de pouvoir travailler avec les données géographiques. En outre, les utilisateurs disposent de plusieurs volets pour afficher et manipuler les informations géographiques :

- **Le volet géodatabase** : un SIG correspond à une base de données spatiales contenant des jeux de données qui représentent des informations géographiques selon un modèle de données SIG générique (entités, rasters, topologies, réseaux, etc.).
- **Le volet géovisualisation** : un SIG est un ensemble de cartes intelligentes et de vues qui montrent des entités et leurs relations à la surface de la terre. Il est possible d'élaborer différentes vues cartographiques des informations géographiques sous-jacentes, qui s'utilisent comme des "fenêtres ouvertes sur la base de données" afin d'effectuer des requêtes, des analyses et de modifier les informations.
- **Le volet géotraitement** : un SIG comprend des outils de transformation des informations qui produisent des jeux de données géographiques à partir des jeux de données existants. Les fonctions de géotraitement partent des informations contenues dans les jeux de données existants, appliquent des fonctions analytiques et écrivent les résultats dans de nouveaux jeux de données.

Ces trois éléments ensemble constituent le noyau d'un système d'information géographique Complet et sont utilisés à différents niveaux de toutes les applications SIG tel montré à la Figure 6.3.



FIGURE 6.3 – Les Les trois volets d'un SIG.

6.7 Domaines d'application

Les systèmes d'information géographique sont utilisés dans de nombreux domaines :

- aménagement du territoire et urbanisme ;
- environnement et gestion des ressources naturelles ;
- transport et logistique ;
- santé et épidémiologie ;
- agriculture de précision ;
- gestion des risques naturels.

6.8 Conclusion

Les systèmes d'information géographique (SIG) permettent de collecter, organiser et analyser des données spatiales pour faciliter la compréhension de phénomènes et la prise de décision. Leur utilisation s'étend à de nombreux domaines allant de l'aménagement du territoire à la gestion des risques. Le chapitre suivant présente les systèmes d'aide à la décision (SAD).

7

Systèmes d'Aide à la Décision (SAD)

Table des matières

7.1	Introduction	68
7.2	Évolution de l'aide à la décision	68
7.2.1	Définitions de base de la décision	69
7.2.2	Processus décisionnel	71
7.2.3	Processus de décision collaborative	73
7.3	Le travail collaboratif	76
7.4	Intelligence collective	78
7.5	Domaines d'application	79
7.6	Conclusion	79

7.1 Introduction

Dans une organisation ou une entreprise, la prise de décision influence le déroulement des activités et la qualité du service. Elle devient complexe lorsque plusieurs informations doivent être analysées simultanément. Pour faciliter ce processus et choisir la meilleure solution, on utilise des systèmes d'aide à la décision (SAD). Ce chapitre présente leurs principes, composantes et principales applications.

7.2 Évolution de l'aide à la décision

Au fil du temps, les outils et méthodes d'aide à la décision ont considérablement évolué, intégrant successivement des concepts mathématiques, des systèmes informatiques, l'intelligence

artificielle et les technologies collaboratives. Nous pouvons résumer cette évolution à travers les étapes suivantes :

- Au 20ème siècle, des outils mathématiques sont introduits : leurs algorithmes s'appuient sur des concepts tels que les probabilités, la théorie des graphes, la recherche opérationnelle.
- 1970 les systèmes informatiques d'aide à la décision sont apparus et ont pris une place croissante dans certains processus de décision, il s'agit des systèmes informatiques intégrés, conçus spécialement pour la prise de décisions, et qui sont destinés plus particulièrement aux dirigeants d'entreprises
- Apparaît par la suite l'intelligence artificielle, exploitant des formes de programmation originales cherchant généralement à simuler des réalités de manières virtuelles pour prendre des décisions suivant les résultats obtenus grâce à ces simulations.
- Les systèmes d'information Géographique (SIG) se sont beaucoup développés depuis les années 1970, avec l'avantage de présenter visuellement et de manière cartographique certains éléments d'aide à la décision.
- Par la suite et avec l'informatique sont également arrivés des Systèmes Interactifs d'Aide à la Décision (SIAD).
- Avec l'internet , on a vu apparaître de nouveaux moyens de consultation et d'expression ainsi que des outils de travail collaboratif permettant de nouveaux types de processus de décision et d'aide à la décision. Il y a également l'apparition de Big Data pour le traitement de très grandes bases de données.

7.2.1 Définitions de base de la décision

Nous partons de la définition du dictionnaire qui définit l'action de décider, comme le fait de « porter un jugement, adopter une conclusion définitive sur un point de litige » ou « arrêter, déterminer ce qu'on doit faire ». Le dictionnaire québécois de la langue française complète cette première définition en indiquant que la décision est une « action volontaire de faire le choix d'une solution face à un problème défini » . La recherche scientifique a également proposé quelques définitions, regroupées dans la Table 7.1.

Malgré les différences qui marquent les définitions de la table 7.1, nous avons noté les points communs suivants :

- Une décision a pour objectif de résoudre un problème ;

Définitions	Références
Décider, c'est « identifier et résoudre des problèmes que rencontre toute organisation ».	(Moigne, 1974)
La décision est synonyme de « choix entre plusieurs solutions existantes, ayant chacune des conséquences différentes, le sera fait selon des critères précis de la sélection ».	(Kast, 1993)
La décision correspond à « l'ensemble des actions que le décideur effectue pour prendre une décision et la mettre en œuvre ».	(Barabel, 1996)
Une décision est une sélection d'une éventualité dont on pense qu'elle permettra d'obtenir des avantages optimaux et des inconvénients minimaux.	(Berchet, 2000)
Une décision est un processus qui conduit un acteur à répondre à une question posée.	(Cardinal, 2000)
Décider, c'est de choisir l'action à entreprendre, c'est à dire choisir une action parmi toutes celles qui sont possibles. Les actions sont aussi appelées des solutions ou des options.	(Bellut, 2002)
La décision est un processus de transformation d'informations. Il conduit un acteur ou un groupe d'acteurs de l'organisation à répondre à une question donnant lieu à une action (par la réponse).	(Longueville, 2003)

TABLE 7.1 – Définitions de la décision

- Elle doit se concrétiser par une action ;
- Elle peut être individuelle ou collective et, parfois, distribuée ;
- Elle peut être construite à partir de solutions existantes partielles parmi lesquelles le décideur choisit une solution pouvant répondre au problème posé.

Le choix à opérer lors d'une prise de décision n'est pas totalement arbitraire, il doit reposer sur une certaine rationalité. Pour cela, la définition et l'utilisation de critères de sélection sont nécessaires. Les critères permettant de faire un choix dépendent principalement de l'objectif de la décision. Si l'objectif est unique, la décision sera mono-critère. Dans la plupart des cas, la décision sera associée à plusieurs objectifs et devra donc être multi-critères. Contrairement à la situation mono-critère qui peut être résolue assez facilement, la décision multi-critère nécessite des méthodes de résolution plus élaborées. L'analyse multi-critères nécessite un travail préparatoire des critères, visant à les identifier, les formaliser explicitement, les classer selon une priorité et, même, les agréger afin de combiner les critères et ainsi rendre possible la prise de décision. Les techniques multi-critères peuvent être linéaires, itératives ou encore hiérarchiques. Chaque

décision est un cas particulier qui nécessite l'utilisation d'une méthode adaptée correspondant, dans la plupart des cas, à une solution possible mais pas forcément optimale.

7.2.2 Processus décisionnel

Les définitions de la décision du paragraphe précédent ont montré qu'il s'agit d'une activité complexe, basée sur un enchaînement de mécanismes conduisant à une prise de décision. Il est possible de matérialiser cet enchaînement par un processus décisionnel composé de plusieurs étapes. La littérature propose plusieurs modélisations du processus décisionnel. Herbert Alexander ; Moigne sont parmi les premiers à proposer une formalisation de la décision comme une succession d'étapes intermédiaires formant un processus. Initialement, le processus proposé par Simon était composé de trois étapes :

- **Renseignement** : recherche d'information, de renseignement sur le problème à résoudre, ses contraintes et son environnement ;
- **Conception** : conception, développement et analyse des solutions possibles pouvant répondre au problème posé ;
- **Choix** : évaluation et choix de la solution, en appliquant des critères d'évaluation, puis mise en pratique de la décision choisie, des actions à mettre en place pour résoudre le problème initial ;
- Une quatrième étape a été ajoutée au modèle initial de Simon : « Révision » traduisant la remise en cause du processus ou d'une de ses étapes. Cette quatrième étape peut être vue comme une étape d'évaluation du processus afin de valider la décision à appliquer. Le processus décisionnel de Simon est présenté sur la Figure 7.1.

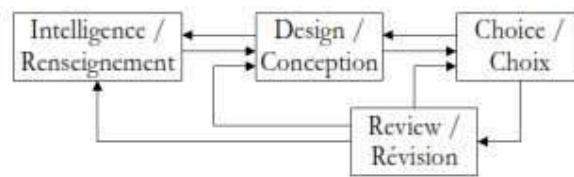


FIGURE 7.1 – Processus décisionnel de Simon.

Courbon a également proposé une modélisation du processus décisionnel, sous forme cyclique (la Figure 7.2), permettant un rebouclage et un réajustement au fur et à mesure de l'évolution du processus.

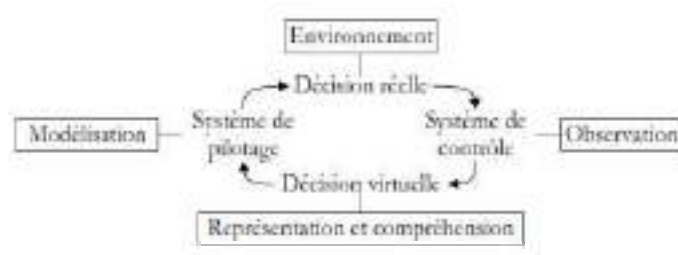


FIGURE 7.2 – Cycle de la décision Courbon.

Cardinal a proposé une autre décomposition des étapes conduisant à une décision. Dans ses travaux, il a identifié six étapes traduisant le processus décisionnel, représentées sur la Figure 7.3.



FIGURE 7.3 – Processus décisionnel de Lavergne.

Lavergne a contribué également à la modélisation du processus de décision pour étudier les dysfonctionnements lors du choix des acteurs. Son modèle intitulé DTL (Decision Time Line) repose sur la perception de la décision comme un processus inscrit dans le temps, se déroulant en plusieurs étapes (voir la Figure 7.4).

Plus récemment, d'autres auteurs se sont intéressés à la modélisation de la prise de décision, en considérant la capitalisation des connaissances (voir la Figure 7.5) comme, par exemple (Cauvin, 2005).

Les modèles que nous venons de présenter, malgré leurs particularités, possèdent plusieurs points communs :

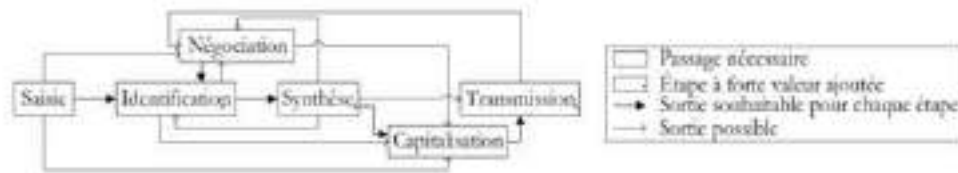


FIGURE 7.4 – Modèle DTL (Cardinal, 2000)

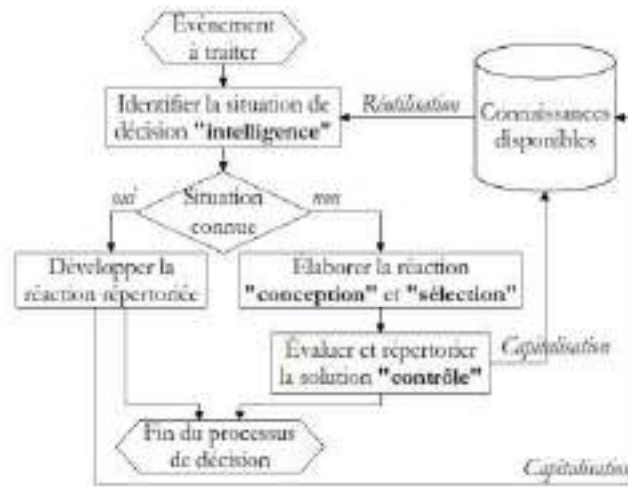


FIGURE 7.5 – Processus décisionnel de Cauvin.

- Une analyse du problème à résoudre et de son environnement ;
- Une recherche d'information sur ce problème afin de construire des solutions possibles ;
- Une évaluation de ces solutions, selon des critères de sélection ;
- Une sélection de la solution la plus adaptée, au vu des critères et du problème considéré ;
- Une application de cette solution ;
- Un suivi et une capitalisation de cette solution après son application.

Ces modèles de processus décisionnel montrent bien la succession d'étapes nécessaires à une prise de décision et la complexité des activités décisionnelles qui font appel à différents mécanismes.

7.2.3 Processus de décision collaborative

Le processus de décision collaborative traduit le déroulement d'une activité assez complexe, pour laquelle il est difficile de trouver une définition consensuelle. Les modèles de processus existants concernent, pour la plupart, l'activité de conception. Voici deux exemples suivis d'un autre modèle de décision collective plus générique.

Kvan et Vera ont proposé un modèle de processus de conception collaborative. Ils décomposent l'activité de conception en une succession d'étapes individuelles ou collectives (voir la Figure 7.6). Ces auteurs considèrent que les acteurs de la collaboration peuvent travailler individuellement sur des portions du problème.

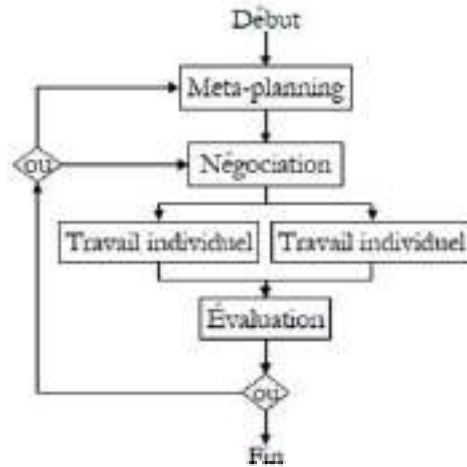


FIGURE 7.6 – Modèle du processus de conception collaborative.

Chiu a travaillé également sur les processus de conception collaborative et a proposé un processus cyclique permettant de reboucler tant que l'objectif recherché n'est pas atteint comme cela est montré à la Figure 7.7.

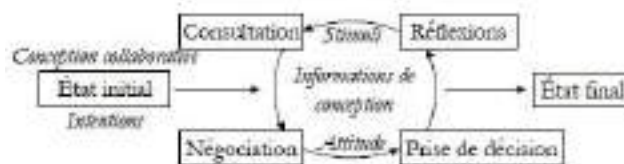


FIGURE 7.7 – Modèle du processus de conception collaborative.

Laborie a proposé également un modèle du processus de conception du concept de salles de décisions collectives appliqué au contexte aéronautique, inspiré notamment du processus de Kvan et Vera mais en lui conférant un caractère plus générique. Il a décliné le concept de salle selon quatre dimensions : organisationnelle, humaine, physique et technologique. Ces dimensions permettent de comprendre le processus de décision collective à partir d'études de cas. Il abstrait un unique modèle, définissant les phases majeures de la décision collective. L'approche de la conception d'une salle est ainsi guidée par le modèle qui illustre les aspects principaux (besoins, enjeux) à considérer. Il applique le concept à des cas d'études, au travers de la conception de

salles. Pour cela il a adopté une méthode de conception itérative et participative, appliquée aux contextes opérationnels. Les « salles » conçues ont été évaluées dans leurs contextes de décisions collectives respectifs. Sa représentation du processus de décision collective est résumée sur la Figure 7.8.

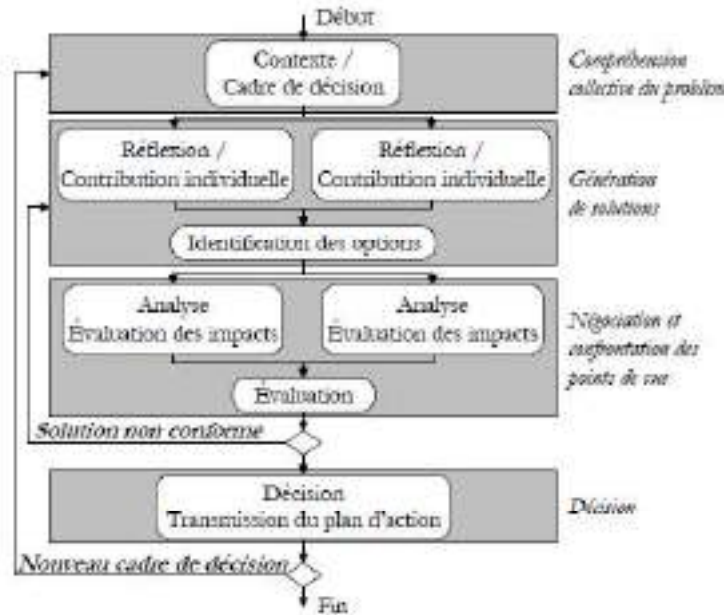


FIGURE 7.8 – Modèle du processus de conception collaborative.

Adla a proposé un processus de prise de décision collective cyclique en trois phases principales : (1) Pré-décision ; (2) Décision ; et (3) Post-décision. Il permet de remettre en cause la liste des solutions et/ou leur organisation en s'appuyant sur le modèle de Simon illustré à la Figure 7.9.

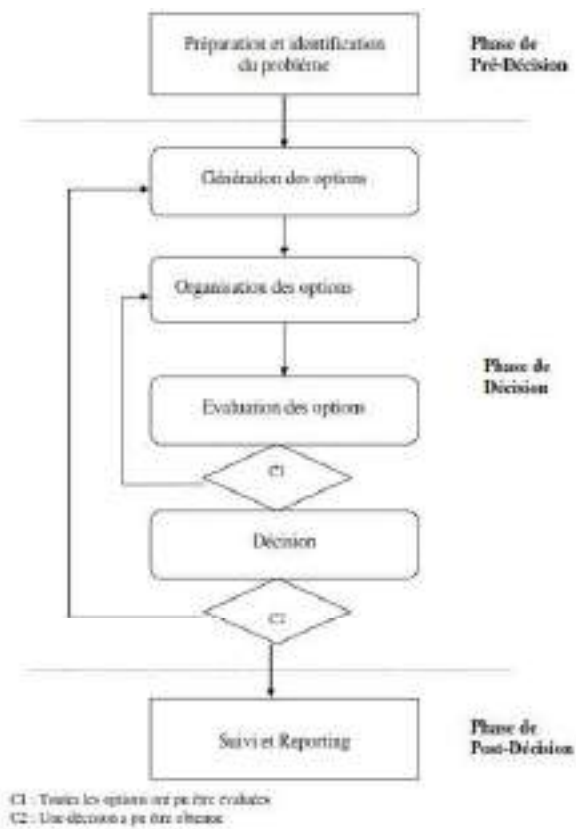


FIGURE 7.9 – Modèle du processus de prise de décision collective.

Benkaddour a proposé aussi un processus d'aide à la décision collaborative qui exploite la connaissance représentée dans une ontologie de domaine du non-tissé et qui utilise un raisonnement à base de cas (voir la Figure 7.10).

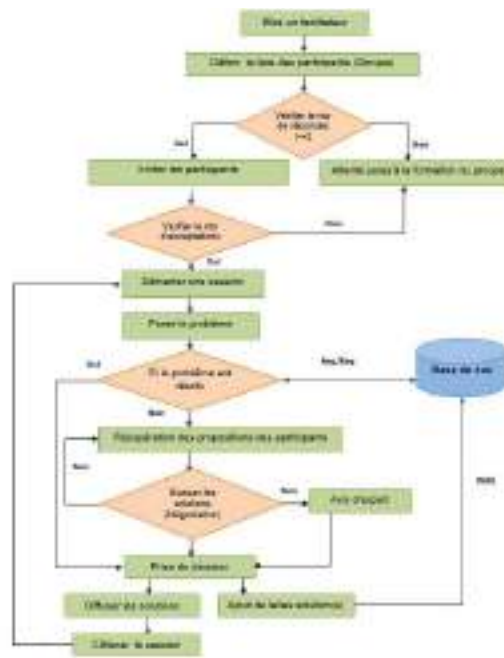


FIGURE 7.10 – Processus du processus de prise de décision collaborative basé sur la connaissance.

7.3 Le travail collaboratif

Le travail collaboratif ne relève pas d'une répartition a priori des rôles. La collaboration s'entend en fait par une situation de travail collectif où tâches et buts sont communs. Chacun des membres du groupe travaille ainsi sur les mêmes points. Il va en effet plutôt s'agir ici de fusionner les contributions individuelles dans l'action. La responsabilité est donc ici collective et incombe au groupe en tant que totalité. Ce mode de travail collectif engage par conséquent une communication régulière entre les membres du groupe et une connaissance précise de la progression de l'action collective. Les interactions interpersonnelles sont donc permanentes afin d'assurer une cohérence globale, condition nécessaire de l'efficacité de l'action et donc de l'atteinte de l'objectif fixé.

La collaboration dans le cadre d'un travail collectif est donc une modalité d'action qui va au-delà de l'action individuelle en s'inscrivant explicitement dans une dynamique d'action collective.

Les apports individuels n'ont donc de sens que par leur intégration, leur fusion à tous les autres et non par leur seule juxtaposition. Autrement dit, il s'agit pour chaque acteur d'un projet d'alimenter ses contributions individuelles par celles des autres. Le travail collaboratif implique de ce fait un engagement mutuel des individus dans un effort coordonné pour effectuer une même tâche, résoudre un ensemble du même problème. Le travail collaboratif nécessite ainsi

des membres d'une équipe, une plus forte interactivité et requiert plus de motivation.

Le facteur humain (confiance, motivation, solidarité, respect, etc.) devenant central dans ce mode de travail collectif, il peut être parfois plus difficile à mettre en œuvre que dans le cadre d'un travail coopératif. Cela étant dit, ce mode d'action collaboratif laisse à penser qu'il permet d'atteindre de meilleures performances en termes de réactivité, d'adaptation à l'inattendu ou encore d'utilisation optimisée des moyens mis à disposition. Par ailleurs, cette capacité d'un groupe à valoriser son capital humain participe d'une intelligence collective. En ce sens, le travail collaboratif incite les individus à mener des activités orientées vers les autres (attention, reformulation et négociation) pour contribuer à un résultat collectif qui est autre chose que la somme des résultats individuels. Une part du succès de la collaboration tiendra notamment à la construction d'une compréhension commune (cf. conscience de groupe) autour des activités menées et donc de connaissances partagées issues des contributions de chacun des membres de l'équipe projet.

Enfin, la collaboration peut être vue comme un processus cyclique qui enchaîne des sous processus de co-réflexion, de co-décision, de co-conception, de co-production, de co-pilotage, de co-apprentissage, etc. (voir la Figure 7.11).

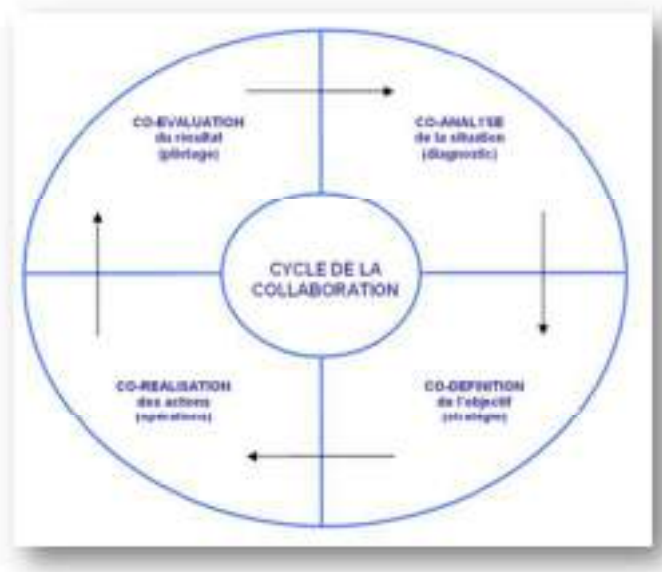


FIGURE 7.11 – Cycle de collaboration.

Ce processus cyclique de la collaboration présente ainsi les phases suivantes :

- **La Co-analyse** : un travail de diagnostic qui permet à un groupe de comprendre une situation donnée et de construire un référentiel cognitif ;
- **La Co-définition** : dans cette phase les interactions collectives permettent au groupe de formuler l'objectif à atteindre en construisant une vision partagée ;
- **La Co-réalisation** : la fixation des règles du jeu, la détermination d'un plan d'action et d'un calendrier pour atteindre l'objectif poursuivi ;
- **La Co-évaluation** : les formes d'interaction qui permettent à un groupe de juger des résultats mais aussi de la valeur de la Co-décision et du Co-apprentissage tout au long du processus.

7.4 Intelligence collective

D'après (Simon, 1969), il s'agit d'une phase de recherche d'information, d'interprétation et de « construction d'une vision » de l'environnement à usage collectif. Selon Courbon, on appellera intelligence collective d'une organisation, sa capacité à dégager des différentes visions des activités de cette organisation et de son insertion dans son environnement (qui sont perçues par les individus et les groupes) une intersection commune lui permettant de diriger d'une manière consciente son évolution. En effet, nous pouvons constater, qu'il existe deux référentiels de mise en place de l'intelligence collective liés à la situation. D'une part, d'une manière quotidienne pour résoudre une situation de travail déjà présente. D'autre part, d'une manière prévisionnelle, pour résoudre une situation de travail planifiée ou imprévue dans le temps et donc à venir.

Ainsi un certain nombre de points semblent faire l'objet d'un consensus et plusieurs hypothèses émergent : L'intelligence collective est un système, on peut dire que c'est la somme des intelligences individuelles des membres d'une équipe plus leurs relations.

- L'intelligence collective est propre à un collectif de travail restreint ;
- L'intelligence collective est un processus. Elle va donc se transformer et évoluer en différents stades ;
- La création de l'intelligence collective est liée à la situation, au contexte de travail ;
- L'intelligence collective est un indicateur d'efficacité d'une équipe de travail.

Autre définition de l'intelligence collective : ensemble des capacités de compréhension, de réflexion, de décision et d'action d'un collectif de travail restreint issu de l'interaction entre ses membres et mis en œuvre pour faire face à une situation donnée présente ou à venir complexe.

7.5 Domaines d'application

Avec le développement des NTIC (Nouvelles Technologies de l'Information et de la Communication), et des besoins, les systèmes collaboratifs sont plus en plus nécessaires et utilisés. Ils sont appliqués dans nombre de domaines comme :

- Les jeux, plus particulièrement des jeux en réseaux ;
- Les Workflows, systèmes qui permettent la coordination. Ce type de système aide plusieurs participants à gérer un processus de travail ;
- L'enseignement où des systèmes collaboratifs reposent sur des médias et des technologies du réseau en vue de faciliter l'apprentissage ;
- La communication où systèmes collaboratifs permettent à plusieurs membres d'un groupe de communiquer et de consulter via des supports audio, vidéos ou textuels ;
- L'espace de travail partagé où des systèmes collaboratifs fournissent un espace partagé dans laquelle des participants peuvent manipuler des artefacts ;
- L'échange d'information, et de résultats entre des chercheurs est important pour le succès d'une action de recherche. Souvent, les chercheurs ne travaillent pas en un même lieu (parfois, ils se situent dans des pays différents). Dans ce contexte, les systèmes collaboratifs sont un bon moyen pour pallier aux contraintes de temps et de distances auxquelles chaque chercheur est confronté. Dans les domaines de la Télémédecine, de l'E-maintenance et de la simulation, quelques outils collaboratifs existent.

7.6 Conclusion

En conclusion, les systèmes d'aide à la décision (SAD) fournissent un cadre structuré pour analyser les informations et prendre des décisions éclairées dans des contextes complexes. Ils améliorent la qualité et la fiabilité des choix organisationnels. Le chapitre suivant s'intéresse aux interfaces homme-machine (IHM), qui facilitent l'interaction entre l'utilisateur et les systèmes informatiques pour rendre ces décisions plus accessibles et efficaces.

Troisième Partie :

INTERFACE HUMAIN-MACHINE

« Cette partie met l'accent sur l'importance de l'ergonomie et de la convivialité pour optimiser l'expérience utilisateur et faciliter l'accès aux fonctionnalités du système ».

8

Interfaces Homme-Machine

Table des matières

8.1	Introduction	81
8.2	Définitions de bases	82
8.3	Objectifs d'une interface homme-machine	82
8.4	Typologie d'une interface homme-machine	82
8.5	Composants d'une interface homme-machine	83
8.5.1	L'utilisateur	83
8.5.2	La tâche	83
8.6	Techniques et styles d'interaction	84
8.6.1	Modes d'interaction	84
8.6.2	Styles d'interaction	84
8.7	Grands modèles d'architecture d'interface	85
8.7.1	Le modèle Seeheim	85
8.8	Interfaces Web	85
8.8.1	Ergonomie des interfaces Web	86
8.9	Conclusion	87

8.1 Introduction

L'interface homme-machine (IHM) constitue un élément fondamental de tout système informatique interactif, car elle conditionne l'efficacité, l'utilisabilité et l'acceptation du système par l'utilisateur. Une interface bien conçue doit être naturelle, fiable et adaptée aux capacités cog-

nitives humaines. L'évolution des technologies a conduit à des interfaces de plus en plus riches et interactives, répondant à des exigences accrues en matière de performance et de convivialité.

Ce chapitre présente les concepts fondamentaux des interfaces homme-machine. Il introduit les notions d'interaction homme-machine, définit les composants et objectifs d'une IHM, puis propose une typologie des interfaces. Enfin, il met en évidence les principaux critères d'évaluation des performances du dialogue homme-machine.

8.2 Définitions de bases

Cette section présente les définitions nécessaires à la bonne compréhension du concept IHM.

Une interface permet l'échange d'informations entre un utilisateur humain et une machine. Afin de faciliter cette communication, différents éléments sont utilisés :

- les **périphériques d'entrée** (clavier, souris, scanner), permettant à l'utilisateur de transmettre des informations ou des commandes ;
- les **périphériques de sortie** (écran, imprimante, diodes), permettant à la machine de restituer des informations.

Une interface homme-machine est un ensemble de composants graphiques organisés selon une structure donnée et capables de gérer les événements issus des interactions utilisateur. Certaines approches modernes remplacent l'objet classique par des agents logiciels, permettant une adaptation dynamique de l'interface et du contrôle des événements.

8.3 Objectifs d'une interface homme-machine

Les principaux objectifs d'une IHM sont :

1. concevoir des systèmes utiles et utilisables ;
2. satisfaire l'utilisateur ;
3. faciliter l'apprentissage.

8.4 Typologie d'une interface homme-machine

L'interaction homme-machine étudie la manière dont les humains interagissent avec les ordinateurs, ainsi que la conception de systèmes informatiques ergonomiques, efficaces et adaptés à leur contexte d'utilisation.

8.5 Composants d'une interface homme-machine

Une interface homme-machine repose sur l'interaction coordonnée de plusieurs composants essentiels permettant d'assurer un dialogue efficace entre l'utilisateur et le système informatique. Ces composants prennent en compte à la fois les caractéristiques de l'utilisateur, la nature des tâches à accomplir, le contexte d'utilisation et les contraintes du système d'information. Leur analyse est indispensable pour concevoir des interfaces ergonomiques, performantes et adaptées aux besoins réels des utilisateurs (voir la Figure 8.1).

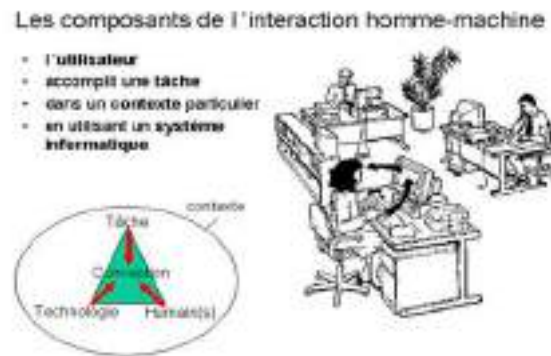


FIGURE 8.1 – Composants d'une interface homme-machine

8.5.1 L'utilisateur

Un système ne peut être adapté à l'ensemble des utilisateurs potentiels. Pour qu'un logiciel soit adapté à une population donnée, il est nécessaire de connaître les caractéristiques de ses utilisateurs (capacités, expérience, formation). Parmi les facteurs liés à l'utilisateur, nous citons :

- **Physiques** : taille, latéralité (droitier/gaucher) ;
- **Psychologiques** : vitesse d'apprentissage, mémorisation ;
- **Socioculturels** : niveau d'étude, origine, pays ;
- **Professionnels** : expérience, compétence dans la tâche.

Nous pouvons classer les utilisateurs comme suit :

- **Novices** : connaissances syntaxiques et sémantiques limitées ;
- **Occasionnels** : bonnes connaissances du domaine, peu de maîtrise de l'outil ;
- **Spécialistes** : besoin de rapidité, de raccourcis et de feedback immédiat.

8.5.2 La tâche

Une tâche est une suite d'actions conduisant à l'atteinte d'un objectif.

L'analyse de la tâche repose sur :

- des organigrammes et scénarios ;
- des simulations et évaluations intégrées dans un processus itératif.

Pour réaliser une tâche, il faut :

1. l'établissement du but ;
2. la formation de l'intention ;
3. la spécification des actions ;
4. l'exécution ;
5. la perception de l'état du système ;
6. l'interprétation ;
7. l'évaluation par rapport au but.

8.6 Techniques et styles d'interaction

Les techniques d'interaction désignent les moyens opératoires de l'échange, tandis que les styles d'interaction caractérisent la manière dont ces échanges sont conduits et interprétés.

8.6.1 Modes d'interaction

Un mode est un état de l'interface dans lequel les actions sont interprétées de manière homogène. On distingue les modes spatiaux et temporels (voir la Figure 8.2).



FIGURE 8.2 – Les modes d'interaction

8.6.2 Styles d'interaction

En interaction homme-machine, les styles d'interaction correspondent aux différents modes de dialogue entre l'utilisateur et le système. Ils jouent un rôle déterminant dans l'utilisabilité et l'expérience utilisateur en structurant la manière dont les commandes et les informations sont échangées. Parmi les principaux styles d'interaction, on trouve les métaphores, les menus, les formulaires, la manipulation directe, les langages de commande et les langages naturels. À ces éléments s'ajoutent la gestion des erreurs et la documentation, qui contribuent à guider l'utilisateur et à assurer une interaction plus fluide et efficace avec le système.

8.7 Grands modèles d'architecture d'interface

Les architectures d'interface ont historiquement été structurées selon plusieurs modèles permettant de séparer les préoccupations liées à l'interaction. On distingue principalement le modèle du dialogue (dit modèle Seeheim), qui organise l'interface en couches séparant la présentation, le contrôle et la gestion des données. Le modèle entrées-sorties se concentre quant à lui sur la gestion des flux d'information entre l'utilisateur et le système. Enfin, le modèle multi-agent réactif repose sur des entités autonomes capables de réagir aux événements et de coopérer pour assurer l'interaction. Ces différents modèles illustrent l'évolution des architectures d'interface afin d'améliorer l'adaptabilité et l'efficacité des systèmes interactifs.

8.7.1 Le modèle Seeheim

Dans ce cours, nous étudierons le modèle Seeheim, qui structure l'interface selon une architecture en trois composantes (comme illustré à la Figure 8.3) : la présentation, chargée de l'affichage et de l'interaction visuelle avec l'utilisateur ; la gestion du dialogue, qui assure la coordination des échanges et le contrôle du flux d'interaction ; et l'interface du noyau fonctionnel, qui établit la communication avec les fonctionnalités principales du système. Cette décomposition vise à séparer clairement les responsabilités afin d'améliorer la modularité et l'évolutivité des interfaces.

8.8 Interfaces Web

Une interface Web est une interface homme-machine reposant sur des pages Web accessibles via un navigateur.

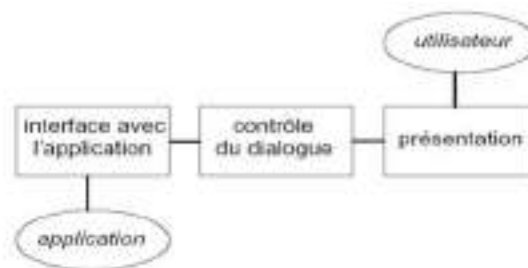


FIGURE 8.3 – Le modèle Seeheim

8.8.1 Ergonomie des interfaces Web

La compatibilité est un critère ergonomique essentiel dans la conception d’une interface web. Malgré les efforts de normalisation du W3C, la diversité des environnements techniques impose de concevoir des sites accessibles au plus grand nombre. Contrairement à un site Intranet, un site Internet doit fonctionner pour la majorité des utilisateurs.

La compatibilité concerne les navigateurs, les versions, les plateformes et les modes d’affichage. Un site ne doit pas être optimisé uniquement pour les versions récentes des navigateurs, ni pour une configuration matérielle spécifique. Il doit également être testé sur différents systèmes et privilégier des technologies largement supportées afin d’éviter toute restriction d’accès.

La navigation constitue l’essence même d’un site web. Sans elle, le Web ne serait qu’un vaste ensemble de pages indépendantes, dépourvues de cohérence. Il est donc essentiel de lui accorder une attention particulière afin de garantir une navigation claire, logique et efficace.

Les menus jouent un rôle central dans l’organisation de la navigation. Ils doivent être structurés par ordre d’importance et informer explicitement l’utilisateur sur le contenu auquel ils donnent accès. Afin de respecter les capacités de la mémoire à court terme, le nombre d’éléments d’un menu ne devrait pas dépasser sept liens. Par ailleurs, il est important de rester constant dans la langue utilisée et d’éviter l’introduction de termes étrangers lorsque le menu est rédigé dans une autre langue.

Les modes de navigation doivent prendre en compte la diversité des profils d’utilisateurs. Certains internautes naviguent de manière exploratoire en suivant les liens proposés, tandis que d’autres souhaitent accéder rapidement à une information précise. Il convient donc de proposer à la fois des menus de navigation pour les premiers et un module de recherche efficace pour les seconds.

La qualité du contenu constitue un facteur déterminant de l’efficacité d’un site web. Un site repose avant tout sur la pertinence et la valeur des informations qu’il propose. Pour capter et maintenir l’attention de l’utilisateur, le contenu doit être structuré, pertinent et utile.

Une attention particulière doit être portée à l'orthographe et à la grammaire. Les erreurs linguistiques nuisent à la crédibilité du site et peuvent entraver la compréhension des informations. Une relecture rigoureuse est donc indispensable avant la mise en ligne.

La validité du contenu est également essentielle. Dans un environnement où l'information est abondante et de qualité variable, il est recommandé de s'appuyer sur des sources fiables et de les citer explicitement. L'indication des dates de création et de mise à jour des pages contribue également à renforcer la confiance de l'utilisateur.

Sur le Web, la concision et la clarté sont primordiales. L'internaute ne lit pas de manière linéaire, mais parcourt rapidement les pages, la durée moyenne de consultation dépassant rarement trente secondes. Il est donc recommandé de présenter l'information selon le principe de la pyramide inversée, où les éléments essentiels apparaissent avant les détails, et d'éviter l'usage de termes ou de libellés peu explicites.

L'identité visuelle joue un rôle croissant dans l'expérience utilisateur. L'aspect graphique d'un site doit être soigné afin d'améliorer son attractivité et sa lisibilité. Pour garantir une cohérence visuelle, la structure des pages, notamment l'emplacement des menus et du logo, doit rester identique sur l'ensemble du site.

Le logo constitue un élément central de l'identité visuelle d'un projet. Il doit être présent sur toutes les pages, tout en restant sobre afin de ne pas alourdir l'interface ni le temps de chargement.

Enfin, **l'aide à l'utilisateur** ne doit pas être négligée. Tous les internautes ne disposent pas du même niveau de maîtrise des outils informatiques. Il est donc nécessaire de proposer des dispositifs d'assistance, notamment lors d'interactions inhabituelles. La mise en place d'une foire aux questions (FAQ) permet de répondre aux interrogations les plus fréquentes et doit être accompagnée d'un moyen de contact avec les webmasters. L'utilisateur doit en effet toujours pouvoir entrer en contact avec les responsables du site, par le biais d'un lien de contact ou d'une adresse électronique accessible depuis chaque page.

8.9 Conclusion

En conclusion, le dialogue homme-machine a considérablement évolué, donnant naissance à des interfaces plus conviviales et efficaces. L'ergonomie repose sur la compatibilité, la clarté de la navigation, la qualité du contenu et la cohérence visuelle. Un choix judicieux des styles d'interaction et une structuration adaptée de l'information permettent à l'utilisateur d'accéder rapidement aux fonctionnalités, de réduire les erreurs et d'améliorer son efficacité, faisant de

l'ergonomie un facteur clé de l'expérience utilisateur. La prochaine partie sera consacrée aux examens et exercices, permettant de mettre en pratique les concepts étudiés dans ce cours.

Quatrième Partie :

EXAMENS TYPES ET EXERCICES

« Cette partie rassemble examens types, corrigés et exercices pour réviser et appliquer les concepts du cours ».

9

Examen Type 1

Université/ Ecole supérieure :Ecole Normale Supérieure d'Oran

Faculté / Département : Sciences exactes / Informatique

Module : Techniques avancées

Enseignant : Dr. BENKADDOUR Fatima Zohra

Niveau : 5ème PES

Année universitaire : 2021 / 2022

Nom et Prénom :

Groupe :

Examen

Exercice 1 (14 points)

1) Les algorithmes génétiques donnent la solution exacte d'un problème :

- ☐ Toujours
- ☐ Jamais
- ☐ Quelquefois

2) Un algorithme génétique ne peut résoudre qu'un seul type de problèmes :

- ☐ Oui
- ☐ Non

— ☐ Je ne sais pas

3) La définition correspondant à un **réseau de neurones** est :

- ☐ Représentation multidimensionnelle des données importantes
- ☐ Arbre graphique de règles de décision
- ☐ Assemblage d'objets informatiques inspiré du cerveau humain
- ☐ Outil logique de modélisation de choix complexes

- 4) La(les) proposition(s) exacte(s) :
 - ☐ Utilisés pour prédiction, classification et analyse
 - ☐ Chaque neurone traite l'information indépendamment
 - ☐ Le cerveau humain contient environ 1 milliard de neurones
 - ☐ Inspirés du cerveau humain et implémentés sur ordinateur
- 5) Année d'apparition des algorithmes de fournis :
 - ☐ Années 1910
 - ☐ Années 1940
 - ☐ Années 1991
 - ☐ Années 2000
- 6) Proposeurs du Ant System (AS) :
 - ☐ Dorigo, Colomi et Maniezzo
 - ☐ Minsky & Papert
 - ☐ Laurel & Hardy
 - ☐ Hebb & Hopfield
- 7) Termes associés aux réseaux de neurones :
 - ☐ Perceptron
 - ☐ Cartes auto-organisatrices de Hopfield
 - ☐ Adaline
 - ☐ Réseaux mono-couche
- 8) Composants d'un neurone biologique :
 - ☐ Dendrites
 - ☐ Synapses
 - ☐ Axone
- ☐ Corps cellulaire
- 9) Composants d'un neurone artificiel :
 - ☐ Poids
 - ☐ Couches
 - ☐ Signaux d'entrée
 - ☐ Signal de sortie
- 10) Les réseaux de neurones sont utilisés en :
 - ☐ Datawarehouse
 - ☐ Datamart
 - ☐ Database
 - ☐ Datamining
- 11) Représentation de connaissances incorrecte :
 - ☐ Règles de production
 - ☐ Objets structurés
 - ☐ Réseaux sémantiques
 - ☐ Représentation graphique
- 12) Intérêt du Datamining :
 - ☐ Conservation des données
 - ☐ Amélioration de la rentabilité
 - ☐ Avantage concurrentiel
 - ☐ Augmentation du ROI
- 13) Types de fouilles spécialisées :
 - ☐ Fouille d'image
 - ☐ Fouille de vidéo
 - ☐ Fouille de texte
- 14) Qu'est-ce qu'une méta-heuristique ?
Quels types de problèmes traite-t-elle ?

Exercice 2 (6 points)

Choisir trois (03) questions.

A. SIG

- A.1 Pourquoi stocker les données SIG dans une base de données ?
- A.2 Définir les données manipulées par les SIG.
- A.3 Exemple réel des SIG dans les Jeux Méditerranéens 2022.

B. Aide à la décision

- B.1 Différence entre SAD et SIAD
- B.2 Objectifs d'un SIAD
- B.3 Application d'un SAD dans le domaine médical (COVID-19)

C. Systèmes embarqués

- C.1 Définition
- C.2 Caractéristiques
- C.3 Exemple réel

Bon courage !

10

Corrigé Type 1

Exercice 1 (14 points)

1. Les algorithmes génétiques donnent la solution exacte d'un problème :

Quelquefois.

Justification : Ce sont des méthodes approchées, pas garanties optimales.

2. Un algorithme génétique ne peut résoudre qu'un seul type de problèmes :

Non.

Justification : Ils sont applicables à de nombreux problèmes d'optimisation.

3. Définition correcte d'un réseau de neurones :

Un assemblage d'objets informatiques s'inspirant du fonctionnement du cerveau humain.

4. Proposition(s) exacte(s) :

- Utilisés pour la prédiction, la classification et l'analyse ;
- Inspirés du cerveau humain et implémentés sur ordinateur.

5. Apparition des algorithmes de fourmis :

Année 1991.

6. Proposeurs du Ant System (AS) :

Dorigo, Colorni et Maniezzo.

7. Termes associés aux réseaux de neurones :

- Perceptron ;
- Adaline ;
- Réseaux de neurones mono-couche ;

8. Composants d'un neurone biologique :

- Dendrites ;
- Synapses ;
- Axone ;
- Corps cellulaire.

9. Composants d'un neurone artificiel :

- Poids de connexion ;
- Signaux d'entrée ;
- Signal de sortie.

10. Domaine d'utilisation des réseaux de neurones :

Datamining.

11. Représentation ne correspondant pas aux systèmes à base de connaissances :

Représentation graphique.

12. Intérêt du Datamining :

- Exploitation des données pour améliorer la rentabilité
- Avantage concurrentiel
- Augmentation du retour sur investissement

13. Types de fouilles par spécialisation

- Fouille d'image
- Fouille de vidéo
- Fouille de texte

14. Méta-heuristique :

Une méta-heuristique est une stratégie générale d'optimisation guidant des heuristiques pour résoudre des problèmes complexes (NP-difficiles), tels que l'optimisation combinatoire, la planification ou l'ordonnancement.

Exercice 2 (6 points)

Réponses (selon les 3 questions choisies)

SIG

Les données SIG sont stockées en base pour assurer la cohérence, le partage, la mise à jour et l'analyse spatiale efficace des informations géographiques.

SAD / SIAD

Un SAD aide à la décision, tandis qu'un SIAD intègre l'informatique pour analyser les données et proposer des solutions aux décideurs.

Systèmes embarqués

Un système embarqué est un système informatique dédié à une fonction spécifique, intégré dans un dispositif plus large (ex : ABS automobile, pacemaker, drone).

11

Examen Type 2

Université/ Ecole supérieure :Ecole Normale Supérieure d'Oran

Faculté / Département : Sciences exactes / Informatique

Module : Techniques avancées

Enseignant : Dr. BENKADDOUR Fatima Zohra

Niveau : 5ème PES

Année universitaire : 2021 / 2022

Nom et Prénom :

Groupe :

Examen

Exercice 1 (14 points)

1. Expliquer qu'est-ce qu'un problème du voyageur et par quoi peut-on le résoudre (selon les méthodes vues en cours).
2. Qu'est-ce qu'un algorithme bio-inspiré ? Expliquer en donnant des exemples.
3. Quelle est la différence entre les termes suivants : **données**, **information** et **connaissance** ?

4. Résumer en quelques lignes l'utilité et la démarche de la gestion de connaissances dans une entreprise. Quels outils peuvent être utiles ?
5. Qu'est-ce qu'une fouille de texte ? Donner un exemple concret.
6. Quels sont les types d'utilisateurs d'une interface graphique ?

Exercice 2 (6 points)

Répondre aux questions suivantes selon le thème de votre exposé.

1. Donner deux champs d'application du système choisi dans l'exposé, expliquer.
2. Résumer en quelques lignes :
 - les entrées du système,
 - le traitement effectué,
 - les sorties (résultats attendus),tout en précisant les éléments indispensables à son bon fonctionnement.
3. Comment l'ENSO peut-elle intégrer votre système ? Expliquer.

Bon courage !

12

Corrigé type 2

Exercice 1

1. **Expliquer qu'est-ce qu'un problème du voyageur et par quoi peut-on le résoudre.**

- Le problème du voyageur de commerce est un problème d'optimisation qui consiste à déterminer, étant donné une liste de villes et les distances entre toutes les paires de villes, le plus court circuit passant par chaque ville une et une seule fois.
- Il peut être résolu à l'aide de méta-heuristiques, notamment les algorithmes bio-inspirés.

2. **Qu'est-ce qu'un algorithme bio-inspiré ? Donner des exemples.**

- L'informatique bio-inspirée (ou naturo-inspirée) s'inspire des systèmes naturels pour concevoir des solutions informatiques.
- Exemples : algorithmes génétiques, colonies de fourmis, réseaux de neurones.

3. **Différence entre données, information et connaissance.**

- **Donnée** : description élémentaire d'une réalité. Exemple : 01/11/1954.
- **Information** : donnée interprétée et porteuse de sens. Exemple : 01/11/1954 correspond à une date.
- **Connaissance** : interprétation humaine de l'information. Exemple : reconnaître le 01/11/1954 comme une date historique.

4. **Utilité et démarche de la gestion des connaissances dans une entreprise.**

- a) **Repérer** : identifier les experts et les connaissances clés.

- b) **Capitaliser** : stocker les connaissances pour éviter leur perte.
- c) **Transférer** : partage direct ou indirect des savoirs.
- d) **Actualiser** : mise à jour continue des connaissances.

5. **Qu'est-ce qu'une fouille de texte ? Donner un exemple.**

- Le text-mining permet d'extraire des connaissances utiles à partir de documents textuels grâce à des techniques statistiques et de machine learning.
- **Exemple** : détection automatique des spams.

6. **Types d'utilisateurs d'une interface graphique.**

- **Novices** : faibles connaissances informatiques.
- **Occasionnels** : bonnes connaissances du domaine mais peu techniques.
- **Spécialistes** : utilisateurs experts exigeant rapidité et efficacité.

Exercice 2

Les réponses sont traitées au cas par cas. Toute réponse logique, cohérente et bien argumentée sera acceptée.

— Fin du corrigé —

13

Travail Dirigé

Exercice 1 : Modélisation du problème

Selon l'algorithme génétique vu en cours, donner d'une manière générale une modélisation du problème du voyageur de commerce (*Travelling Salesman Problem : TSP*) en détaillant toutes les étapes de l'algorithme génétique (AG).

1. Quel est l'objectif de cette étude ?
2. Structurer le problème du TSP selon les paramètres de l'algorithme génétique.
3. Initialiser la population.
4. Définir le codage :
 - **Codage direct** : l'individu représente directement la solution du problème à résoudre.
 - **Codage indirect** : l'individu n'exprime pas directement la solution, mais une manière de la construire.
5. Appliquer l'opérateur de croisement à un point.
6. Appliquer la mutation aléatoire.
7. Arrêter l'algorithme à la génération 5.

Exercice 2 : Opérateur de croisement

Croisement à deux points :

En choisissant aléatoirement deux points de coupure, chaque parent est dissocié en trois fragments. Les deux fragments aux extrémités du Parent 1 (respectivement Parent 2) sont copiés dans l'Enfant 1 (respectivement Enfant 2).

La partie restante de l'Enfant 1 est complétée par les éléments du Parent 2, et celle de l'Enfant 2 par les éléments du Parent 1, en balayant de gauche à droite et en ne reprenant que les éléments non encore transmis.

- Appliquer ce type de croisement à l'exercice 1.

Exercice 3 : Opérateur de mutation

- **Opérateur d'insertion** : consiste à sélectionner au hasard un gène et une position dans le chromosome, puis à insérer le gène sélectionné dans la position choisie.
- **Opérateur d'échange réciproque** : consiste à sélectionner deux gènes et à échanger leurs positions.
- Appliquer ces opérateurs de mutation à l'exercice 1.

Références

- Baesler, S. (1997). *Recommandations ergonomiques pour la création de pages web*. (Bureau qualité)
- Barabel, M. (1996). Un style de décision à la française? *Revue française de gestion*, 159-170.
- Bellut, S. (2002). *Les processus de la décision : Démarches, méthodes et outils* (Paris éd.). AFNOR.
- Benkaddour, F. Z. (2017). *Intégration d'une ontologie et des technologies web 2.0 dans la conception d'un système d'aide à la décision collaborative (sadc) : Cas de l'industrie du non-tissé* (Thèse de doctorat non publiée). Université Oran 1 Ahmed Ben Bella.
- Benkaddour, F. Z., & Benhacine, F. Z. (2013). *Réalisation d'une interface d'un site web pour une agence de voyage* (Mémoire de fin d'étude). Université d'Oran 1 Ahmed Ben Bella.
- Benkaddour, F. Z., Taghezout, N., & Ascar, B. (2016). Towards a novel approach for enterprise knowledge capitalization utilizing an ontology and collaborative decision-making : Application to inotis enterprise. *International Journal of Decision Support System Technology (IJDSST)*, 8(1), 1–24.
- Berchet, C. (2000). *Modélisation pour la simulation d'un système d'aide au pilotage industriel* (Thèse de doctorat non publiée). INP de Grenoble.
- Bernoussi, S. (s. d.). *Cours c2si – mathématiques pour l'informatique*. Consulté sur <http://www.fsr.ac.ma/DOC/cours/maths/Souad%20Bernoussi/Cours%20C2SI.pdf> (Consulté le 29/01/2026)
- Bordage, S., Thévenon, D., & Dupaquier, L. (2005). *Conduite de projet web*. Groupe Eyrolles.
- Cardinal, J. S. L. (2000). *Étude des disfonctionnement des processus de décision, application au choix d'acteur* (Thèse de doctorat non publiée). École Centrale Paris.
- Cauvin, A. (2005). *Analyse, modélisation et amélioration de la réactivité des systèmes de décision dans les organisations industrielles* (Thèse de doctorat non publiée). Université Paul Cézanne Aix-Marseille III.
- Chaton, A.-C. V. (2008). *La gestion des connaissances*. Consulté le 08/04/2016, sur <http://>

- www.memoireonline.com/07/08/1199/m_la-gestion-des-connaissances8.html
- Contensin, M. (2004). *Base de données et internet avec php et mysql*. Paris : Dunod.
- Courbon, J.-C. (1979). *Aide à la décision et intelligence collective*. Grenoble, France : CERAG.
- Courbon, J.-C. (1982). Processus de décision et aide à la décision. *Économies et Sociétés*, 16(12), 1455-1476.
- Cours optimisation et métaheuristiques*. (s. d.). Consulté sur https://elearning-facsci.univ-annaba.dz/pluginfile.php/9632/mod_resource/content/2/Cours%20Optimisation%20%20m%C3%A9taheuristiques.pdf (Consulté le 29/01/2026)
- Courteille, A., Allot, P., Tarditi, J.-P., Ermine, J.-L., & Marc, L. (2001). Ingénierie des connaissances et innovation. application dans le domaine automobile. *Extraction des connaissances et apprentissage*, 1(4), 203-220.
- Djeffal, A. (s. d.). *Cours sur les réseaux de neurones artificiels*. Consulté sur http://www.abdelhamid-djeffal.net/web_documents/coursrna.pdf (Consulté le 29/01/2026)
- Dominé, C.-H. (1988). *Techniques de l'intelligence artificielle : un guide structuré*. Bo-Pré.
- Ermine, J.-L. (2001). Les processus de la gestion des connaissances. In *Egc* (p. 17-29).
- Greselle, O. Z. (2007). Vers l'intelligence collective des équipes de travail : une étude de cas. *Management et Avenir*, 4(14), 41-59.
- Grundstein, M. (2002). Gameth : un cadre directeur pour repérer les connaissances cruciales pour l'entreprise. *Lamsade Université Paris-Dauphine*.
- Grundstein, M., et al. (2002). De la capitalisation des connaissances au renforcement des compétences dans l'entreprise étendue. In *Conférence invitée, 1er colloque du groupe de travail "gestion des compétences et des connaissances en génie industriel"*, nantes (Vol. 103).
- H, K. (1985). Un outil d'analyse pour les langages de représentation de connaissances. In *5ème congrès rfa* (p. 1099-1112).
- Herbert Alexander, S. (1960). *The new science of management decision* (Harper and Row éd.). AFNOR.
- Herbert Alexander, S. (1977). *The new science of management decision* (revised éd.). Prentice-Hall.
- Imene, B., & Hassiba, T. B. (2007). *Réalisation d'une interface graphique et animée du projet siamd* (Mémoire de fin d'étude). Université d'Oran.
- Introduction aux algorithmes génétiques*. (s. d.). Consulté sur http://ressources.unit.eu/cours/EnsR0tice/module_avance_thg_voo6/co/gaintro.html (Consulté le 29/01/2026)
- IONOS. (2026). *Text mining : Définition, techniques et exemples*. Consulté sur <https://>

www.ionos.fr/digitalguide/sites-internet/developpement-web/text-mining/
(Consulté le 31 janvier 2026)

- Kast, R. (1993). La théorie de la décision. In (La Découverte éd.).
- Kvan, T. (2000). Collaborative design : What is it ? *Automation in construction*, 9(4), 409-415.
- Lajnef, M. A., Ayed, M. B., & Kolski, C. (2005). Convergence possible des processus du data mining et de conception-évaluation d'ihm : adaptation du modèle en u. In *Proceedings of the 17th conference on l'interaction homme-machine* (pp. 243–246).
- Lamidieu, E. (2007). *Travail collaboratif assisté par ordinateur*. Consulté le 14/01/2014, sur http://www.eric.lamidieu.org/public/Pdf/TCA0_Travail_Collaboratif_Assist_par_Ordinateur.pdf
- Lavergne, J.-P. (1983). *La décision : psychologie et méthodologie : connaissance du problème, applications pratiques*. ESF-Entreprise moderne d'édition : Librairies techniques.
- Leake, D. B., & Kinley, A. (1996). Acquiring case adaptation knowledge : A hybrid approach.
- Levan, S. K., & Vickoff, J.-P. (2004). *Travail collaboratif sur internet : concepts, méthodes et pratiques des plateaux projet*. Vuibert.
- LIRIS. (2003). *Notion d'ontologie*. Consulté le 24/01/2015, sur http://liris.cnrs.fr/amille/enseignements/DEA-ECD/ontologies/notion_ontologie.htm
- Longueville, B. (2003). *Des processus de décision dans les projets d'innovation : Application à l'automobile* (Thèse de doctorat non publiée). École Centrale Paris.
- Moigne, J.-L. (1974). *Les systèmes de décision dans les organisations* (Paris éd.). Presses Universitaires de France.
- Nonaka, I., & Takeuchi, H. (1997). *La connaissance créatrice : la dynamique de l'entreprise apprenante*. De Boeck Supérieur.
- Paul, R., Debove, J. R., & Rey, A. (2004). *Le nouveau petit robert : dictionnaire alphabétique et analogique de la langue française* (Paris éd.). Dictionnaires le Robert.
- Piquet, A. (2009). Guide pratique du travail collaboratif : Théories, méthodes et outils au service de la collaboration. *Document destiné au «Groupe Communication» du réseau Isolement Social*, 17(2), 7-9.
- Poyet, F., & Drechsler, M. (2009). *Impact des tic dans l'enseignement : une alternative pour l'individualisation ?* Consulté le 15/04/2016, sur <http://ife.ens-lyon.fr/vst/LettreVST/41-janvier-2009.php>
- Prax, J.-Y. (2003). Le manuel du knowledge management(une approche de 2 ème génération). *Collection Stratégies et management*.
- Prax, J.-Y. (2012). *Manuel du knowledge management-3ème édition : Mettre en réseau les*

- hommes et les savoirs pour créer de la valeur*. Dunod.
- Quillian, M. R. (1968). *Semantic memory. semantic information processing* (Marvin Minsky éd.). Cambridge, MA : MIT Press.
- québécois de la langue française, O. (2014). *Grand dictionnaire terminologique*. Consulté le 30/04/2014, sur <http://gdt.oqlf.gouv.qc.ca/>
- Radet, S. (s. d.). *Document de soutenance (informatique)*. Consulté sur <http://deptinfo.unice.fr/twiki/pub/Linfo/PlanningDesSoutenances20032004/Radet-Souquet.pdf> (Consulté le 29/01/2026)
- Ramzi, A., & Amel, B. (2009). *Création d'un site web dynamique pour les opérateurs oranais* (Mémoire de fin d'étude). Université d'Oran.
- Rouen, I. (2026). *Introdmbearer.pdf – introduction aux méthodes et à la notion de data mining*. Consulté sur https://moodle.insa-rouen.fr/pluginfile.php/1336/mod_resource/content/4/Parties_1_et_3_DM/IntroDMBearer.pdf (Consulté le 31 janvier 2026)
- Salzberg, S. L. (1994). C4.5 : Programs for machine learning by j. ross quinlan. *Machine Learning*, 16(3), 235-240.
- Seguy, A. (2008). *Décision collaborative dans les systèmes distribués – application à la e-maintenance* (Thèse de doctorat non publiée). L'Institut National Polytechnique de Toulouse.
- Simon, H. (1969). *Sciences of the artificial* (Cambridge : Massachusetts éd.). London, England : The MIT Press.
- Slides de cours sys843*. (s. d.). Consulté sur https://cours.etsmtl.ca/sys843/SLIDES/acetA1_sys843.pdf (Consulté le 29/01/2026)
- Thèse ou document de recherche*. (s. d.). Consulté sur <https://tel.archives-ouvertes.fr/tel-00093143/document> (Consulté le 29/01/2026)
- Toussile, Q. (2026). *Introduction au text mining*. Consulté sur https://toussile.quarto.pub/learn-text-mining/01-introduction_au_text_mining.html (Consulté le 31 janvier 2026)
- Vergnaud, G. (2011). Au fond de l'action, la conceptualisation. *Education et formation*, 275-292.
- Wikipedia. (2015). *Savoir-faire*. Consulté le 14/01/2015, sur https://fr.wikipedia.org/wiki/Gestion_des_connaissances