

Support de Cours de
Pr H.SERIDI-BOUCHELAGHEM

SYSTEME D'INFORMATION

Ce cours est une introduction au cours sur les bases de données et au cours de génie logiciel dans la mesure où les bases de données et les SGBD sont des composants essentiels de tout système d'information moderne et le Génie logiciel présente les méthodologies de développement des logiciels.

Baptisées dépôts communs pour les données, les bases de données sont utilisées par les systèmes d'information qui offrent à une organisation un moyen de partage de données structurées.

Nous présenterons les étapes d'analyse et de conception pour le développement d'un système d'information et nous survolerons la méthode Merise qui reprend le modèle Entité/Association et le langage UML qui définit les diagrammes de cas d'utilisation, d'activité, de séquence et de diagramme de classe et d'autres diagrammes.

Méthodes d'analyse et de conception des Systèmes d'Information

Table des matières

Chapitre 1 : Méthodologie des systèmes d'information : Définition et Organisation

- 1.1 Notion de Système et de Système d'Information
- 1.2 Les différentes formes de structures de l'entreprise
- 1.3 Les principales fonctions de l'entreprise
- 1.4 Les différents sous-systèmes de l'entreprise
- 1.5 Place d'un système d'information
- 1.6 Exemple
- 1.7 Automatisation d'un Système d'information
- 1.8 Cycle de développement d'un Système d'information automatisé
- 1.9 Les diagrammes UML pour l'analyse et la conception d'un SI automatisé (diagramme de cas d'utilisation, d'activités et Séquence).

Chapitre 2 : Les outils d'analyse

- 2.1 Définir l'information, sa classe et sa représentation
- 2.2 La codification de l'information
- 2.3 Contrôle de l'information
- 2.4 Stockage de l'information (Buffer, Block, RAID technologie)
- 2.5 Du fichier traditionnel au fichier numérique : Définition
- 2.6 L'approche Base de Données

Chapitre 3 : Les fichiers

- 3-1 Enregistrement, type d'enregistrement, enregistrement de taille fixe et variable
- 3-2 Blocs de fichiers, allocation des blocks sur disque et en-tête de fichiers
- 3-3 Opérations sur les fichiers
- 3-4 Fichiers à enregistrements triés et non triés
- 3-5 Techniques de hachage
- 3-6 Organisations primaires et B-arbre
- 3.7 Exercice

Chapitre 4 : Notion de méthode d'analyse et de conception d'un Système d'Information

- 4.1 Pourquoi une méthode ?
- 4.2 Historique des méthodes
- 4.3 Catégorisation des méthodes
- 4.4 Les méthodes cartésiennes
- 4.5 Les méthodes systémiques
- 4.6 Les méthodes orientées objet
- 4.7 Exemples de méthode

Chapitre 4 : Modélisation des données par niveau d'abstraction 5-1**Niveau conceptuel : Le modèle Entité-Association****Construction d'une liste d'informations****Qu'est ce qu'une donnée ?****Formalisme du modèle E/A et du M.C.D.****Les Contraintes d'Intégrité Fonctionnelle (C.I.F.)****La vérification du Modèle conceptuel de données****Présentation de la méthode descendante et ascendante****5-2 Niveau Logique : Le modèle Relationnel****5-3 Niveau physique****Chapitre 5 : Modélisation des traitements par niveau d'abstraction****6-1 Niveau conceptuel : Le modèle conceptuel des traitements****6.1.1 Généralités****6.1.2 Acteur****6.1.3 Evénement, Résultat****6.1.4 Opération****6.1.5 Règles d'émissions****6.1.6 Synchronisation****6.1.7 Règles de Syntaxe****6.1.8 La vérification du modèle****6-2 Niveau Logique : Le modèle organisationnel des traitements****6-3 Niveau opérationnel des traitements****Ce cours s'appuie pour l'essentiel sur les ouvrages :****Analyse et conception de systèmes d'information, par Satzinger,
Jackson, Burd, Simond, Villeneuve.**

Chapitre 1 : Méthodologie des systèmes d'information : Définition et Organisation

1.1 Notion de Système et de Système d'Information

Les entreprises sont confrontées à de nouveaux problèmes de types organisationnels et informationnels qui les amènent à revoir et perfectionner leurs méthodes de travail et d'organisation. Pour promouvoir l'entreprise et assurer un bon fonctionnement de ses structures, il faut améliorer son système d'information et de communication. En effet, la connaissance et le contrôle de l'information ainsi que l'exploitation et la mise à disposition de l'information et de façon optimale constituent le challenge de toute entreprise.

Un système d'information d'une entreprise est constitué des éléments matériels ou immatériels qui transforment un flux d'informations en entrées en flux d'informations en sortie.

Dans ce cours nous allons définir et expliquer les outils nécessaires et suffisants pour comprendre et modéliser le fonctionnement d'un système d'information dans une entreprise, afin de proposer et développer des solutions informatiques en vue d'une automatisation de ces systèmes d'information.

1.2 Les différentes formes de structures de l'entreprise

1.2.1 Définition du mot entreprise

Le mot entreprise caractérise une opération de commerce et il fût utilisé après pour décrire une organisation de production de biens ou de service à caractère commerciale.

L'entreprise est définie comme étant une organisation économique et social dans laquelle des décisions sur la combinaison de moyens humains, matériels et informationnels sont prises en vue de créer une valeur ajoutée selon des objectifs fixés.

L'entreprise doit avoir une structure qui répond à une organisation et différentes structures ont été adoptées par les entreprises et nous détaillerons dans ce qui suit les différentes structures et les avantages et les inconvénients de chacune d'elles.

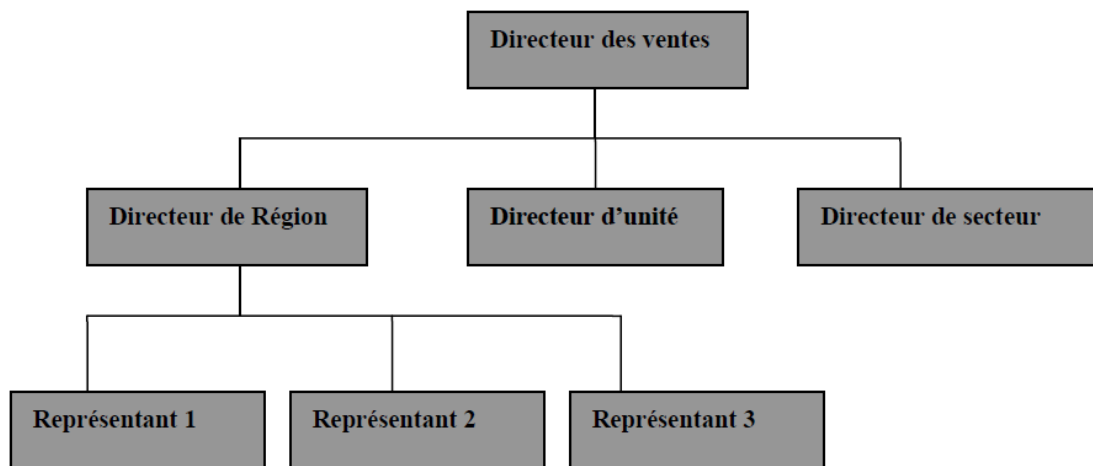
1.2.2 Les structures traditionnelles

La structure représente les rapports entre les tâches et les individus qui occupent des postes et ce depuis la naissance des entreprises. Les structures adoptées par les entreprises n'ont pas cessé de se modifier et deux catégories de structures émergent : les structures traditionnelles et les nouvelles structures.

Parmi les structures traditionnelles connues, nous pouvons citer : les structures hiérarchiques, les structures fonctionnelles et les structures Staff and Line.

La structure hiérarchique

Ce type de structure est le plus ancien. Il s'agit d'une organisation pyramidale de type militaire. Tous les subordonnés n'ont qu'un chef qui a une délégation générale sur tous les aspects de leurs actions. Les découpages des responsabilités peut se faire par unité de production, par zone géographique ou par produit.



Exemple de structure hiérarchique dans le service de ventes

Les avantages

- ⊗ **La simplicité** : ce type d'organisation propose une définition claire de l'autorité et des responsabilités de chacun des employés.
- ⊗ **L'efficacité dans la solution des conflits** : seules les personnes les plus compétentes sont habilitées à intervenir pour résoudre les conflits

Les limites

- ⊗ **L'absence de spécialistes fonctionnels** : il est supposé que chaque supérieur a une compétence générale sur tous les problèmes posés.
- ⊗ **La rigidité face aux changements** : le centralisme rend difficile la circulation de l'information.

Les structures fonctionnelles

L'idée de base repose sur l'association de l'autorité et de la compétence. Cette association est souvent impossible, ce qui induit une spécialisation dans un nombre de tâches afin de pouvoir commander ces subordonnés. Chaque subordonné va recevoir des ordres non pas d'un seul mais de plusieurs supérieurs hiérarchiques, chacun dans son domaine de compétence.

Les avantages

- ⊗ **Adaptée aux gestions complexes** : ce type d'organisation est appropriée pour des entreprises à gestion complexe.
- ⊗ **Adaptée pour des entreprises à leurs débuts** : centrées sur les grandes fonctions : production, vente, finance, personnel.

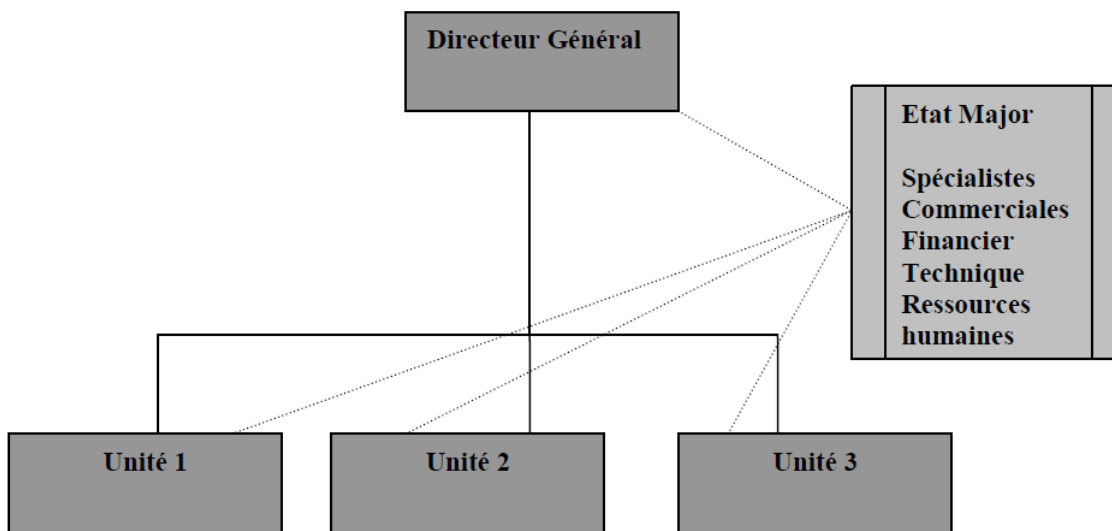
Les limites

- ⊗ **La multiplicité du commandement peut être source de conflits perpétuels.**

La structure STAFF and LINE

Il s'agit de concilier entre la structure hiérarchique et la structure fonctionnelle. Pour protéger l'unité de commandement et la spécialisation des

tâches, deux types de relations sont créés : une **relation LINE** qui exprime le commandement direct sur tous les subordonnés et la relation **STAFF** qui fournit les informations nécessaires aux décideurs.



Les avantages

Cumul des avantages de l'unité de commandement et de la spécialisation.

Les limites

La **cohabitation de deux groupes de personnes** aux fonctions différentes peut engendrer un certain nombre de conflits.

1.2.3 Les nouvelles structures

Après la deuxième guerre mondiale et devant l'évolution des techniques, la croissance de la taille de l'entreprise et le changement des idées, les entreprises ont été contraintes à envisager des structures nouvelles. Deux structures nouvelles ont été créées : la **décentralisation** et la **structure par projet**.

La décentralisation consiste à répartir les tâches de direction et de contrôle entre les niveaux hiérarchique inférieur : c'est une délégation permanente et général d'autorité et de responsabilité. Il existe deux types de décentralisation :

- ⊗ La décentralisation fonctionnelle qui est une délégation d'autorité qui porte sur certaines fonctions précises de l'entreprise (production, achat,...)
- ⊗ La décentralisation fédérative qui consiste à déléguer l'autorité par toute une division ou toute une usine par exemple.

La structure par projet consiste à faire des regroupements dans l'entreprise, non plus par fonction mais par projet en constituant des équipes de spécialistes d'origines diverses qui sont ensemble chargées de la réalisation complète d'un projet.

1.3 Les principales fonctions de l'entreprise

a- La fonction de production est la production de la marchandise et des services pour les échanger dans le marché.

- b- La fonction de commercialisation** permet de commercialiser les produits de l'entreprise afin de créer une valeur ajoutée à l'entreprise.
- c- La fonction administrative** pour fournir des services non commerciale tel que l'administration publique.

Pour que l'entreprise puisse assurer ces fonctions, elle a besoin de certains moyens tel que les moyens humains, les moyens matériels, les moyens financiers et les moyens informationnels.

1.4 Les différents sous-systèmes de l'entreprise

a- Définition d'un système

Un système est un ensemble d'éléments en interaction dynamique organisés en fonction d'un but. D'après cette définition, l'entreprise est donc un système, composé d'un ensemble d'éléments (services, département, etc....). Ces éléments de l'entreprise sont en interaction dynamique les uns avec les autres et avec le monde extérieur (fournisseur, client, ...,etc).

b- La structure d'un système d'entreprise

L'entreprise étudiée comme un système peut être décomposée en trois sous-systèmes :

⊗ **Sous système de pilotage**

C'est le système nerveux de l'entreprise. Il prend les décisions, fixe les objectifs et les moyens de les atteindre. Dans ce sous système, les informations sont perçues, analysées et combinées suivant un processus définis pour enfin générer de nouvelles informations (ou décision).

⊗ **Sous système opérant**

C'est la partie la plus évidente de l'entreprise qui réalise toutes les tâches de l'entreprise. Exemple : les employés dans une chaîne de production.

⊗ **Sous système d'information**

Le système d'information sert à traiter l'information et à la véhiculer entre le système de pilotage et opérant. Le (SI) fournit et reçoit des informations du monde extérieur.

1.5 Place d'un système d'information

Le système d'information (S.I.) d'une organisation étant l'ensemble des éléments chargés de stocker et de traiter les informations (ordinateurs, postes de travail, règles et méthodes) est en interface entre les deux autres systèmes de l'organisation.

Rôle d'un système d'information au sein d'une entreprise

D'après les définitions précédentes, on peut dégager quatre fonctions d'un système d'information.

1. **Collecter** les informations qui proviennent des autres sous-systèmes de l'entreprise et du monde extérieur.
2. **Mémoriser** les données manipulées par le système.
3. **Traiter** les informations mémorisées.
4. **Transmettre** les informations vers les autres composants du système d'entreprise et le monde extérieur.

Liaisons entre les trois types de systèmes

- A. Le sous-système d'information perçoit et stocke les informations provenant du système opérant et du monde externe, ces informations sont des données élémentaires d'usage multiple généré par des événements internes ou externes. Exemple : absence d'un ouvrier comme événement interne et l'arrivée d'un bon de commande qui est événement externe.
- B. Le système d'information fournit des informations au système de décision (le système de pilotage). Ces informations sont des messages de gestion servant d'entrées pour les décisions. Exemple : situation de la trésorerie.
- C. Le système de décision fournit des informations au (SI). Ces informations appartiennent à deux catégories différentes : des ordres à transmettre et des messages particuliers destinés au (SI) lui-même pour la construction des messages d'exécution.
- D. Le (SI) fournit des informations au système opérant, il s'agit des messages d'exécutions indiquant les opérations à réaliser et la condition de leur réalisation. Exemple : un bon de sortie magasin.

Le système d'information fournit des informations à l'extérieur, se sont des messages qui traduisent les opérations effectuées avec l'extérieur. Exemple : la facture qui traduit l'opération de vente.

1.6 Les fonctions d'un Système d'information

Les fonctions d'un SI sont diverses :

- ☞ Fonction de communication,
- ☞ Aide à la décision en utilisant les données mémorisés il est utiles pour les décideurs d'avoir des statistiques, l'évolution des activités,...toutes les informations pertinentes et fiables.

Un Système d'information doit permettre un accès rapide à l'information pour garantir des décisions pertinentes, fiables et cohérentes.

1.7 Exemple d'application

Soit une entreprise commerciale qui met en vente une liste de produits :

En entrée Les produits achetés, les commandes, les factures de paiements,

En sortie Les produits vendus, les factures, chèques de règlements

Nous devons distinguer entre les deux flux :

Flux physiques produits achetés, produits vendus

Flux d'information paiements clients/fournisseurs

1. Quelles sont les informations qui sont manipulées par le système d'information de l'entreprise en question.
2. Donner une description succincte de l'entreprise.

Corrigé

1. Le système d'information comprend les informations :

Sur les flux (produits en stock, produits commandés,...)

Sur l'environnement extérieur (clients, fournisseurs, ...)

Sur l'organisation de l'entreprise (que se passe t-il entre l'enregistrement d'une commande et sa livraison ?)

Sur les règles de gestion et contraintes légales (gestion, lois, règlement, paramètres financiers,...)

2.

☺ Le système " entreprise " est composé d'éléments tels que " employés ", " services " et " articles ". Les propriétés décrivant ces éléments peuvent être le " matricule de l'employé ", son " nom ", la " référence de l'article ", sa " désignation ", ...

☺ Un système est également caractérisé par son environnement. Il subit de la part de celui-ci un certain nombre de contraintes qui viennent le perturber et l'obligent à réagir en déclenchant des activités tendant à le ramener à un état stationnaire, dans l'attente d'une nouvelle perturbation.

☺ Le système " entreprise " reçoit des " commandes " de clients qui doivent être traitées jusqu'à leur aboutissement soit sous forme de " rejet " soit sous forme de " livraison " et de " facturation.

1.9 Automatisation d'un Système d'information

Avant d'aborder l'automatisation des Systèmes d'information nous expliquons l'évolution qu'a connue ces derniers qui est passée par quatre étapes dont :

- la première étape fût celle des main-frames ou les systèmes développés étaient des systèmes centralisés avec des applications indépendantes qui utilisaient des fichiers pour le stockage des informations (peu de structuration et redondances des données). Comme applications nous pouvons citer celles de la gestion de la paie, de comptabilité, de facturation,....

- la deuxième étape des systèmes décentralisés avec des applications centralisées dans des serveurs avec un accès distant. Les données sont aussi centralisées et utilisation des SGBD.

- la troisième étape des systèmes distribués trois tiers préconise les traitements des données stockés indépendamment des programmes et de leurs présentations le schéma suivant :



- la quatrième étape est celle du Web et des serveurs Web. Les SI sont beaucoup plus ouverts vers l'extérieur avec un développement plus rapide. Nous notons l'utilisation des serveurs de pages Web, des serveurs de services et des navigateurs.

Avant de développer un SI pour une organisation, nous énumérons les raisons qui peuvent pousser à le faire :

- ≡ une organisation peut ne pas avoir de système d'information informatisé,
- ≡ une organisation demande à étendre ou compléter son système d'information (SI),
- ≡ une organisation qui veut moderniser son SI avec les nouvelles technologies.

1.8 Cycle de développement d'un Système d'information automatisé

Le développement d'un SI doit impérativement passer par un cycle de développement qui se résume en plusieurs étapes que le développeur devra suivre pour l'automatisation du système d'information.

Les différentes étapes sont :

- ≡ **Analyse des Besoins** : le développeur doit prendre connaissance des besoins des utilisateurs pour cerner le domaine de ses investigations.
- ≡ **Analyse Préalable ou d'opportunité** : le développeur doit prendre connaissance des détails des structures de l'entreprise, des procédures de travail, des supports d'information (documents, registres, fichiers,...), des moyens de traitements des informations,...
 - o **Analyse de l'existant,**
 - o **Critiques de l'existant,**
 - o **Solutions apportées,**
- ≡ **Analyse détaillée ou conception** : le but de la conception informatique est de développer des applications qui sont adéquates aux besoins des utilisateurs et qui sont capables d'évoluer, qui sont sécurisées et documentées.

Une méthode de conception apporte un ensemble de directives ou suite d'activités ou enchaînement à suivre et apporte des modèles fondés sur des langages ou formalismes.

≡ **Réalisation ou Implémentation** : cette étape porte sur le codage et l'écriture des programmes,

≡ **Tests Mise en Œuvre** : une fois tous les programmes écrits, il est nécessaire de tester tous les programmes écrits et de mettre en oeuvre chez les utilisateurs le logiciel réalisé.

≡ **Maintenance**.

Pour aider les développeurs dans les étapes précédentes, il a été proposé d'utiliser la **modélisation** dans les trois premières étapes du moins.

La modélisation pourquoi ?

Ecrire des modèles aide à comprendre, évaluer, spécifier, communiquer, promouvoir et réutiliser. Les modèles aident à maîtriser la complexité et nous pouvons proposer une modélisation par niveau d'abstraction qui consiste à présenter un modèle pour chaque vue du système d'information.

Nous citons les méthodes de conception qui utilisent le concept de modélisation :

▫ **Merise** : Données et Traitement

▫ **OMT** : Modèle Statique/Dynamique/Fonctionnel

▫ **UML** : Langage unifié de modélisation

Le langage de modélisation apporte une syntaxe, une notation, un formalisme et des règles de construction.

Le langage de modélisation apporte cohérence et interopérabilité (cad partager une même notation, une même interprétation et un même formalisme).

Que doit on modéliser ?

L'organisation, le domaine, le système d'information, les procédures de travail doivent être modélisés.

1.9 Analyse Préalable d'un SI : Modèles utilisés (diagrammes UML de cas d'utilisation, activités et séquences)

L'analyse préalable d'un système d'information consiste à étudier et analyser les **postes de travail concernés par l'application** ainsi que les **procédures de travail et les documents utilisés dans l'organisation**.

Plusieurs méthodes ont été utilisées ces dernières années et le langage UML est utilisé ces dernières années comme un standard. Nous étudierons les diagrammes de cas d'utilisation et les diagramme d'activité dans ce cours.

- Le langage UML

UML (Unified Modeling Language), traduit par **langage de modélisation unifié** est une notation standard permettant la modélisation d'un problème. Plusieurs diagrammes sont proposés dont chacun est dédié à la modélisation d'un aspect du problème.

Un diagramme est une représentation graphique unifiée et dans ce cours nous allons étudier deux diagrammes utiles pour l'étape d'analyse : Diagramme de cas d'utilisation et d'activité.

Les vues statiques représentées par les diagrammes suivants :

≡ Diagramme de classe

≡ Diagramme d'objet

≡ Diagramme des cas d'utilisation

≡ Diagramme de composant

≡ Diagramme de déploiement

Les vues dynamique est la notion qui permet de changer les types des objets de façon dynamique.

≡ Diagramme d'activité

≡ Diagramme de séquences

≡ Diagramme d'état de transition

≡ Diagrammes de collaboration

- Diagramme de cas d'utilisation pour l'étape d'analyse des besoins

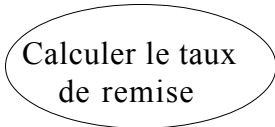
Le diagramme de cas d'utilisation représente les **cas d'utilisation** et les **acteurs** qui interagissent avec ces cas.

Le **cas d'utilisation** dénote et explicite les fonctions des applications des systèmes d'information.

Vous devez identifier les cas d'utilisation et pour chacun vous devez donner un titre ou intitulé sous forme de verbe et complément.

Par exemple 'Payer le fournisseur', 'Calculer le taux de remise'.

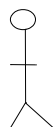
Le formalisme Graphique :



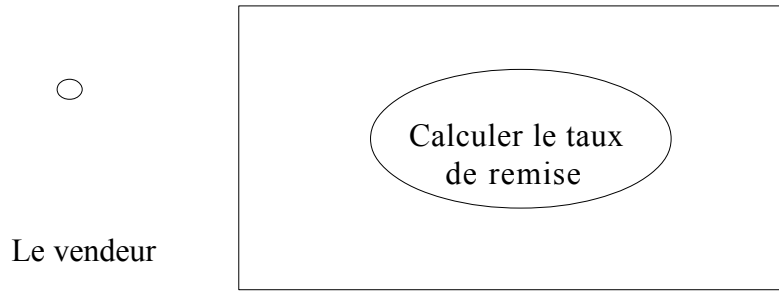
Calculer le taux
de remise

Les **acteurs** appartiennent au système d'information et représente l'entité matérielle ou humaine qui interagit avec le cas d'utilisation.

Le formalisme Graphique :



Nom- Acteur



Le diagramme de cas d'utilisation : Vue Fonctionnelle

Les relations entre les cas d'utilisation peuvent exister : **Extend et Include**

Extend : pour spécifier qu'un cas peut étendre un autre cas.

Include : pour spécifier qu'un cas inclut un autre cas.

Héritage : pour spécifier les relations d'héritage entre les cas d'utilisation.

Un Chapitre sur le langage UML sera présenté dans un chapitre à part.

Chapitre 2 : Les outils d'analyse

2.1 Définir l'information, sa classe et sa représentation

Nous introduirons dans ce chapitre les éléments de base pour mener à bien une analyse d'un système d'information informatisé.

2.1.1 Notion d'information : Définition

Une information est une donnée isolée ou un ensemble de données isolées ou un ensemble de données relatives à un fait, à un objet ou à un être vivant.

Exemple : Aujourd'hui, nous sommes le xx/xx/2007, la température est de 18°C et nous faisons le cours de SI.

2.1.2 Rôle de l'information

L'information joue un rôle fondamental dans les relations entre les éléments du système. L'activité de service est guidée par les ordres de la direction.

2.1.3 Classe de l'information

Il y'a plusieurs classes d'informations :

- ⊗ Information de nature hiérarchique : représente les ordres ou les instructions qui sont destinées vers les subordonnées de la part d'un supérieur (Inventaire).
- ⊗ Informations de conseils : certains postes de travail ne reçoivent pas uniquement des ordres mais aussi des informations destinées à l'idée d'accomplissement des tâches.
- ⊗ Informations d'interactions : c'est lorsqu'il y'a échange d'informations entre les services.
- ⊗ Informations de représentation : sont généralement des résultats ou réponses aux informations de décision. Exemple : L'état d'inventaire.

2.1.4 Représentation de l'information : Notion d'entité

Définition

Une entité est une chose concrète ou abstraite appartenant au réel perçu à propos duquel nous voulons enregistrer les informations. Exemple : Etudiant (concret) ; Module (Abstrait).

Les caractéristiques d'une entité sont :

- ⊗ Elle possède une existence autonome.
- ⊗ Elle est distinguée des autres entités.
- ⊗ Elle peut posséder des propriétés.
- ⊗ Elle possède une durée de vie.

Définition de l'association

Une association est définie par une correspondance (une relation) entre deux ou plusieurs entités. Une association peut posséder une ou plusieurs propriétés.

CLIENT (Entité 1) passer (Association) COMMANDE (Entité2)

2.1.5 Notion de propriété

Une propriété est l'élément d'information de base, le plus fin. C'est une donnée élémentaire qui est utilisée pour décrire les entités et les associations.

2.1.6 Notion d'occurrence

Représente la valeur à attribuer à une propriété. Exemple : 12/12/1981 [≠] valeur d'une date de naissance.

Etudiant (Mat, nom, prénom, date-naissance)

[≠] **Occurrence de l'Entité. (81/200/01, Benali, sami, 22/10/1989).**

2.1.7 Notion de fichier

Un fichier est un ensemble d'informations se rapportant à un même sujet.

Un fichier est un ensemble structuré d'informations décrivant des entités de même nature à l'aide de renseignements similaires. Chaque entité est décrite par un article du fichier regroupant les informations indivisibles qui s'y rapportent, et que l'on appelle rubriques de l'article. Une ou plusieurs rubriques, permettant de différencier les articles, constituant l'indicatif du fichier.

Exemple

Un fichier « personnel » rassemble autant d'articles que d'employés, chaque article décrit une personne à l'aide des données utiles : nom, prénom, adresse, date de naissance, matricule, emploi, ... : le matricule peut constituer l'indicatif de ce fichier.

- ⊗ Structure d'un fichier : c'est un ensemble d'informations, il est à vrai dire un ensemble d'ensembles d'informations. Le deuxième ensemble est un groupe de propriétés relatives à une même entité. Ce groupe de propriétés constitue ce que l'on appelle **un enregistrement**. Nous pouvons dire qu'un fichier est un ensemble d'enregistrements. Exemple : Soit le fichier étudiant constitué d'un ensemble d'informations relatives aux étudiants. « Matricule, nom, prénom, date-naissance, adresse, année-cursus, année-Bac ». Cette suite d'informations constitue la structure de l'enregistrement du fichier étudiant. La structure d'un fichier est la structure de l'enregistrement qui le constitue.
- ⊗ Notion de champ : chaque information constitue l'enregistrement du fichier qui s'appelle champ (zone, rubrique, propriété). Exemple : Mat est un champ, nom est un champ.

Remarque : chaque champ doit avoir un type (le type de l'information qu'il contient : numérique, caractère, date) et une taille (c'est le nombre de caractères qui composent le champ).

Exemple :

Mat	Type	Taille
99/0145	Caractère	07.

Chaque fichier est identifié par son nom et chaque élément de fichier 'enregistrement' est identifié grâce à une zone particulière appelée **Clé d'accès ou Index ou Indicatif**.

Cette clé d'accès doit avoir une valeur unique (qui ne doit pas être partagée par les autres enregistrements). Le champ matricule du fichier étudiant est sa clé d'accès.

2.2 La codification

Les informations saisies par les organisations se présentent parfois sous des formes difficiles à exploiter : elles sont volumineuses, verbeuses, la même information peut apparaître sous des présentations différentes. La codification permet dans de nombreux cas, en réduisant et en normalisant la forme de ces données, d'en faciliter l'utilisation, d'éviter un coût de traitement prohibitif et de limiter les erreurs. Elle consiste à associer un ensemble de symboles aux informations.

Exemple : Dans une entreprise, les clients réguliers peuvent se voir attribuer un numéro de code. Un code est représenté par une ou plusieurs **tables de codage**

Code	Client
0001	Dupont Jean
0002	Dupont George
0003	Durand

2.2.1 Les qualités possibles d'un code

Les qualités d'un code doivent s'apprécier selon les circonstances de création du code. Nous citerons les qualités suivantes :

- ⊘ **La non-ambiguïté** : à un code correspond alors un renseignement et un seul ;
- ⊘ **La concision** : un code de dimension réduite permet un gain de place sur les supports d'enregistrement.
- ⊘ **La durabilité** : il devra répondre au besoin sur une durée suffisante pour éviter la modification des supports où il apparaît déjà. A cet effet, il permettra, sans modification de la codification existante, la représentation de nouvelles entités. Ceci peut se traduire par l'insertion de symboles entre ceux qui existent déjà, ou par l'extension de la liste des codes après son dernier élément.

- ⊘ **L'adaptation au problème à traiter** : notamment aux moyens utilisés pour le codage et le décodage de l'information.
- ⊘ **La facilité de contrôle qu'il présente** : la validité d'un code contrôlable peut être en partie vérifiée grâce à son propre contenu afin de limiter les erreurs de codage, de retranscription. Exemple : un numéro minéralogique tel que 203RZ802 sera immédiatement réputé faux dans notre pays.
- ⊘ **La logique** : la correspondance établie entre symboles et entités représentées doit être naturelle. Exemple : pour représenter des masses de 10, 20, 30 kg, mieux vaut choisir des codes 1, 2,3 que 4, 6,1.
- ⊘ **La transparence** : dans le cas où le code peut être décodé sans recours à une table de codage.

2.2.2 Les différents modes de codification

- ⊘ **Code séquentiel compact**
- ⊘ **Code séquentiel à trous**
- ⊘ **Code séquentiel par tranches**
- ⊘ **Code significatif : descriptif articulé et mnémorique**
- ⊘ **Code à niveaux**
- ⊘ **Code combiné**
- ⊘ **Code contrôlable par clé**

2.3 Les contrôles

Les contrôles détectent les informations erronées qui risquent de devenir sources de problèmes pour le système. Les lieux de contrôles sont les postes de saisies, les centres de calcul informatique et les postes de diffusion des résultats.

Comment contrôler les informations ?

Le contrôle comporte la détection des erreurs et la correction des erreurs détectées. Il existe deux types de contrôles : directe et indirecte.

- ⊘ Les contrôles directes : mettre en jeu uniquement l'information à contrôler. Les principaux types de contrôles directs : **contrôle de présence, contrôle de cadrage de l'information, contrôle de type, contrôle sur la plage de valeur.**
- ⊘ Les contrôles indirectes : ils sont réalisés en comparant ou en rapprochant les valeurs de plusieurs informations. Les opérateurs de comparaison (=, >, <) peuvent être utilisés pour les contrôles indirectes pour détecter les erreurs. Les opérations de calcul peuvent être utilisées aussi.

2.4 Stockage de l'information (Buffer, Block, RAID technologie)

Le stockage central ou primaire considère les medias de stockage directe dans le CPU comme la mémoire centrale et la mémoire cache qui sont des mémoires à petite capacité mais très rapide. Les mémoires de stockage primaires fournissent un accès rapide des données mais avec taille limité.

Les mémoires secondaires ou ternaies qui sont utilisées pour le stockage permanent des données qui ne peuvent pas être traités par le CPU mais copiées avant traitement sur un support de stockage primaire.

Les supports de fichiers ont évolué durant ces dernières années de la bande magnétique aux mémoires adressables : disquettes, disque dur, CD ROM, Clé USB, Disque externe,....

Dans les supports de type disque, les informations sont enregistrés sur la surface du disque le long des pistes concentriques. En plus des données la piste doit contenir un certain nombre d'informations supplémentaires nécessaires au système d'exploitation, tel que l'adresse de la piste, l'adresse et la position de chaque enregistrement, la marque de secteur et l'espace entre enregistrements.

La capacité de stockage d'une mémoire centrale ou primaire (GO) est nettement inférieure à celle des mémoires secondaires et ternaies (TO et GO).

Le temps d'accès de la mémoire primaire est de l'ordre de nanosecondes et celui des mémoires secondaires est millisecondes.

Le disque peut être à une seule face ou à double faces. La **piste** est partagé en **secteurs** alors que la division de la piste en blocks disques égaux en taille ou **pages** est faite par le système d'exploitation lors du formatage.

Pour repérer une donnée sur disque il faut introduire le **numéro du cylindre, numéro de la piste et numéro de block ou page ou bien le numéro du premier secteur.**

L'adresse de la piste comporte les informations suivantes :

Co : caractère de contrôle ; Cy : cylindre ; Nt : Numéro de la tête de lecture et d'écriture ; De : numéro de caractère des données.

Les caractéristiques des unités de disque sont : la densité d'enregistrement, la vitesse de rotation, le temps de positionnement d'un bras, la vitesse de transfert, le temps de lecture et d'écriture.

Le facteur de blocage représente le nombre d'enregistrements bloqués dans un enregistrement physique (voir enregistrement logique).

A chaque invocation d'une donnée localisée sur disque, elle doit être repérée avec un minimum d'accès à la mémoire secondaire sur le disque puis copiée sur la mémoire primaire.

L'adresse de la piste comporte les informations suivantes :

Co : caractère de contrôle ; Cy : cylindre ; Nt : Numéro de la tête de lecture et d'écriture ; De : numéro de caractère des données

Les caractéristiques des unités de disque sont : la densité d'enregistrement, la vitesse de rotation, le temps de positionnement d'un bras, la vitesse de transfert, le temps de lecture et d'écriture.

Le facteur de blocage représente le nombre d'enregistrements bloqués dans un enregistrement physique (voir enregistrement logique).

Un **système RAID** (Redundant Array of Independent Disks) peut être défini comme un ensemble de disques configurés pour être utilisés et interrogés rapidement comme un seul disque. Cette technique se base sur le fait que des données sont disparates sur plusieurs disques et au moment de l'interrogation les requêtes sont prise en compte en parallèle ce qui confère une grande vitesse de transfert. Les données peuvent être redondantes et ceci confère une fiabilité si un disque tombe en panne.

2.5 Du fichier traditionnel au fichier numérique

Les données sur le disque sont organisées en fichiers qui sont structurés en enregistrements. Le système de gestion de fichiers du système d'exploitation est chargé du suivi des fichiers.

Un fichier a **un nom**, a un **ensemble d'enregistrements**. Un enregistrement est composé de **champs (attributs)**.

Deux types d'enregistrements (taille fixe où la taille des champs est fixe où taille variable ou la taille des champs est variable).

Les enregistrements sont stockés sur disque dans les pages.

2.6 L'approche Base de Données

Une base de donnée est une collection de fichiers reliés entre eux. Un système de gestion de base de données est utile pour gérer les données dans la base de données.

