

Chapitre 3

Les outils d'Analyse

Plan du chapitre

- **Généralités**
- **La codification des informations**
- **Les contrôles**
- **La compression des données**
- **Le Cryptage de l'information**

Généralités



Information?

- Moyen de communication.
- Tout ce qui peut se présenter, s'écrire ou se dire
- correspond à deux notions distinctes :
 - c'est un renseignement qui améliore notre connaissance sur un sujet quelconque.
 - C'est une représentation du renseignement à l'aide de symboles.

Rôle de l'information?

a/ Information base de décision :

- Exemple :
- La décision de réapprovisionnement s'appuie sur quelles informations ?
- Réponse :
 1. Etat du stock
 2. Délai de livraison du fournisseur
- Remarque : l'efficacité de la décision dépend de la qualité de l'information disponible.

Rôle de l'information?

b/ information moyen de coordination :

- la circulation de l'information dans l'entreprise est un moyen de coordination.

Rôle de l'information?

C/information obligatoire à l'égard de l'entreprise

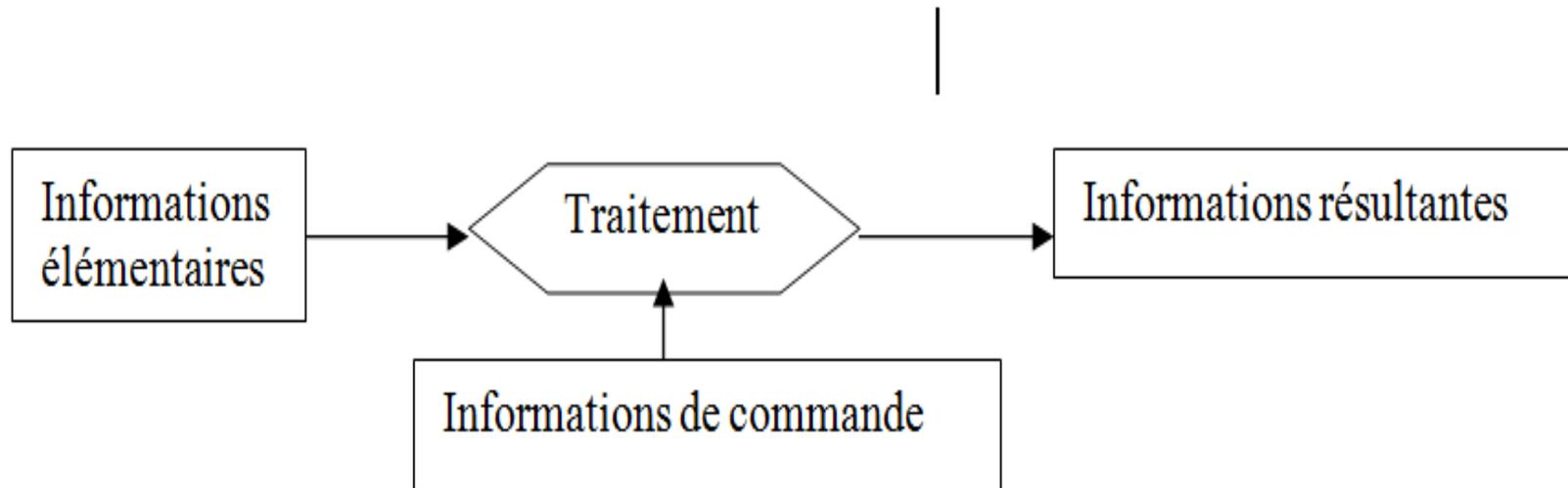
Exemple : obligation d'établir une facture pour toute vente

Conclusion :

- L'entreprise doit disposer d'information précise contrôlée et présentée sous une forme imposée.

Classes d'information?

- **informations élémentaires** (entrées) :
- **informations résultantes** (sorties) :
- **informations de commande** (instructions) :



Représentation de l'information :

Concepts : information comme renseignement

- **a/ L'entité(individu, objet) : c'est l'objet, le concept, ou l'événement qu'on veut décrire.**
- Exemples:
- Entité Etudiant
- Entité Client
- Entité Commande

Représentation de l'information :

Concepts : information comme renseignement

- **b/l'attribut** : (la propriété ou la rubrique liée à l'individu) c'est la caractéristique que l'on retient pour décrire l'entité ou l'individu.

Entité	Attributs
Etudiant	N étudiant, nom, prénom, date -naissance,....
Client	nclient, nomcl, adrcl , nfiscal,
commande	Qté_commandée, prix unitaire, N° produit

Entité Etudiants

Nom	Prénom	Adresse
Ali	Ben Ali	C.Laid
Ammar	Ben Ammar	Grarem
Mohamed	Ben Mohamed	Mila
Meriem	Ben Ahmed	Mila
Iness	Ben Moubarek	Ferdjioua

Entité Véhicule

Matricule	Année circulation	Marque	Type
00120.102.43	2002	KIA	Picanto
00025.112.43	2012	VW	Passate
12300.111.16	2011	Citroen	Ibiza

Entités

Attributs

**Les valeurs
Occurrences**

Nom de l'entité

Entité Commande

3 champs= attributs= rubrique

<u>N°Commande</u>	<u>DateCommande</u>	Montant
27	13/2/2007	120
65	12/1/2008	34
2	10/06/2006	27
34	14/12/2007	1500

4 occurrences= valeurs= enregistrements

Représentation de l'information :
Concepts : information comme renseignement

- **c/ Valeur(occurrence)** : c'est la réalisation particulière concrète de chaque attribut ou propriété.
- **Exemple :**

<u>Attribut</u>	<u>Occurrence</u>
Solde d'un client	500.000DA
Nom client	Mohammed

Information= entité +attributs + Valeurs.

Représentations de l'information

- ensemble conventionnel de symboles (signifiants).
- Une correspondance entre l'ensemble des renseignements (signifiés) et l'ensemble de ses symboles (signifiants).
- Un support capable de rendre ses symboles perceptibles en pratique,
- On distingue deux types de représentation :
 - représentation Analytique.
 - représentation par langage

Codification des informations



Codification des informations (Motivation)

Pourquoi est-il nécessaire de codifier les informations

- Besoin de structurer l'information.
- Cette structuration passe par l'association de codes aux différentes informations manipulées.
- Ces codes vont permettre de désigner chaque information de manière claire et unique.

Codification des infos (Exemple)

exemple : soit le document suivant :

BON DE COMMANDE

Numéro de commande :.....

Date commande :.....

Numéro client :.....

Nom client :.....

Adresse client :.....

Référence produit	Désignation produit	Quantité commandée
-----	-----	-----
-----	-----	-----
-----	-----	-----

Codification des infos (Exemple)

Les données à extraire ?

- Numéro commande
- Date commande
- Numéro client
- Nom client
- Adresse client
- Référence produit
- Désignation produit
- Quantité commande

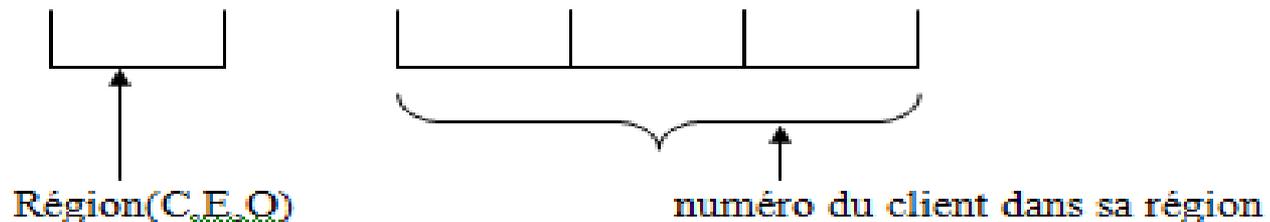
Exemple :

- | | | |
|-------------------|--------|--------|
| ▫ Numéro commande | —————> | Num_C |
| ▫ Date commande | —————> | Date_C |
| ▫ Numéro client | —————> | Num_CL |
| ▫ Nom client | —————> | Nom_CL |

•

Codification des infos (Exemple)

- **Codifier** : Associer un code à chaque donnée en utilisant un nom abrégé unique.
- Num-CL, selon le numéro de la commande peut prendre différentes valeurs par exemple :
- La 1^{ère} commande est effectuée par le 3^{ème} client se trouvant dans la région centre.
- La seconde commande est passée par le 15^{ème} client se trouvant dans la région Est.
- Pour différencier ces deux (02) clients, ont doit affecter à leur codes Num_CL deux valeurs différentes :
 - valeur1 : 3^{ème} client dans la région centre
 - valeur2 : 15^{ème} client dans la région Est.
- données longues et lourdes à manipuler → les abréger.
- Pour cela, on va représenter ces données comme suit :

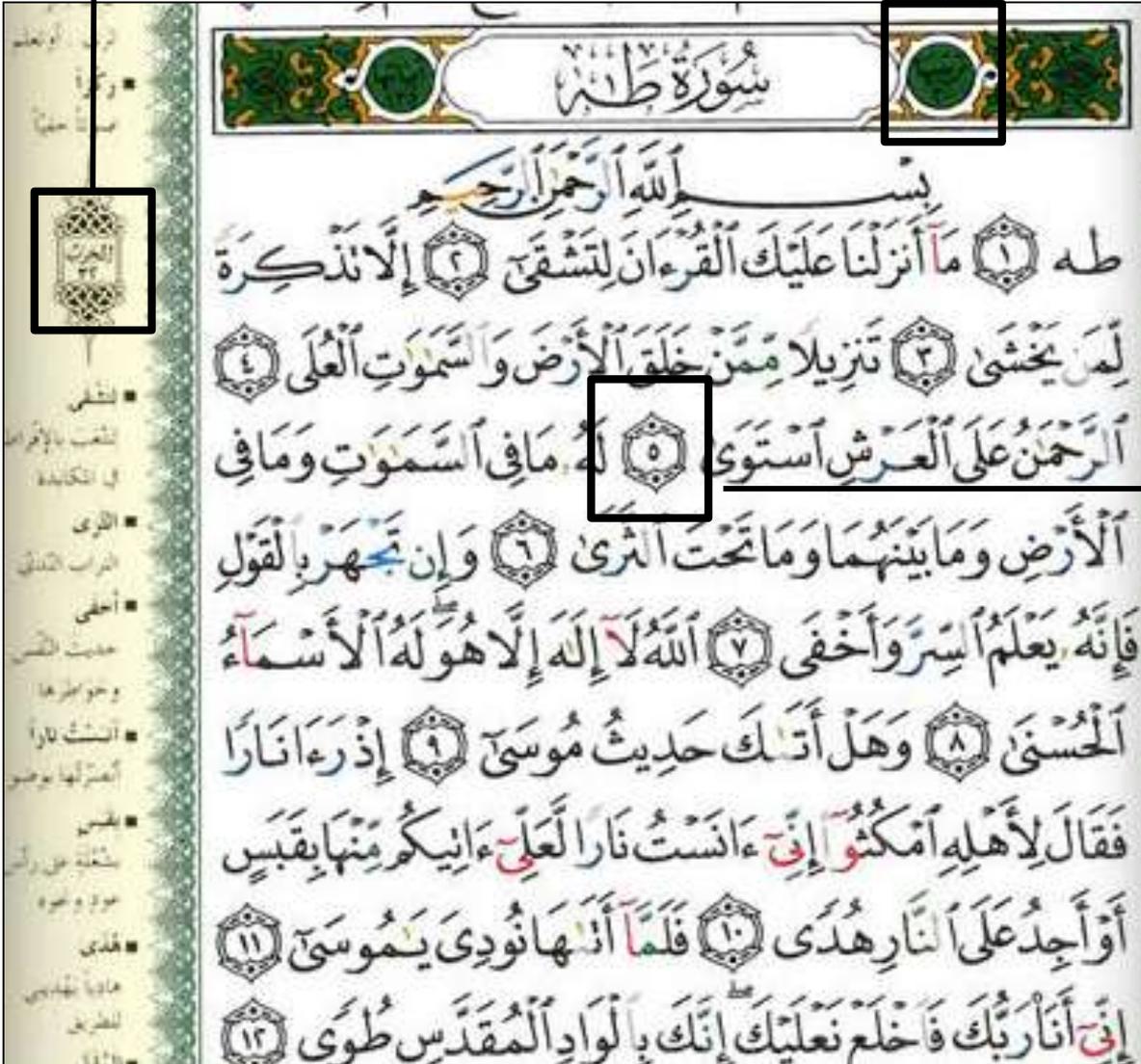


- Pour le premier client : Num_CL=C003
- Pour le second client : Num_CL =E015

رقم الجزء

رقم السورة

رقم الآية



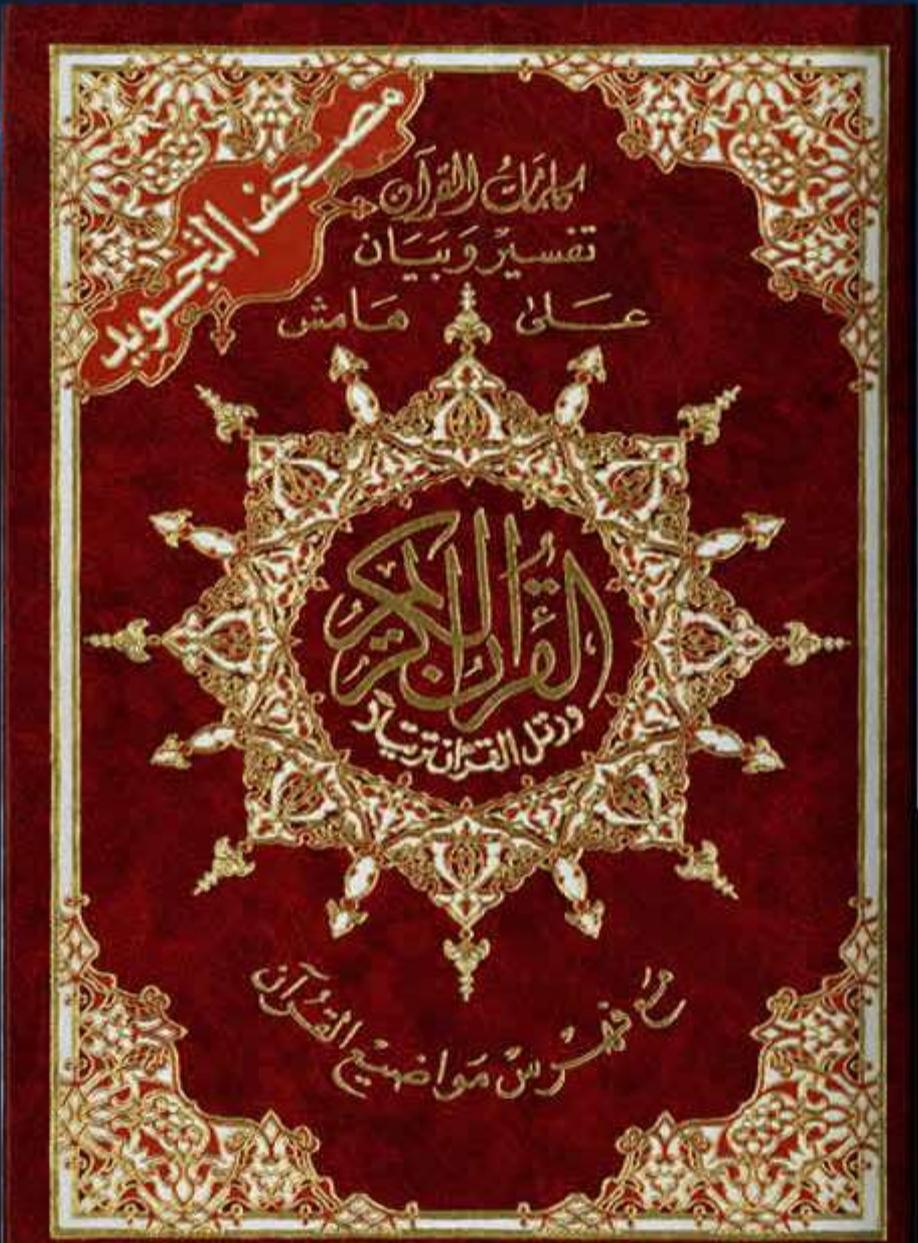
الآية 5 من سورة طه

020005

رقم الآية رقم السورة

أدوات التحكم

- عرض كل السور
- ٣٠
- ٢٩
- ٢٨
- ٢٧
- ٢٦
- ٢٥
- ٢٤
- ٢٣
- ٢٢
- ٢١
- ٢٠
- ١٩
- ١٨
- ١٧
- ١٦
- ١٥
- ١٤
- ١٣
- ١٢
- ١١
- ١٠
- ٩
- ٨
- النساء
- المائدة
- الأنعام
- الأعراف
- الأنفال
- التوبة
- يونس
- هود
- يوسف
- الرعد
- إبراهيم
- الحجر
- النحل
- الإسراء



quranflash.com

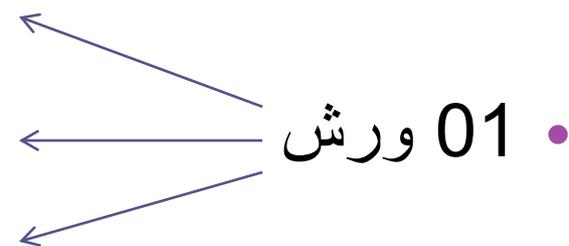
Question de réflexion

القراءات

ياسين الجزائري

مشاري العفاسي

لعيون الكوشي



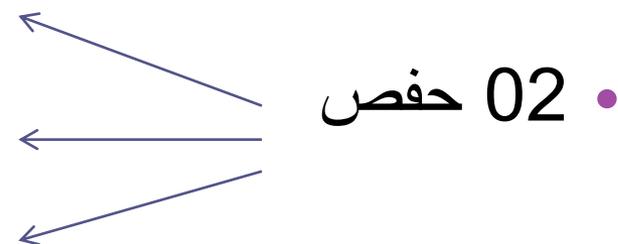
الآية 5 من سورة طه

020005

سعد الغامدي

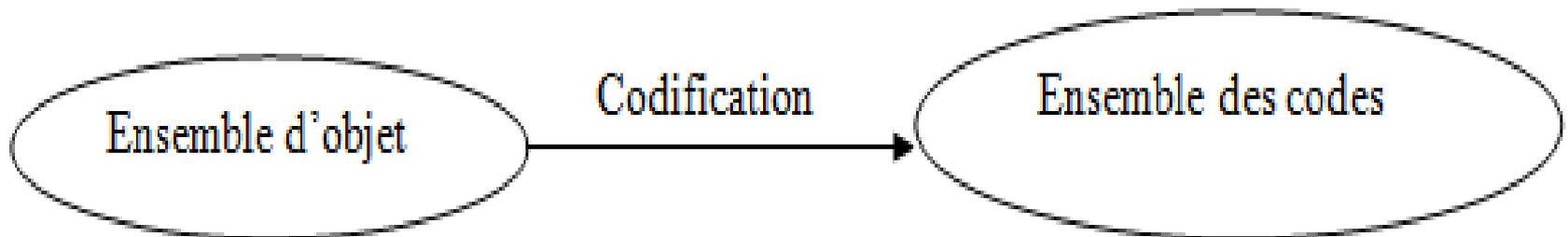
فارس عباد

مشاري العفاسي



Définitions

- **Un code** : est un nom abrégé ou une représentation de l'information permettant de désigner un objet ou un concept de manière claire et unique.
- **La codification** : est l'opération qui consiste à remplacer une information sous sa forme naturelle par un code clair qui serait mieux adapté aux besoins de l'utilisateur de l'information.
- La codification porte sur le nom de l'information (ou concept) à codifier, mais aussi sur sa valeur.



Principales caractéristiques d'une codification

Une codification remplace une information par un code unique et clair.

Le code doit répondre aux besoins de l'utilisateur et lui faciliter les tâches de codification et d'interprétation des codes établis.

La codification mise en place doit être aussi stable que possible,

Caractéristiques essentielles pour la codification :

Elle ne doit pas être ambiguë

Elle doit s'adapter aux besoins des utilisateurs.

Elle doit permettre l'insertion de nouvelles infos.

Elle doit être concise.

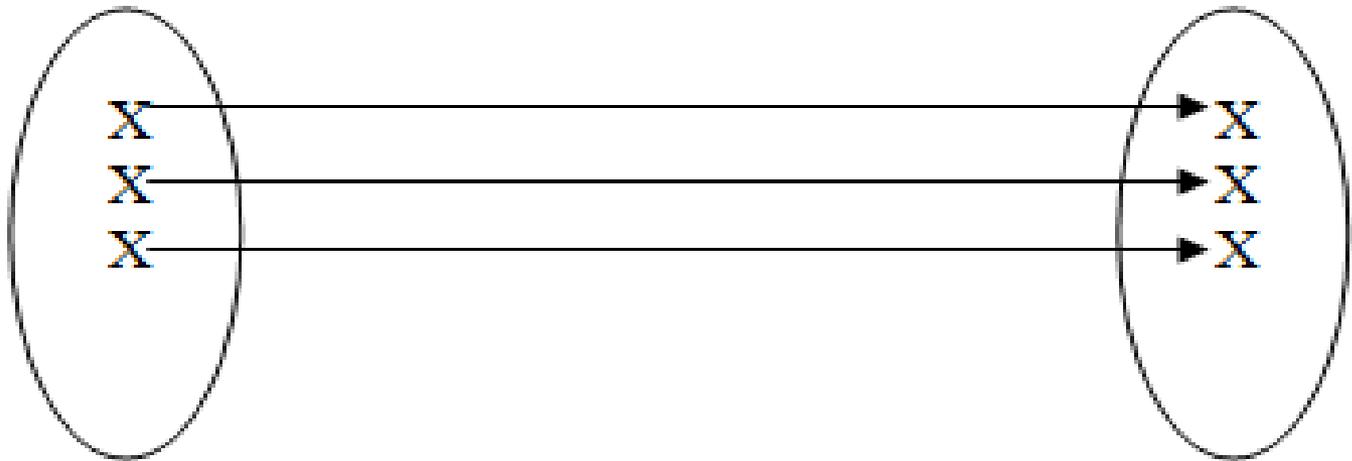
Elle doit être aussi significative que nécessaire

Non-Ambiguïté:

Une codification doit associer un code et un seul à chaque information.

Chaque code doit désigner et un seul objet.

Et chaque information doit être associer à un seul code.



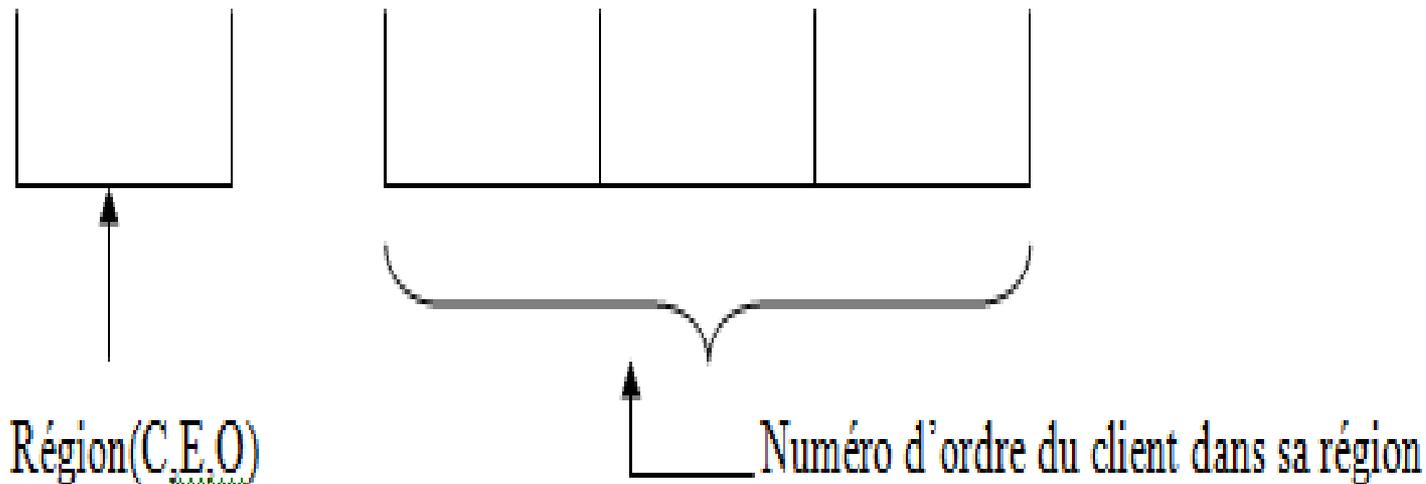
Exemple: Bon de commande, on peut choisir code

Num - C → numéro commande.

Num - C → numéro client

Facilité d'utilisation

On considère le code associé au numéro client.



Si l'entreprise vient à avoir un nouveau client dans la région Ouest, sa codification est très facile :

Numéro du dernier client dans la région Ouest : 301, le nouveau client sera le 302^{ème} dans la région Ouest.

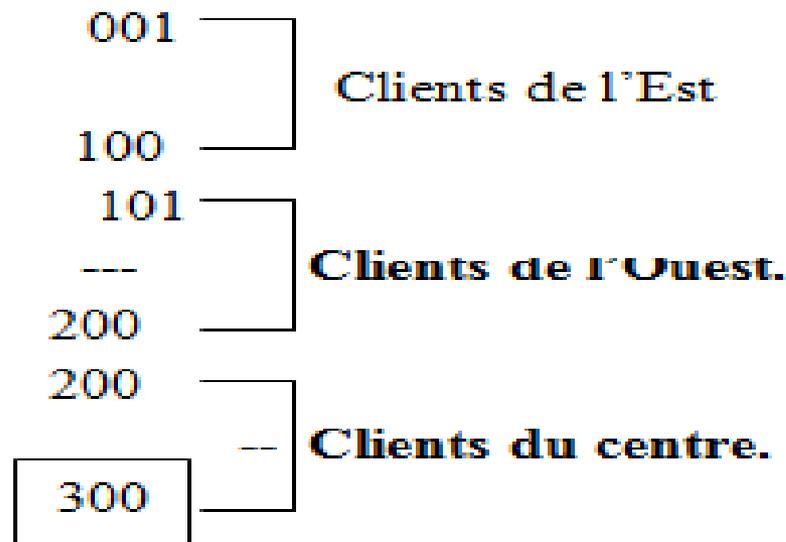
Alors le numéro du nouveau client sera : 0302. De même, l'interprétation de ce code est chose aisée.

La possibilité d'extension et d'insertion

Exemple :

La codification des clients d'une entreprise se fait comme suit :

- les numéros de 1 à 100 clients originaires de l'Est.
- les numéros de 101 à 200 clients originaires de l'Ouest.
- les numéros de 201 à 300 clients originaires du Centre.



Ce code ne permet pas d'avoir plus de 100 clients par région
Ce code ne permet pas l'insertion.

La possibilité d'extension et d'insertion

C.001 Dans cet exemple, le code proposé permet de faire l'insertion.

C.002

.....

C.085

E.001

E.002

.....

E.060

O.001

O.002

O.101

.....

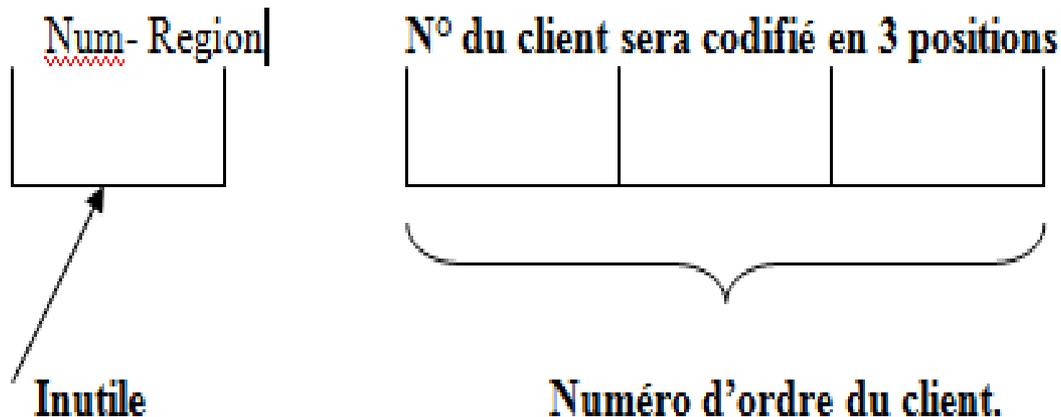
Insertion du nouveau
code E.061 entre les
deux codes E.060 et
O.001

La concision :

La concision traduit le fait qu'un code doit être clair et court, sans pour autant négliger la possibilité de l'évolution de l'ensemble des informations à codifier.

- **Exemple** : Si le fait d'indiquer la région du client dans le code ne sert à rien pour les utilisateurs du code, il serait inutile de la préciser

:



Si on est sûr que le nombre de clients ne dépasse pas 99, deux positions suffisent pour le numéro du client.

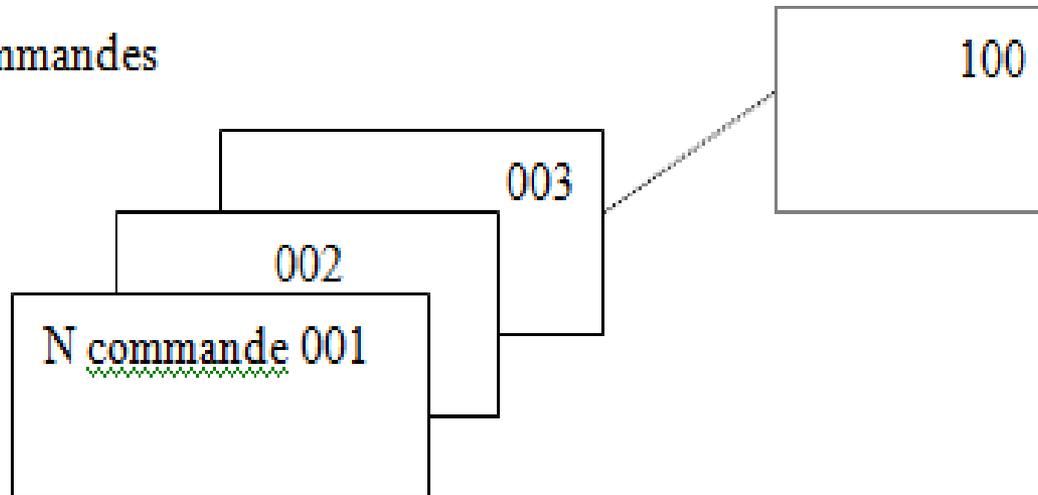
Les différents types de codification



1- La codification séquentielle:

- Attribuer à l'information un numéro de sorte que les numéros associés soient successifs (1,2,3,.....).

Exemple : les commandes



Avantages

- Non ambiguë
- Simple(dernier N+1)
- Extension possible

Inconvénients

- Insertion impossible
- Non significative.

La codification Par tranche

- Attribuer une tranche de codes à chaque catégorie d'objets à codifier (les codes sont séquentiels dans une tranche).
- **Exemple** : dans une bibliothèque, les ouvrages sont classés par catégories
 - Technologie 001 à 100
 - Littérature 101 à 200
 - Sociologie 201 à 300
 - Médecine 301 à 400
 - Culture générale 401 à 500

La codification Par tranche

Avantages

- Non ambiguë
- Simple : il suffit de connaître la catégorie de l'objet et lui affecter un numéro incrémenté de 1.
- Insertion possible si le nombre d'infos à codifier ne dépasse pas l'intervalle
- Extension possible par ajout d'une nouvelle tranche de codes si une nouvelle catégorie d'objets entre en jeu

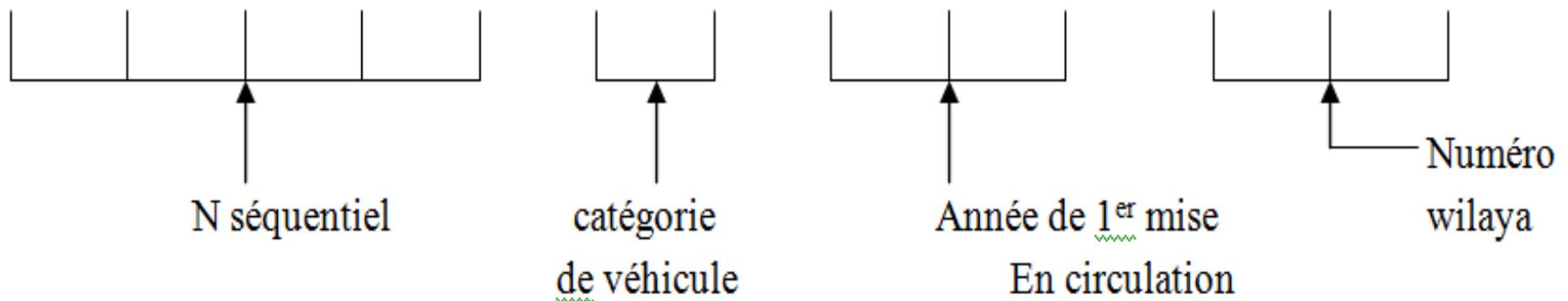
Inconvénients

- Non significative.
- Nombre de code dans une tranche difficile à fixer
- la répartition des objets en catégories n'est pas toujours évidente.
- Insertion impossible si le nombre d'infos à codifier dépasse l'intervalle de la tranche

3. La codification articulée :

- Attribuer des codes découpés en zones (zone ou descripteur)

Exemple1 : Immatriculation d'un véhicule



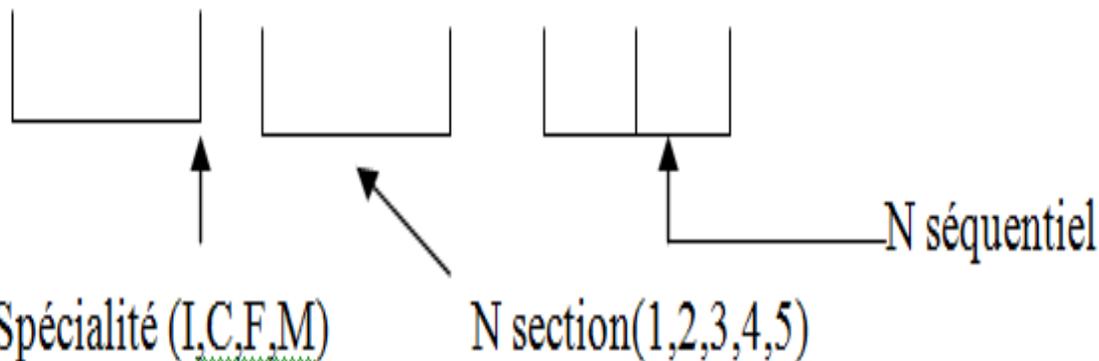
3. La codification articulée :

Exemple 2 : codifier les numéros des étudiants d'une école,

Spécialités enseignées : Informatique (I), Comptabilité (C), Fiscalité (F), Marketing (M)

Dans une spécialité, on peut avoir plusieurs sections (<5)

Une section contient moins de 100 étudiants



Le code C315 : étudiant N15 de la 3^{eme} section de la spécialité comptabilité.

Le code F210 : étudiant N10 de la 2^{eme} section fiscalité.

3. La codification articulée :

- **Avantages**

- Non ambiguë
- Insertion et extension possible
- très utilisée
- Possibilité de regrouper les objets selon un critère donné

- **Inconvénients**

- code trop long donc lourd à manipuler
- Possibilité de saturation d'une zone
- Instabilité : si une zone change tout le code change

d) Codification par niveau :

Cas particulier de la codification articulée :

Les zones(descripteurs) sont des niveaux :

Exemple 1 : Code postal 16010

Wilaya

Daïra

Commune

5- Codification mnémonique

5-1 Codification mnémonique

Associer au nom de l'objet, un nom abrégé qui rappelle l'objet codifié.

Numéro commande —————> Num C

Code postal —————> CP

Compte courant postal —————> CCP

Avantages

- Facile à mettre en œuvre
- Significative

Inconvénients

- Porte sur le nom et non sur la valeur
- Utiliser uniquement pour coder les variables dans un Programme informatique.

Le contrôle de données



Contrôle de données: Définition

■ Contrôler une information, c'est vérifier sa justesse et sa conformité à la réalité de l'organisation.

Exemple : numéro d'employé : F123 : Erroné car les codes-services existants dans l'organisation sont : P,D,S,C

Contrôle de données Par ordinateur?

- Le contrôle peut être effectué par l'homme (observation, calcul) comme il peut être réalisé par la machine(l'ordinateur).
- C'est à l'utilisateur de spécifier à la machine les caractéristiques de l'information qui doivent être contrôlées.
- **Est-ce que l'ordinateur peut effectuer ces contrôles?.**
- L'ordinateur n'exécute que les commandes qui lui sont introduites par l'homme.
- De ce fait, les contrôles doivent être prévus et exigés par l'utilisateur, pour que la machine puisse les reproduire.

Principaux types de contrôle :

I) Les contrôles directs :

- Contrôle direct: les contrôles qui s'effectuent sur l'information elle-même, sans tenir compte des autres informations existant dans le système.
- Exemple : le prix unitaire d'un produit doit être supérieur à 0,
- Deux types de contrôles directs :
 - ❖ Les contrôles de présence et de non présence.
 - ❖ Le contrôle de type

Contrôle direct:

a) Contrôle de présence ou de non présence :

- Vérifier l'existence ou non d'une information sur le support où elle devrait se trouver. Le support peut être un document ou un fichier.
- **Exemple1** :(présence) lors de prêt d'un ouvrage, l'abonné doit exister.

Ferhat BENFERHANE?????

- **Exemple2** :(non présence) lors de l'inscription d'un nouvel abonné celui-ci ne doit pas exister déjà dans le fichier des abonnés.

Contrôle direct:

B) Contrôle de type :

- Vérifier que le type d'une information correspond à ce qu'il doit être.
- **Exemple** : Code postal = 16I90
- Code erroné car le caractère I n'est pas un chiffre.(Le code postal est numérique.)

Principaux types de contrôle :

II) Les contrôles indirects :

- permettent de vérifier la conformité d'une information par rapport à l'ensemble des informations se trouvant dans le système.
- Exemple : la date de recrutement d'un employé doit être supérieure à la date de création de l'entreprise.
- Consiste à vérifier la conformité d'une information par rapport à d'autres informations (il y a comparaison entre les informations).

Types de contrôles indirects:

le contrôle de cohérence interne.

le contrôle de cohérence externe.

le contrôle de vraisemblance.

Contrôles indirects:

a) Contrôle de cohérence interne :

- S'applique généralement à la codification articulée. Il s'agit de vérifier l'exactitude d'une partie de l'information par rapport à d'autres parties de la même information.
- **Exemple1** : soit la date : JJ/MM/AAAA = 30/02/1997.
- le contrôle de cohérence interne détecte une erreur : MM= 02 donc $JJ \leq 29$
- **Exemple2** : Num_Etud= LA/5/035 (filiera, année d'étude, nseq)
- La filiera LA(lettres Arabe) est enseignée sur 3 ans et non sur 5 ans.

Contrôles indirects:

a) Contrôle de cohérence externe :

- Consiste à vérifier la conformité d'une information par rapport à d'autres informations.
- **Exemple** : soient les informations d'un employé. Année de recrutement : 1995.
Ancienneté : 60 ans
- Le contrôle de cohérence externe détecte une erreur (ancienneté = année en cours – 1995)

Contrôles indirects

c) Contrôle de vraisemblance :

- Il s'agit de s'assurer que l'information est vraisemblable, c'est-à-dire possible et concevable en fonction de son sens.
- **Exemples :**
- Date de naissance = 5/12/2215 (non vraisemblable car date future)
- Code postal 49010 non vraisemblable car $1 \leq \text{wilaya} \leq 48$

Compression des données



Compression des données

- **But:** réduire la taille des fichiers pour optimiser les temps de recherche et réduire les coûts de stockage et de transport.
- **Définition :** La compression des données est une technique qui consiste à réduire la taille de l'information à stocker ou à transmettre, sans qu'elle ne soit changée ou perdue.

Les différentes techniques de compression

a) Suppression des blancs :

- Consiste à balayer le flot initial des données, en remplaçant les blancs consécutifs par un caractère particulier :

Caractère spécial

Compteur des blancs

Exemple :

- Flot initial : ABCD Ъ Ъ Ъ Ъ Ъ x z
- fichier comprimé : ABCD λ 5 x y

Les différentes techniques de compression

b) Suppression des répétitions de caractères :

- Elle consiste à supprimer la répétition de n'importe quel caractère (seuil pour répétition >3) selon le format suivant :
- on remplace la répétition de caractères par le triplet :

Caractère spécial	Compteur de répétition	Caractère répété
-------------------	------------------------	------------------

Exemple : A x x x x x x x B Y Y Y Y C

- Texte comprimé A#7xB #4YC

Les différentes techniques de compression

c) Codage de Huffman :

- Méthode mise en œuvre en 1952 mais elle est mise au goût du jour.
- C'est la 1^{ère} technique à ne pas coder les caractères sur le même nombre de bits.
- Selon la fréquence d'apparition du caractère dans le texte, elle offre au caractère le plus fréquent un nombre de bits faible et au caractère le moins fréquent, un nombre de bits élevé.

Codage de Huffman

intérêt :

- Pour les données rares, nous perdons quelques bits regagnés pour les données répétitives.
- Par exemple, dans un fichier ASCII le "w" apparaissant 10 fois aura un code très long: 0101000001000. Ici la perte est de 40 bits (10 x 4 bits), car sans compression, il serait codé sur 8 bits au lieu de 12.
- Par contre, le caractère le plus fréquent comme le "e" avec 200 apparitions sera codé par 1. Le gain sera de 1400 bits (7 x 200 bits).

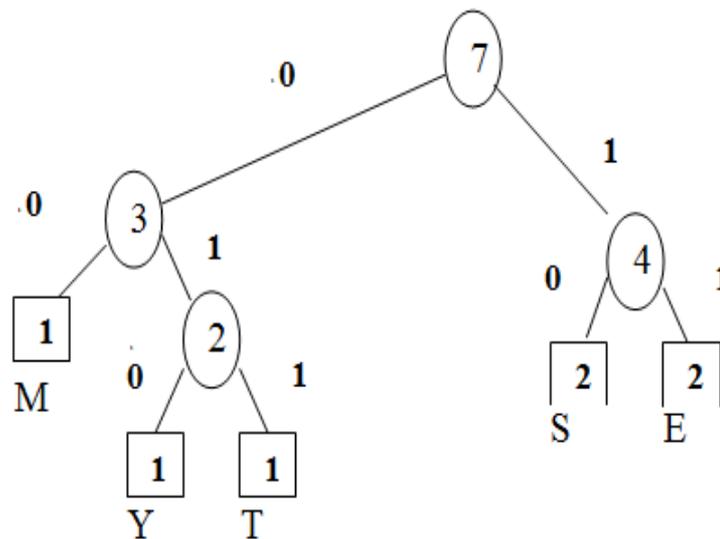
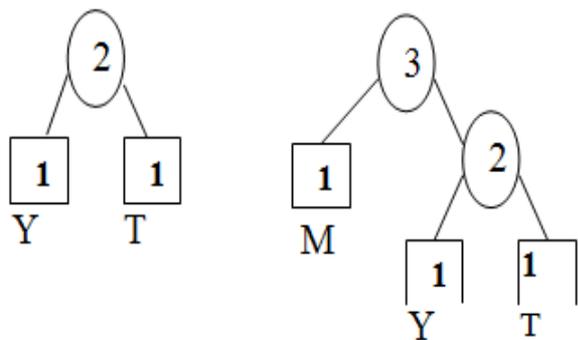
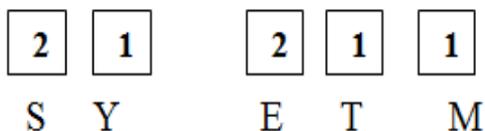
Codage de Huffman

Algorithme:

1. calculer les fréquences de tous les caractères du texte.
2. Arranger les fréquences.
3. Générer le père des 2 feuilles minimales (poids du père = somme des valeurs des feuilles)
4. Répéter le même processus pour toutes les feuilles restantes et sur les nœuds générés jusqu'à ce que la racine soit générée. On obtient l'arbre de Huffman.
5. Parcourir l'arbre de Huffman en mettant 0 pour la branche gauche et 1 pour la branche droite.
6. le caractère correspondant à chaque feuille aura comme code : la suite des 0 et des 1 qui ont été mémorisés à partir de la racine.

Codage de Huffman ; exemple:

Exemple : le texte : "SYSTEME "



M	_____	00
Y	_____	010
T	_____	011
S	_____	10
E	_____	11

texte initial : ' système '

⇒ texte après compression :

"10 010 10 011 11011"

Codage de Huffman

Suite de l'exemple:

Longueur du texte avant compression = $7 \times 8 = 56$ bits \Rightarrow

Longueur du texte après compression : 16 bits

$$\text{Taux de compression} = \frac{\text{Taille du texte comprimé}}{\text{Taille du texte avant compression}} = \frac{16}{56} = 0,28 = 28\%$$

\Rightarrow 68 % de gain d'espace dans le stockage

Codage de Huffman

exercice:

- Construisez un codage de Huffman du message "ceciestuncodagedehuffman" (on a supprimé les espaces et la ponctuation pour simplifier la construction). Il y a plusieurs codages de Huffman possibles.