

Techniques de Représentation de l'Information

Représentation - Codification – contrôle

Badiâa DELLAL-HEDJAZI

USTHB- Faculté d'informatique

Information : Signifiant, signifié, référence

1. **Le signifiant** : mots, codes, symboles, alphabet
2. **Le signifié** : ce qu'on comprend : idées, sens, pensée.
3. **La référence** : ce dont on parle : choses, faits, monde réel

Exemple : Fiche client contient les informations « **code client** » et « nom du client ».

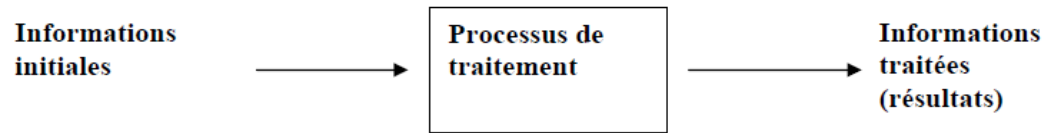
Signifiant : code client **G01516**

Signifié : **15^{ième} grossiste de la wilaya 16**

Référence : le client (personne physique ou morale).

Manipulation de l'information

1. **Création** : générer une nouvelle information. *Exemple* : saisie de nouvelles informations
2. **Recopie** : passer l'information d'un support vers un autre sans rien lui ajouter. *Exemple* : Flash-disk vers disque
3. **Traduction** : Recopie avec modification de l'information.
Exemple : Traduire un fichier sur disque en un texte imprimé sur papier.
4. **Mémorisation** : Conservation de l'information sur un support. *Exemple* : Fichiers sur disque, flash-disk, CD, papier, ...
5. **Transport** : Déplacer l'information par réseaux de transmission, bus d'E/S, ondes.
6. **Traitement** : Le processus de traitement peut se traduire par un algorithme.



Remarque: Information traitée n'est pas information créée.

Exemple: $2+3 = 5$; 5 n'est pas une information nouvelle, on la retrouve à partir des informations 2 et 3.

7. **Destruction** : suppression.
8. **Compréhension** : Processus complexe et difficile (reconnaissance : parole, formes).
9. **Résumé** : Réduire son volume en gardant son sens. *Exemple* : Compression

Structures de Données

- Les tableaux
- Les fichiers
- Les bases de données

Dans ce qui suit, nous nous intéressons aux **fichiers**.

Classes d'information et occurrence de classe

Une **classe** est un ensemble d'éléments ayant les *mêmes caractéristiques*.

Rubrique

Matricule	Nom	Prenom	Datenais	Adresse
32142	Saidi	Mohamed	14/10/1990	5 Place Martyrs
33113	Lounisi	Omar	21/05/1991	2 Rue Belle Vue

Article {

Classe

Une **occurrence** d'une classe est un élément particulier à cet ensemble.

Exemples:

- **La classe Etudiant:** chaque étudiant est une occurrence de cette classe.
- **La classe Enseignant:** chaque enseignant est une occurrence de cette classe.

Rubriques et occurrences de rubriques

Une **rubrique** est le plus petit élément d'information manipulé lors d'un traitement.
On dit aussi propriété et occurrence de propriété.

Exemples:

- Le matricule d'un étudiant à l'USTHB.
- Une rubrique prend à un instant t , une et une seule valeur appelée occurrence.
- L'occurrence 33113 de la rubrique matricule.

Classification des rubriques

1) **Rubriques primaires** : caractéristiques naturelles. *Exemple*: Nom d'une personne, numéro d'un compte.

2) **Rubriques calculées** : occurrences obtenues par un calcul lors d'un traitement.

- **Compteur** : la nouvelle occurrence est obtenue en ajoutant une constante appelée « pas » à l'ancienne occurrence.

Exemple : NPERS := 0 ; NPERS := NPERS + 1

- **Cumuls** : la nouvelle occurrence est obtenue en ajoutant à l'ancienne occurrence l'occurrence d'une autre rubrique.

Exemple: QTE_EN_STOCK (EN_COURS) = QTE_EN_STOCK (PRECEDENTE) + NOUVELLE_QTE_ENTREE

On cumule à chaque fois la nouvelle quantité à la valeur précédente du stock.

- **Choix** : chaque occurrence est obtenue en appliquant un certain nombre de conditions sur les occurrences d'autres rubriques.

Exemple : Si Salaire_Brut <= 5000 alors IRG = 0

- **Calculs** : par expressions arithmétiques et/ou logiques. *Exemple*: NET_A_PAYER = Σ INDEMNITÉS - Σ RETENUES + Σ AVANTAGES SOCIAUX.

3) **Rubriques de situation** : mise à jour suite à un événement. *Exemple*: Situat. familiale : C: Célibataire, M: Marié, D: Divorcé, V: Veuf.

Important : préciser leurs états à l'instant initial t0.

4) **Rubriques libres** : indépendantes du système d'information, facilitent la lecture des documents et générées par le système d'édition.

Exemples: n° d'ordre, numéros de page, signatures, entêtes et pieds de pages, etc.

Format d'une rubrique (Nature, longueur)

1 - Nature de la rubrique (type) :

Numérique : chiffres (0, ...,9), les symboles de signe + et -

Alphabétique : toutes les lettres de l'alphabet

α-numérique : chiffres, lettres, caractères spéciaux

Date: 15/11/2018

2 - Longueur de la rubrique : C'est le nombre de caractères de l'occurrence la plus longue.

Exemple: Le matricule étudiant : 7 caractères numériques , Nom étudiant : 30 caractères.

3 - Champ d'application ou domaine : l'ensemble des occurrences possibles pouvant être prises par cette rubrique.

4 - Occurrences vides : attention: vide ≠ nulle

Exemple: Un employé en congé ne perçoit pas de PRI (Prime de Rendement Individuel).

Une valeur nulle entraîne un abattement de 5% sur le salaire de l'employé.

La valeur vide peut être symbolisée par une valeur α-numérique spéciale par exemple.

5 - Occurrences inconnues :

Exemple: à la fabrication d'un produit, on le prévoit dans le catalogue alors que son prix de vente non encore fixé.

6 - Rubrique élémentaire / rubrique groupée :

Elémentaire : on ne peut pas la décomposer en sous rubrique. *Exemple*: Nom employé

Groupée : qui peut être subdivisée en sous rubriques élémentaires. *Exemple*: Date de naissance

Description de classe

Regroupe toutes les rubriques décrivant une classe d'informations.

Exemples: Les descripteurs de classes suivants :

AFFECTATION_ETUDIANT {Matricule, Filière, Année, Groupe}

CLIENT {CodeClient, NomClient, PrenomClient, **Adresse** {Rue, CodePostal, Ville}}

ETUDIANT {matricule, Nom_Pren, **MODULE**{Code_Module, Coefficient, **EMD**{NumEMD, NOTE}}}

Adresse est une rubrique **groupée**.

C'est une rubrique « **groupée** » et « **répétitive** » si le client a plus d'une adresse.

Module est une classe subordonnée à la classe Etudiant

EMD est une classe subordonnée à la classe Module

Article : est l'ensemble ordonné des occurrences des rubriques d'une description de classe.

Client : {G01516, Younsi Khaled, {Rue de l'indépendance, 16023, Alger}}

Notion de fichier

Un fichier est l'ensemble de tous les articles Correspondant à une classe donnée.

Exemple : Le fichier Client.

CodeClient	NomClient	PrénomClient	Adresse		
G00116	Ben Said	Mohamed	Rue	Codepostal	Ville
			L'indépendance	16023	Alger
			Cité 5 juillet	16311	Alger
D12042	Mahmoudi	Khaled	Rue	Codepostal	Ville
			Rue d'Alger	42400	Tipaza
			Cité 8 mai 45	42000	Tipaza
D14335	Younsi	Faouzi	Rue	Codepostal	Ville
			Rue de la gare	35000	Boumerdas
G04709	Bentaleb	Bahidja	Rue	Codepostal	Ville
			Place 1 ^{er} Novembre	9000	Blida

Remarque:

Seule la classe générale (non subordonnée) donne lieu à un fichier, les classes subordonnées donnent lieu à des rubriques à occurrences répétitives.

Référence et indicatif

Une **référence** : rubrique qui identifie *sans ambiguïté* chaque occurrence de classe.

Exemples:

Code client, Matricule étudiant, N° sécurité sociale, N° permis de conduire, ...

Les occurrences d'une rubrique de référence sont appelées **indicatifs**.

Exemple: Matricule étudiant :

584230, Aziz, Faouzi, 10/11/1985, Alger

584301, Benabdellah, Mahdi, 15/09/1984, Constantine

2 occurrences distinctes de classe d'information **<==>** **2** indicatifs distincts

A toute classe il faut attribuer une rubrique de référence soit :

- En choisissant une parmi les rubriques de description de la classe.
- En ajoutant une nouvelle rubrique de référence.

Schéma d'information

C'est l'ensemble des descripteurs de classe du système d'information étudié noté : $S : \{DC1, \dots, DCn\}$.

Deux niveaux :

- **Niveau conceptuel** (conception) : aspect structurel des informations (schéma conceptuel du SI).
- **Niveau interne** (réalisation) : aspect implémentation ou réalisation (Base de données + logiciels).

Niveau Conceptuel	Niveau Interne
Classe Entité/Association/Événement	Fichier
Occurrence de classe	Article
Propriétés	Occurrence de rubrique
Référence	Indicatif
Schéma conceptuel du SI	SI objet artificiel « Base de données + Logiciel »

Etude de document

Société DELTA Informatique Place du centre, BARAKI 16210		BULLETIN DE PAIE		
		Mois de : Octobre 1999 (Paye)		
Matricule : 150 Fonction : CHEF DE DEPT. ADMIN ET FINANCES Affectation : Service Finances H/trav : J/Trav : 26		Nom : TOUMI Farès Cat : 15 49, Rue Hassan Khodja, Sec : 2 El-Harrach, Alger. Ind : NSS : 5623212456 S/F : M Nmut : M/p : V N° CNI : 25631/45		
Code	Libellé	Taux/Nbr	Gains	Retenues
01	SALAIRE DE BASE	26	6730.00	664.25
04	INDEMNITE DE NUISANCE	10 %	673.00	
05	I.E.P	8.00 %	538.40	
06	P.R.I	8 %	538.40	
12	P.R.C	15 %	1009.50	
30	SALAIRE DE POSTE		9489.30	
31	INDEMNITE DE PANIER	22	693.00	
32	INDEMNITE DE TRANSPORT	26	60.00	
60	SALAIRE BRUT		10242.30	
61	RETENUE S.S			
62	SALAIRE IMPOSABLE		9578.05	742.30 0.00
63	RETENUE I.R.G	9578.05		
64	RETENUE C.S.N	8835.75		
73	SALAIRE UNIQUE		5.50	
75	REMBOURSEMENT DIVERS		560.00	
90	NET A PAYER		9401.25	
Salaire de Poste		Salaire Imposable		Gains
9489.30		9678.05		10807.80
				Retenues
				1406.55
Net à Payer : 9401.25				

Classement rubriques:

Employeur {Nom employeur, Adresse employeur} on ne la crée pas car elle contient une seule occurrence.

Employé {matricule, Nom employé, Adr employé, fonction, affectation, S/F, Cat,Sec, Ind, M/p, NSS, Nmut, N° CNI}

Fiche de paye {mois, année, typecalcul, jtrav, htrav}

Element de paye{code, libellé, Taux/Nbr, type élément, formule de calcul, valeur}

Nom rubrique	Occurrence	Nature	Règle de calcul
Nom de l'employeur	Société Delta Informatique	A/P	∅
Adresse employeur	Place du centre, BARAKI, 16210	A/P/G	∅
Nom du document	BULLETIN DE PAIE	A/L	∅
Mois de paye	Octobre	D/P	∅
Année de paye	1999	D/P	∅
Type de calcul	Paye	A/P	∅
Matricule	150	N/P	∅
Fonction	CHEF DE DEPT. ADMIN ET FINANCES	A/S	∅
Affectation	Service Finances	A/S	∅
H/Trav	∅	N/P	∅
J/Trav	26	N/P	∅
Nom-prénom	TOUMI Farès	N/P/G	∅
Adresse employé	49, Rue Hassan Khodja, El-Harrach, Alger	N/P/G	∅
Cat	15	N/S	∅
Sec	12	N/S	∅
Indemnité nuisance	673,00	N/C	Salaire de base * Taux nuisance
IEP	538,40	N/C	Salaire de base * Taux IEP
PRI	538,4	N/C	Salaire de base * Taux PRI
PRC	1009,50	N/C	Salaire de base * Taux PRC
Salaire de poste	9489,30	N/C	Salaire de base + indemnités (Nuis. + IEP+PRI+PRC)
Indemnité panier	693,00	N/P	∅
Indemnité transport	60,00	N/P	∅
Salaire brut	10242,30	N/C	Salaire de poste + indemnités (panier + transport)
Salaire imposable	9578,05	N/C	Salaire brut – Retenue SS
Net à payer	9401,25	N/C	Salaire imposable - Retenue IRG + Indemnités + Remboursements divers

Codification de l'information

Intérêts de la codification:

- Améliorer la désignation des informations les caractériser d'une manière unique .
- Eviter de manipuler des informations trop longues, donc gain de temps et d'espace.
- Structurer l'information pour son traitement automatique et sa définition de manière claire et précise

Définitions:

- Un **code** est une représentation abrégée d'une information.
- La **codification** consiste à remplacer une information par un code clair et adapté aux besoins de l'utilisateur du S.I.

La codification d'un rubrique ou variable peut concerner son **nom** ou/et sa **valeur**.

Exemples : Le numéro client : Num_Cl ; le nom client : Nom_Cl; le numéro facture : Num_F;

Num_Cl = G 025 16 ; Num_F = 0137 20

Caractéristiques d'une codification

1 - Non ambiguë : 1 objet \Leftrightarrow 1 code

2 - Bien adapté : aux besoins des utilisateurs.

3 - Interprétable :

Un code doit être facile à décoder.

EX : Situation familiale = M, C, D, V facile à codifier et à décoder.

= 1, 2, 3, 4 facile à codifier, mais difficile à décoder

4 - Extensible : Offre la possibilité d'être étendu selon la classe d'objet à codifier. Ex: 1, 2, 3, 4, 5,

5 - Avec Insertion possible : prévoir suffisamment de valeurs pour l'ensemble des occurrences de la classe.

Il est, parfois, nécessaire de faire des études statistiques pour définir par exemple des plages de valeurs possibles.

6 - Concis : court et facile à utiliser et donc gain de temps et d'espace.

7 - Stable : interchangeable

Exemple 1:

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 client originaires du Centre.

001

100

Clients de l'Est

← Impossible d'insérer 101^{ème} client de l'Est car le numéro 101 est déjà pris.

101

100

Clients de l'Ouest

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

Exemple 2 :

Par contre un code ayant la forme suivante, est un code qui permet l'insertion

COO1
COO2
--
CO85
EOO1
EOO2
--
--
EO60
0001
0002
--
0101

Insertion du nouveau code
EO61 entre les deux codes
EO60 et 0001

Si un nouveau client arrive dans la région Est un code sera EO61. Ce nouveau code sera inséré entre les deux EO60 et 0001.

Types de codification

- 1) Codification séquentielle
- 2) Codification par tranches
- 3) Codification articulée
 - Codification juxtaposée
 - Codification hiérarchisée par niveaux
- 4) Codification mnémonique
- 5) Codes contrôlables

1- Codification séquentielle

Attribuer en séquence un numéro aux objets au fur et à mesure qu'ils se présentent.

Avantages	Inconvénients
Non ambiguë Simple Permet des extensions	Ne permet pas les insertions Non significative Aucun regroupement sur les objets

Exemple : Les employés codifiés selon l'ordre chronologique de leur recrutement: 001, 002,..., 112.

Les nouveaux recrutés auront les codes 113, 114,...

2 - Codification par tranches

Consiste en l'attribution de tranches de valeurs contiguës.

Avantages	Inconvénients
Non ambiguë Simple	Nécessité d'un tableau de correspondance entre les tranches et les classes d'objets. Ne permet pas des extensions Ne permet pas les insertions

Exemple:

La codification des ouvrages dans une bibliothèque peut se faire comme suit :

De 001 à 100 : Technologie

De 101 à 200 : Littérature

De 201 à 300 : Sociologie

De 301 à 400 : Médecine

De 401 à 500 : Culture générale

3 - Codification articulée

Le code est divisé en plusieurs zones ayant chacune une signification. On distingue deux variétés de cette codification :

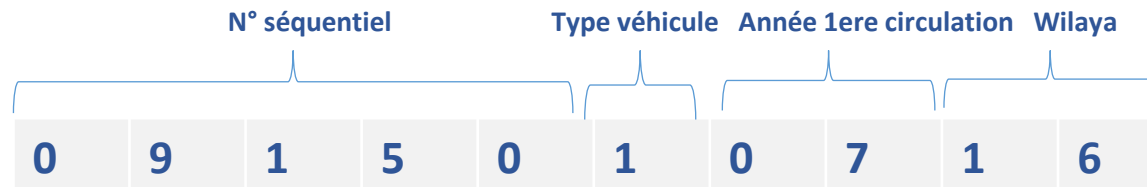
1. Codification juxtaposée :

001213 115 16

Le code se présente sous forme d'une série de champs (zones / descripteurs) les uns à côté des autres.

Exemple:

L'immatriculation d'une voiture :



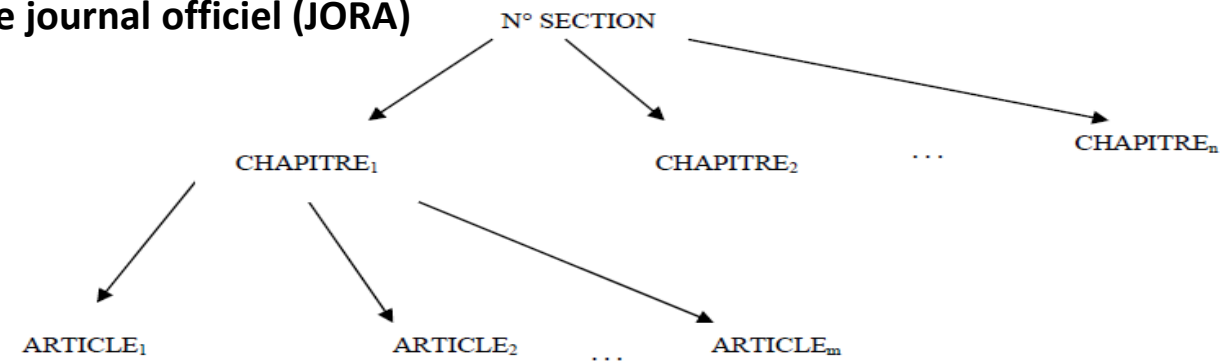
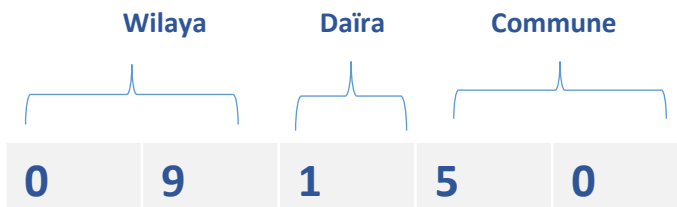
2. Codification hiérarchisée à plusieurs niveaux :

Le code peut être représenté sous forme d'une hiérarchie arborescente où chaque niveau de l'arbre dépend de son prédécesseur.

Exemples:

Codification des documents juridiques tels que le journal officiel (JORA)

Code postal :



4 – Codification mnémonique

Le code est une abréviation de la désignation littéraire d'un objet.

Exemple : le numéro de compte est désigné par NUMCOMPT.

Avantage	Inconvénient
Très pratique car près du sens de l'objet.	Ne se prête pas toujours au décodage automatique.

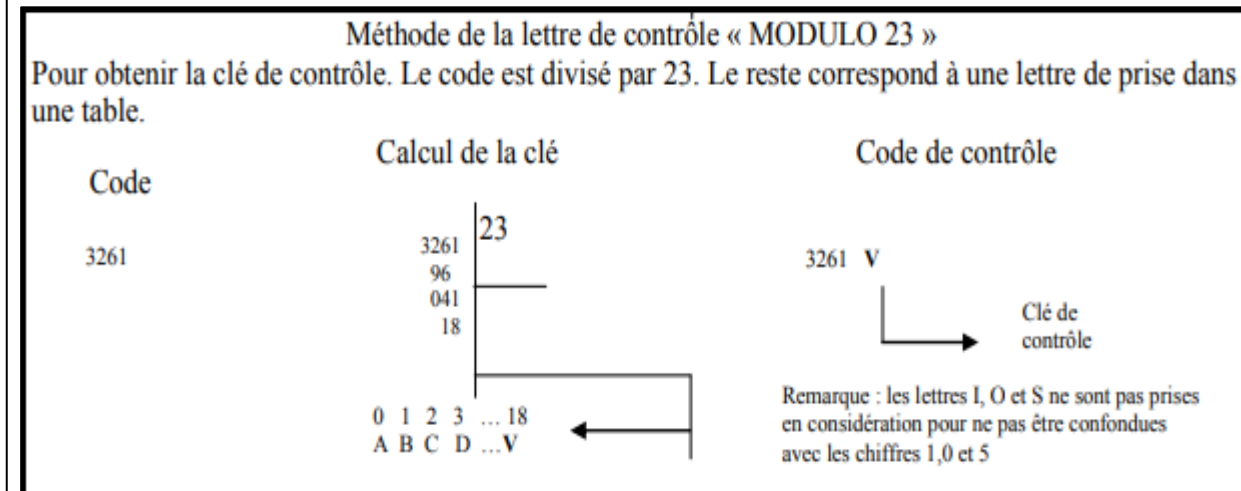
5 – Codes contrôlables

La structure permet de déceler les erreurs de transcription.

Exemple: Compte bancaire Clé : fonction de contrôle

La fonction de contrôle « lettre modulo 23 », et le code dont le format est «99999X».

L'occurrence 95121**T** : $95121 \text{ Modulo } 23 = 16 \rightarrow$ la lettre **T**
→ le code est correct.



Choix d'une codification

- 1) Étudier l'utilisation future des codes à définir.
- 2) Déterminer le nombre d'objets à codifier ainsi que la croissance de ce nombre.
- 3) Étudier la répartition statistique des objets à définir.
- 4) Rechercher s'il existe déjà des codifications pour les objets étudiés.
- 5) Définir une codification avec ses utilisateurs et par souci de standardisation tenir compte des codes déjà utilisés.
- 6) Tester la codification définie et l'adapter avant de l'utiliser définitivement.

Contrôle de l'information

L'information est vitale pour la prise de décision dans une organisation.

La présence d'une erreur dans l'information peut avoir de graves conséquences sur l'organisation. → Nécessité de contrôle

Le contrôle d'une information est la vérification de sa justesse et sa conformité par rapport à la réalité de l'organisation.

2 types de contrôles :

- **Contrôles directs**
- **Contrôles indirects.**

Contrôles directs

S'effectuent sur l'information elle-même sans tenir compte des autres informations existant dans le système.

1. Contrôle de présence ou de non présence : Vérifier l'existence ou non d'une information sur le support (document , fichier).

Exemples:

- Avant d'insérer les notes d'examen d'un étudiant, on vérifie que ce dernier existe (présence).
- Avant d'insérer un nouvel étudiant, on vérifie que ce dernier n'existe pas dans le fichier des étudiants (non présence).
Cela évite d'avoir des doublons de l'information.

2. Contrôle de type : Consiste à vérifier que le type de l'information est conforme à ce qu'il devrait être.

Exemple : N° de téléphone = 02145T566 ==> Erreur : « T » n'est pas un chiffre et le numéro de téléphone est numérique.

3. Contrôle de cadrage : Désigne la position d'une information dans une zone de saisie ou de remplissage.

Exemples: Les informations numériques sont cadrées à droite.

Les informations alphabétiques et alphanumériques sont cadrées à gauche.

Remarque:

Les outils de développement actuels (Delphi, Windev, ...) permettent la prise en charge automatique de ce type de contrôles.

Contrôles indirects

Vérifient la conformité d'une information par rapport à l'ensemble des informations se trouvant dans le système.

1. Contrôle de cohérence interne : Vérifie une partie d'une information par rapport à d'autres parties de la même information.

Exemples:

- Date : cohérence du jour par rapport au mois: la date 31/05/2007 est fausse car le mois de mai compte 30 jours.
- Soit la codification sur des véhicules roulants : [Code Véhicule] [Nombre de roues].
L'occurrence [CA] [02] signifie « Camion à 2 roues », or ceci est faux car il n'existe pas de camion ayant seulement 2 roues.

2. Contrôle de cohérence externe : vérifier la conformité d'une information par rapport à d'autres informations externes.

Exemples:

- Soit la date de naissance = 25/10/2005 d'un étudiant.
- Le matricule d'un étudiant est structuré comme suit : [Durée de formation] [Année du BAC] [N° séquentiel].

Soit l'occurrence de matricule : [5] [2007] [3569]. L'année d'obtention du BAC = 2007.

L'âge de l'étudiant = 2 ans ➔ impossible d'avoir le BAC à cet âge ==> erreur sur la date de naissance.

3. Contrôle de vraisemblance : s'assurer qu'une information est possible et concevable en fonction de son sens.

Exemples:

- Le matricule étudiant [6] [2007] [8974] porte une information invraisemblable car il n'existe pas de formation sur 6 années à l'USTHB.
- L'occurrence de code postal [60 210] est invraisemblable car il y a 58 wilayas seulement en Algérie, d'où la valeur 60 n'est pas correcte.

Exercice

Une école maternelle dispose d'une capacité d'accueil de 150 enfants. En effet, elle reçoit chaque année 50 nouveaux enfants qui arrivent à l'âge de 3 ans. Ils quittent l'école maternelle à l'âge de 6 ans pour entamer l'école fondamentale. Pour constituer le fichier informatique d'identification des enfants, l'école maternelle souhaite adopter une codification qui doit respecter les critères suivants :

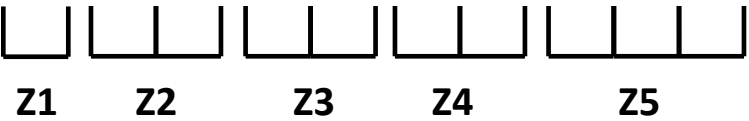
- Pérennité (durabilité du code), Détermination de l'âge en nombre d'années et de mois, Concision (le plus court possible)
- On supposera que la répartition des enfants suivant l'âge ou le genre est équilibrée.

Questions :

- 2. Proposer une codification adaptée à ce système.
- 3. Quels sont les contrôles pouvant être effectués sur cette codification ?

1) Type de codification : Codification articulée juxtaposée.

1ère solution: 5 zones (descripteurs)



- Z1 : Genre : Alphanumérique : 1 position : { M, F }
- Z2 : Année inscription : numérique : 2 positions
- Z3 : Année naissance : numérique : 2 positions
- Z4 : Mois naissance : numérique : 2 positions
- Z5 : N° séquentiel : numérique : 3 positions

Exemple:

M	0	5	0	2	1	1	0	1	4
---	---	---	---	---	---	---	---	---	---

Le 14 ième enfant garçon inscrit en 2005. Il est né en novembre 2002.

3) Contrôles directs: Type ; Présence ; Cadrage.

Contrôles indirects:

- **Vraisemblance:** Z1 ∈ {M,F} ; Z2 ∈ [00..99] ; Z3 ∈ [00..99] ;
Z4 ∈ [01..12] ; Z5 ∈ [000..150] ← extensible à 999

• **Cohérence interne**

Entre année inscription et année naissance

$3 \leq \text{année inscription} - \text{année naissance} \leq 5$

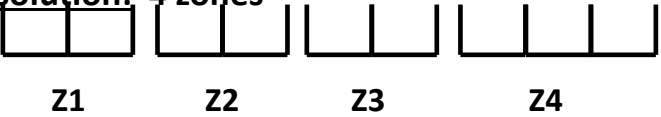
• **Cohérence externe:**

$\text{Année actuelle} - \text{année naissance} \geq 3$

$\text{Année actuelle} \geq \text{année inscription}$

$\text{Année ouverture école} \leq \text{année inscription}$

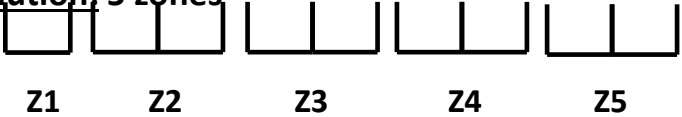
2ième solution: 4 zones



Z1: année inscription; Z2: année naissance; Z3: mois naissance; Z4: n° séquentiel

Pour Z4: codification par tranches: $150/2 = 75 \Rightarrow [001..075] \rightarrow M$ et $[076..150] \rightarrow F$

3ième solution: 5 zones



Z1: Genre; Z2: année inscription; Z3: année naissance; Z4: mois naissance; Z5: n° séquentiel

Pour Z5: codification par tranches : $150/2 = 75 \Rightarrow [01..75] \rightarrow M$ et $[01..75] \rightarrow F$

Ex.

F	0	6	0	3	1	0	2	5
---	---	---	---	---	---	---	---	---