

CHAPTER 2

CODIFICATION OF INFORMATION

Introduction

When studying an application, a large amount of information is examined and defined for computer processing. The key advantage of computers is their ability to process large amounts of information automatically, at high speed, and under excellent security conditions. However, this crucial benefit can only be achieved through significant prior work in organizing this information. In fact, as the volume of information to be processed grows, the first issue that may arise is the unambiguous identification of an object within any given set.

Objectives

- Understand the concept of codification
 - Learn about different types of codification
 - Apply the acquired concepts to real-world cases
-

Definitions and Examples

Consider student information: it's often possible to find two or more students with the same first and last name. It would be impossible to distinguish them if we only relied on their name for identification. The best solution seems to be assigning them codes. The process of assigning codes is called "codification."

Definition 1: Codification is an optimization process that replaces naturally formatted information with a conventional symbol (code) better suited for the user's objectives and for automatic processing.

Definition 2: A code is a conventional, generally abbreviated, representation of an object.

Example: Assign codes to different students: University entry year + sequential number, e.g., 201900154.

Objectives of Codification

- Unambiguously identify an individual within a set
 - Save space and time
 - Represent certain properties of an object
 - Allow certain format checks on the information
-

Expected Qualities of Codification

- **Uniqueness:** The code should uniquely identify an object within a set.
- **Flexibility and Durability:** A code should typically be used for a long time. Flexibility allows it to adapt to changes in the represented population and retain its value over time. It's important to foresee extensions and insertions due to the growth of objects over time.

Example: If a company has 95 employees and a two-digit code is chosen, the code becomes invalid if the number of employees increases to 100.

- **Ease of Use:** This includes conciseness, ease of control, and ease of exploitation.
-

Different Types of Codification

Codification possibilities are theoretically unlimited. Only a few well-known types of codes are presented here.

Sequential Codification

This involves assigning consecutive numbers to the objects to be coded within a set.

Example: Employees in a company are coded based on their chronological recruitment: 001, 002,..., 112. New recruits will be assigned codes 113, 114, etc.

Block Codification

A block codification example divides set of objects into categories, with each category assigned a specific range of codes. Here's an example for a **bookstore inventory**:

- **001 to 099:** Fiction books
- **100 to 199:** Non-fiction books
- **200 to 299:** Children's books
- **300 to 399:** Magazines

Within each block, the codes are sequential:

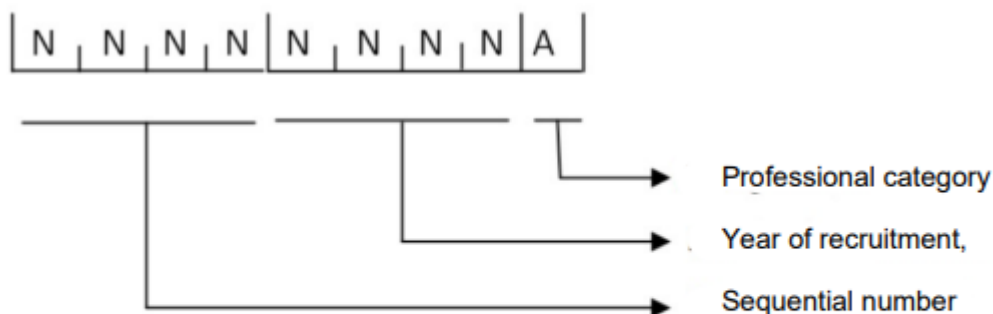
- Fiction books: 001, 002, 003, etc.
- Non-fiction books: 100, 101, 102, etc.

This system allows easy identification of the category and keeps each block organized for future additions.

Articulated Codification

The codes are divided into multiple zones, each with a specific meaning. These zones can be of the same type or different types.

Example: Coding employee IDs in a company.



Exemple: 4087 2005 S "S": Senior manager

Hierarchical Codification

This is used when there are inclusion relationships between different sets. The code is divided into several zones (levels), each representing a set of objects. From left to right, these zones represent increasingly restricted sets. This can be considered a particular case of articulated codification.

Example: Hierarchical Codification of Fruits

1. Food

1.1. Fruits

1.1.1. Citrus Fruits

- Orange
- Lemon
- Lime

1.1.2. Berries

- Strawberry
- Blueberry
- Raspberry

1.1.3. Stone Fruits

- Peach
- Cherry
- Plum

Explanation

In this example:

- Food is the highest level.
- Fruits is a subcategory of food.
- Citrus Fruits, Berries, and Stone Fruits are specific types of fruits.
- Each type includes examples of individual fruits.

This structure organizes fruits into categories, making it easier to understand and locate different types.

Mnemonic Codification

This involves abbreviating the name of an object using a reduced set of characters that evoke the coded object. There are two types:

- **Consonant Codes:** Created by removing vowels, e.g., "FACTURE" becomes "FCTR."
- **Abbreviative Codes:** Created by abbreviating the object's name, e.g., "FACTURE" becomes "FACT."

Conducting Codification Operations

Codification is a crucial step in information preparation, as errors at this stage can have serious consequences and are often hard to detect. Therefore, the following approach is recommended:

- Define the population to be coded (type of individual, number of individuals, growth, maximum, etc.)
 - Study statistical distributions, if necessary
 - Check if codifications have already been proposed by specialized organizations
 - Ensure the stability of the properties upon which the code is based
 - Define the codification with the users who will later work with it
 - Implement the code (provide usage instructions and test it with users)
-