

Computer Aided Design for Soil Classification Relational Database and Retrieval Techniques

J.E. Atajeromawwo, E.U. Tibi (PhD), O. Ugboh (PhD)

Abstract

The paper focuses on the problems associated with classification, storage and retrieval of information on soil data, such as the incompatibility of soil data semantics; inadequate documentation, and lack of indexing; hence it is pretty difficult to efficiently access large database. Consequently, information on soil is very difficult to retrieve hence, modification and update of soil data become very difficult. The relations supported by the system for the purpose of storage and retrieval of soil classification are outlined and translated into graphical presentation. The paper also presents algorithms showing the procedure for generating various soil classifications, retrieval techniques for efficient and quick retrieval of information on soil data.

Introduction

Classification is the means whereby we order knowledge. There are rules and regulations guiding classification of objects in many fields of endeavour. In object – oriented design, recognizing the sameness among things allows us to expose the commonality within key abstractions and mechanisms, and finally leads us to simpler designs. In reality, there is no golden path to classification. To the reader accustomed to finding cookbook answers, we unequivocally state that there are no simple recipes for identifying classes and objects. There is no such thing as the “perfect” class structure, nor the “right” set of objects. In engineering discipline, for instance, design choices are a compromise, shaped by many competing factors.

In the same vein, at a conference on software engineering, several developers were asked what rules they applied to identify classes and objects, reported by Swaine (1988). Stroutup, the designer of C++, responded, “It’s a holy Grail.

There is no panacea". Gabriel, one of the designers of CLOS, cited by Booch(1991) stated, "That is a fundamental question for which there is no easy answer. Luckily, there thus exists a vast legacy of experience with classification in other fields (disciplines)". From more classification approaches, the techniques of object-oriented analysis, domain analysis, and various hybrid methods have emerged, (Swaine 1988).

The Importance of Proper Classification

Good classification is based on specialization, generalization and aggregation hierarchies. Intelligent classification is actually a part of all good science. As Michalski and Stepp (1983) observed:

An omnipresent problem in science is to construct meaningful classifications of observed objects or situations. Such classifications facilitate human comprehension of the observations and subsequent development of a scientific theory

Therefore, classification is relevant to every aspect of object-oriented design and helps to identify generalization, specialization and aggregation hierarchies among classes. Classification also guides in making decisions about modularization; that is, it supports modular design. This means that certain classes and objects may be placed together in the same module or in different modules, depending upon the sameness we find among their declarations; coupling and cohesion are simply measures of this sameness; classification also plays a vital role in allocating a processor or processors. We place certain processes together in the same processors or different processors, depending on how these processes are related functionally.

Soil classification entails the grouping of soils on the basis of their common properties and the formation of classes of soils. (Brady, 1984). Soil classification helps to deal with complexity of soil relationships with one another, which involves aggregation and generalization. Soils have been classified based on certain properties and the classes of soil include that based on:

Genesis:

The concept of natural bodies as introduced in Russia (North cote, 1979) and the emphasis was on the relationship among climate, vegetation and soil characteristics. Some of the terms used until recently in the United States to describe soils were Russian origin; example being chemozerm (black earth) and podzol (under ash). The soil classification based on natural body concept has a strong genetic bias, in that classification of a given soil was dependent to a considerable degree on the supposed process by which it was formed. The difficulty of ascertaining the genetics of a given soil was one of the weaknesses of this system (Atajeromevwo, 2005)

Soil Taxonomy:

A comprehensive soil classification system called soil taxonomy which maintains the natural body concept and has two other major features that make it most useful, is presented in soil taxonomy (soil survey staff, 1975). It was a system of arrangement into groups or categories on the basis of their characteristics. First, the primary basis for identifying different classes in the system is the properties of soils, that can be measured quantitatively either in the field or in the laboratory. Furthermore, others can verify the measurements so obtained. This reduces the likelihood of controversy over the place of a given soil classification system which is common when scientists deal with systems where genesis or presumed genesis is the basis for classification.

Another significant feature of soil taxonomy is the nomenclature employed, especially for the higher classification categories. The name give a definite connotation of the major characteristics of the soil which can be easily understood in many languages where words were derived from Latin or Greek. Consideration used for naming has been given to the nomenclature used after brief reference is made to the major soil properties. (Tabor 2001).

Soil taxonomy is based on the properties of soils as they are found today. One of the objectives of the system is to group

soils similar in genesis and identified properties such as chemical, physical and biological. The advantages of this system over those based primary on soil genesis or presumed soil genesis are as follows:

- It permits classification of soils not on the basis of soil forming process.
- It focuses on the soil rather than related sciences such as geology and climatology.
- It permits the classification of soils of unknown genesis.
- It permits greater uniformity of classification as applied by a large number of soil scientists. Differences in interpretation of how a soil was formed do not influence its classification.
- It incorporates diagnostic horizon to give credibility to this soil classification system because soil properties are the primary ingredient or criteria for classifying soils. (Soil survey Staff, 1999).

United State Soil Classification based on Soil Taxonomy:

The first formal system of soil classification was introduced in the United States. The system, however, had some serious limitations, and by the early 1950's, the classification was improved upon. The process of development of the new system took nearly a decade to complete. By 1960, the review process was completed and the Seventh Approximation soil classification system has undergone numerous minor modifications. It was later put under the control of Natural Resources Conservation Service (NRCS), which is a branch of the Department of Agriculture. Many diagnostic properties used in the classification were so narrowly defined that laboratory analysis was necessary before a soil class could be established (F.A.O. 1990). The provision and relative absence of overlapping ranges of property values had strong criticism from some pedologists but the classification became useful in arranging soils for which such data were available.

The seventh approximation was based on pedogenesis. The criteria of categories being soil properties indicative of or produced by particular pedogenic process. However, in this case, the primary soil orders were not associated with climate or geographic environments but are based on internal characteristics of the diagnostic horizon. This has been regarded as one of the merits and advances of the seventh approximation, despite the departure from the environmental approach. The current version of the system has six levels of classification in its hierarchical structure (USDA, 1975).

Other soil classification systems include: Canadian system of soil classification, Food and Agricultural Organisation (FAO) system of soil classification, Australians soil classification and classification of soil of tropical regions (FAO, 1998).

Relational Model Of Soil Classification Database

The general form of the relational representation is given by $R[A_1, A_2, A_3... A_k, A_{k+1}, \dots A_{n-1}, A_n]$ [Date 2000]

Where "R" represents the name of the relation, the set

$\{A_i\}$, $i=1, 2, \dots n$ represents the attributes of the relation "R" and the set of attribute(s) underscored constitutes the unique key of the relation R. The major features of relational database model are as follows:

- a. A relation is identical to the conventional file
- b. A relation is viewed as a two dimensional array, that is, a matrix of 'n' rows and 'm' columns
- c. A row of relation, which is technically described as a tuple, represents a record.
- d. A column of a relation describes the attribute of a record
- e. An attribute draws its value from values set described as a domain. For example, the domain of the attribute 'horizon' of a soil record is the $\{O, A, B, Be\}$.
- f. Two or more attributes of a record may draw their values from the same domain. For example, the attributes 'PBC' and 'EAR' of the record of a soil type in a soil database

- draw their values from number digits with two decimal places.
- g. Each record or tuple has a unique identifier or primary key. For example, the unique identifier of a diagnostic surface horizon in a soil database is the 'general soil classification (GST).
 - h. A tuple may have one or more candidate key, which may form an alternate primary key. For example, 'diagonistic subsurface horizon' may serve as alternate key to a horizon in a soil database.
 - i. A tuple may have secondary keys and such secondary keys do serve as indices to a cluster of related records. For example, 'classification code and 'general classification code' may serve as the secondary keys in a soil database. With the secondary key 'classification', say, it is possible to index and report all the soil that belong to the same classification system.
 - j. The logical connection of one record in a relation to another relation is done by use of a foreign key. When the key of a record in relation 'A', say is replicated in another record in relation B at run time. Thus, the foreign key serves as a logical pointer which is established at the database creation. For example 'soil classification code' which is a primary key in diagnostic horizon is a foreign key in Soil Classification.
 - k. Each attribute of a relation has single value. Thus, multiple valued attributes such as vector or repeating group are not allowed. Consequently, the length of a record in a relational database is fixed and not variable.
 - l. The output report of a relational database processing is a relation. Thus, in a relational database environment, the following two types of relations are commonly found.
 - i. Base relation
 - ii. Derived relations

The base relations are the relations specified at the database design and creation time. They are static in nature. The

derived relations on the other hand, are the relations, which are obtained at the run time of database; consequently, they are dynamic in nature.

- m. The processing operations are drawn from the mathematical operations characteristics of algebra and calculus. For example, there are operations such as:
 - i. Selection
 - ii. Projection
 - iii. Intersection
 - iv. Union
 - v. Join

The algebraic relational database processing is procedural in nature, hence fairly difficult to learn, understand and use. The calculus relational database processing is declarative in nature, hence very easy to learn, understand and use.

The following are the relations considered in the soil classification database:

- a. Diagnostic Surface Horizon [soil classification code, diagnostic surface horizon-id], name of soil order, surface horizon, colour, saturation, appearance, organic content, mineral].
- b. Diagnostic Subsurface Diagnostic Horizon: [Soil classification code, diagnostic subsurface horizon id subsurface horizon, horizon, colour, mineral content, soil particle, types of weathering].
- c. Soil Diagnostic Horizon-File [soil diagnostic horizon acronym, soil diagnostic horizon-description].
- d. EAR-PBC [equilibrium activity ratio potassium buffer capacity id soil type, horizon, clay percentage, EAR, PBC]
- e. Soil classification [soil classification code, name of soil order, soil classification description]
- f. Soil vegetation, [soil general id, soil classification code, diagnostic surface horizon id name of soil order, vegetation, characteristic]

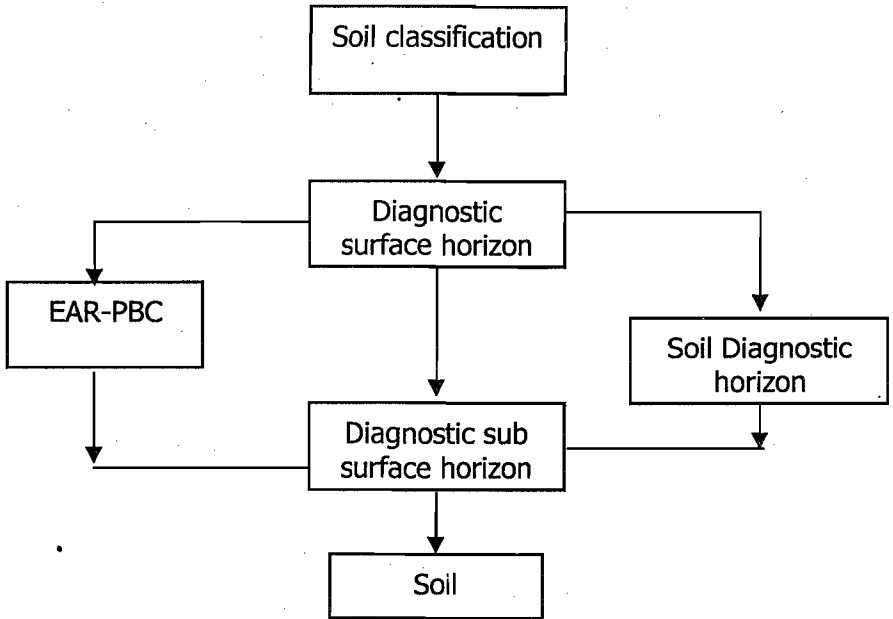


Figure 0.1 The Graphical Representation of Relational Model of Soil Classification Database.

Figure 0:1 shown above represents integrated view of the soil data. A rectangle describes a relation and an arrow describes the logical relationship between two relations.

Transactions on Soil Classification Database

Typical transactions that can be carried out on the database are:

- a. Generate the diagnostic soil surface horizon for a given soil classification such as Generate soil taxonomy, USA, Canada, Australia, FAO,
- b. Generate the diagnostic soil surface horizon for a given soil classification.
- c. Generate a report on soil diagnostic horizon description
- d. Generate a matrix of equivalence of all soil classification
- e. Given diagnostic surface horizon, generate the EAR/PBC
- f. Report the vegetation that characterized a given soil surface horizon.

Algorithm for Generating Soil Surface Horizon

The logic of the transactions required for generating soil surface horizon is presented as follows:

- a. Generate Taxonomy Soil Classification
Open Diagnostic Soil surface horizon file
On the status field, SELECT, [Diagnostic Surface horizon]
[Classification Code], [Diagnostic Code], [Diagnostic
Surface] [horizon] [name of soil order] annexid,
[Diagnostic Surface Horizon] appearance, [Diagnostic
Surface horizon] [organic Content] FROM [Diagnostic
surface]
WHERE ([Diagnostic Surface Horizon] [classification]) ="
general taxonomy"
Display the record of Diagnostic Subsurface horizon in
general taxonomy
Else
Display "no record of Diagnostic Surface Horizon in general
taxonomy found"
End if
Close the diagnostic surface soil horizon file
- b. USA Soil Classification
Open Diagnostic Soil Surface Horizon
On the status field, SELECT, [Diagnostic Surface horizon]
[Classification Code], [Diagnostic Code], [Diagnostic
Surface][horizon] [name of soil order] annexid, [Diagnostic
Surface Horizon] appearance, [Diagnostic Surface horizon]
[organic Content] FROM [Diagnostic surface]
WHERE ([Diagnostic Surface Horizon] [classification]) ="
general taxonomy"
Display the record of Diagnostic Subsurface horizon in
general taxonomy
Else
Display "no record of Diagnostic Surface Horizon in general
taxonomy found"
End if
Close the diagnoses of the surface soil horizon file

- b. General Taxonomy Soil Subsurface Horizon
 Open Diagnostic Subsurface horizon file
 On the status field, SELECT, [Diagnostic Surface horizon]
 [Classification Code], [Diagnostic Code], [Diagnostic
 Surface] [horizon] [name of soil order] annexid,
 [Diagnostic Surface Horizon] appearance, [Diagnostic
 Surface horizon] [organic Content] FROM [Diagnostic
 surface]
 WHERE ([Diagnostic Surface Horizon] [classification
 System]) = "General taxonomy"),
 If found;
 Display Diagnostic subsurface horizon in General taxonomy
 classification system.
 Else
 Display "no record of Diagnostic Subsurface horizon in
 General taxonomy found"
 End if
 Close the diagnostic subsurface horizon file
- c. Canada Soil subsurface Horizon
 Open Diagnostic Subsurface horizon file
 On the status field, SELECT, [Diagnostic Surface horizon]
 [Classification Code], [Diagnostic Code], [Diagnostic
 Surface horizon]. [Diagnostic Surface] Colour, [Diagnostic
 Horizon] [type of weathering], [Diagnostic Subsurface
 Horizon] [Classification System]
 FROM [Diagnostic Subsurface Horizon]
 WHERE ([Diagnostic Surface Horizon]
 [classification System]) = "Canada"),
 If found;
 Else
 Display "no record of Diagnostic Subsurface horizon in
 Canada found"
 End if
 Close the diagnostic subsurface horizon file

Algorithm for Generating Diagnostic Horizon Description

The Diagnostic Soil Surface Horizon Description is a general characteristic that is common to all existing Soil Classification.

The logic for generating it is as follows:

Open the diagnostic soil surface horizon description file

```
SELECT [soil diagnostic description] [soil diagnostic horizon acronym], [soil diagnostic description] [sdh-description]
```

```
FROM [soil diagnostic description]
```

If found,

Display the record of diagnostic soil surface horizon

Else

Display "no record of diagnostic soil surface horizon found"

End if

Close the diagnostic soil surface horizon file

Algorithm for Generating The Equivalence of All Soil Classification

In practice, it is desirable to compare the features of all soil classification. In this paper, an attempt is made to present the logic for generating the matrix of equivalence of all soil classification.

```
SELECT [soil classification] [classification system], [soil classification] [name of soil order], [Soil Classification] classification code/Appearance]
```

```
From [Soil Classification]
```

If found;

Display the record of all soil classification system

COMPARE

The features of found classification systems.

End if

Close the soil classification file

Algorithm for Generating Ear-Pbc of a Given Surface Diagnostic Horizon

Open EAR-PBC file

```
SELECT [EAR-PBC] horizon, [EAR-PBC] [soil type], [EAR-PBC] location, [EAR-PBC] [clay percentage in group two], [EAR-PBC]
```

```

[clay percentage in group three], [EAR-PBC] [clay percentage in
group four]
FROM [EAR-PBC0;
If found;
    Display the record of EAR-PBC
Else
    Display "no record of EAR-PBC found"
End if
Close the EAR-PBC file
Algorithm for Retrieving Soil Classification
Open Soil Classification file
SELECT [Soil Classification] Classification system], [Soil
Classification] [name of soil order,] [Soil Classification]
[Classification code]
FROM [Soil Classification]
If found;
    Display the record of Soil Classification system
Else
    Display "no record of Soil Classification system found"
End if
Close the Soil Classification file
Algorithm for Generating Soil Vegetation
Open the soil vegetation file
SELECT [Soil Vegetation] [classification code], [Soil Vegetation]
[name of soil order], [Diagnostic surface horizon] [surface
horizon], [Soil Vegetation] characteristics
FROM [Diagnostic surface horizon] INNER JOIN [Soil Vegetation]
ON ([Diagnostic surface horizon] [name of soil order] = [Soil
Vegetation] [name of soil order]) AND ([Diagnostic surface
horizon] [dsh id] = [Soil Vegetation] [dsh id] AND ([Soil
Vegetation] [dsh id] = [Diagnostic surface horizon] [dsh id]) AND
([Diagnostic surface horizon] [classification code] = [Soil
vegetation] [classification code]) AND ([Diagnostic surface horizon]
[dsh id] = [Soil Vegetation] [dsh id])
WHERE ([Diagnostic surface horizon] [surface horizon] = "A");
If found;
    Display the record of soil vegetation
Else

```

Display "no record of the soil vegetation found"

End if

Close the soil vegetation file

Conclusion

The basic objectives of the Computer-Aided Design for Soil Classification relational database and retrieval techniques are simply: to provide a framework for the creation of soil classification database to facilitate the storage and retrieval of soil data in various classification system as posited in the paper. The archival system of soil data storage is faced with myriad problems of retrieval and this has been solved with the Computer-Aided Design approach, which gives rooms for proper documentation of soil classification database as it will provide acquisition, storage display and analysis of soil data. The use of computer-aided design will assist in making it possible for soil data to be properly indexed to make accessibility to large soil database, smaller and efficient for modification. It also provides declarative processing techniques, involves structural query, processing language such as SQL; which supports intelligent, user friendly and interactive processing of soil classification database.

References

- Akinyokun O.C. and Arekete (1996): A case study of knowledge Based Server for Evaluation of Assets. *Ife Journal of Technology*. Published by Obafemi Awolowo University Press: Volume 6 No. 1 pp. 11-24.
- Atajeromavwo E.J. (2005) Conceptualization and implementation of soil classification Database; Unpublished M. Tech Thesis, Department of Computer Science, Federal University of Technology, Akure, Nigeria.
- Avery B.W. (1980); Soil Classification for England and Wales: *Higher Categories*. Soil.
- Brady, N (1984) The Nature and Properties of Soils New York: Macmillan Publishing Co. pp. 24-50.

- Brady and Weil (1996) Soil Classification in USA Vol 856 of Agricultural Research Report.
- Booch, G. (1991) Object Oriented with Application: New York Benjamin/Cumming Publishing Company Inc. Canada Expert Committee on Soil Survey (1987). The Canadian system of soil Classification. Publications 1646.
- Canadian Government Publishing Centre. Ottawa, 3rd Edition.
- Codd E.F. (1990): The Rational Data Model of Data Management. Version 2 Reading MA Addison Wesley.
- Date C.J. (1986): Relational Database Selected Writings Reading. Addison Wesley pp. 461.
- Date C. J. (2000): An Introduction to Database Systems. 7th Edition Addison Wesley.
- Darwin C. (1984): The Origin of Species. Vol. 49 of Great Books of the Western World Chicago, II Encyclopedia Britannica. pp.207.
- F.A.O. (1998); Soil Map of the World: Revised Legend. World Soil Resources Report 60. *Food and Agriculture Organization of the United Nations*, Rome.
- FAO. (1998). World reference base for soil resources. World Soil resources Report 84. *Food and Agriculture Organization of the United Nations*, Rome.
- International Institute of Tropical Agriculture Research, Ibadan, (1976): Annual report, pp.71-73.
- Michalski, R. and Stepp R. (1983): Learning from Observation: Conceptual Clustering, In Machine learning: An Artificial Intelligence Approach. Palo Alto, California
- May R. (1988): How many Species Are There on Earth? Science Vol. 241 pp. 1441.
- Ojeniyi O.O (2004). A study on the development of Automated Herbarium Databank; Unpublished M.SC Thesis, Department of

Computer Science and Engineering, Obafemi Awolowo University, Ile- Ife, Nigeria.

Soil Survey Staff. Keys to soil Taxonomy (1998) US Government Printing Office, Washington, DC, 8th edition.

Soil Survey Staff. (1999) Soil taxonomy a basic system of soil classification for making and interpreting soil surveys. Agricultural Handbook 436. US Department of Agriculture Soil Conservation Service, Washington, DC, 2nd edition.

Soil Survey Staff. (1975) Soil Taxonomy: a Basic System of Soil Classification for Making and Interpreting Soil Surveys. USDA Handbook No. 436. U.S. Government Printing Office Washington, D.C.

Swaine M. (1988). "Programming Paradigms". Dr. Dobb's Journal Software Tools, No 140, page. 110.

Tabor, J. 2001. Soil Classification Systems. <http://ar.Arizona.edu/OALS/IALC/soils/classifdtems.html>

Wegner (1987): The object oriented classification Paradigm, in Research Directions in Object oriented Programming, Cambridge: The MIT Press P. 480