DATA MODELLING AND CONCEPTUAL MODELLING: A COMPARATIVE ANALYSIS OF FUNCTIONALITY AND ROLES

T.William Olle
T.William Olle Associates Ltd.,
Walton on Thames,
Surrey, United Kingdom

ABSTRACT

This paper reviews the functionality associated "data modelling" and "conceptual modelling". The history and origins of each term are considered, together with the current interpretation of each term. The term "information modelling" is also taken into account. Alternative representation forms are presented and reviewed. The merit of diagrams as a basis for a dialogue with a subject area expert is indicated. The paper suggests that a clear distinction is needed between data modelling and conceptual modelling. Both analytic modelling and prescriptive modelling are reviewed. The requirements for a conceptual schema modelling facility over and above the functionality provided by currently available data modelling facilities are presented. The need is emphasized for a conceptual schema modelling facility to support a representation form easily assimilatable by a subject area expert not familiar with information system. Based on the distinctions made, the paper suggests a way in which a data modelling facility and a conceptual schema modelling facility can be positioned in an information systems life cycle.

INTRODUCTION TO DATA MODELLING

Data modelling is a widely practised skill. It is usually associated with the preparation of diagrams showing entity types and the relationships between pairs of entity types. There are innumerable variations on the basic theme of what in this paper are given the generic name of data structure diagrams. The term "entity type" is not always used and is sometimes replaced by the term "object type" or some related variant.

The term "data model" is used in two ways. One way refers to a model for a specific subject area (such as claims handling in an insurance company) as a data model. The other way follows the practice established by Date (1975) of referring to the "relational data model" and the "hierarchical data model". Some very confusing discussions can take place if this conflict is not resolved.

This paper distinguishes by using two terms. The first term is "application data model" which relates to a subject area. Some people regard an application data model and a data structure diagram as essentially synonymous. In this paper the argument is put forward that any given application data model may be represented as a diagram or possibly in some other form. Hence the diagram is merely one of a number of options.

The second term is "data modelling facility" which is an approach used to prepare an application data model.

These two terms are based on the ISO Reference Model of Data Management (1993) which makes considerable use of the latter and introduces the former merely to clear up any possible confusion. This is discussed later in this paper.

INTRODUCTION TO CONCEPTUAL MODELLING

The term "conceptual model" is much less widely used. In some cases, it seems to have the same meaning as "application data model" introduced above. In other cases, it is clearly attempting to be much more. The term "conceptual schema" was first introduced in the "draft interim report" prepared by an ANSI/X3/SPARC Study Group (1975). Their work was picked up and elaborated some two years later by an ISO Committee (TC97/SC5/WG3) which prepared a seminal report entitled "Concepts and Facilities for the Conceptual Schema and the Information Base" (1977). (The use of the definite article in the title was presumably not intended to imply global uniqueness.)

When the idea of a conceptual schema was first introduced in 1975, it was seen as a solution to a practical problem. It became a springboard form much research. The practical problem was that there was considerable debate about the relative merits of the three main approaches in use at that time to structuring a database. These approaches were identified as network, hierarchical and relational.

The ANSI Study Group rightly recognized that data in a subject area had an inherent structure which was quite independent of the type of database management system (DBMS) which might

subsequently be used to structure and manipulate the data. This DBMS independent structure could be represented in a "conceptual schema". The DBMS dependent schema was called an "internal schema".
In passing, it should be noted that the 1977 document borrows a term used in philosophy by Wittgenstein, namely "universe of discourse". The more down to earth term "subject area" is preferred in this paper.
One of the most important added value concepts in the 1977 document is that of the 100% principle. This is defined as follows:

> "All relevant static and dynamic aspects, i.e. all rules and laws, etc. of the universe of discourse should be described in the conceptual schema. The information system cannot and should not be held responsible for not meeting those described elsewhere, including in particular those in application programs."

The clarification in the second sentence is clearer than the statement of the principle in the first sentence. The 100% principle may be paraphrased as follows.

> The business rules applicable to a given subject area must be specified declaratively. The place where they are captured is called a conceptual schema. The business rules should not be embedded in application programs.

In absence of the widespread recognition of the 100% principle and the availability of commercially available products based on this principle, there has typically been no choice available to the designer of an information system. Even those who recognize and accept the principle have had to fall back on representing many of the business rules for a subject area using a programming language.
There have been numerous attempts to design and promote facilities for preparing a conceptual schema. In essence, these amount to modelling approaches which go somewhat further than the more widely used data modelling facilities towards meeting the 100% principle. It is a matter for evaluation and conjecture whether the 100% principle is indeed fully attainable using such approaches.

INFORMATION MODELLING

The origin of the term "information modelling" is obscure and it is impossible to point to any first usage. It is certainly not defined in the 1977 document.
The distinction between the term "data" and the term "information" as established by IFIP (1966) ought to help clarify the useful distinction between "information modelling" and "data modelling".
The terms "data" and "information" were defined as follows:

> "Data: A representation of facts or ideas in a formalized manner capable of being communicated by some process. Note: The representation may be more suitable either for human interpretation (e.g. printed text) or for interpretation by equipment (e.g. punched cards or electrical signals)."

> "Information: In automatic data processing, the meaning that a human ascribes to data by means of the known conventions used in its representation. Note: the term has a sense wider than that of ordinary information theory and nearer to that of common usage."

One may be able assume from these definitions that data modelling has something to do with modelling the representations of the data and that information modelling has something to do with modelling the meaning ascribed to the data. However, the definition of "information" makes it clear that the meaning is ascribed "by means of the known conventions used in its representation."
The crux of the distinction between information and data, appears to be found in the phrase "the meaning that a human ascribes to data". If no known human ascribes any meaning to the data, should

it still be included in a model? If different humans ascribe different meanings to the same data, should both (or all) be modelled?
Looking at the "data modelling facilities" which use the term "data model" (SSADM4 (1990), Rochfeld & Morejon 1989), and those which use the term "information model" (Veryard 1992, Jardine & van Griethhuysen 1989), not to mention those which use the term "entity relationship modelling" (Barker 1990), there does not appear to be a significant distinction in functionality. In other words, the choice of term used to qualify the kind of modelling involved is essentially arbitrary. An in-depth comparative analysis of functionality provided by each approach would certainly provide valuable insight. However, it is not clear that it would justify the disparity of terminology.
In this paper the terms "data modelling" and "data modelling facility" are preferred to other variants.

MORE ON DATA MODELLING

The ISO Reference Model of Data Management (1977 clause 5.4) describes a data modelling facility as follows:

> "A schema is prepared according to a set of **data structuring rules**. Each set of data structuring rules may have an associated set of **data manipulation rules** which define the processes which may be performed on data structured according to the data structuring rules.
>
> The data structuring rules and data manipulation rules together are called a **Data Modelling Facility.**"

This description of a data modelling facility clearly sets bounds around the concept of a data modelling facility to data and the associated processes.
The ISO Reference Model of Data Management (1977 clause 8.4.2) also identifies three ways in which a data management standard can involve a data modelling facility as follows

> "a) a data modelling facility may itself be the subject of a standard
> b) a data modelling facility may be used explicitly as a definition tool in the specification of another standard
> c) the data structuring rules of a data modelling facility may be implicit in another standard (such as a standard for a database language)."

EXPRESS (1993) is an example of alternative a). The IRDS Services Interface (1992) is an example of alternative b). SQL92 (1992) is an example of alternative c).
An underlying thought here is that an application data model prepared using a data modelling facility may have different representation forms. Some typical representation forms are the following

> a) diagrams
> b) language (based on a clearly identified character set)
> c) narrative proses
> d) structured tables

This also implies that one may consider an application data model independently of any particular representation form. The same application data model may be represented as a diagram, in linguistic form and in tables. The following is a trivial but abstract example of the first three representation forms
By far the most common kind of relationship which needs to be expressed using a data modelling facility is the well known one to many relationship (where many is to be interpreted as zero, one or more)
In narrative prose, it might be represented as follows:

> For each A, there exists zero, one or more B's and for each B, there must exist exactly one A.

In diagrammatic form, it might be represented in a data structure diagram as shown in Figure 1.

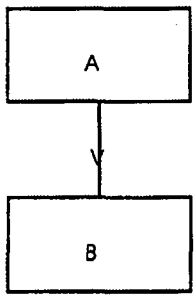

Figure 1. Diagrammatic representation of a one to many relationship

Using the Database Language SQL it might be defined as follows

```
CREATE TABLE A
(A1 SKEY PRIMARY KEY)

CREATE TABLE B
(B1 SKEY PRIMARY KEY,
 B2 SKEY NOT NULL

CONSTRAINT NAME BA
REFERENCES A
```

Figure 2. Representation of a one to many relationship in SQL92

Each of these three representation forms has advantages and disadvantages. The purpose of using them here is merely to illustrate that alternative representation forms are possible. In addition to these three, it is possible to use a mathematically oriented form such as predicate logic, which would fall in the category of a language representation.

VIEWS OF CONCEPTUAL MODELLING AND CONCEPTUAL SCHEMAS

If the 1977 document is regarded as the authoritative source of the term "conceptual schema" then one might assume that "conceptual schema" is a term for the meaning of which there is some kind of consensus. Over the years since the term "conceptual schema" has come to mean many things to many people. The following list of alternative interpretations is not exhaustive. The alternatives are not mutually exclusive interpretations. The list is intended as an indication of the differing emphases placed on the term.

A. The results of an analysis of the data and possibly also the processes perceivable in some subject area.

B. A repository of "meta-data" in which it is possible to specify declaratively 100% of the semantics of the data in a computerized information system (the 100% principle of the 1977 document).

C. A data definition which has the property that it is independent of its representation in storage.

D. A data definition which is common to the collections of data at two separate sites, such that it can be used as a common frame of reference when exporting data from one site and importing it at another site.

E. A data modelling facility which is different from and therefore "neutral" with respect to broadly similar data modelling facilities used in commercially available data base management systems.

These alternative interpretations may be analyzed as follows.

Results of analysis

As indicated earlier in this paper, data analysis facilities are widely used and there is both disparity and commonality among the techniques in use in various parts of the world. Some techniques are being energetically promoted by minority groups. The situation is complicated by the recent rapid developments in CASE (Computer Aided Software Engineering) tools.
The protagonists of some data analysis facilities choose to use the terms containing the word "conceptual" rather than "data" to refer to the approach used. In terms of the definition of "conceptual schema" cited earlier in this paper, it would be difficult to provide a justification for the terminology chosen.

100% principle

The 100% principle expounded in the 1977 document is an important principle which indirectly has had a major influence in the development of Database Language standards such as the 1989 and 1992 versions of the ISO Database language SQL standard (1992). This language specifications make it possible to specify declaratively a very large percentage of the constraints on the data which a database designer is ever likely to want to define.
While SQL is never thought of or promoted as a means of defining a conceptual schema, it is, in this very important respect, functionally superior to many of the approaches labelled conceptual modelling.

Storage independent data definition

Some groups, whose view of information systems in general and data concepts in particular, is normally on the level of stored representations, have adopted the term conceptual schema to refer to some kind of representation of the data definition which is above the level of stored representations. They tend to think that SQL92 is concerned with stored representations of data because it is used for defining databases.
SQL92 is in fact a language which enables the definition of data in a way which is independent of how the data is represented in storage. The traditional term "logical data structure" is used in the standard (rather than the term "conceptual") in order to emphasize the difference with "physical data structure" which is used to connote both representation in storage and also the inclusion of data constructs such as indices and pointers which are included for performance reasons but of which the user does not need to be aware.
The designers of the SQL92 standard (1992) have in fact made a point of excluding from the standard certain aspects, such as the definition of indexes, which are seen as belonging to the physical data structure. The fact that such facilities are to be found in commercially available products is misleading.

Common view for electronic data interchange

It is clear that in electronic data interchange (EDI), one needs a definition of the data to be interchanged which is common to all sites involved in a set of interchanges. Much of the EDI work has been concerned with the specification of standard formats for an industry area, such as banking or travel. As EDI work tends to adopt a more generalized approach to its standardization, the need for a common definition facility becomes apparent.
It is important to note the distinction between the requirement to have a data definition which is common to several sites and the requirement to be able to prepare such a definition easily. The representation form suitable for one is not suitable for the other. More specifically, to support the person to person dialogue necessary during the preparation, it is normal to use some form of diagram.

However, to support the computerized interchange, some form of computerized representation form is needed and it is not likely to be a graphics file.

Neutral data definition

The relative merit of various data modelling facilities has proved to be a somewhat emotive topic among data base researchers and practitioners. The debate goes back to the early seventies. Various advantages have been claimed for each approach.
When it is difficult to evaluate competing approaches, there is a tendency to develop yet another "neutral" approach. An interesting and potentially far reaching example of this kind of thinking is to be found in the international standard for Directory Services. To avoid getting involved in what was seen at the time as a domestic wrangle among database experts, a "neutral" data modelling facility was developed which was claimed to be more suitable for the task than those used for commercial information systems.
Since the term "conceptual schema" was originally introduced for this purpose, the choice of label to indicate neutrality is perhaps understandable.

ANALYTIC AND PRESCRIPTIVE USE OF A MODELLING FACILITY

An important distinction needs to be made at this point between the analytic and the prescriptive modes of use of a modelling facility. The term "modelling facility" is used in this paper when discussing aspects which are common to the use of both a data modelling facility and a conceptual schema modelling facility.

Analytic models

A modelling facility may be used to prepare an analytic model of a subject area. Such an analytic model is a statement by an analyst (or team of analysts) of how the subject area is to be interpreted. Unless the analyst is an expert in the subject area, then the analyst has to conduct a dialogue with one or more subject area experts, on the basis of which the analytic model can be prepared.
In practice, some subject area experts describe the subject area as it is and other experts describe it as they would like to see it at some time in the future. To be purely analytic, a model must avoid any statement about the information system which may be needed to support organizational and personal activities in the subject area.

Prescriptive models

A modelling facility may be used to prepare a prescriptive model for a subject area - or indeed of a subject area. A prescriptive model for a subject area is a prescriptive for an information system to be built to support some or all of the activities being performed in the subject area. A prescriptive model of a subject area is a model of how the subject area ought to be - quite independently of any information system which may be necessary to support the activities which are to be performed in the subject area.
The situation of more interest here is the former. A prescriptive model in effect prescribes, in subject area terms, the requirements for an information system to support the activities being performed in the subject area. This model is independent of how the information system is to be constructed.

Data modelling and conceptual modelling

As indicated, the discussion of analytic models and prescriptive models applies equally to both data modelling and conceptual modelling. However, a data modelling facility has to be seen (in terms of this paper) as functionally more restricted than a conceptual modelling facility. A data modelling facility is restricted to data and the processes which may be performed on the data.
While many self styled conceptual schema modelling facilities do no more than this, it is surely desirable for future progress in the filed of modelling to regard a conceptual schema modelling facility as not restricted to data and the processes to be performed on the data.

In order to pursue this analysis further, it is useful to examine the role of modelling facilities in an information systems life cycle.

ROLE OF MODELLING FACILITIES IN INFORMATION SYSTEM LIFE CYCLE

There is no widely accepted standard for an information system life cycle. Numerous alternative breakdowns into stages have been proposed. The breakdown into the following major stages is used in this paper and is based on the one used in the IFIP framework for understanding information systems methodologies (Olle, Hagelstein, MacDonald, Rolloand, Sol, Van Assche & Verrijn-Stuart 1991).

planning
analysis
system design
construction design
construction
operation
evolution

Outline of stages

The planning stage is a preliminary stage in which decisions can be made about the subject area such as establishing boundaries for the ensuing analysis.
The analysis stage is a stage in which a subject area (which may in practice be a complete enterprise) is examined with the aim of building a formalized understanding of what activities are performed in the subject area and the kinds of data which are needed to support such activities.
It is necessary to distinguish between two stages containing the word "design". "System design" is the stage in which the externally visible aspects of the information system are prescribed, quite independently of how the system is to be constructed and which construction tools are to be used. This could also be called logical design or even conceptual design.
The construction design stage is that which takes into account the construction tools to be used and prescribes how the system is to be constructed using these tools. Where appropriate, this sub-stage may consider such aspects of design as the breakdown into programming modules and the designation of performance enhancing indexes to be used when accessing the database. This stage could also be called implementation design.

Role of a conceptual schema modelling facility

A conceptual schema modelling facility used in analytic mode is used only in the analysis stage. In the prescriptive mode, it may subsequently used in the system design stage. It is not useful to consider a conceptual schema modelling facility in a purely prescriptive role (to the exclusion of the analytic role).

Conceptual schema in evolution stage

The role of a conceptual schema in the evolution stage is significant and indicates that there are different kinds of evolution.

1. new or changed business rules - revise conceptual model developed in analysis stage
2. improved conceptual design based on unchanged business rules
3. conceptual design unchanged, but performance enhancing changes required to construction design
4. construction design unchanged but construction errors to be corrected

These requirements emphasize the need for the conceptual schema, in both analytic and prescriptive use, to be easy to modify.

Role of a data modelling facility in different stages

A data modelling facility may also be used in the same stages as those in which a conceptual schema modelling facility is used. However, a data modelling facility would need to be supplemented in various ways.
In addition to these uses, a data modelling facility may be used in the construction design stage and in the operational stage. A construction tool may be based on a Database Language standard. An important component in the operation of a computerized information system would normally be a Database Server which interprets and enforces specifications expressed in that language.
A useful distinction between a conceptual schema modelling facility and a data modelling facility can be made in terms of the potential role of each in an information systems life cycle. However, where there is an overlap, as in the analysis stage a more important distinction can be made in terms of how much of the complete analysis necessary can be done using a data modelling facility (as currently understood) and how much needs to be done in terms of a conceptual schema modelling facility (as perceived to be needed).
There are several requirements for a conceptual schema modelling facility over and above the functionality provided by a typical data modelling facility and these are presented in detail in the following section.

REQUIREMENTS FOR A CONCEPTUAL SCHEMA MODELLING FACILITY

This section itemizes requirements for a conceptual schema modelling facility and relates these where applicable to the capability provided by data modelling facilities.

Provide user friendly representation form

A conceptual schema modelling facility must be usable as an elicitation technique for determining the business rules in a given subject area. This means that it is necessary that one form in which a conceptual schema may be represented is assimilatable by subject area experts who are not familiar with informatics oriented representation forms and certainly not with forms such as predicate logic. Typically, this representation form would be a graphic representation form. As such, the form would have a prescribed syntax and semantics.
It is also a requirement that a conceptual schema modelling facility should provide three or four "capability levels" to enable dialogue between a modeller and subject area experts when eliciting business rules of varying complexity. In addition, the level of skill for understanding conceptual models which is available among subject area experts may vary considerably.
There may be other standardized representation forms for the conceptual schema in addition to the user friendly form. In order to preserve the user friendliness, it may be possible to represent more business rules in the non-graphic form than in the graphic form. The standard must clarify this difference.

Business activities

A conceptual schema modelling facility must be able to model business activities. A business activity is an activity performed in a business area. It may be computerizable, computer supported or a purely human activity. One aim of the use of the conceptual schema modelling facility is to determine in which category to place a business activity with respect to computerization.
A conceptual schema modelling facility must provide for the identification and representation of business activities, namely activities which are performed (somehow or another) in a given business area. Depending on the scope of the business area, it must be possible to represent a hierarchical decomposition of each business activity into two or more lower level business activities. There should be no limit on the number of levels of decomposition allowed.
A conceptual schema modelling facility must be able to model other inter-relationships between business activities as well as decomposition. Such inter-relationships include the following

a) satisfactory completion of one business activity may impact the initiation of another business activity
b) one business activity may not be in progress at the same time as another business activity

A conceptual schema modelling facility must also be able to model (prescriptively) the users' perception of the user tasks which are required to be performed by an information system to be based on the conceptual schema. It must also be possible for a conceptual schema modelling facility to be able to model the work patterns in the business area in order to determine the most appropriate grouping of user tasks to support such work patterns.
The is also a requirement for a conceptual schema modelling facility to be able to relate the business processes modelled with the data concepts. This would normally be in the form of a bidirectional cross reference table (also called a matrix).
The representation forms used to depict business activities must include a graphic option (to satisfy the user friendly requirement) and may have another representation form.

Handle analysis of distribution considerations

A conceptual schema modelling facility must be able to analyze the distribution aspects of the business area. These include the following

a) what locations are relevant to the business area
b) what business activities are performed at each location
c) what business events can happen at each location
d) what data is required at each location to support business activities

It is necessary to achieve an understanding of the distribution aspects of the business in order to be able to prescribe the best approach to distribution for an information system to support the business area.
As far as can be ascertained, none of the conceptual schema modelling facilities in use address the problem of distributed systems (or the more demanding problems of inter-operable systems to be discussed in the next section). In the case of distributed systems, the lack of existing facilities is due to the fact that conceptual schema modelling facilities traditionally avoid implementation issues and whether to distribute or not is typically regarded as an implementation issue.
It is reasonable to expect that a conceptual schema modelling facility should be able to represent the concept of distribution. In analytic mode, representing the status quo for a business area, it should be possible to represent how various activities are performed at different physical locations. In prescriptive mode, it should be possible to define where certain activities should be performed in a reorganization of the way in which business activities are to be performed.
If the distribution aspects of the business activities in a business area can be represented using a standard conceptual schema modelling facility, then this would provide a distinct benefit to users of products based on the standard.
A conceptual schema modelling facility must provide for different representations of the different physical locations at which business activities are performed and also which activities are performed at each location. Possible representations of location might be as follows

latitude and longitude
name of town, city or country
identification of work site

This aspect of "representation" is orthogonal to that considered in section 4 of this paper.

Temporal considerations

A conceptual schema modelling facility must be able to model business events. A business event is an event which can happen in a business area. Examples are

1. Send job offer to applicant
2. Receive insurance claim
3. Dispatch goods.

The difference between business activities and business events should be perceived as follows. A business activity is on-going and consumes measurable resources. A business event occurs at a specific point in time, can be date time stamped and does not consume measurable resources. A business event may be initiated by human action, calendar clock or a situation occurring in the business.
The aim of the use of the conceptual schema modelling facility is to determine which business event are of sufficient importance to merit consideration in the design of a computerized information system.
A conceptual schema modelling facility must provide for the representation of temporal aspects of the business area. These aspects include, but are not limited to, the following

a) definition of business events (as distinct from business activities)
b) definition of temporal constraints between business activities
c) definition of temporal constraints between business events
d) definition of time zones associated with locations at which business activities are performed and business events happen

Both a graphic and non-graphic representation form are needed.

Handling data concepts

While data concepts are handled to varying degrees of completeness in existing standards for data modelling facilities (EXPRESS and SQL92), it is necessary to address the added requirements for a conceptual schema modelling facility standard and relate these requirements to the functionality provided by these two standards.
The requirement for a user friendly representation of data concepts is of major importance. Both EXPRESS and SQL92 are standards for a concrete syntax. However, it should be noted that the introduction to the EXPRESS standard contains the following statement (page xiv)

> "EXPRESS is consistent with the requirements specified in ISO TR 9007 for a conceptual schema language."

It is also clear from the document that use of the EXPRESS "data specification language" is intended to precede use of possible selected "implementation methods", one of which is "database implementations". In addition to the EXPRESS data specification language, there is also included EXPRESS-G which is a graphical representation form.
The requirements for the handling of data concepts in a standard schema modelling facility are one of the more complex issues associated with the proposed standard and the following factors must be taken into account.
One of the major aims is to be able to achieve the 100% principle stated in TR9007. This means that business rules which can be represented as constraints on data need to be represented declaratively rather than embedded in application programs. (It is noted that not all business rules can necessarily be stated in terms of constraints on data.) With the inclusion of the CHECK clause and ten types of predicate (for example comparison, like, match, overlaps, exists, unique), SQL92 has gone a long way towards allowing a user of its data modelling facility to achieve the 100% principle.
A requirement could be stated for a standard conceptual schema modelling facility that it should be able to represent the superset of all data constraints expressible in SQL 92 and all data constraints in EXPRESS. However, it is not clear that it is desirable to use scarce standards development resources to reinvent a new concrete syntax for this purpose. It is also not clear that existing graphic standards (such as EXPRESS-G) can be extended to enable representation of the full set of data constraints necessary to achieve the 100% principle.
The requirement for handling data concepts in the conceptual schema modelling facility standard is therefore stated as follows. The requirement must be to provide facilities for a user friendly graphical representation for sets of carefully selected "constructs" (such as a zero or one to many relationship

or an exclusivity constraint). The subsetting should be standardized in terms of difficulty levels (possibly three or four such levels) in order to place bounds on the dialogue which can take place between the conceptual schema modeller and a subject area expert.
It would be useful to provide a technique in the graphic representation form for designating the **existence** of a data constraint on data in the entity type rather than attempting to designate the possible multiplicity of other entity types which are referenced in formulating the constraint.

Handling process concepts

As indicated earlier, the conceptual schema modelling facility standard should provide facilities for modelling business activities (sometimes called business processes or business functions). It is also noted that a data modelling facility "may have an associated set of data manipulation rules". The term "manipulation" is intended to refer to basic processes performed on the data (create, retrieve, update, delete) rather than to any application process which makes use of these basic processes.
It is not a requirement for a standard conceptual schema modelling facility to be able to specify application processes which can be regarded as handled in terms of application programming.
It is a requirement for a conceptual schema modelling facility to be able to specify the users' view of the processes which the information system can perform. The proposed name for these processes is "user tasks. The concept of a user task is between a business activity and an application process. A user task must be meaningful to a user in the sense that a user may be able to select it from a menu containing several such tasks. Alternatively a user who makes use of the same user task repetitively may be able to press a function key or enter a short code identifying the task.
How a user task is initiated is clearly an implementation issue and not appropriate for a conceptual schema modelling facility standard. However, which user tasks may be initiated is part of a conceptual schema modelling facility when it is used in the prescriptive mode, but not when it is used in the analytic mode (see analytic models - earlier).

Access control

The requirement for handling access control in a conceptual schema modelling facility is to be able to develop an analytic model of the different user roles and to identify which user roles are authorized to perform which business activities. The business activities which need to be considered for this purpose are normally low level activities identified as a result of decomposing higher level activities.
When a conceptual schema modelling facility is used in a prescriptive mode, then the consideration of authorization of user roles would need to relate to the user tasks which may be initiated by a user at the computer human interface.
The conceptual schema modelling facility standard used in analytic mode must provide a means of analyzing the requirements for access control in the business area. These requirements would need to be expressed in terms of business activities on the lowest level in the decomposition hierarchy.
The conceptual schema modelling facility used in prescriptive mode must provide a means of prescribing the requirements for access control in the information systems to be designed for use in the business area. These requirements would need to be expressed in terms of the user tasks relevant to the information system to be designed for use in the business area.

Provide both analytic and prescriptive mode of use

A conceptual schema modelling facility is intended for use in an early stage of the information system life cycle. It may be used to prepare an analytic model of a business area (universe of discourse). It may also be used to prepare a prescriptive model of a business area. An analytic model of a business area may be used as a basis for specifying a prescriptive model of an information system for the business area. A prescriptive model of a business area may be used for reorganizing the business activities to be carried out in a business area.
Certain conceptual schema modelling facilities are already in use in this way. However, these appear to have no common basis and there are perceivable incompatibilities among them.

CONCLUSIONS

The following points developed in this paper can be formulated as conclusions.

1. In terms of approaches currently available, there appears to be no significant difference among the following - a data modelling facility, an information modelling facility and a conceptual schema modelling facility. The traditional term "data model" and its associated derivatives are preferred.

2. A data model (or a conceptual model) may have more than one representation form - diagram, language, narrative.

3. Since the term "conceptual schema" was introduced, it has come to mean different things to different people.

4. It is useful to recognize a clear distinction between a data modelling facility and a conceptual schema modelling facility, based on the overlapping and differing roles that each can fulfil in an information systems life cycle.

5. For a conceptual schema modelling facility to be useful in practice, it is required to be able to embrace a number of aspects of a subject area such as activities, temporal consideration and distribution, as well as data.

ACKNOWLEDGEMENT

The work on this paper has been motivated by the author's participation in the ISO SC21 Special Working Group on Modelling Facilities which was formed with a finite life span to determine (among other things) whether an international standard for a conceptual schema modelling facility would be beneficial. The group completed its work in June 1993 and drafted a New Project Proposal for such a standard.
The views expressed in this paper are the author's own and should not be construed as representing a consensus view of the Special Working Group.

REFERENCES

ANSI/X3/SPARC (1975) **Study Group on Data Base Management Systems:** Interim Report 75 02-08. ACM SIGMOD Newsletter, FDT, Vol 7 No 2

Barker R. (1990) **CASE*Method Entity Relationship Modelling** Addison-Wesley.

Date C.J. (1975) **Introduction to Database Systems** Addison Wesley

IFIP ICC (1966) **Vocabulary of Information Processing** North Holland

IS 10027 (1993) **Reference Model of Data Management**

IS 10303-11 (1993) **Product Data Representation and Exchange - Part 11: The EXPRESS Language Reference Model.**

IS TR 9007 (1977) **Concepts and Facilities for a Conceptual Schema Language**

IS 10032 (1992) **IRDS Services Interface**

IS 9075 (1I992) **Database Language SQL**

IS 9594 (also known as X500) **Directory Services**

Jardine D,.A. and van Griethuysen J.J. (1989) **Infomod Volume 1 Concepts and Information Structure** MMJ Information Engineering

Olle, T.W., Hagelstein, J., Macdonald, I.G., Rolland, C., Sol, H.G., Van Assche, F.J.M., Verrijn-Stuart, A.A. (1991) **Information Systems Methodologies - A Framework for Understanding** Addison- Wesley

Rochfeld A. and Morejon J. (1989) **La Methode Merise Tome 3 Gamme Operatoire** Les Editions d'Organisation.

SSADM Version 4 Reference Manuals. Published 1990 by National Computing Centre and Blackwells

Veryard R. (1992) **Information Modelling: Practical Guidance** Prentice Hall