TECHNICAL REPORT 9007

Published 1987-07-01

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION • МЕЖДУНАРОДНАЯ ОРГАНИЗАЦИЯ ПО СТАНДАРТИЗАЦИИ • ORGANISATION INTERNATIONALE DE NORMALISATION

Information processing systems — Concepts and terminology for the conceptual schema and the information base

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work.

The main task of ISO technical committees is to prepare International Standards. In exceptional circumstances a technical committee may propose the publication of a technical report of one of the following types :

— type 1, when the necessary support within the technical committee cannot be obtained for the publication of an International Standard, despite repeated efforts;

— type 2, when the subject is still under technical development requiring wider exposure;

— type 3, when a technical committee has collected data of a different kind from that which is normally published as an International Standard ("state of the art", for example).

Technical reports are accepted for publication directly by ISO Council. Technical reports types 1 and 2 are subject to review within three years of publication, to decide if they can be transformed into International Standards. Technical reports type 3 do not necessarily have to be reviewed until the data they provide is considered no longer valid or useful.

ISO/TR 9007 was prepared by Technical Committee ISO/TC 97, *Information processing systems*.

The reasons which led to the decision to publish this document in the form of a technical report type 3 are explained in the Preface.

UDC 681.3.02 **Ref. No. ISO/TR 9007 : 1987 (E)**

Descriptors: data processing, information interchange, factual data bases, basic concepts.

Printed in Switzerland

Price based on 120 pages

PREFACE.

0.1. PURPOSE OF THE REPORT.

It is expected that future data base management systems will include a component for handling conceptual schemata. This Report explains the roles and concepts for a conceptual schema with the intention of providing a framework for discussion and for the design of conceptual schema languages. The rules described in the conceptual schema control to a large extent what may or may not happen in an information system and a data base. Therefore this Report is not limiting its attention to the conceptual schema alone, but also considers basic concepts for the mechanisms involved in manipulating a conceptual schema and a data base.

This Report is aimed at designers of information systems and data bases as well as suppliers of conceptual schema facilities. The provided framework will prepare the way for eventual standardization in the area of data base management. It does not, however, describe any particular method for using such facilities. In the meantime, the general principles in this Report can be used to evaluate emerging DBMS facilities.

The approaches and associated languages described in appendices to the Report are intended to be explanatory only and are not ipso facto candidates for a standard conceptual schema language.

0.2. STRUCTURE OF THE REPORT.

The main body of the Report (chapters one through four) contains the fundamental concepts and terminology for the conceptual schema, the information base, and the mechanisms involved in manipulating them.

Chapter one gives an introduction to the subject, mentions the origins of some ideas developed in the Report, and discusses some major topics. In particular, it explains what a conceptual schema is used for, its roles, and requirements for a conceptual schema facility.

Chapter two explains fundamental concepts, provides definitions of the concepts and terms, and develops some of the consequences of those concepts and definitions. Both static and dynamic aspects of the information system are considered and explained. Some readers may wish to skip this chapter on the first reading.

Chapter three discusses some aspects of implementation. In particular, principles are formulated for the contents and scope of a conceptual schema, and an information system architecture based on three levels is outlined.

Chapter four reviews some approaches to information modelling and manipulation for data bases. The approaches selected for illustration are outlined in more detail in appendices to the Report.

Several appendices have been added to the Report as follows:

Appendix A gives a glossary of the terms and definitions.

Appendix B provides an example situation to be described in information modelling approaches.

Appendix C gives a syntax notation to be used for defining grammars of example conceptual schema languages.

Appendix D outlines Entity-Attribute-Relationship approaches.

Appendix E demonstrates Binary and Elementary N-ary Relationship approaches.

Appendix F discusses Interpreted Predicate Logic approaches.

Appendix G elaborates on expressing dynamic rules and constraints in conceptual schemata.

Appendix H presents some thought on interacting with information systems and examples of permissible action descriptions.

0.3. STATUS OF THE REPORT.

This Report is an ISO Technical Report of type 3. It is the Working Group's first response to item 1 of its Program of Work. As such it is a statement of the Working Group's current view on concepts for conceptual schemata and information bases. Considering the rapid development in data base technology and applications possible, also taking into account the requirements for distributed data base systems and related data communication facilities, periodic revisions of the Report are to be expected.

0.4. REFERENCES.

[1] MURRAY, J.A.H. et al.
'The Oxford English Dictionary' with supplements,
Clarendon Press, 1933 - 1977.

TABLE OF CONTENTS.

Chapter 1. INTRODUCTION TO THE CONCEPTUAL SCHEMA AND THE INFORMATION BASE. 9
1.1. The ANSI/SPARC framework. 9
1.2. The universe of discourse. 10
1.3. Describing the universe of discourse. 11
1.4. Static and dynamic aspects of a conceptual schema and information base. 13
1.5. Interaction between the real world and an information system. 13
1.6. The roles of users and information processors. 14
1.7. Guidelines for the description of a universe of discourse. 15
1.8. Guidelines for the contents of a conceptual schema. 16
1.9. Roles for a conceptual schema. 17
1.10. Requirements for a conceptual schema facility. 18
1.11. References. 19

Chapter 2. FUNDAMENTALS FOR A CONCEPTUAL SCHEMA AND AN INFORMATION BASE. 21
2.1. General concepts and definitions. 21
2.2. Basic concepts and definitions for actions on the conceptual schema and information base. 28
2.3. The behaviour of an information processor. 31
2.4. Inserting a conceptual schema - the minimal conceptual schema. 33
2.5. Behaviour rules for the environment. 34
2.6. Static and dynamic rules and constraints. 34
2.7. Expressing rules and constraints. 36
2.8. Co-ordination of permissible actions. 38
2.9. References. 44

Chapter 3. SOME CONCEPTS AND PRINCIPLES FOR IMPLEMENTATION. 45
3.1. Principles for the contents and scope of a conceptual schema. 45
3.2. Principles for the description of a universe of discourse. 47
3.3. Abstract syntax for a conceptual schema and information base. 50
3.4. Semantics of a conceptual schema and information base. 51
3.5. Principles for the composition of conceptual schemata. 51
3.6. The Three Level Architecture. 53
3.7. Information Resource Dictionary Systems (IRDS) Model. 58
3.8. The conceptual schema in the context of current DBMS implementation. 59
3.9. Correspondence of the Three Level Architecture for information systems and the Reference Model for Open Systems Interconnection. 60
3.10. References. 61

Chapter 4. OVERVIEW OF SOME MODELLING APPROACHES. 63
4.1. Introduction. 63
4.2. Review of some approaches. 64
4.2.1. Entity attribute relationship approaches. 66
4.2.2. Binary and elementary n-ary relationship approaches. 66
4.2.3. Interpreted predicate logic approaches. 67
4.3. Translation of approaches to current data base technology. 68
4.4. References. 70

APPENDICES.

Appendix A. GLOSSARY OF TERMINOLOGY AND DEFINITIONS. 71

Appendix B. EXAMPLE UNIVERSE OF DISCOURSE. 77
B.1. Introduction. 77
B.2. Rules, etc. for the universe of discourse. 77
B.3. Some things and happenings in the relevant entity world. 79

Appendix C. THE PASCAL SYNTAX NOTATION. 81

Appendix D. THE ENTITY - ATTRIBUTE - RELATIONSHIP APPROACHES. 83
D.1. Emphasis of the approaches. 83
D.2. Primitive concepts of the approaches. 84
D.2.1. The basic concepts. 84
D.2.2. Abstraction concepts. 84
D.2.3. Characteristics of relationships. 86
D.3. Grammar and semantics. 91
D.4. Graphic formalism. 93
D.5. Modelling. 94
D.5.1. Some pragmatic modelling rules. 94
D.5.2. Formal rules for modelling. 95
D.6. Example conceptual schema. 96
D.6.1. Graphic representation. 96
D.6.2. Language example. 97
D.7. Check list for the conceptual schema. 99
D.8. Mapping of an EAR conceptual schema to a network data base schema and a relational data base schema. 102
D.9. References. 104

Appendix E. THE BINARY RELATIONSHIP APPROACHES. 105
E.1. Emphasis of the approaches. 105
E.2. Primitive concepts of the approaches. 107
E.3. Grammar and semantics. 111
E.3.1. The language and its relation to the universe of discourse. 111
E.3.2. Formal syntax. 113
E.3.3. Semantics. 115
E.4. Graphic formalism. 117
E.4.1. Linguistic object types. 117
E.4.2. Binary relationship types. 118
E.4.3. Constraints having a diagrammatic representation. 118
E.4.4. Some examples of the graphic formalism symbols. 119
E.5. Modelling. 121
E.6. Example conceptual schema. 122
E.6.1. Graphic representation. 122
E.6.2. Language example. 123
E.7. Check list for the conceptual schema. 131
E.8. References. 134

Appendix F. THE INTERPRETED PREDICATE LOGIC APPROACHES. 135
F.1. Emphasis of the approaches. 135
F.2. Primitive concepts of the approaches. 136
F.3. Grammar and semantics. 137
F.3.1. Abstract syntax. 137
F.3.2. Concrete syntax. 139
F.3.3. Semantics. 142
F.4. Graphic formalism. 142
F.5. Modelling. 143
F.5.1. Classification of axioms. 143
F.5.2. Constructs. 146
F.6. Example conceptual schema. 154
F.6.1. Graphic representation. 154
F.6.2. Language example. 155
F.7. Check list for the conceptual schema. 166
F.8. References. 169

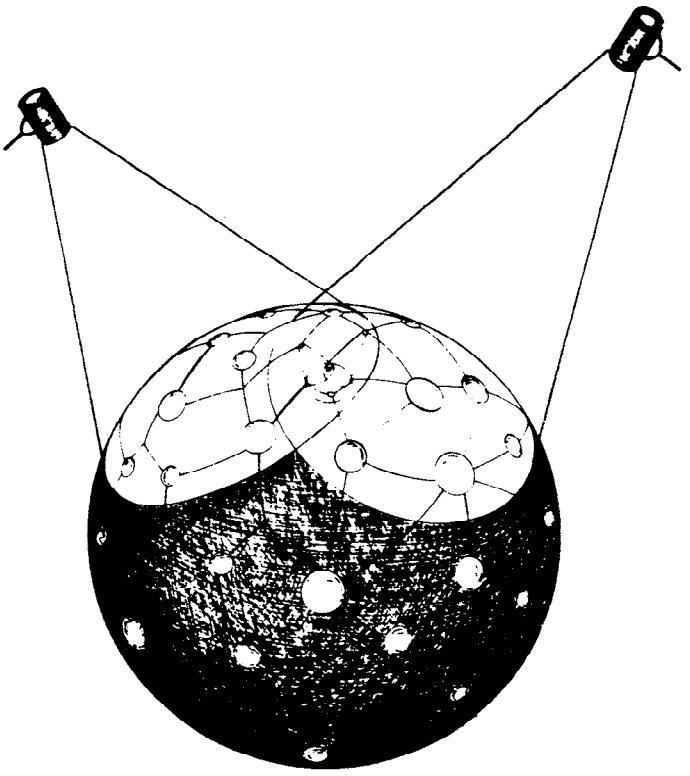
Appendix G. EXAMPLES OF DYNAMIC RULE DESCRIPTION. 171
G.1. Introduction. 171
G.2. State-oriented descriptions. 171
G.3. State-oriented description of rules in the example conceptual schema. 171
G.4. State independent rules in action-oriented descriptions. 172
G.5. State dependent rules in action-oriented descriptions. 173
G.6. Action-oriented description of rules in the example conceptual schema. 173

Appendix H. EXAMPLES OF CO-ORDINATING PERMISSIBLE ACTIONS. 177
H.1. Interaction between environment and information system. 177
H.2. Some implementation considerations for permissible actions. 178
H.3. Describing permissible actions in the example conceptual schema. 179
H.4. References. 184

THE HELSINKI PRINCIPLE.

These utterances are to be interpreted (recursively) as international English utterances [1]:

> Any meaningful exchange of utterances depends upon the prior existence of an agreed set of semantic and syntactic rules. The recipients of the utterances must use only these rules to interpret the received utterances, if it is to mean the same as that which was meant by the utterer.
>
> (ISO TC97/SC5/WG3 - Helsinki 1978)



"THE METAPHOR OF THE SEARCHLIGHTS" on universes of discourse.

This page intentionally left blank

CHAPTER 1. INTRODUCTION TO THE CONCEPTUAL SCHEMA AND THE INFORMATION BASE.

1.1. THE ANSI/SPARC FRAMEWORK.

The reports of the ANSI/X3/SPARC DBSG [1, 2] identified the need for a conceptual schema in the context of a three-schema framework for data base management systems.

Subsequent papers [3, 4, 5, 6, 7] have emphasized the importance of a conceptual schema to users and designers of data base systems. In this context, a conceptual schema comprises a unique central description of the various information contents that may be in a data base. This includes the description of what actions, such as changes and retrievals, are permissible on the information content. The data base itself may be implemented in any one of a number of possible ways. Users and application programs may view the data in a variety of ways, each described by an external schema. Each external schema is therefore derived from the common conceptual schema. The physical storage structure that may be in use at any given time is described by an internal schema that is also derived from the conceptual schema.

The conceptual view, as meant by ANSI/SPARC, concentrates on the meaning of the information. It is the conceptual schema that describes this view. The external views concentrate on the forms - the data - that represent the information to the outside. These are described in the external schemata. The internal view concentrates on the internal physical representation of the data inside the computer system and is described in the internal schema.

Such a three-schema framework is widely, but not yet universally, accepted. It is assumed in this report. Furthermore, it may be noted that the conceptual schema concept is valuable in other environments than a three-schema framework.

It is widely acknowledged that the conceptual schema also plays a key role in systems analysis and data base design. One may therefore ask whether it should be biased to one or the other. Should the conceptual schema be primarily an enterprise model, resulting from the systems analysis, or should it serve as a focal point between user views and the physical data base design? We believe that it should play both roles in the next generation of DBMS.

We believe the data base user will benefit from the clear separation of the information meaning from the external data representation and the internal physical data storage layout. A clear methodology for producing a conceptual schema would help the implementor of an information system to improve his systems analysis, even if a manual step of translating it into data base design in terms of an existing DBMS were then required.

The ANSI reports introduced the conceptual schema in broad terms. Besides, the term "conceptual schema" is sometimes used for data base aspects which are not at all conceptual. Therefore, elaboration of the conceptual schema's objectives, roles, form, and content is needed. What a conceptual schema must include, which appropriate modelling concepts are to be used in it, and the exact role it plays in data bases, are the major subjects of this Report.

1.2. THE UNIVERSE OF DISCOURSE.

In the past, data processing systems were often designed so as to provide all users with the same set of capabilities or functions. However, this uniform functional view is not adequate to construct today's data base systems. A single data base may support quite different functional requirements concurrently, or at different times, during its existence.

The prime characteristic of the data base environment is that common data is shared between many users of a single system. By sharing common data, these users establish a dialogue with each other through the system. Clearly, if this communication is to be useful and reliable there must be some common understanding of the information represented by the data. Since it may happen that two users never meet, this common understanding must refer to something external to both of them. This common understanding must be recorded and in order to establish a dialogue a common predefined established grammar is needed.

We will call those things and happenings to which the common understanding of the represented information refers the universe of discourse. Universes of discourse may be concrete like an inventory, or abstract like the organizational structure of an enterprise. They even may be hypothetical like Wonderland which was visited by Alice.

In this Report we will take an (informal) naive realism approach to universes of discourse.

The typical universe of discourse is perceived as containing real and abstract objects, which we will call entities. It can be perceived as also containing classes of entities, e.g. persons, departments, and dates. This classification is based on similarity and takes into account characteristics common to several entities. The selection of characteristics for grouping the entities into classes is arbitrary; the choices will be made pragmatically, based on the purpose of the universe of discourse.

Some general properties to which entities adhere, that classify entities, that associate entities, etc., in the universe of discourse are also perceived (e.g. persons are not departments, a person may be assigned to no more than one department). These may be informally described as "classifications", "rules", "laws" or "constraints" about the state of affairs and behaviour of entities in the universe of discourse.

In general, what is considered to be part of the universe of discourse will be time-dependent, that is, the selected things and happenings may change with time. This will be equally true for the classifications, rules, laws, etc; however, it is likely that the rate of change of these will be relatively slow compared with that of the former.

1.3. DESCRIBING THE UNIVERSE OF DISCOURSE.

There are in fact two systems of interest: the universe of discourse and a data processing system which contains a linguistic representation of that universe of discourse. Following common usage we say that information about the universe of discourse "describes" or "models" that universe of discourse. We want, however, to emphasize that the description process may be in fact a very complex task calling for creative analysis and iterative refinement.

Without prejudging its physical representation we consider that the information contained in the data processing system describes the universe of discourse. A concrete physical representation of this information will be called a data base. We will use the term data base system for a data processing system dealing with a data base. It is possible for the data base system itself to be one of the subjects being described, in which case the data base system would be included in the universe of discourse. However, to simplify the discussion, we will generally assume that the data base system is disjoint from the universe of discourse, although this is not necessarily the case.

It is the classifications, rules, etc., that are of primary interest to a systems designer designing a data base system. In analysing the universe of discourse, it is these things he will want to identify, discuss with users and describe. In recording them he will actually create a "skeleton" description of the universe of discourse, the conceptual schema. In this way the conceptual schema describes which entities can possibly exist in the universe of discourse, that is, which entities exist, have existed, or might ever exist. In the same sense it describes what facts and happenings are possible for those entities or, if relevant, are required for them. We assume it will be held in a formal representation within the data base system.

We also want to record all other relevant information which describes the entities that are considered to be of interest and their actual state of affairs at a specified instant or period of time (usually "now"). We call this further information the information base.

Although each description necessarily will have a representation form to make the description communicable, it is the interpretation of this representation (the meaning of the description) which interests us in the first place. The representation form, although not irrelevant, is considered to be of secondary importance. We will use the term "information" when we want to emphasize our interest in this interpretation. We will use the term "data" when we want to concentrate on the representation forms of the information.

At this point it may be useful to consider the information describing a universe of discourse within the context of an ANSI/SPARC framework: We consider both the conceptual schema and the information base to be at the conceptual level, providing a conceptual view of the information about the universe of discourse.

The data base or parts thereof as seen by a user of the system (the strings of data) we consider to be at the external level giving an external view on the information about the universe of discourse. The internal storage forms within a computer we consider to be at the internal level being the internal view of the information about the universe of discourse. For the external and internal views the representation forms are of primary interest. The interpretation of those forms is, of course, the interpretation meant in the conceptual view.

Summarizing we have now identified:

- o Universe of discourse:
  The collection of all objects (entities) that ever have been, are, or ever will be in a selected portion of a real world or postulated world of interest that is being described.

- o Conceptual schema:
  The description of the possible states of affairs of the universe of discourse including the classifications, rules, laws, etc., of the universe of discourse.

- o Information base:
  The description of the specific objects (entities) that in a specific instant or period in time are perceived to exist in the universe of discourse and their actual states of affairs that are of interest.

Precise definitions for the above concepts will be given in chapter 2.

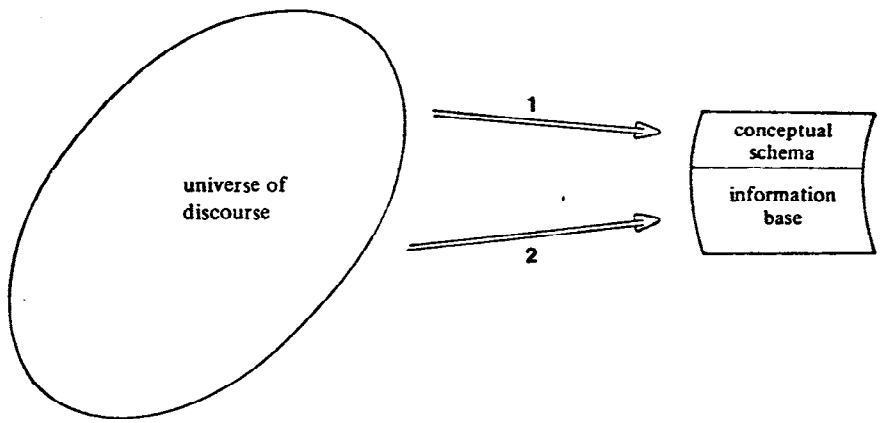


Figure 1.1. Describing the universe of discourse.

The description process is illustrated in figure 1.1; the two numbered processes are:

1: Classification, abstraction, generalization, establishing rules, etc, about the universe of discourse and recording them. This is a human process, describing a (shared) mental model of the universe of discourse.

2: Recording facts and happenings about the universe of discourse including what entities actually are of interest.

The conceptual schema describes the general rules, etc, of the universe of discourse, which, to a great extent, may govern its behaviour. These rules described in the conceptual schema therefore also control what may or may not occur in the information base. It is for this reason that we do not limit our attention to the conceptual schema and information base; we will also consider the mechanisms involved in manipulating the contents of the information base and the conceptual schema.

1.4. STATIC AND DYNAMIC ASPECTS OF A CONCEPTUAL SCHEMA AND INFORMATION BASE.

Much of the past work on concepts for the conceptual schema has been concentrated on the static aspects, that is, on defining the concepts to be used to describe valid states of a conceptual schema and information base.

However, the set of concepts for the conceptual schema should also cover the dynamic aspects. Firstly, the conceptual schema may change to correctly reflect changes in the selected portion of a real or postulated world. Secondly, dynamic aspects are involved in describing those manipulations which are needed to make known part or all of the conceptual schema and information base.

In some cases, the time scales of changes within the universe of discourse and the corresponding changes in the conceptual schema and information base need not be tied closely together: changes in the universe of discourse may be recorded in the conceptual schema and information base retrospectively and even in a different sequence. In other cases, the time scales are so closely related to each other that the conceptual schema and information base necessarily become part of the universe of discourse; especially in these cases the description of this interaction must also be part of the dynamic aspects.

No clear boundary has been defined between static and dynamic aspects, and the boundary may well be found to vary between different approaches or even to be non-existent in some cases. Some of the ideas introduced on this subject in the present Report have not yet been the subject of wide debate, but may at least serve to indicate areas deserving further study. In particular it is not clear whether different sets of concepts should be used to describe static aspects and dynamic aspects, or whether, at least for some approaches, the same set of concepts may fulfil both purposes.

1.5. INTERACTION BETWEEN THE REAL WORLD AND AN INFORMATION SYSTEM.

A conceptual schema and information base is totally static unless something operates on it to cause change. That something we will call an information processor. We will define an information system as consisting of a conceptual schema, an information base, and an information processor.

An information processor operates to produce change in the information base or conceptual schema only on receipt of a message. A message contains information and/or expresses commands. Messages originate from a part of the real world referred to as the environment, which may be disjoint from, or overlap with, the universe of discourse. On receipt of an appropriate message containing a command an information processor may also operate to make known, by means of a message, information present in the conceptual schema and information base. For further details see chapter 2.

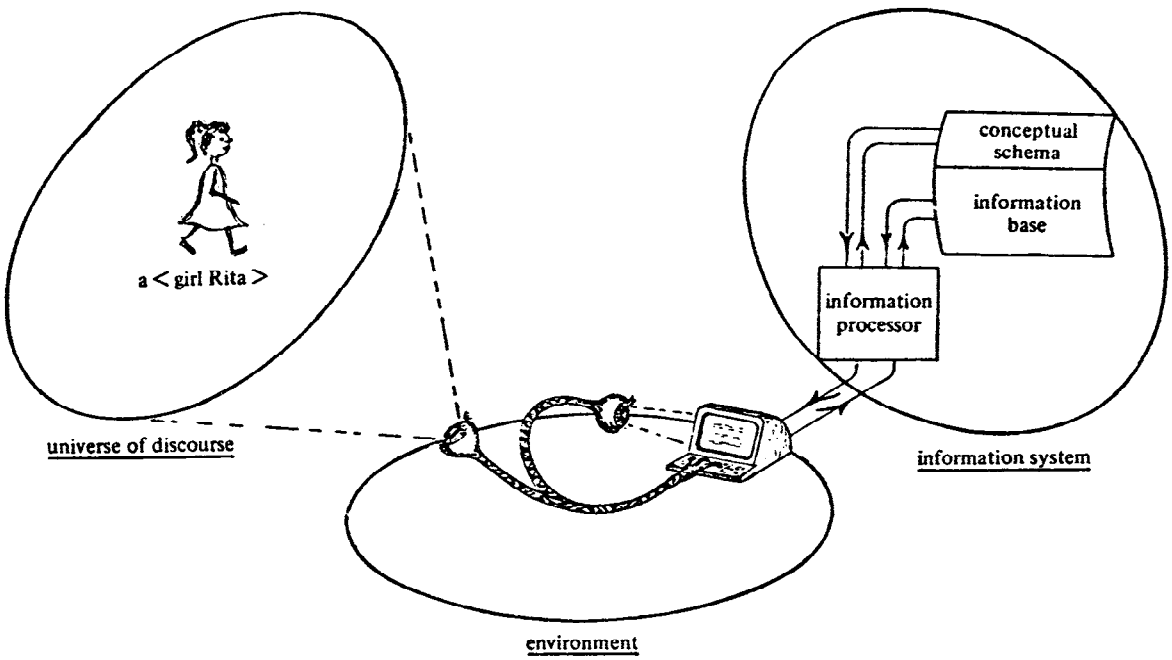


Figure 1.2. Information system and environment.

The information system is distinguished from the environment in the following way:

- o The information system is a formal system, the environment as a whole is not so.
- o The behaviour of the information system is completely defined by behaviour rules and constraints which are established, directly or indirectly, by the environment. The information system on its own initiative never establishes rules for the environment.
- o An information system, being fully predictable, is unable to deviate from the rules or constraints. The environment can deviate from its rules.

Although we may consider the information system together with the environment to be parts of an encompassing system, this latter system may not be formal or fully predictable. Therefore we use the term information system as above, excluding the environment - the users of an information system.

1.6. THE ROLE OF USERS AND INFORMATION PROCESSORS.

The users of an information system can be machines or other systems as well as human beings. A user is anybody or anything that issues commands and messages to the information system and receives messages from the information system. As such they are part of the environment. Some users also have the authority to establish behaviour rules or constraints for the information system.

An information processor transfers messages between the environment and the information base or conceptual schema, as explained above. In doing this it has no initiative of its own; it can only behave exactly as specified by the rules, the whole rules, and nothing but the rules.

Normally an information processor will be a computer system or some parts thereof, but human beings can also play the role of an information processor, provided they do not deviate from prescribed rules or act on their own initiative. Computer systems, on the other hand, can act as users of an information system. An example is a network of information systems communicating with each other. If each has a set of rules which is independent of the others, then each plays the role of user of the other information systems. We therefore conclude, that the role determines whether something must be regarded as a user or as an information processor.

The above formulation of users and information processors in terms of roles implies that the environment and the information system need not necessarily be disjoint. Similarly, if the information base contains information about the users of the information system, the environment and the universe of discourse will not be disjoint. However, even if they are not disjoint they will always be distinguishable from each other.

1.7. GUIDELINES FOR THE DESCRIPTION OF A UNIVERSE OF DISCOURSE.

Sometimes in the literature on various modelling methods for information systems and conceptual schemata, no clear distinction is made between the things and the description of the things, nor between the information meaning and the data representation.

This stems partly from the origins of some approaches which are in effect the data modelling techniques of the early seventies. Partly, however, the reason for this is a debate on fundamentals, which is still going on. The difference is found in whether the conceptual schema must be defined in terms of entities in the universe of discourse itself and states of affairs about them, or in terms of descriptive constructs found in the information base describing the entities of interest in the universe of discourse. Either view is possible and can be presented systematically.

It is most important to note though that the two alternative views above cannot be indiscriminately mixed in the same discussion without leading to confusion, paradox, and error. It is unfortunate but true that many variants of modelling approaches, both in practice and as described in the literature, suffer from precisely these problems.

The current work of WG3 is based on the assumption that the conceptual schema and the information base should describe the conceptual view. This implies that the conceptual schema is defined in terms and constructs referring to things in the universe of discourse itself and expressing states of affairs about those things.

The constructs used in a conceptual schema and information base should be based on the fundamentals of formal logic as theoretical foundation. However, how closely they must follow the spirit and notation forms of these fundamentals is a subject of further investigation. It is quite possible to limit them to elementary constructs expressing these fundamentals. However, it will be

always possible to define, upon these fundamentals, a variety of more complex constructs ("macro constructs"), that may be more convenient or efficient for describing various aspects of a universe of discourse.

The choice of specific macro constructs is based on practical arguments such as ease of understanding and use. That choice is considered to be dependent on the application area of the information system for which a conceptual schema and information base has to be provided.

As already stated, it is important to distinguish carefully between the entities and their descriptions. In these descriptions entities are usually identified by names that refer to those entities. This includes synonyms - different names refering to the same entity - and homonyms - identical names refering to different entities. The relevance of this distinction, not only for information systems in particular, but for human communication in general, has been well-known in language philosophy and linguistics for a long time. Therefore the constructs should provide for synonyms and possibly cope with homonyms.

The conceptual schema not only describes the static aspects and dependencies of the universe of discourse, but also the dynamic aspects. This determines what manipulations of the descriptions are allowable as well as what descriptions may be present in the conceptual schema and information base. Therefore it may be clear that constructs have to be available both for the descriptions and for their manipulation in the information system.

The subject is elaborated further in chapter 3.

1.8. GUIDELINES FOR THE CONTENTS OF A CONCEPTUAL SCHEMA.

Since the selection of what are considered to be general classifications, rules, etc. of the universe of discourse is to a certain extent arbitrary, it follows that the choice of which should be described in the conceptual schema and which in the information base is arbitrary to a similar extent. In practice, however, the systems designer might consider various factors in deciding the boundary of the conceptual schema. These might include:

- describing classes (types, variables) in the universe of discourse rather than individuals (instances),
- describing concepts that are less subject to change rather than concepts that are changing more frequently,
- inclusion of rules or constraints having wide influence on the behaviour of the universe of discourse (and therefore on the behaviour of the conceptual schema and information base) rather than narrow influence.

At all times the following general principles for the conceptual schema should be observed:

* 100 Percent principle:
  All relevant general static and dynamic aspects, i.e. all rules, laws, etc., of the universe of discourse should be described in the conceptual schema. The information system

cannot be held responsible for not meeting those described elsewhere, including in particular those in application programs.

* Conceptualization principle:
A conceptual schema should only include conceptually relevant aspects, both static and dynamic, of the universe of discourse, thus excluding all aspects of (external or internal) data representation, physical data organization and access as well as all aspects of particular external user representation such as message formats, data structures, etc.

A more detailed discussion may be found in chapter 3.

1.9. ROLES FOR A CONCEPTUAL SCHEMA.

A fundamental impact of a conceptual schema is that the concepts used harmonize - and to a certain level make possible - human communication. Moreover, these concepts will influence the methods and results of analysing organizations and their information needs. In a way, a conceptual schema constitutes a general agreement concerning how to perceive a universe of discourse. This agreement may alter over time, but supports the evolution of applications over their life cycles as well as changes of this agreement itself (cf. The Helsinki Principle).

A conceptual schema is intended to properly describe the behaviour of a universe of discourse. Therefore the rules given therein, naturally, restrict possible evolutions and manipulations of the description of the universe of discourse, i.e. of the conceptual schema itself as well as of the information base.

Mainly for reasons of human convenience and efficiency, different users within the environment of a common information system will use different forms of external data representing the information. At the same time, for reasons of machine and storage handling efficiency, internal data organizations will be designed and used that may or may not differ from those external forms. In this context, the conceptual schema enforces preservation of meaning in transformations between the various data representations and defines the interpretation of these representations.

Therefore, and considering what is outlined in the previous sections, the following fundamental roles for a conceptual schema have been identified:

1. To provide a common basis for understanding the general behaviour of the universe of discourse;

2. To define the allowed evolution and manipulation of the information about the universe of discourse;

3. To provide a basis for interpretation of external and internal syntactical forms which represent the information about the universe of discourse;

4. To provide a basis of mappings between and among external and internal schemata.

1.10. REQUIREMENTS FOR A CONCEPTUAL SCHEMA FACILITY.

It is anticipated that the data base management systems developed in the future will include a component for handling a conceptual schema definition which will fulfil the roles mentioned in the previous section. In the course of time, provision of such a component should become a standard requirement.

To fulfil the roles indicated above, a conceptual schema facility must satisfy the following requirements:

1. It must provide basic concepts which are suitable for adequately describing both the static and dynamic aspects of a universe of discourse and ipso facto its description in terms of a conceptual schema and information base.

2. It must provide a language in which a conceptual schema can be expressed so as to be readily understandable to a user of the facility.

3. It must provide a language for precisely communicating a conceptual schema to a computer.

4. It should provide for easily modifying the conceptual schema to reflect changes in the general classifications, rules, laws, etc. of the universe of discourse, and for predicting the direct consequences of such changes.

5. The views of the information that different users wish to see are limited to those which do not contradict the assertions in the conceptual schema. If such external schemata are subject to change, the facility should be such that this should not affect the conceptual schema.

6. The conceptual schema should be kept invariant by the facility with respect to changes in the internal (physical) representation of the data within a computer.

The two languages mentioned in 2 and 3 may be the same but are not necessarily so. For the former purpose an additional graphic notation may be helpful.

These six requirements as a minimum must be met by any candidate conceptual schema facility.

1.11. REFERENCES.

[1] ANSI/X3/SPARC, 'Study Group on Data Base Management Systems: Interim Report 75-02-08', In: ACM SIGMOD Newsletter, FDT, Vol 7, No. 2, 1975.

[2] TSICHRITZIS, D., and KLUG, A. (eds.) 'The ANSI/X3/SPARC DBMS Framework. Report of Study Group on Data Base Management Systems', AFIPS Press, Montvale NJ, 1977.

[3] BSI/DPS13/WG1, 'Role and Composition of a Conceptual Schema in an Overall DBMS Architecture', 1976.

[4] AFNOR/Z6/SC5, 'Study of Users Requirements', 1977.

[5] NIJSSEN, G.M. 'The Next Five Years in Data Base Technology', In: Infotech State of the Art Conference, London, 1977.

[6] NIJSSEN, G.M. 'A framework for advanced mass storage applications', In: Medinfo 80, Proceedings of the Third World Conference on Medical Informatics, Tokyo 1980, North-Holland Publishing Company, 1980.

[7] FALKENBERG, E. 'On the conceptual approach to data bases' In: International Conference on Data Bases, Aberdeen 1980, Heyden and Son, London, 1980.