FIDEL, Raya: **Database Design for Information Retrieval: a Conceptual Approach.** New York: Wiley 1987. xv, 232p. ISBN 0-471-82786-X

Databases have become part and parcel of information business in recent years; moreover, they have almost become a common feature of privately based computing. Very much like system software, database management systems can be purchased off the shelf at tremendously attractive prices. Yet the brave new world of decreasing prices and rapidly improving technical performance appears to be but one aspect of database construction, leaving important questions to be answered by individual database designers themselves: what about the contents of the database to be designed; indeed, what about the concepts?

While a mass of literature has been devoted to technical matters of database design, Dr Fidel's book focuses explicitly on the conceptual approach which may easily be underrated if primary attention is drawn to technicalities. Her approach is to be understood as a general method applicable to the design of databases of any kind. Although distinguished internal, external, as well as conceptual levels of database design, Fidel only mentions the first and second in passing, if at all. Matters of entry format, indexes information display and so on are taken to be basically independent from conceptual ones. Fidel's objectives, then, are different. She sets out to demonstrate, in minute detail, the very first step among four major ones of database design which is crucial to every kind of conceptual approach: data requirements analysis. Subsequent steps, namely hardware and software considerations, coding (actual implementation of the design), and testing of performance are only mentioned in introducing the book.

The book, intended for novices and advanced readers alike, is written in a clear and jargon free style. It is well illustrated, referencing both to graphic displays as well as to the basic example which is repetitively drawn upon to demonstrate the various steps of an intricate process called conceptual design in a non-structural information environment. Most helpful, a glossary and an index are provided. Reading is not impaired by a mass of references. The example mentioned, the design of a database on restaurants, may be felt, at times, to divert some readers' interests from abstract conceptual considerations to rather more fundamental attractions. Yet before readers turn to recipes and conventionally published restaurant guides, Fidel manages to direct the attention back to more information related business.

Following the introduction, the process of database design is displayed in four stages, dealing with "study of the problem", "representing the data in formal terms", "selection of rules", and "evaluation" respectively. It is most welcome that from the very beginning the information needs of potential database users are taken into account. Elaborating her very telling example, Fidel distinguishes, first of all, two basic concepts of database design, namely "enterprise" (restaurants) and "environments" (possibly varying from administrative to culinary ones). As a consequence, a strong plea is made to start user-oriented database design by finding out about users' attitudes and likely information needs.

The recommended method is rooted in a thorough problem study: survey of environments and interviewing potential database users (considering, of course, varying environments), followed by a complete documentation of their statements. Although these statements include information on all levels of design, conceptual ones are most important for the present study. The interview data are then subjected to linguistic analysis (elimination of redundancy, ambiguity) and to interpretation of information provided only implicitly. If successful, all this will result in a data dictionary, listing precise definitions of all data required to describe the enterprise of the database, accompanied by an operations dictionary, comprising information on actions involved with these data. At this stage, environments are still to be kept separately.

In order to facilitate the transformation of that information into a database, a formal representation of the elements in the data dictionary is required. It has to display all entity types contained in the data dictionary, the relationships among them as well as the attributes found out necessary to describe both entities as well as relationships. Due to a surprisingly straightforward method, generating a clear mini-diagram from each data dictionary entry, complex structures can be built up step-by-step and vividly displayed in an "entity-relationship diagram" (as outlined by P.P.-S. Shen, "The Entity-Relationship Model: Towards a Unified View of Data", *ACM Transactions on Database Systems*, 1(1976)No.1, p.9-36). In its final version, such a diagram will include all environments specific to that database. As Fidel explains, integration of schemata will involve the resolution of conflicts arising from inconsistent names or incompatible or missing diagram structures. Modifications of formal structures must be met by correspondent alterations of the data dictionary.

Before embarking on data collection, however, some stringent rules have to be dealt with. These pertain to the mandators or optional character of relationships and attributes, their degrees of cardinality (e.g. one-to-one or one-to-many), specifications of attributes, the selection of authorized sources, dealing with fuzzy or borderline cases etc. Further to the rules, it is necessary to control the rigorous design and the quality of the database concept as far as that can be achieved before the database is actually in operation. The list of control features is long and quite likely to point to a somewhat cyclical character of database design, for all weaknesses outlined in that evaluation process prior to database implementation will unavoidably refer the designer back to the diagram and probably to the data dictionary to amend the conceptual scheme itself.   Heiner Schnelling

Dr. H. Schnelling, Universitätsbibliothek
PF 5560, D-7750 Konstanz 1

THARP, A.L.: **File Organization and Processing.** New York: John Wiley 1988. ISBN 0-471-60521-2

The book by Alan L. Tharp covers a very interesting subject which – for some – may not be expected under this title. An alternate title might be "All about Signatures, Bloom Filters, Hashing and other magic representations of information you wanted to know but were afraid to