This paper provides an integrated approach for research related to the problem of database design. The process of database design is classified into five phases: requirements analysis, view modeling, view integration, view restructuring, and schema analysis and mapping. The input, processing steps, and output for each phase are described. The problems associated with each phase are pointed out. Existing approaches to database design are reviewed and related to these five phases. The significance of this integrated approach for the development of computer-aided methodologies for database design is discussed.

1. INTRODUCTION

At the current state-of-the-art, the methods used in the design of database applications are essentially trial-and-error, supported by neither a scientific foundation nor an engineering discipline. The ad hoc approach to design frequently leads to inflexible solutions that do not meet the prescribed requirements. Costly remedial measures often produce more delay in operation without a tangible improvement. Much of the existing information on system design is presented in the form of individual analyses. These analyses do provide valuable insight, but they can hardly be adequate substitutes for a systematic design discipline.

It is generally accepted that there are two levels in the design of a database a) the logical design, defining and combining the views of many applications into a centrally controlled and maintained logical databases, with provisions for data sharing and security; and b)
the physical design, including all the implementational details and considerations of a particular database system.

In this paper, we will address mainly the design issues which apply at the logical design level. Since realistic databases involve thousands of data elements, and the evaluation of an enormous amount of structured information is implied (e.g. see Raver and Hubbard [A5]), it is desirable to develop computer aided tools to aid the design. In what follows, a conceptual framework is presented within which current research in logical database design is reviewed. It is also suggested as to how these seemingly unrelated approaches may be integrated into a computer aided design system.

2. THE DATABASE DESIGN PROCESS

The problem of database design is rich; it ranges from system-independent analysis to system-dependent optimization. Existing research tends to concentrate on only a few aspects of the design process. Consequently, each approach has its own view of the design process. It is desirable to define a general design process in order to compare and integrate existing approaches.

The process of database design can be divided into five general steps:

1) Requirement Analysis. The problem or environment in the real world must be analyzed to make the necessary components of the database explicit and to elicit both the data and processing needs of all potential database users.

2) View Modeling. Using the results of step 1 as input, abstract representations must be developed that correspond with each user's view of the real world. This step both verifies the previous step and lays a basis for the next.

3) View Integration. The several (and perhaps conflicting) user views must be integrated into one global or community view of the database. This global view must continue to support all user views.