5-6-1985

Fuzzy data and logical data base design

Karen Steelman

FUZZY DATA

AND

LOGICAL DATA BASE DESIGN

Rochester Institute of Technology
School of Computer Science and Technology

Fuzzy Data and Logical Data Base Design

A thesis submitted in partial fulfillment of
Master of Science in Computer Science Degree Program

By: Karen Steelman

Approved By:

Chris Comte

Professor Chris Comte, Advisor

Jeffrey Lasky

Professor Jeffrey Lasky

Rayno Niemi

Professor Rayno Niemi

Date: May 6, 1985

Title of Thesis: Fuzzy Data and Logical Data Base Design

I, Karen Steelman, hereby deny permission to the Wallace Memorial Library, of Rochester Institute of Technology, to reproduce my thesis in whole or in part.

Date: May 24, 1985 Signature: Karen Steelman

Karen Steelman

TABLE OF CONTENTS

CHAPTER ONE

INTRODUCTION

The goal of this thesis is to gain understanding of the conceptualization phase of the data base design process for fuzzy data and to determine characteristics of the data base design process which are particularly appropriate for deriving the conceptual model for fuzzy data. Several potentially useful design methodologies will be evaluated. Genealogical research data will be used as an example of fuzzy data.

The data base design process has three phases: (1) analysis, (2) conceptualization and (3) accomodation (Buchmann and Dale 1981). The conceptualization phase is the focus of this thesis. The analysis phase and the accomodation phase will not be discussed.

The purpose of the conceptualization phase is to describe the logical structure of the information of interest by means of a data model. Other names for the data model are 'conceptual model' or 'logical data model' For the purposes of this paper 'logical data base design' refers to the derivation of the logical data model. The derivation process for the data model should not be influenced by the physical storage of the data, data access by a data language, or usage of the data in a specific application. Therefore, physical storage, data languages and data usage will not be discussed in this paper

The proposed approach to the problem of deriving a logical data model includes these activities:

(1) describe fuzzy data

(2) identify criteria for a 'good' logical data model

(3) describe approaches to logical data base design which have potential for representing fuzzy data

(4) derive logical data models using each of the approaches described in (3) to model a representative subset of genealogical research data

(5) evaluate the processes described in (3) and the data models derived in (4) using the criteria identified in (2)

(6) summarize results of (5) to identify criteria for logical data base design methodologies for fuzzy data.

The following considerations are beyond the scope of this thesis: (1) exhaustive evaluation of existing methodologies, (2) recommendations of modifications to existing methodologies to meet the criteria with the possible exception of the selected techniques and (3) derivation of a new data model.

CHAPTER TWO

2.1 FUZZY DATA

Webster's New Collegiate Dictionary defines 'fuzzy' as 'indistinct' or 'not clear'. In the data base environment 'fuzzy' refers to data which may be (1) imprecise (Babad and Hoffer 1984, Bouille 1978, Gaines 1981, Gray 1981), (2) inapplicable (Bouille 1978, Kambayashi et al. 1982, Vassiliou 1979, Wiederhold 1983), (3) incomplete (Babad and Hoffer 1984, Kambayashi et al. 1982, Vassiliou 1980), (4) inconsistent (Babad and Hoffer 1984, Bouille 1978, Vassiliou 1978, Vassiliou 1980) and (5) not known (Babad and Hoffer 1984, Bouille 1978, Gaines 1981, Gray 1981, Kambayashi et al. 1982, Vassiliou 1979, Vassiliou 1980, Wiederhold 1983). A model of the real world needs to reflect the inexactness of fuzzy data. The type of fuzziness associated with data provides a context in which to interpret and to use the data. Often the 'null value' has been used to represent all of these cases. 'Null' means 'none' or 'not any'. Information is lost when a single descriptor, 'null', is used for all of these cases as there is no means to distinguish between the different types of fuzzy data.

Much of the literature relative to null values and fuzzy data in the data base environment is oriented to physical implementation considerations, not logical design methodology. A computer search of the literature did not identify articles of value in selecting a logical data base design methodology for fuzzy data. Logical data base design is recognized as an element indispensible to the success of data base systems. The absence of design techniques to apply when data are fuzzy indicates that this is an area of high research potential.

This chapter considers five types of fuzziness which may be associated with data. Examples of each type are described. The design implementations carried out in Chapter Five include data with these types of fuzziness.

2.2 EXAMPLES OF FUZZY DATA

The five types of fuzziness identified in the first section of this chapter are found when the data of interest include dates or names. The most informative representation of a date is year, month, and day.

IMPRECISE DATA

Assume that it is possible to determine that the date of birth falls within a range of years. The year of birth is not precisely defined. The exact year of birth cannot be distinguished in the known range. The range is the most precise information avaliable. The year of birth exists but is imprecise. The information that the year of birth is included in the range of years should not be lost.

INAPPLICABLE DATA

The value of a middle name for an individual who has no middle name, only a single given name and a surname, is an example of inapplicable data.

INCOMPLETE DATA

If the year and month of birth are known, the birthdate is incomplete. The day is not known. The information that the birth took place in the given month of the given year should be preserved even if the exact day of birth is not known.

INCONSISTENT DATA

If different sources give different dates for the date of birth and it is not possible to identify the more accurate source, the values for the date of birth are inconsistent. It would be helpful to have both pieces of information available. This prevents having less information at hand than, in fact, is known.

UNKNOWN DATA

Assume that the year, month, and day of birth are unknown. The information exists (the individual does have a birth date) but the value is not known.

CHAPTER THREE

CRITERIA FOR EVALUATION OF LOGICAL DATA BASE MODEL

3.1 INTRODUCTION

There is no one optimal way to derive a logical data model. Selection of a design process from which to derive the logical data model depends on (1) area of application, (2) problem structure, (3) expertise of designer and (4) preference of user (Falkenberg 1983). The area of application and the problem structure are particularly relevant considerations for the fuzzy data which are associated with the genealogical example referenced in this paper Criteria should be identified to evaluate the selected design process and the resulting logical data model.

The design process has an impact on the quality of the resulting data model. The design process describes how the perceptions of the environment are translated into the final logical data model. A process which permits assumptions leads to inconsistencies and misrepresentations in the logical data model. This is particularly true as the environment becomes more complex, i.e. as the number of objects and interrelationships increases. An incorrect data model may result from an erroneous translation of correct or incorrect perceptions of the environment. In addition, an error-free translation of incorrect perceptions of the environment will also result in an incorrect logical data model.

In this Chapter criteria are identified for the evaluation of a design process and the final logical data model. The criteria are used in Chapter Six to evaluate the design processes and the final logical data models.

3.2 CRITERIA FOR THE DESIGN

The following criteria are used in Chapter Six to evaluate the logical data designs derived by application of the processes described in this paper:

(1) convenient
(2) expressive
(3) minimally redundant

CONVENIENT

A convenient design is easy to use. The user has easy access to the data. The right data are available to the right people at the right time. The design can be understood by all users. The design is clear, not ambiguous. Data are represented in simple and easily comprehensible form. Non-essentials are removed. No unnecessary complexity is introduced. The number of logical constructs is minimal (Atre 1980; Biller and Neuhold 1977; Borkin 1980; Bracchi 1983; Curtice and Jones 1982; Date 1981; Deen 1981; Kahn 1982; Martin 1981b; Nijssen 1977; Robinson 1981; Smith and Smith 1982; Ullman 1980; Vetter and Maddison 1981).

EXPRESSIVE

An expressive design models the real world and represents all entities and relationships for the application. All appropriate user views are part of the global view. Everything in the universe of discourse can be described. This includes fuzzy data. The design should be correct, that is, it should describe the environment of interest to the user. Its usefulness

depends on the objectives to be met. A correct design conforms to fact and to truth. It is error-free. A design is accurate if it reflects the essential properties of the data. It is precise and exact. An expressive design is not confusing or misleading. The design is meaningful. It is effective for the user The design is called 'natural' if the descriptions are meaningful to the user. It is consistent. The design is described in terms which are familiar to the user The design has simplicity. There are a reasonable number of constructs, a limited number of ways to express the same thing. The design is not overloaded, that is, the same constructs are not used to specify too many different kinds of objects. The design is complete (Atre 1980; Biller and Neuhold 1977; Bracchi 1983; Curtice and Jones 1982; Date 1981; Deen 1981; Falkenberg 1983; Kahn 1982; Martin 1981a; Martin 1981b; Nijssen 1977; Robinson 1981, Smith and Smith 1982; Vetter and Maddison 1980).

MINIMALLY REDUNDANT

Redundancy may occur for data items or for associations between data items. Redundancy may result in multiple states of the same data. The several states may be inconsistent. An attribute which is derived from other attributes represents one form of redundancy. This type of redundancy can be avoided by deriving the attribute when it is needed rather than by including the derived attribute in the logical data model. For example, age is an attribute derived from the current date and the date of birth. Age will not be correct unless new calculations are made as the years pass. If the date of birth is part of the logical data model, age can be calculated as needed. If facts occur once, the design is consistent (Atre 1980; Bracchi 1983; Curtice and Jones 1982; Date 1981; Hubbard 1981; Kahn 1982; Martin 1981a; Robinson 1981).

3.3 CRITERIA FOR THE DESIGN PROCESS

The following criteria are used in Chapter Six for the evaluation of the three selected design processes:

(1) complete
(2) consistent
(3) efficient
(4) error-free
(5) implementation independent
(6) precise
(7) theoretically sound
(8) useable

COMPLETE

A complete design process is thorough. Careful attention is paid to details. The steps of the process are described in a way such that there is no reason to omit any step or to make assumptions. The process is comprehensive, that is, it can be used to model a range of data problems. It is not limited to a specific problem environment. A variety of types of data can be modeled. Redundant data can be identified and a decision made about its inclusion in the design. A complete design process results in an accurate design, a model of the real world (Bracchi 1983; Curtice and Jones 1982; Hubbard 1981, Nijssen 1977).

CONSISTENT

If a design process is consistent, there is agreement as to what rule or usage should be followed for a given step in the process. Guidelines are known. Different designers may produce logical designs for

a given problem which are equivalent, not necessarily identical. Consistency in the design process results in designs which are not ambiguous. The classification of entities should be consistent. Objects which are similar in some way can be classified appropriately and objects which are not identical to each other can be differentiated as necessary. Special attention should be given to the classification boundaries so that a given entity is classified in the same way each time it is considered in the process. Ambiguity should not be introduced through inproper choice of words. It is important to use the same word or phrase for the same entity or association each time it is referenced in the process (Curtice and Jones 1982; Date 1981).

EFFICIENT

The process should use resources efficiently. Time is needed to understand the process as well as to implement it. The number of iterations necessary to produce the design should be as few as possible. The time needed for each iteration should be minimal (Bracchi 1983; Curtice and Jones 1982; Date 1981; Hubbard 1981; Robinson 1981).

ERROR-FREE

The design process should be free of errors. The process should not introduce errors into the data model. The process has integrity. The details of implementation of the process are consistent and not contradictory. The logical data model should not be a false description of the environment as a result of including or excluding inappropriate objects or relationships when following the design process (Curtice and Jones 1982; Hubbard 1981; Robinson 1981).

IMPLEMENTATION INDEPENDENT

Development of the logical data model should be completed with no constraints imposed by prior selection of an implementation method for the physical data base. The process is independent of hardware and software considerations. The result of the design process is a logical data model which describes the environment of interest to the user. This design is useful even if a change is made in the implementation method (Atre 1980; Borkin 1980; Date 1981, Martin 1981b; Nijssen 1977; Olle 1983; Robinson 1981; Smith and Smith 1982; Vetter and Maddison 1981).

PRECISE

There is nothing vague or inexactly defined in a precise process. The rules and their usage are clear There is no ambiguity about how to proceed from the beginning to the end of the process (Nijssen 1977; Vetter and Maddison 1981).

THEORETICALLY SOUND

A design process is theoretically sound if it is based on scientifically acceptable principles. Examples of such principles include set theory, functions, graph theory and normalization (Date 1981; Vetter and Maddison 1981).

USEABLE

A useable methodology is (1) easy to understand, (2) easy to use, (3) simple and (4) straight-forward. A process which is easy to understand is clear. It is not confusing or misleading to the designer. Use of graphical or other visual representation may increase clarity. Concepts and terminology are clearly defined. The vocabulary is not excessively large. A process which is easy to use is not tedious. Ease of use is related to familiarity with other processes. A more complex process which resembles a known process will seem easier to use than it would be if it were the first process learned. The process can be modified to fit a particular need of the application without destroying its integrity. A simple methodology is not complex. The number of constructs and rules is not excessive. A straight-forward process is free of unnecessary elements, is easier to understand and is easier to use. The number of partitions and the number of levels of detail are adequate to model the data but not so great that the designer becomes confused (Biller and Neuhold 1977; Bracchi 1983; Curtice and Jones 1982; Date 1981; Kahn 1982; Nijssen 1977).

CHAPTER FOUR

APPROACHES TO DESIGN OF LOGICAL DATA MODELS

4.1 INTRODUCTION

There are a number of approaches to deriving a logical data model. Some appear to have potential for representating fuzzy data. In this chapter three logical data models are discussed. The process needed to develop each model is also described. An application state of each logical data model is presented in Chapter Five. The logical data models and design procedures are evaluated in Chapter Six. The three approaches selected are (1) Associative Data Model, (2) Semantic Relation Data Model and (3) Vetter / Maddison Relational Model.

The Associative Data Model was developed to provide the designer with a technique for creating logical structures to define the environment of interest. Particular attention is given to identification of the important data items and to defining them at an appropriate level of detail. The notation of the Associative Data Model permits identification of optional variables and specification of domain types. These features may be helpful when modeling fuzzy data.

The Semantic Relation Data Model was selected as representative of semantic models. The structure of the semantic model is such that terminology which is relevant to the application becomes part of the data model. Constraints are an essential part of the data model. The effect of the specification of constraints for fuzzy data is shown in Chapter Six.

Vetter and Maddison have designed a systematic procedure for deriving a logical data model. The result is at least one list of non-redundant elementary relations. The design can best be described as a special case of the 'relational model'. For convenience, in this paper the model is referred to as the 'Vetter / Maddison Relational Model'. This method was selected because it is possible that some, if not all, of the features included in the process may be especially useful in modeling fuzzy data.

4.2 ASSOCIATIVE DATA MODEL

4.2.1 ASSOCIATIVE DATA MODEL - DESIGN

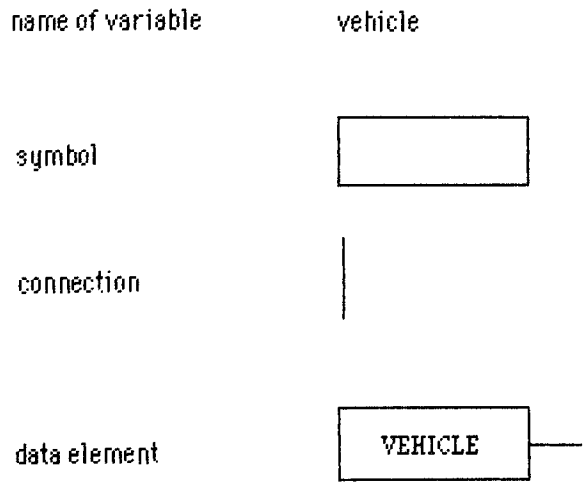
The Associative Data Model is a set of assertion templates expressed in the special notation of the model. The assertion templates describe the logical structure of the environment of interest to the user. Additional information is included in the key-level overview and the data element definitions.

An entity is an object of interest. An entity identifier is associated with each entity. For example:

ENTITY	ENTITY IDENTIFIER
truck	T12345

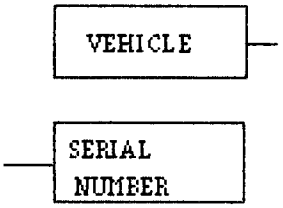
The entity identifier belongs to a domain of truck identification numbers.

Data elements are basic units in an assertion template. They are a representation of the entity identifier in an assertion template. A data element includes the name of a variable, a symbol, and the connection to another element in an assertion template. A data element takes its values from a particular domain. For instance:

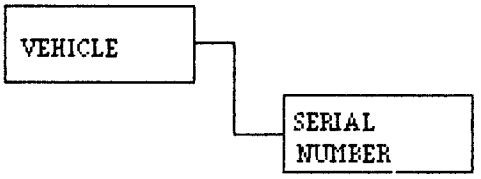


Data elements are grouped to become assertion templates. For example:

data elements



assertion template



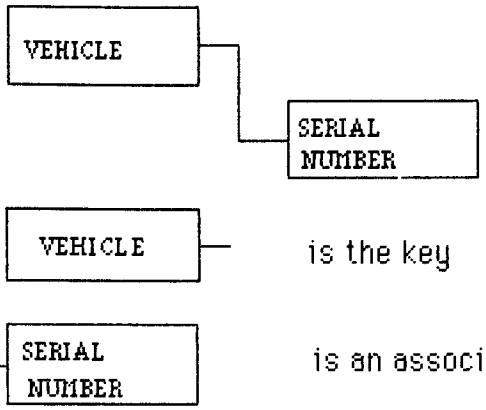
Typical values of vehicle are sedan, station wagon, truck, or van.

Each assertion template contains detailed information about one simple data element called the key or subject of the template. A particular variable name may appear only once as a key in any design. The symbol for the key is  The key appears on the left side of a connection in a template. The bar at the right end of the symbol connects the key to the remainder of the assertion template.

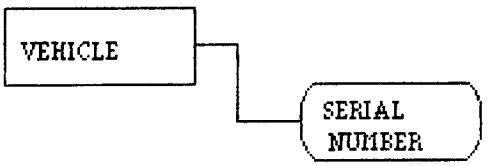
All other data elements in the template are targets or objects. Targets appear on the right side of the connection in a template. The symbol for the target includes a bar at the left end to connect the target to the rest of the template. There are two types of targets, associators and non-associators. An associator is a cross-reference to the key of another assertion template. The symbol for an associator is 

A non-associator does not reference the key of another assertion template.

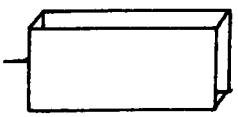
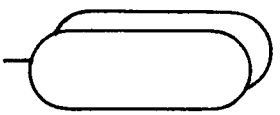
The symbol for a non-associator is  In the example



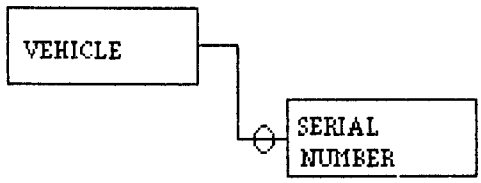
A vertical line separates the subject (vehicle) from the target (serial number). An interpretation of the template is that a particular vehicle has a serial number Serial number is identified as an associator because it is the key of another template. If serial number were not the key of another template, serial number would be a non-associator and the template would be



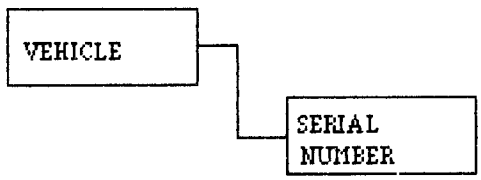
The determination of whether a data element is an associator or a non-associator is made by the designer after analysis of the environment of interest.

Enhancements to the target symbols provide additional information. Solid figures,  and 

indicate that more than one value of the data element is possible. These sets are unordered and do not contain duplicate values. Any of the targets may be mandatory or optional. The value of an optional target does not have to be included. A value of a mandatory target must be present for each value of the key. If the target is optional, a small circle appears on the horizontal line which connects the symbol to the template. For example:

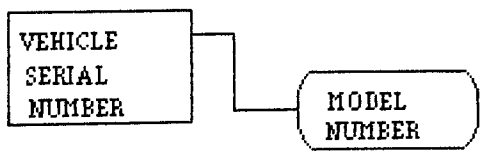


This template indicates that each vehicle may have up to one serial number. Any target which is not optional is mandatory. For example:



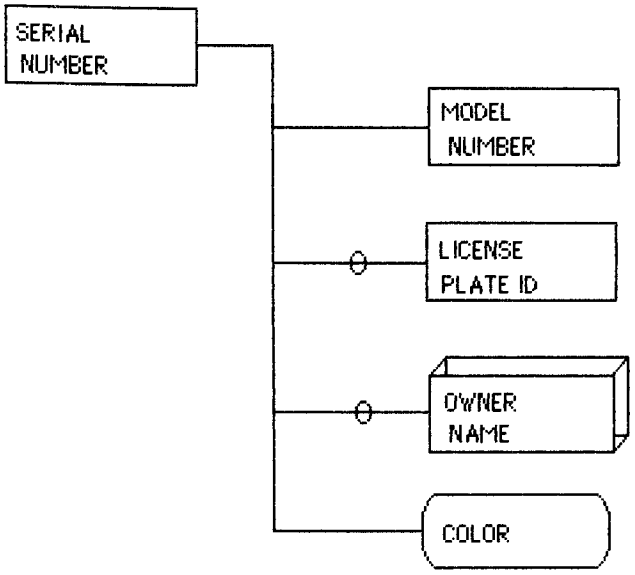
This is to say, each vehicle must have a serial number.

Assertion templates are classified as primitive or compound. A primitive assertion template represents a relationship between the key and a simple data element. For example:



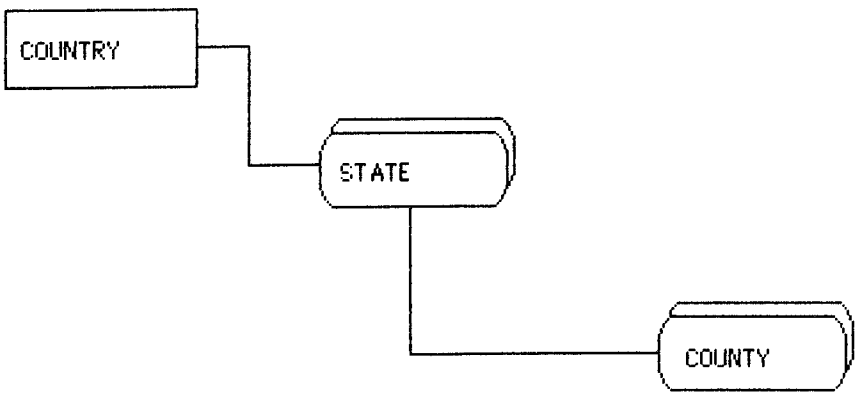
A compound assertion template results when multiple data elements are attached to the key.

For example:



The subject (serial number) appears to the left of the vertical line. The targets (model number, license plate id, owner name, color) appear to the right.

When more than two data elements are part of one relationship, the key and targets form a cascade. The vertical line appears at the lower center rather than at the right center to increase readability of the template. For example:



A vertical line identifies a subject. Each subject has a relationship to the target at the next level.

For example:

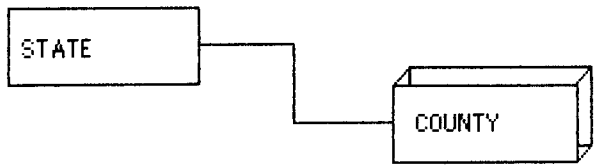
SUBJECT	TARGET
country	state
country-state	county

This assertion template indicates that a country is made up of states. A state in a country is made up of counties. A particular county depends on the combination of country and state. This combination (country and state) is a secondary subject in a relationship with county.

The Associative Data Model requires that each associator apprears as the key of an assertion template such that the inverse of the first relationship is expressed by the second relationship. This is known as the Associator Reversal Rule. For example:

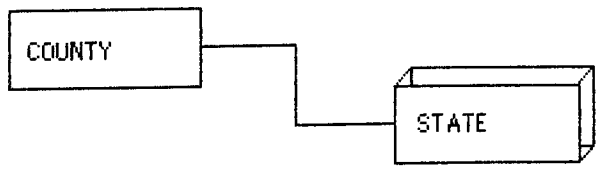
ORIGINAL RELATIONSHIP

KEY	TARGET
state	county



TEMPLATE CREATED BY APPLICATION OF ASSOCIATOR REVERSAL RULE

KEY	TARGET
county	state

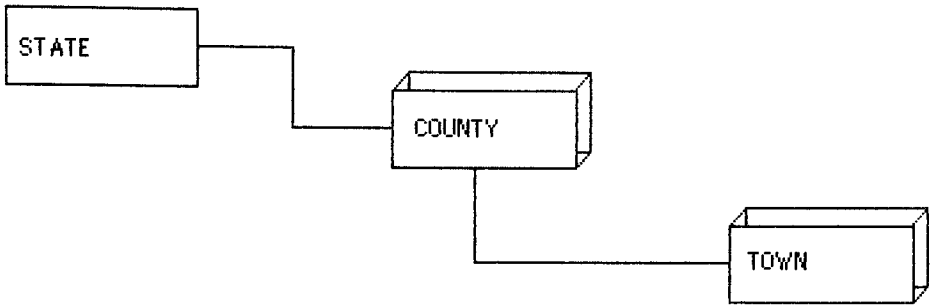


Reversal of a cascade is from the target to the entire subject.

For example:

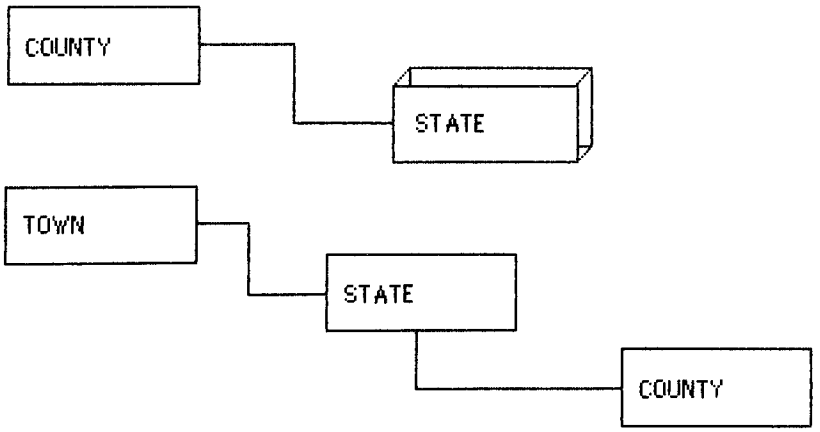
ORIGINAL TEMPLATE

KEY	TARGET
state	county
state-county	town



TEMPLATES AFTER APPLYING ASSOCIATOR REVERSAL RULE

KEY	TARGET
county	state
town	state



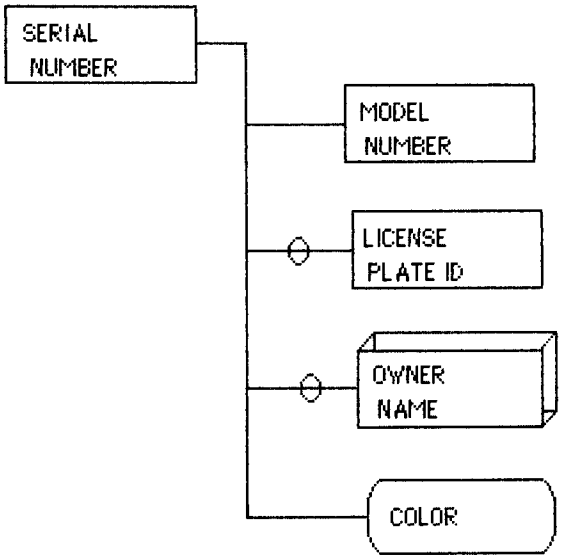
The COUNTY-STATE template indicates that several states may have a county with a given name.

The assertion templates are summarized in a key-level overview. All keys in the design are shown. The relationship between two keys is indicated by a line connecting the corresponding boxes in the overview. A straight line (——) indicates that the relationship between the keys is

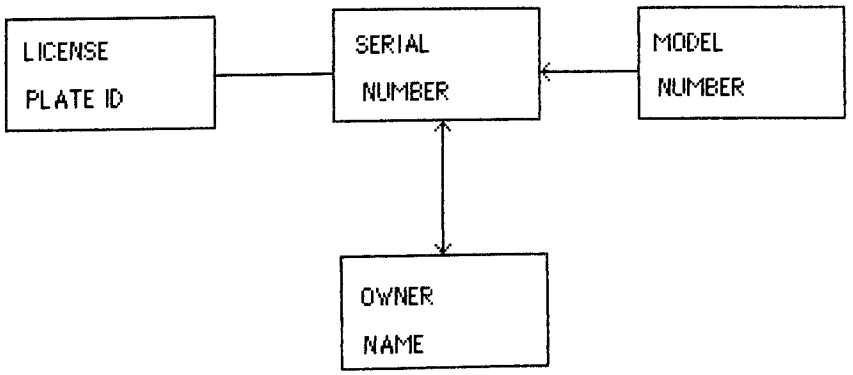
one-to-one. An arrowhead at the end of the line (——>) signifies a one-to-many relationship. A many-to-many relationship is described by (<——>). For example:

ORIGINAL TEMPLATE

key	serial number
associator	model number
	license plate id
	owner name
non-associator	color



KEY-LEVEL OVERVIEW



The keys in the design are (1) model number, (2) license plate id, (3) owner name and (4) serial number

The relationships in the key-level overview are summarized as follows:

KEY PAIR	RELATIONSHIP
model number - serial number	1:N
license plate id - serial number	1:1
owner name - serial number	M:N

Additional information about the design is included in the definition of each data element. This definition should include one or more of the following:

(1) function of variable named in data element
(2) description of conditions under which data element is optional
(3) description of conditions under which relationship exists
(4) domain

For example:

COLOR

indicates exterior paint color of vehicle
domain: black, blue, red, silver, white

LICENSE PLATE ID

indicates license plate identification for vehicle
optional conditions: vehicle not assigned identification (inapplicable)
domain: A-Z, 0-9

MODEL NUMBER

indicates name specified by manufacturer to indicate group of vehicles with specific characteristics
domain: three integers (0-9) followed by a single alpha character (A-Z)

OWNER NAME

indicates name of vehicle owner other than original manufacturer
optional condition: vehicle not sold by manufacturer (inapplicable)
domain: names of persons or organizations

SERIAL NUMBER

indicates unique identification for each vehicle
domain: A-Z, 0-9

4.2.2 ASSOCIATIVE DATA MODEL - DESIGN PROCESS

The design process for the Associative Data Model is:

(1) identify objects of interest
(2) identify relationships between objects
(3) draw assertion templates
(4) apply Associator Reversal Rule
(5) draw key-level overview
(6) write data element definitions

(1) IDENTIFY OBJECTS OF INTEREST

All of the objects of interest in the environment of interest need to be identified. Objects which are the same object but are called by different names should be treated as a single object in the design. Conversely, objects which are called by the same name may be different objects and should be treated as such in the design. Each object has a unique name. Compound objects are divided into simple pieces and the separate pieces are considered for inclusion in the design.

(2) IDENTIFY RELATIONSHIPS BETWEEN OBJECTS

Pairs of objects which are related in some way are identified. Compound relationships are separated into simple relationships.

(3) DRAW ASSERTION TEMPLATES

Those objects which have a substantial amount of information to be recorded in the data base are identified as keys of the assertion templates. Objects related to the keys become targets in the assertion

template for that key. Each target is identified as an associator or as a non-associator. The name of the target may reflect the nature of the relationship between the key and the target. Special characteristics of the relationship (multiple valued, mandatory, optional) are identified on the templates.

(4) APPLY ASSOCIATOR REVERSAL RULE

The Associator Reversal Rule is applied for all associators of all of the templates. Multiple primitive templates with the same key should be combined into a compound assertion template.

(5) DRAW KEY-LEVEL OVERVIEW

The key-level overview can be drawn after all of the assertion templates are completed. The steps to create the key-level overview are:

(1) represent each assertion template key by a labeled rectangle
(2) represent each special associator by a labeled rectangle
Note: A 'special' associator is one which is immediately subordinate to the key and has other associators subordinate to itself.
(3) connect key and associator at immediately subordinate level by a line
Note: If associator is not connected directly to a key, connect it to the next higher level associator. Label duplicate lines. Label line drawn from a key to itself.
(4) add arrowheads on ends of the line to indicate relationships

(6) WRITE DATA ELEMENT DEFINITION

A definition is written for each data element in each assertion template. One or more of the following is included for each of the data elements: (1) function of variable named in data element, (2) description of condition under which data element is optional, (3) description of conditions under which relationship exists and (4) domain.

4.3 SEMANTIC RELATION DATA MODEL

4.3.1 SEMANTIC RELATION DATA MODEL - DESIGN

The Semantic Relation Data Model views the data base as sets of statements describing an existing application state. A particular statement contains information about an entity. An entity is something about which information is to be recorded, an object or concept which exists and can be distinguished. An entity may be associated in some way with another entity. An entity type is the name of an entity set.

A characteristic is a descriptor of an entity type. One or more characteristics describe an entity type.

Example:

entity	truck in my driveway
entity type	truck
characteristic	serial number
characteristic	color

An entity and its characteristics may be represented in tabular format.

truck	
serial number	color

A characteristic is essential if it is necessary and sufficient to define a unique entity. There may be several statements with the same value of the essential characteristic. The value of the essential characteristic does not uniquely identify the tuple. This differs from the relational key which is an attribute with unique values which can be used

to distinguish a particular tuple from all others within a relation.

A column-group is the set of all columns of characteristics of an entity type in a relation. If a column-group includes all of the essential characteristics of an entity type, then the essential characteristic name(s) are italicized or highlighted in some way. None of the characteristic names of the column-group are italicized or highlighted if one or more of the essential characteristics of the entity is missing from the column-group. In this paper '_' is used to highlight the essential characteristic. For example, the column-group for truck is made up of the serial number and color columns. The serial number is an essential characteristic. This information is displayed as follows:

truck	
serial number	color

A domain is a set of values from which the value of the characteristic is taken. The null value is a part of every domain. The null value means 'no value' and is represented by '---' If one characteristic of an entity is null, then all characteristics of the entity are null. For example, the characteristics of a truck are serial number and color. If the color is not known (null), then the serial number must also be null.

Associations and entities are described by predicate: case pairs. In languages the subject and the predicate are the two principle elements of a sentence. A subject is that about which anything is said or done. The predicate is the verb and its modifiers, object, or complement. An object is that which is acted on or accomplished by the verb. 'Case' shows the

relation between a noun or pronoun and some other element in the sentence.

Example:

sentence	John drives a truck.
subject	John
predicate	drives a truck
verb	drives
object	truck
case of truck	objective

In the Semantic Relation Data Model the predicate is a verb phrase. 'Case' is a noun (entity) associated with the predicate as the subject or the object. 'Agent' is often used instead of 'subject' Predicate: agent and predicate: object are predicate: case pairs which describe an association.

Example:

association	drive
predicate	to drive
to drive: agent	person
to drive: object	truck

Each entity type is described by a 'be' predicate, 'be entity type: object'. This special form of the predicate: case pair is used to state the existence of an entity whether or not it is part of an association.

Example:

entity	person
entity type	person
'be' predicate	be person: object

Organized true statements form relations. A relation is specified by (1) predicate: case pairs, (2) entity types, (3) characteristics of

entities and (4) domains of characteristics. Two or more relations can be combined to form a new, more complex relation. The new relation has multiple predicates. At least one of the predicates must not have the null value.

TRUCK

be truck: object
truck
serial number
number
x12345

PERSON

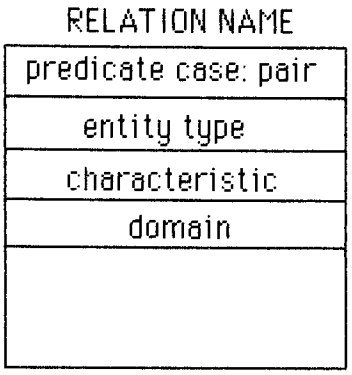
be person: object
person
name
name
John A. Smith

DRIVE

be truck: object to drive: object	be person: object to drive: agent
truck	person
serial number	name
number	name
x12345 ----- y67891	----- John A. Smith James B. Good

A 'be' predicate and one or more of the predicate: case pairs in an association form a set of predicate: case pairs in a relation. A particular predicate: case pair cannot appear in two different predicate: case sets in a given relation. Relations can be named for easier reference. There cannot be duplicate statements (rows) in a relation. The order in which rows and columns appear is not significant.

A relation is displayed as follows:



Constraints are an essintial part of the Semantic Relation Data Model. They give added meaning to the statements. Constraints specify conditions which are always true of the relational state which describes a particular application state. Restricted and mandatory characteristics of an entity or association are described in the constraints. An example of a constraint for a mandatory characteristic is 'each truck must have a serial number' A restriction of the characteristic is 'each truck must have a serial number which is unique'

Logical statements about the application environment can be derived from the specification of the relation. The interpretation of a relation is based on the information in the relation and in the constraints.

A relation for a particular application is:

TRUCK

be truck: object	
truck	
serial number	color
number	black/white/red

Constraint: serial number is unique

The interpretation of this relation is: 'A truck is black, white or red in color and has a unique serial number'

4.3.2 SEMANTIC RELATION DATA MODEL - DESIGN PROCESS

The design process to derive the Semantic Relation Data Model is:

1. identify entity types
2. identify associations
3. identify predicate: case pairs
4. form relations
5. identify characteristics of entity types
6. define domains of characteristics
7 identify constraints

The items of interest in the final data base design are the entity types and the associations between entity types. All of these items should be identified as early as possible in the design process. Each entity type is described by a predicate: case pair called a 'be' predicate. Two predicate: case pairs, predicate: case agent and predicate: case object, are developed for each association. Relations are formed from one or more predicate: case pairs. It is useful to name relations as they are formed although it is not mandatory that relations be named. The design process is iterative. Review of the first four steps in the design process before moving to the final three steps may reduce the number of iterations necessary to complete the design.

The characteristics of each entity type are identified. An appropriate label is selected for each characteristic. In this step it is useful to consider the likelihood of the occurence of 'null' as a value of the characteristic and its effect on the use of the design. Essential characteristics are identified. The constraints are listed. Review of the design should include consideration of the possibility of generating duplicate tuples. A design which results in duplicate tuples should be modified since duplicate tuples are not allowed. Entering sample data in the relations may indicate design weaknesses for the particular application and the need to modify the design.

4.4 VETTER / MADDISON RELATIONAL MODEL

4.4.1 VETTER / MADDISON RELATIONAL MODEL - DESIGN

The Vetter / Maddison Relational Model is a list of non-redundant elementary relations describing the environment of interest to the user. Several such lists may be generated. Each of the lists fully describes the environment of interest. One of these models is selected for implementation. This paper considers the generation of the list(s) of elementary relations. It does not discuss the selection and implementation of one of the models.

An entity is something about which information is recorded. An occurrence of an entity is called a tuple. There are n elements in an n-tuple. If an n-tuple is ordered, each element is selected from a separate set of possible values. The first element belongs to the first set. The nth element belongs to the nth set. Each of the n sets is sometimes called a domain. Vetter and Maddison, however, use the term domain for the values which occur at a particular instant in time. An example of a 4-tuple is:

truck 20L A579T red

where:	element of tuple	belongs to set
	truck	vehicle type
	20L	model number
	A579T	serial number
	red	color

A relation is composed of ordered n-tuples. N is the degree of the relation. A relation of degree n can be displayed as a table with n columns. An alternate format for display is R[A(1),A(2),...,A(N)] where R is the

relation name and A(N) is a set. An example of a relation is:

VEHICLE(TYPE, MODEL ID, SERIAL NUMBER, COLOR)

VEHICLE

TYPE	MODEL ID	SERIAL NUMBER	COLOR

The project operation forms a subset of the original relation by extracting specified columns of the relation and removing duplicate rows. For example:

Original Relation

TYPE	MODEL ID	SERIAL NUMBER	COLOR
truck	20L	A579T	red
van	50M	D432V	blue
sedan	35N	G771S	black
truck	20L	B688T	green

Apply project operation to vehicle type

TYPE
truck
van
sedan

Two relations can be joined to create a new relation in which each tuple is formed by concatenating a tuple from each of the original relations. Each of the two tuples which is joined satisfies a specified condition.

For example:

Original Relations

VEHICLE

TYPE	MODEL ID	SERIAL NUMBER
truck	20L	A579T
van	50V	D432V
sedan	35N	G771S
truck	20L	B688T

MILEAGE

MODEL ID	MPG
20L	27
50V	24

Apply join operation to two original relations when MODEL ID's are equal

VEHICLE MILEAGE

TYPE	MODEL ID	SERIAL NUMBER	MPG
truck	20L	A579T	27
van	50V	D432V	24

An elementary relation is a relation which cannot be reduced into relations of smaller degree by using the project operation such that the original relation is recreated when the relations resulting from the use of the project operation are joined.

An example of an elementary relation is:

VEHICLE - MODEL

TYPE	MODEL NUMBER
truck	20L
van	50V
sedan	35N

4.4.2 VETTER / MADDISON RELATIONAL MODEL - DESIGN PROCESS

The design process described by Vetter and Maddison includes the following phases:

(I) specify environment of interest
(II) transform description of environment of interest into list of elementary relations
(III) expand list of elementary relations to include all possible elementary relations
(IV) reduce list of all possible elementary relations to a minimal set of elementary relations

There are several steps associated with each phase.

SPECIFY ENVIRONMENT OF INTEREST (PHASE I)

The steps to specify the environment of interest are:

(1) identify entity sets
(2) identify properites of entity sets
(3) identify domains of properties of entity sets
(4) identify key domains of entity sets
(5) identify entity attributes and express as relations with entity set replaced by primary key domains
(6) identify relationship sets and express as relations with entity sets replaced by primary key domains
(7) identify relationship attributes and express as a relation with entity sets replaced by primary key domains

An entity is anything of interest to the user An entity may be a thing, a person, a concept, an event or a relationship. It can be distinguished from other entities in some way.

Examples of entities are:

TYPE OF ENTITY	ENTITY
thing	apple
person	John Brown
abstract concept	July 1, 1984
event	birth
relationship	marriage

Entities of the same type form entity sets. Each entity set has a unique name. For example, the entity 'apple' belongs to the entity set 'fruit'.

An entity is identified and classified by its properties. Properties are named characteristics of the entity. Some properties of a person are name, age and eye color Properties which are relevant to the application area are selected for the model. An occurrence of a property of an entity is a property value. Property values for an entity belong to a specified domain. The association from an entity set to a domain is called an entity attribute. For example, 'age' is an entity attribute for the entity set 'person'. The domain is 'positive integers'.

An attribute which has different values for each occurrence of an entity is called an entity key. If there are several entity keys, then one of them is selected as the primary entity key. A primary entity key may also be called a primary key. A primary key is identified for each entity set. It is generally assumed that the primary key always has a non-null value. The domain underlying the primary key is used to identify the entity set. This domain is called the primary key domain. For example, the attribute 'social security number' is a primary entity key for the entity set 'person' The primary key domain is social security numbers. The attribute 'name'

associates the domain 'social security number' with the domain 'name'. 'Social security number', the attribute, associates the domain 'social security number' with itself.

An association between entities is a relationship. Each relationship has a unique name. For example:

ENTITY

John
coffee

ASSOCIATION

prefers

RELATIONSHIP: BEVERAGE PREFERENCE

John prefers coffee

TRANSFORM DESCRIPTION INTO ELEMENTARY RELATIONS (PHASE II)

The conceptual objects identified in Phase I are transformed into elementary relations. An elementary relation is not reducible. It is a relation with one of the following properties: (1) all key, (2) single attribute is functionally dependent of second single attribute or (3) single attribute is fully functionally dependent on several attributes. Execution of the following steps, in order, leads to elementary relations:

(1) eliminate non-full functional dependencies on candidate keys
(2) eliminate transitive dependencies on candidate keys
(3) eliminate multivalued dependencies

(4) if necessary, determine single or composite primary key and take projections such that each projection has the primary key and one attribute from the complement of the primary key.

For a given relation, an attribute B is functionally dependent on attribute A if, and only if, each value of A is associated with one value of B at any given time. Attribute B is fully functionally dependent on attribute A if attribute B is functionally dependent on attribute A but is not functionally dependent on a subset of attribute A. In the case of dependence on a single attribute, there is no difference between functional dependence and full functional dependence. The dependence is usually shown as functional dependence (----->). Full functional dependence is shown as (=====>). For example:

relation:	bank account
attribute A:	bank account number
attribute B:	account balance

Attribute B is functionally dependent on attribute A. Attribute B is also fully functionally dependent on attribute A.

A transitive dependency exists between attribute X and attribute Z in a relation if there is an attribute Y such that a value of Y cannot be associated with attribute X unless a value of attribute Z is associated with attribute Y. For example:

RELATION: (PRODUCT ID, MACHINE ID, EMPLOYEE ID)

where 1) a product is made on one machine only
2) a machine is operated by one employee only

MANUFACTURE

PRODUCT ID	MACHINE ID	EMPLOYEE ID
P44	M234	E12

EMPLOYEE ID is transitively dependent on PRODUCT ID through MACHINE ID.

A multivalued dependency exists for attribute B if the value of B depends on the value of attribute A and there is more than one value of B for a given value of attribute A in the relation. For example:

RELATION: OPERATE(MACHINE ID, EMPLOYEE ID)

where a given machine is operated by any one of several employees

OPERATE

MACHINE ID	EMPLOYEE ID
M234	E12
M234	E22
M234	E24

There is a multivalued dependence of EMPLOYEE ID on MACHINE ID.

EXPAND LIST OF ELEMENTARY RELATIONS (PHASE III)

The expanded list of elementary relations includes the original list of elementary relations described in Phase II and elementary relations derived from those of the original list. The complete set of elementary relations is called a transitive closure. It is unique for each original list

of elementary relations. The steps to derive the expanded list are:

(1) represent original list of elementary relations by digraph
(2) represent digraph by connectivity matrix
(3) modify connectivity matrix to represent transitive closure
(4) eliminate meaningless elementary relations from transitive closure

The attributes of the elementary relation are represented by the nodes of the directed graph. A cycle of length one occurs at a node which represents an entity key or at a node which represents an elementary relation which is all key. The arcs of the digraph represent the dependencies. Arcs are named by the name of the elementary relation whose dependence is shown by the arc. The different types of dependence (functional, full functional and trivial) are distinguished by different line formats. A single line (----->) represents functional dependence. A double line (=====>) represents full functional dependence. A broken line (- - ->) represents trivial dependence.

A digraph can be represented by a square matrix called a connectivity matrix. The size of the matrix is N x N where N is the number of nodes in the digraph. The first row and the first column are identified with the first node. The Nth row and the Nth column are identified with the Nth node. The connections between the nodes of the digraph are represented by either one or zero. One indicates that there is a connection between the nodes. Zero or a blank indicates that there is no connection. No distinction is made between the several types of dependence in the matrix. Node A may be repeatedly dependent on node B. Node A is represented in the matrix by as many rows and columns as there are

repetitions. A descriptor is added to the node name to distinguish the multiple rows and columns. For example:

STATE is part of HOME ADDRESS and SCHOOL ADDRESS

H.STATE can be used for HOME ADDRESS

S.STATE can be used for SCHOOL ADDRESS

A transitive closure is the set of all elementary relations, original and derived, which describes the environment of interest. The transitive closure is unique. The transitive closure for the original list of elementary relations can be created by identifying transitive dependencies of length 2 until no further such dependencies can be identified. The derived transitive dependencies are added to the connectivity matrix to form a modified connectivity matrix. Any new elementary relations derived in this way must be reviewed. Only those relations which have meaning in the environment of interest are retained in a modified connectivity matrix. This modified matrix defines a digraph which completely describes the environment of interest.

REMOVE REDUNDANT ELEMENTARY RELATIONS (PHASE IV)

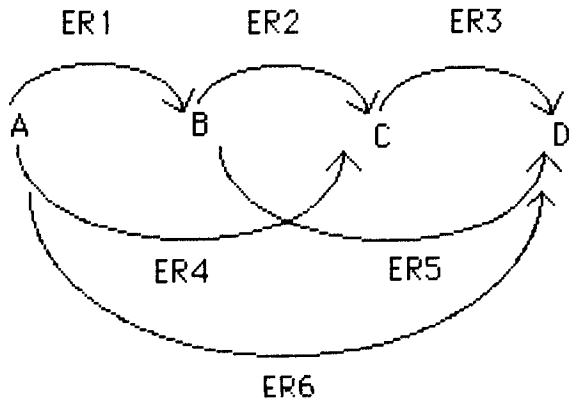
A redundant elementary relation is an elementary relation which is composed of another elementary relation. Removal of redundant elementary relations results in one or more minimal covers. Each minimal cover includes all of the information in the original elementary relations. There are two conditions for the removal of an elementary relation. The first condition requires that the elementary relation to be removed must

be a composition of two other elementary relations. The second condition requires that an elementary relation can be removed only if it is not part of a composition of another elementary relation. For example:

ELEMENTARY RELATIONS

ER1(A,B)
ER2(B,C)
ER3(C,D)
ER4(A,C)
ER5(B,D)
ER6(A,D)

DIGRAPH



REDUNDANT ELEMENTARY RELATION	COMPOSED OF
ER4	ER1, ER2
ER5	ER2, ER3
ER6	ER1, ER5
ER6	ER4, ER3

Consider removal of ER4

Condition 1	yes	ER4 is composed of ER1 and ER2
Condition 2	no	ER4 is part of ER6

Consider removal of ER6

Condition 1	yes	ER6 is composed of ER1 and ER5 ER6 is composed of ER4 and ER3
Condition 2	yes	ER6 is not part of another ER

ER6 can be removed.

Consider removal of ER4 after removal of ER6

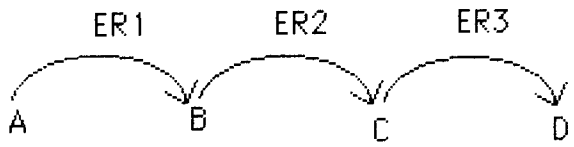
Condition 1	yes	ER4 is composed of ER1 and ER2
Condition 2	yes	ER4 is not part of another ER

ER4 can be removed.

Similarly, ER5 can also be removed. The remaining elementary relations are non-redundant.

NON-REDUNDANT ELEMENTARY RELATIONS

ER1
ER2
ER3



CHAPTER FIVE

THREE LOGICAL DATA MODELS OF THE GENEALOGICAL EXAMPLE

5.1 INTRODUCTION

This chapter includes the implementation of the three designs described in Chapter Four The application selected for implementation is a genealogical example including instances of the types of fuzziness described in Chapter Two.

5.2 NATURE OF GENEALOGICAL DATA

The two problems solved by genealogists are (1) determination of the pedigree of an individual and (2) creation of a family genealogy. The pedigree identifies the direct ancestor of an individual. Basic data in the pedigree are name, date and place of birth, date and place of marriage and date and place of death. The data are often presented on an Ancestor Chart (see Appendix A). The family genealogy includes data on all members of the family. In the simplest form it includes the relevant data for a husband, wife and their children. The data may be entered on a Family Group Sheet (see Appendix A). A family genealogy generally includes related families for many generations. In addition to the basic data, it is often useful to store other information about an individual. One list includes seventy-eight items of possible interest about an individual (Genealogical Computing, July 1982).

For the purposes of this example, names and dates have been selected as representative of fuzzy data often encountered in genealogical research.

Names are necessary to identify persons and places. In addition to the surname, the name of a person includes given name, nickname or alias. The given name is the first name and all middle names. A nickname may be used instead of, or in addition to, the given name. An alias can result from a legal name change, previous marriage and so forth. Titles, such as 'Jr.' and 'Sr.', may or may not be used to identify members of the same family with the same given name. Variations in spelling the surname are not unusual (see Appendix B). Some families followed the practice of naming children after other family members. This results in duplicate names, not only between generations, but also within generations. The name of a deceased child was often given to the next child of the same sex to be born in the family. This can lead to confusion, particularly if both births occurred in the same calendar year.

Place names identify the country, state, county, town, city, village, township or other political subdivision. Place names are usually associated with events in the life of an individual or family. Of interest are place of birth, place of death, place of marriage, place of divorce, place of burial, place of christening or place of residence. It is useful to be able to identify changes as states, counties and towns were created or annexed.

Dates are associated with the events in the life of an individual or family. Such dates are those of birth, death, marriage, divorce or christening. Other dates of interest are those associated with documents such as deeds, military service records or military pension records. Census records, court records, directories and ship passenger lists also include dates of interest to the genealogist. Dates are usually entered as a year, month and day. Use of census records may result in identification of a range of years during which a birth or death occurred. Data from

several sources for the same event may not agree. It is also helpful to know which calendar was in use so that given dates may be identified correctly.

This example will model the information associated with births, deaths, marriages and biological children. An individual may have no marriages, one marriage or more than one marriage. There may be no children, the children of one marriage or the children of several marriages living in the family unit.

5.3 GENEALOGICAL EXAMPLE

Specific examples of fuzzy data are found in the Ancestor Chart and the Family Group Sheet in Appendix B.

IMPRECISE DATA

It is sometimes difficult to know whether the name found is the given name or a nickname. Consider the name 'Sally Batchelder' It is possible that (1) the given name is 'Sally' or (2) the given name is 'Sarah' and the nickname is 'Sally'. To express dates as a range of years is to express dates imprecisely. For instance, by referring to the census records it can be determined that the year of birth of Ares, daughter of Betsy and Oliver Smith, occurred between 1805 and 1810. Another way of expressing imprecise dates is by using a modifier with the year. For example, the records indicate that Chester, son of Oliver Smith, died after the 1880 Census had been completed. This information is often expressed as 'after 1880' Another case of an imprecise date is the birth date of Sarah Webster, wife of Joseph Mygatt, who was born around 1655. This information is stated as 'about 1655'.

INAPPLICABLE DATA

It is possible that no middle names were given to the five children of Emily and Charles Smith. If this is the case, then the middle name is an example of inapplicable data. Chester Smith, son of Oliver Smith, never married. Any data relating to date and place of marriage is inapplicable for Chester

INCOMPLETE DATA

Betsy, wife of Oliver Smith, does not have a known surname. She is referred to as 'wife of Oliver Smith' or 'Betsy' Her name is incomplete. The third child of Lyman Smith and his wife Maria was born in 1834 and named Oliver A. Oliver's name is an example of incomplete data because the complete middle name is not known.

INCONSISTENT DATA

The 1850 Census indicates that Chester Smith was born in Massachusetts. The 1880 Census indicates that the same Chester Smith was born in New Hampshire. These two pieces of information cannot both be true. There may, however, be useful information in both statements. Chester may have lived in both places before moving to Pennsylvania with his family. In other instances, what appears to be an inconsistency may be an inaccuracy. For example, both Crawford County and Erie County were formed from Allegheny County (Pennsylvania) in 1800. An individual who lived in Crawford County after 1800 may always be described as being from Crawford County. In reality, this individual was from Allegheny County prior to 1800. The place names associated with events in his life would have taken place in Allegheny County prior to 1800 and in Crawford County after 1800. There can be several such dates associated with a political subdivision. Allegheny County, for instance, was created from parts of Westmoreland and Washington Counties in 1788. Therefore, information for Crawford and Erie Counties for the years prior to 1788 is found in the records for Westmoreland or Washington Counties (see Appendix B).

UNKNOWN DATA

It is likely that Betsy was Oliver Smith's second wife. The name of his first wife is not known. Samuel Batchelder, son of Molly and Isaiah Batchelder, was married twice. The names of both of his wives are unknown. James Leach was born in New Hampshire in 1780. The town and county of his birth are not known. At least six of the eleven children of Emily and Charles Smith have known middle names. If the remaining five children were given middle names, then these middle names are examples of unknown data.

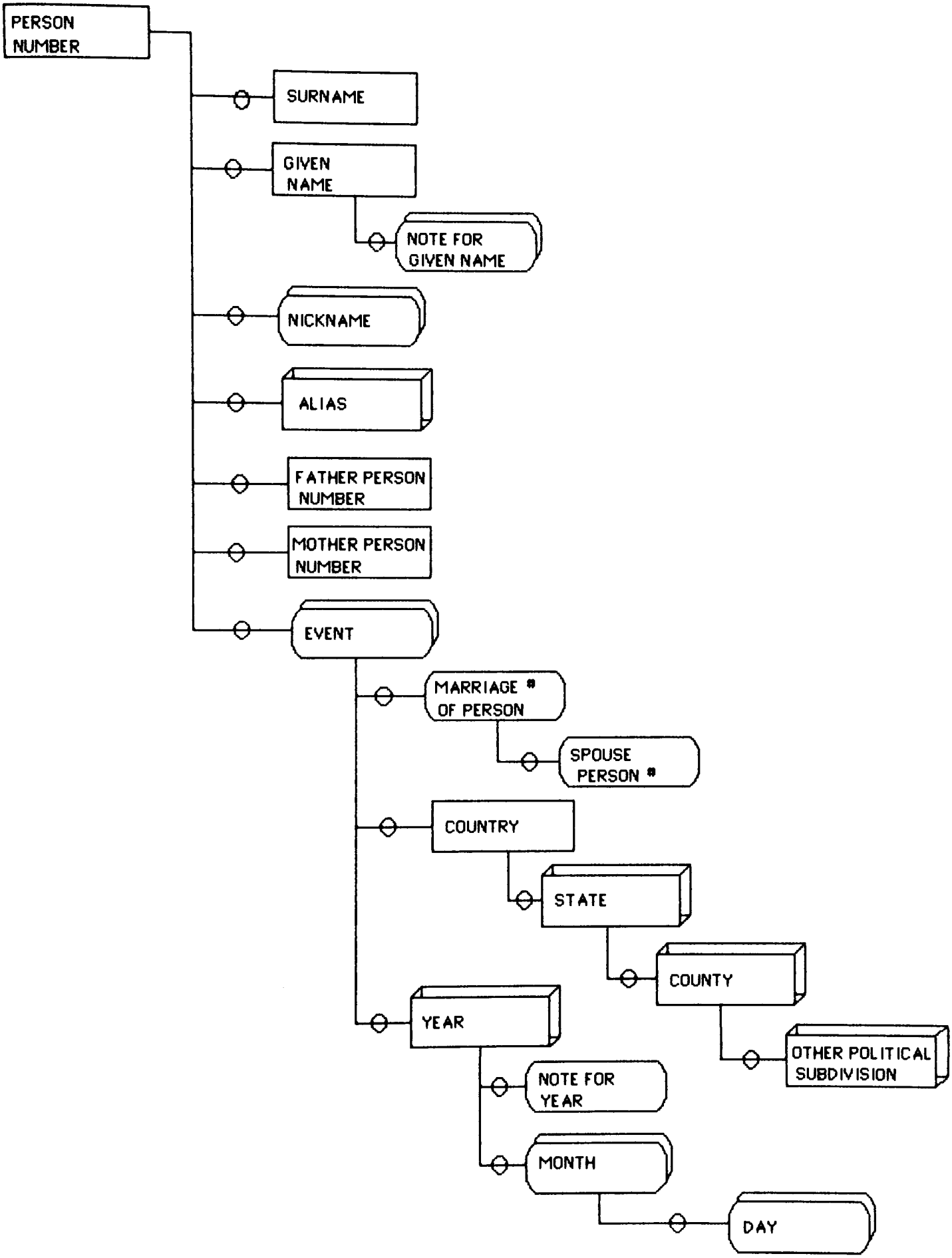
The above are examples of fuzzy data which are a part of a genealogical data base detailed in Appendix A.

5.4 ASSOCIATIVE DATA MODEL

The first part of this section is a description of the Associative Data Model for the genealogical example. The second part is an example of an application state for this model based on the data found in Appendix A.

5.4.1 ASSOCIATIVE DATA MODEL - GENEALOGICAL EXAMPLE DESIGN

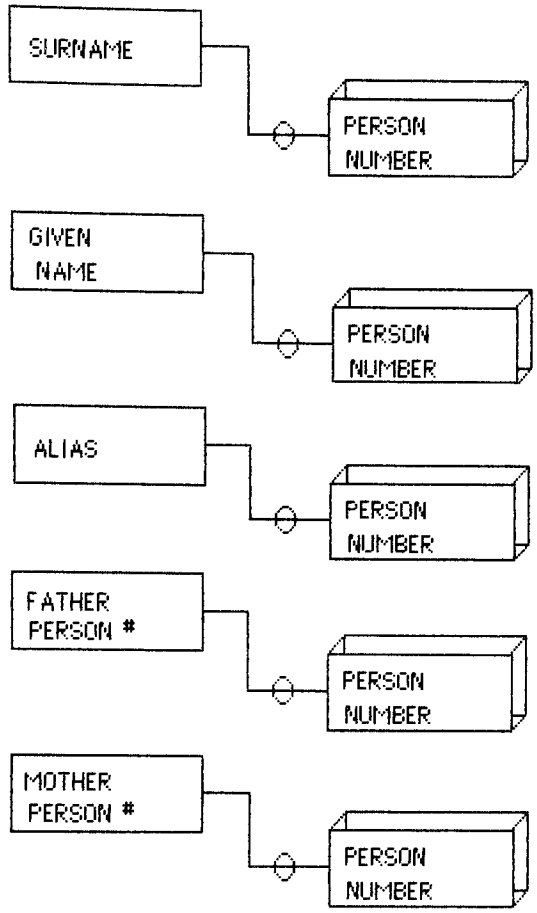
The genealogical example is described by eleven assertion templates, a key-level overview and data element definitions. Each of the assertion templates is described separately. PERSON NUMBER is the key of the original template. The other templates result from application of the Associator Reversal Rule. The template which appears on the next page is the original template.


PERSON NUMBER
SURNAME
GIVEN NAME
NOTE FOR GIVEN NAME
NICKNAME
ALIAS
FATHER PERSON NUMBER
MOTHER PERSON NUMBER
EVENT
MARRIAGE # OF PERSON
SPOUSE PERSON #
COUNTRY
STATE
COUNTY
OTHER POLITICAL SUBDIVISION
YEAR
NOTE FOR YEAR
MONTH
DAY

Associators in the original assertion template are SURNAME, GIVEN NAME, ALIAS, FATHER PERSON NUMBER, MOTHER PERSON NUMBER, COUNTRY, STATE, COUNTY, OTHER POLITICAL SUBDIVISION and YEAR. Non-associators are NOTE FOR GIVEN NAME, NICKNAME, EVENT, MARRIAGE NUMBER OF PERSON, SPOUSE PERSON NUMBER, NOTE FOR YEAR, MONTH and DAY. All associators are optional. Multiple values may occur for NOTE FOR GIVEN NAME, NICKNAME, ALIAS, EVENT, STATE, COUNTY, OTHER POLITICAL SUBDIVISION, YEAR , MONTH and DAY.

It should be noted that multiple values of STATE, COUNTY and OTHER POLITICAL SUBDIVISION accomodate the imprecision of the genealogical data. For instance, assume that it is known that an individual was born in Massachusetts or in New Hampshire. There are two values of STATE to be associated with a single event, BIRTH. It should also be noted that multiple values of YEAR, MONTH and DAY can also be associated with an event. For example, if it is known that an individual was born between 1800 and 1805, there are multiple values of YEAR associated with the event, BIRTH. The fact that a range of years is known can be included in the NOTE FOR YEAR.

Keys of the other assertion templates are SURNAME, GIVEN NAME, ALIAS, FATHER PERSON NUMBER, MOTHER PERSON NUMBER, COUNTRY, STATE, COUNTY, OTHER POLITICAL SUBDIVISION and YEAR. PERSON NUMBER is the target in the templates with the following keys: SURNAME, GIVEN NAME, ALIAS, FATHER PERSON NUMBER, MOTHER PERSON NUMBER. Multiple values of this target may occur

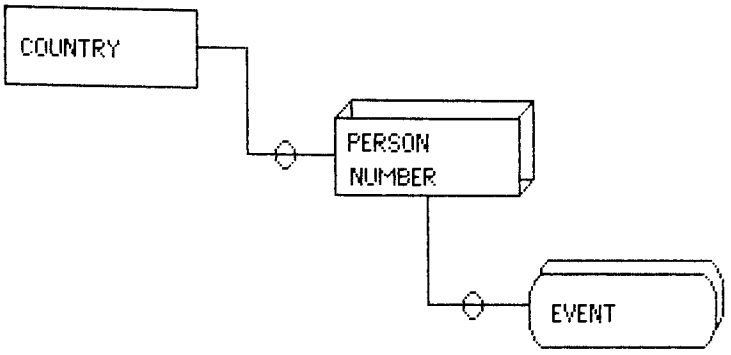


This is to say that there may be more than one individual with the same name SURNAME, GIVEN NAME, ALIAS, FATHER or MOTHER. For example:

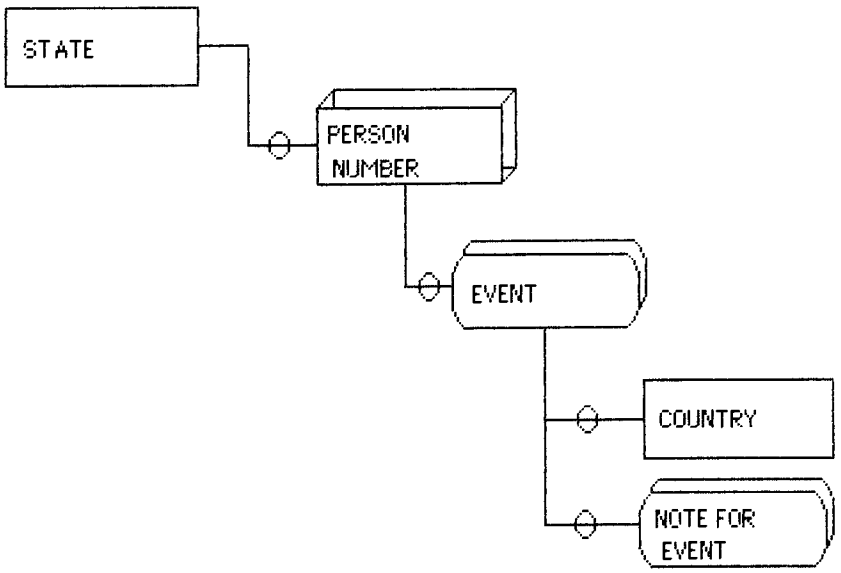
KEY: SURNAME	TARGET: PERSON NUMBER
Webster	3
Webster	6

The target of the subject COUNTRY is PERSON NUMBER. PERSON NUMBER is an associator which is optional and may have multiple values. EVENT is the target of the subject of the subject COUNTRY-PERSON NUMBER. It is a non-associator which is optional and can have multiple values. This template describes the association between COUNTRY,

PERSON NUMBER and EVENT For a particular country there are one or more persons who had one or more events take place in the country.



The optional target of the subject STATE is PERSON NUMBER. STATE and PERSON NUMBER form the subject for the optional target, EVENT STATE, PERSON NUMBER and EVENT form the subject for the optional targets COUNTRY and NOTE FOR EVENT PERSON NUMBER, EVENT and NOTE FOR EVENT may have multiple values. EVENT and NOTE FOR EVENT are non-associators. PERSON NUMBER and COUNTRY are associators.



One or more persons may optionally be the target of the key STATE. Each STATE-PERSON NUMBER key may have one or more EVENTS as optional targets.

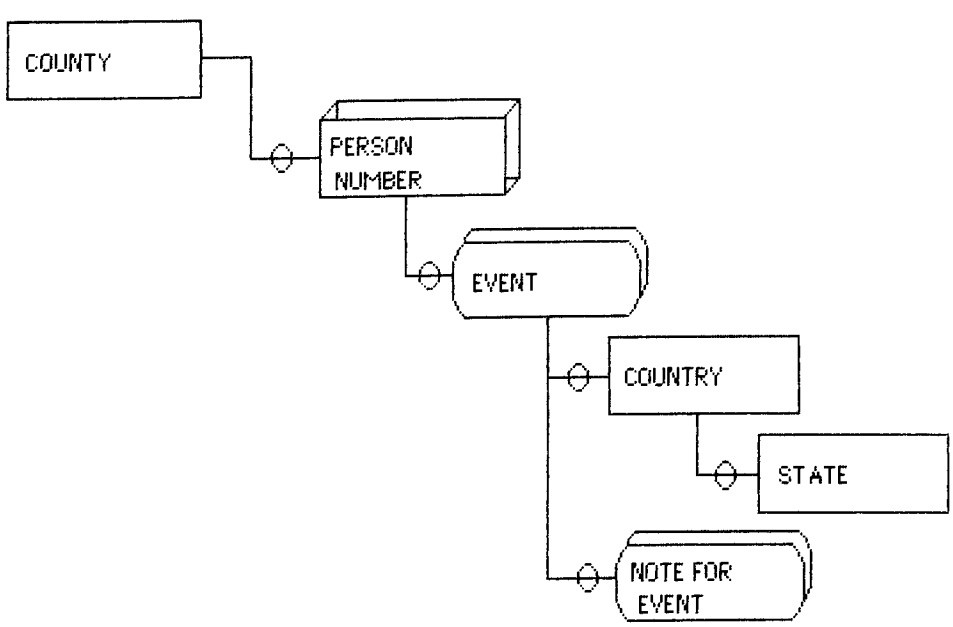
For example:

KEY : STATE	TARGET: PERSON NUMBER
NY	1
NY	2

KEY: STATE-PERSON NUMBER	TARGET: EVENT
NY 1	birth
NY 2	birth

KEY: STATE-PERSON NUMBER-EVENT	TARGET COUNTRY
NY 2 death	Canada

The optional target of the subject COUNTY is PERSON NUMBER. COUNTY and PERSON NUMBER form the subject for the optional target EVENT COUNTY, PERSON NUMBER and EVENT form the subject for optional targets COUNTRY and NOTE FOR EVENT. STATE is the optional target of the subject COUNTY-PERSON NUMBER-EVENT-COUNTRY. PERSON NUMBER, EVENT and NOTE FOR EVENT may have multiple values. PERSON NUMBER, COUNTRY and STATE are associators. EVENT and NOTE FOR EVENT are non-associators.



For example:

KEY: COUNTY	TARGET: PERSON NUMBER
Orange	28
Orange	29

KEY: COUNTY-PERSON NUMBER	TARGET: EVENT
Orange 28	birth
Orange 28	marriage
Orange 29	marriage

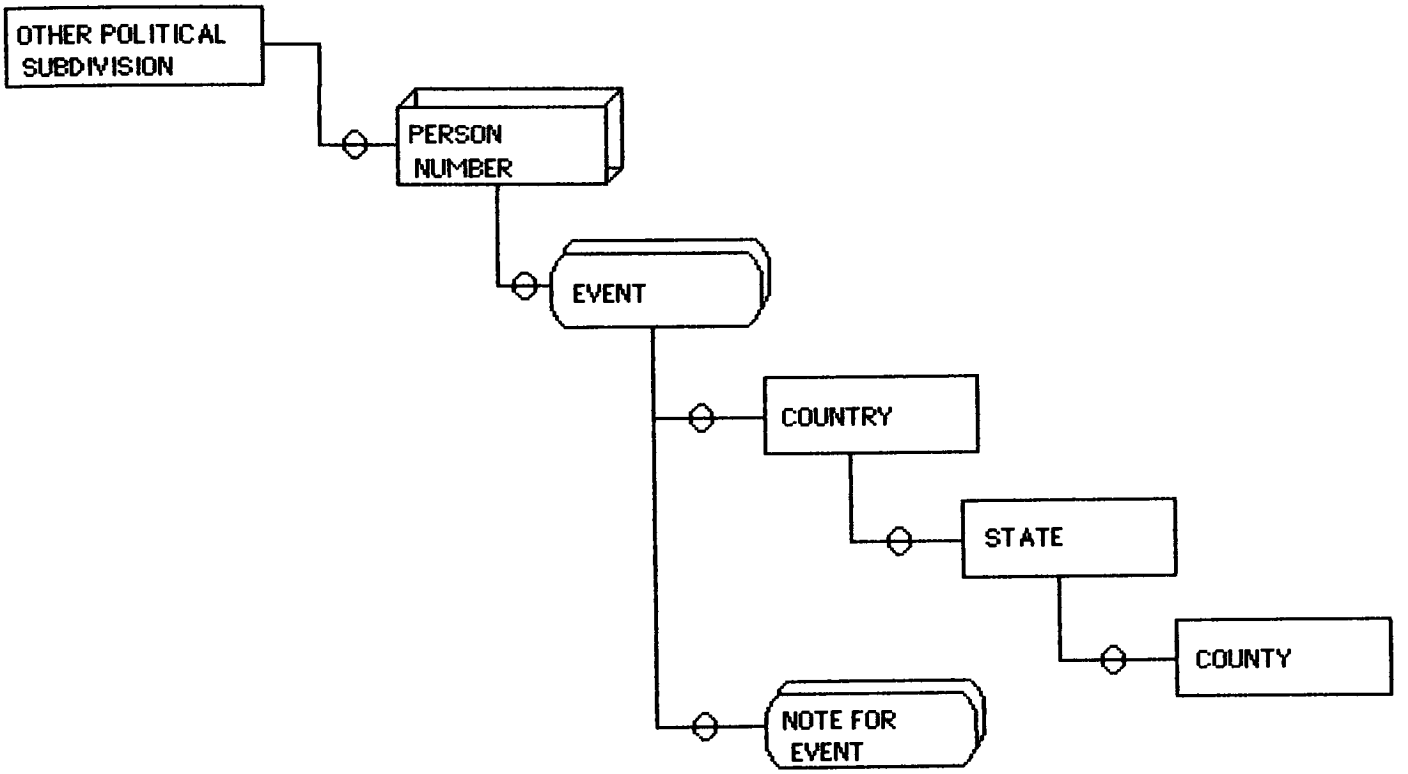
KEY: COUNTY-PERSON NUMBER-EVENT	TARGET: COUNTRY
Orange 28 birth	USA

KEY: COUNTY-PERSON NUMBER-EVENT-COUNTRY	TARGET STATE
Orange 28 birth USA	VT

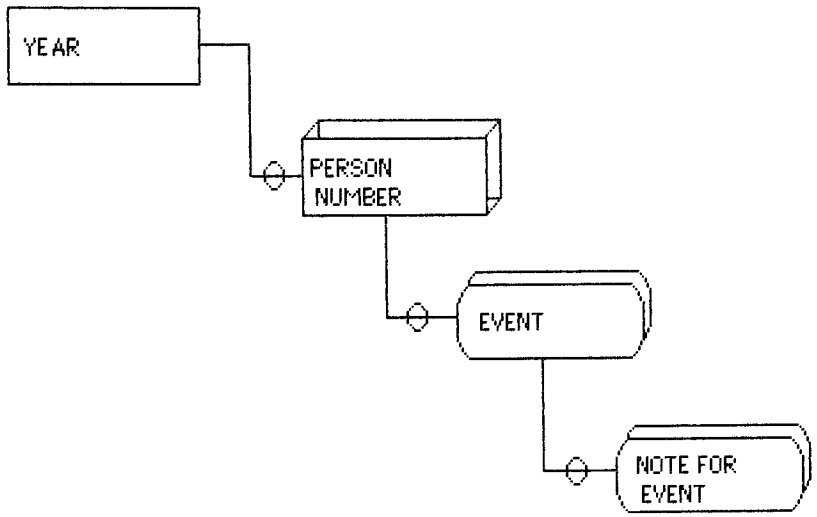
OTHER POLITICAL SUBDIVISION is the subject for the optional associator target PERSON NUMBER. OTHER POLITICAL SUBDIVISION-PERSON NUMBER is the subject for the optional non-associator target EVENT. OTHER POLITICAL SUBDIVISION, PERSON NUMBER and EVENT form the subject for the optional associator COUNTRY and the optional non-associator NOTE FOR EVENT. STATE is the optional associator target of the subject OTHER POLITICAL SUBDIVISION-PERSON NUMBER-EVENT-COUNTRY. The subject OTHER POLITICAL SUBDIVISION-PERSON NUMBER-EVENT-COUNTRY-STATE has the optional associator COUNTY for a target. PERSON NUMBER, EVENT and NOTE FOR EVENT may have multiple values. Examples of keys and targets follow the pattern of the assertion templates described earlier.

ASSERTION TEMPLATE

OTHER POLITICAL SUBDIVISION

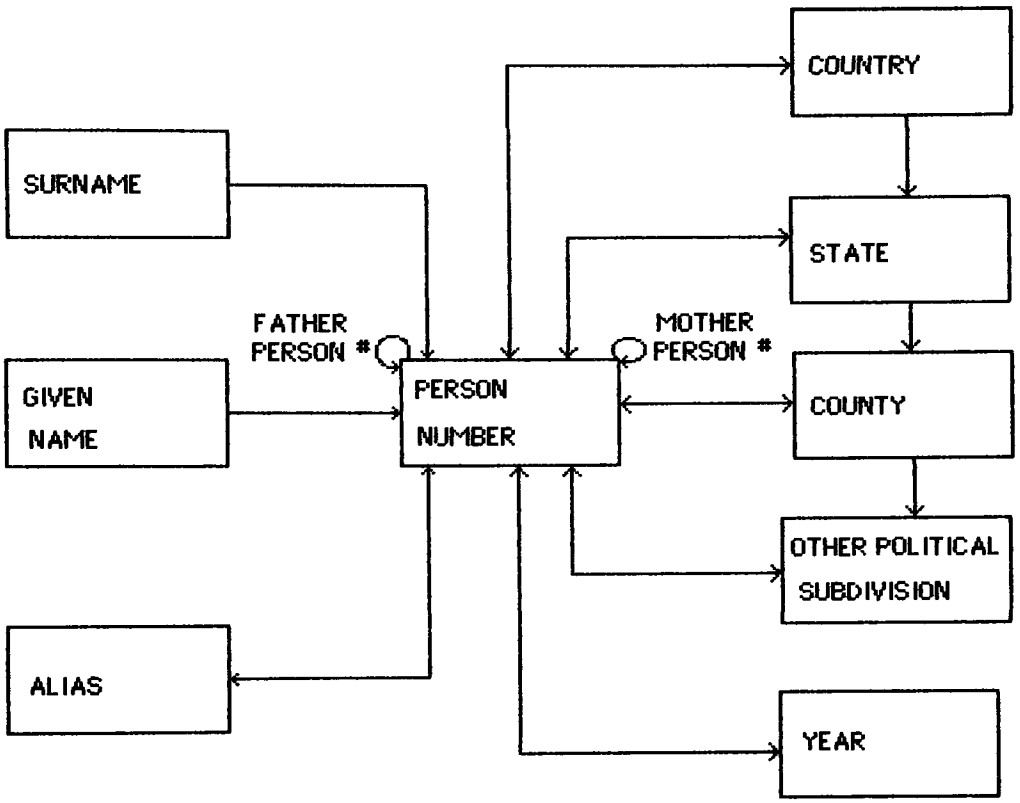


Year is the subject for the optional associator PERSON NUMBER. YEAR and PERSON NUMBER form the subject for the optional non-associator target EVENT. NOTE FOR EVENT is the optional non-associator target of the subject YEAR-PERSON NUMBER-EVENT-NOTE FOR EVENT PERSON NUMBER, EVENT and NOTE FOR EVENT may have multiple values.



The key-level overview summarizes the eleven templates.

KEY-LEVEL OVERVIEW



The relationships indicated by the key-level overview are the following:

PAIR	RELATIONSHIP
surname - person number	1:N
given name - person number	1:N
alias - person number	M:N
father person number - person number	1:N
mother person number - person number	1:N
person number - country	M:N
person number - state	M:N
person number - county	M:N
person number - other political subdivision	M:N
person number - year	M:N
country - state	1:N
state - county	1:N
county - other political subdivision	1:N

The data element definitions include information about the function of the variable named in the data element, conditions for optional inclusion and the domain.

DATA ELEMENT DEFINITIONS

ALIAS

identifies one or more names, other than given name and surname by which person is called or known; not a nickname

example: Mrs. John Jones for wife of John Jones

domain: names

optional conditions: unknown, inapplicable

COUNTRY

names country where event took place; name at time of event

domain: past and current names of countries

optional conditions: unknown

COUNTY

names county, province or other like political subdivision where event took place; name at time of event

domain: past and current names of counties

optional conditions: unknown

DAY

indicates day of month in which event took place

domain: numbers 1 - 31

optional conditions: unknown

EVENT

specifies event about which at least one detail is known and which is associated with a person

domain: events
example: birth, death, marriage

optional conditions: incomplete, unknown

FATHER PERSON NUMBER

indicates person number of father of person of interest

domain: existing person number

optional conditions: unknown

GIVEN NAME

indicates first name and all middle names of a person

domain: names

optional conditions: unknown

MARRIAGE NUMBER OF PERSON

indicates the sequential number of a marriage event of a person

domain: number

optional conditions: inapplicable, unknown

MONTH

indicates abbreviation of month of year in which event took place

domain: Jan, Feb, Mar, Apr, May, Jun, Jul, Aug, Sep, Oct, Nov, Dec

optional conditions: unknown

MOTHER PERSON NUMBER

indicates person number of mother of person of interest

domain: existing person number

optional conditions: unknown

NICKNAME

indicates name given instead of one belonging to a person

example: familiar form of a proper name
Bill for William

name which is descriptive or given in jest
Slim

domain: nicknames

optional conditions: inapplicable , unknown

NOTE FOR EVENT

includes comment about event

domain: comments

optional conditions: inapplicable, unknown

NOTE FOR GIVEN NAME

includes information about middle name(s)

domain: comment

optional conditions: inapplicable, unknown

NOTE FOR YEAR

includes information about imprecision of value of year
example: range, about, after, prior

domain: comment

optional conditions: inapplicable, unknown

OTHER POLITICAL SUBDIVISION

specifies name of town, village, city, township or other comparable political subdivision where event took place

domain: past or current names of local political subdivisions

optional conditions: unknown

PERSON NUMBER

indicates assigned sequential number used to uniquely identify person of interest; assigned when at least one item of information identified

domain: number

mandatory

SPOUSE NUMBER

indicates person number of spouse of person of interest

domain: existing person number

optional conditions: inapplicable, unknown

STATE

specifies name of state, province or other comparable political subdivision where event took place; name at time of event

domain: past and current names of states or provinces

optional conditions: unknown

SURNAME

indicates surname of person of interest; maiden name of married female

domain: surnames

optional conditions: unknown

YEAR

indicates year in which event took place

domain: year no later than current year

optional conditions: unknown

5.4.2 ASSOCIATIVE DATA MODEL - GENEALOGICAL EXAMPLE APPLICATION STATE

The application state is based on the information from the Ancestor Charts and Family Group Sheets in Appendix A. The specific examples of fuzzy data discussed in Chapter 5 Section 3 are included. The format of the information follows:

Line 1: person number
surname
given name
father person number
mother person number

Line 2: note for given name

Line 3: nickname

Line 4: alias

Line 5: birth
birth state
birth county
birth other political subdivision
birth year
birth month
birth day

Line 6: marriage
marriage number
spouse person number
marriage state
marriage county
marriage other political subdivision
marriage year
marriage month
marriage day

Line 7: death
death state
death county
death other political subdivision
death year
death month
death day

It should be noted that all of the above items of information will probably not be available for each individual. The order may vary but the event identifies any particular line.

1 Mygatt Jacob
marriage 1 2

2 Whiting Sarah
marriage 1 1

3 Webster Robert
marriage 1 4

4 Treat Susannah
marriage 1 3

5 Mygatt Joseph 1 2
marriage 1 6 1677 Nov 15
death 1698 ?

6 Webster Sarah 3 4
birth 1655 about
marriage 1 5 1677 Nov 15
death 1743

7 Mygatt Joseph 5 6
birth 1678 Oct 23
death CT Hartford Hartford 1724 Dec 27
marriage 1

9 Mygatt Susannah 5 6
birth 1680 Oct 3
death 1698 after

10 Mygatt Mary 5 6
birth 1682 Dec 4

11 Mygatt Jacob 5 6
birth 1684 Dec 9
death 1685 Jan 29

12 Mygatt Jacob 5 6
birth 1686 Nov 9
death 1687 Nov

13 Mygatt Thomas 5 6
birth 1688 Sept 11
death 1727 May 16

14 Mygatt Sarah 5 6
birth 1691 Mar 9
marriage 1 1721 Nov 6

16 Mygatt Zebulon 5 6
birth 1693 Nov 3
marriage 1

18 Mygatt Dorothy 5 6
birth 1696 Jan 26
marriage 1 1715 May 5
death 1775

20 Batchelder Page
marriage 1 21

21 Hill Elizabeth
marriage 1 20

22 Copp Joshua
marriage 1 23

23 Poore Sarah
marriage 1 22

24 Batchelder Isaiah 20 21
birth NH Rockingham Chester 1749 Mar 2
marriage 1 25 NH Grafton Rumney 1778 June 4
death PA Erie Springfield 1823

25 Copp Molly 22 23
birth NH Rockingham Hampstead 1759 July 15
marriage 1 24 NH Grafton Rumney 1778 June 4
death PA Erie Springfield 1830

26 Batchelder Jonathan 24 25
marriage 1

28 Batchelder Sally 24 25
birth VT Orange Bradford 1784
marriage 1 29 VT Orange Bradford 1802 Mar 4
death PA Erie Springfield 1848 Mar 3

29 Leach James
birth NH 1780
marriage 1 28 VT Orange Bradford 1802 Mar 4
death Canada Ontario Niagara Falls 1814 July 25

30 Batchleder Samuel 24 25
marriage 1
marriage 2

31 Batchelder Susana 24 25
marriage 1

33 Batchelder Betsey 24 25

34 Batchelder Abigail 24 25
marriage 1

36 Batchelder Hannah 24 25
marriage 1

38 Batchelder Mary 24 25
birth 1797
marriage 1
death 1877

40 Batchelder Isaiah 24 25

41 Batchelder Joshua 24 25

42 Batchelder Elmira 24 25
marriage 1

44 Smith Oliver
birth NH 1767
death PA Erie Springfield 1838 May 14
marriage 1
marriage 2 46

45 Smith Chester 44
birth MA 1792
birth NH 1792
death 1880 after

46 Betsy
birth NH 1776
death PA Erie Springfield 1861 Dec 21
marriage 1 44

47 Smith Lyman 44 46
birth NY 1798
death 1875 Sept 3
marriage 1
marriage 2

50 Smith Electa 44 46
birth NY 1799
marriage 1 1822 Apr 25
death 1836 Apr 2

52 Smith Charles Francis 44 46
birth PA Erie Springfield 1803 Mar 20
marriage 1 53 PA Erie Springfield 1825 Nov 24
death PA Erie Springfield 1886 Feb 17

53 Leach Emily Louise 29 28
marriage 1 52 PA Erie Springfield 1825 Nov 24
birth VT 1808 Mar 30
death PA Crawford Linesville 1888 Aug 26

54 Smith Oliver 44 46
birth PA 1805
marriage 1
marriage 2
death 1861 Sept 29

57 Smith Clarissa 44 46
marriage 1
death 1830 prior

59 Smith Miriam 44 46
death 1830 prior

60 Smith Ares 44 46
birth 1805 range
birth 1806 range
birth 1807 range
birth 1808 range
birth 1809 range
birth 1810 range

62 Smith Freeman 44 46
birth PA 1810
marriage 1

63 Harriet
birth 1815
marriage 1 62

64 Smiith John 44 46
birth 1811 range
birth 1812 range
birth 1813 range
birth 1814 range
birth 1815 range
birth 1816 range
birth 1817 range
birth 1818 range
birth 1819 range
birth 1820 range

65 Smith Amos 44 46
birth PA 1815
marriage 1

67 Smith Hiram 47
birth PA

68 Smith Betsy 47
birth PA

69 Smith Oliver A. 47
birth PA 1834
death 1856 Mar 14

70 Smith Harriet M. 47
birth PA

71 Smith Giles L. 47
birth PA

72 Smith George L. 47

73 Leach Betsey 29 28

74 Leach James 29 28

75 Leach Isaiah 29 28

76 Leach Priscilla 29 28

77 Leach Joshua 29 28

78 Smith Betsey 52 53
birth 1827 Jan 28
death 1855 Aug 31
marriage 1

80 Smith Amos 52 53
birth 1829 Aug
death 1900 Sep 19

81 Smith Lathrop Winfield 52 53
birth 1830 May
death 1831 Nov 28

82 Smith Sarah 52 53
birth 1831 Sep 29
marriage 1

84 Smith Aris A. 52 53
birth 1833 Nov 15
death 1852 Apr 22

85 Smith Charles 52 53
birth 1838 Mar 18
death 1923 Feb 11

86 Smith Laura A. 52 53
birth 1840 Apr 8
death 1855 Aug 13

87 Smith Helen 52 53
birth 1842 Aug 11
marriage 1

89 Smith Adelaide Matilda 52 53
Addie
birth 1845 Feb 7
marriage 1 90 1863 Sept 28
death 1876 Dec 12

90 Rundel Theodore Benedict
birth PA Crawford Cussewago 1835 July 24
marriage 1 89 1863 Sep 28
marriage 2 94 1880
death PA Crawford Conneautville 1919 Mar 19

91 Smith James Oliver 52 53
birth 1847 July 22
marriage 1 1873 Apr 21

93 Smith Emily H. 52 53
birth 1851 Jan
death 1852 July 23

94 Nash Flora Augusta
Mrs. Flora Augusta Nash Linder
marriage 1
marriage 2 90 1880

95 Smith ---- E. 62 63
birth PA 1839

96 Smith Cordelia 62 63
birth PA 1843

97 Smith John B. 62 63
birth PA 1844

98 Smith William E. 62 63
birth PA 1847

99 Smith Betsy 62 63
birth CT 1848

SURNAME - PERSON NUMBER

Treat	4
Webster	3
	6
Whiting	2

ALIAS - PERSON NUMBER

Mrs. Flora Augusta Nash Linder 94

FATHER PERSON NUMBER - PERSON NUMBER

1	5	44	45
			47
3	6		50
			52
5	7		54
	9		57
	10		59
	11		60
	12		62
	13		64
	14		65
	16		
	18	47	67
			68
20	24		69
			70
22	25		71
			72
24	26		
	28	52	78
	30		80
	31		81
	33		82
	34		84
	36		85
	38		86
	40		87
	41		89
	42		91
			93
29	53		
	73	62	95
	74		96
	75		97
	76		98
	77		99

MOTHER PERSON NUMBER

2	5	28	53
			73
4	6		74
			75
6	7		76
	9		77
	10		
	11	46	47
	12		50
	13		52
	14		54
	16		57
	18		59
			60
21	24		62
			64
23	25		65
25	26	53	78
	28		80
	30		81
	31		82
	33		84
	34		85
	36		86
	38		87
	40		89
	41		91
	42		93
		63	95
			96
			97
			98
			99

COUNTRY - PERSON NUMBER - EVENT

Canada	29	death	

STATE - PERSON NUMBER - EVENT - NOTE FOR EVENT

CT	7	death	
	99	birth	
MA	45	birth	1850 census, NH 1880 census
NH	24	birth	
		marriage	
	25	birth	
		marriage	
	29	birth	
	44	birth	
	45	birth	1880 census, MA 1850 census
	46	birth	
NY	47	birth	
	50	birth	
PA	24	death	
	25	death	
	28	death	
	44	death	
	46	death	
	52	birth	
		marriage	
		death	
	53	marriage	
		death	
	54	birth	
	62	birth	
	65	birth	
	67	birth	
	68	birth	
	69	birth	
	70	birth	
	71	birth	
	90	death	

PA	95	birth
	96	birth
	97	birth
	98	birth

STATE - PERSON NUMBER - EVENT - NOTE FOR EVENT

VT	28	birth
		marriage
	29	marriage
	53	birth
Ontario	29	death

COUNTY - PERSON NUMBER - EVENT - NOTE FOR EVENT

Crawford	53	death
	90	death
Erie	24	death
	25	death
	28	death
	44	death
	46	death
	52	birth
		marriage
		death
	53	marriage
Grafton	24	marriage
	25	marriage
Hartford	7	death
Orange	28	birth
		marriage
	29	marriage
Rockingham	24	birth
	25	birth

OTHER POLITICAL SUBDIVISION - PERSON NUMBER - EVENT

Bradford	28	birth
		marriage
	29	marriage
Chester	24	birth
Conneautville	90	death
Hampstead	25	birth
Hartford	7	death
Linesville	53	death
Niagara Falls	29	death
Rumney	24	marriage
	25	marriage
Springfield	24	death
	25	death
	28	death
	44	death
	46	death
	52	birth
		marriage
		death
	53	marriage

YEAR - PERSON NUMBER - EVENT - NOTE FOR EVENT

1655	6	birth	about
1677	5	marriage	
	6	marriage	
1678	7	birth	
1680	9	birth	
1682	10	birth	
1684	11	birth	
1685	11	death	

YEAR	PERSON NUMBER	EVENT	NOTE FOR EVENT
1686	12	birth	
1687	12	death	
1688	13	birth	
1691	14	birth	
1693	16	birth	
1696	18	birth	
1698	5	death	?
	9	birth	after
1715	18	marriage	
1721	14	marriage	
1724	7	death	
1727	13	death	
1743	6	death	
1749	24	birth	
1759	25	birth	
1767	44	birth	
1775	18	death	
1776	46	birth	
1778	24	marriage	
	25	marriage	
1780	29	birth	
1784	28	birth	
1792	45	birth	
1797	38	birth	
1798	47	birth	
1799	50	birth	
1802	28	marriage	
	29	marriage	
1803	52	birth	
1805	54	birth	
	60	birth	range 1805-1810
1806	60	birth	range 1805-1810
1807	60	birth	range 1805-1810
1808	53	birth	
	60	birth	range 1805-1810
1809	60	birth	range 1805-1810
1810	60	birth	range 1805-1810
	62	birth	
1811	64	birth	range 1811-1820

YEAR - PERSON NUMBER - EVENT - NOTE FOR EVENT

1812	64	birth	range 1811-1820
1813	64	birth	range 1811-1820
1814	29	death	
	64	birth	range 1811-1820
1815	63	birth	
	64	birth	range 1811-1820
	65	birth	
1816	64	birth	range 1811-1820
1817	64	birth	range 1811-1820
1818	64	birth	range 1811-1820
1819	64	birth	range 1811-1820
1820	64	birth	range 1811-1820
1822	50	marriage	
1823	24	death	
1825	52	marriage	
	53	marriage	
1827	78	birth	
1829	80	birth	
1830	25	death	
	57	death	prior
	59	death	prior
	81	birth	
1831	81	death	
	82	birth	
1833	84	birth	
1834	69	birth	
1835	90	birth	
1836	50	death	
	85	birth	
1838	44	death	
1839	95	birth	
1840	86	birth	
1842	87	birth	
1843	96	birth	
1844	97	birth	
1845	89	birth	
1847	91	birth	
	98	birth	
1848	28	death	
	99	birth	

YEAR - PERSON NUMBER - EVENT - NOTE FOR EVENT

1851	93	birth	
1852	84	death	
	93	death	
1855	78	death	
	86	death	
1856	69	death	
1861	46	death	
	54	death	
1863	89	marriage	
	90	marriage	
1873	91	marriage	
1875	47	death	
1876	89	death	
1877	38	death	
1880	45	death	after
	90	marriage	
	94	marriage	
1886	52	death	
1888	53	death	
1900	80	death	
1919	90	death	
1923	85	death	

5.5 SEMANTIC RELATION DATA MODEL

The first part of this section is a description of the Semantic Relation Data Model for the genealogical example. The second part is an example of an application state for this model based on the data found in Appendix A.

5.5.1 SEMANTIC RELATION DATA MODEL - GENEALOGICAL EXAMPLE DESIGN

Relations for the genealogical example expressed in the format of the Semantic Relation Data Model are (1) Person - Birth - Death - Marriages - Parents and (2) Marriage - Date of Marriage - Place of Marriage. These relations include the information about an individual's birth, marriage, death, parents and biological children.

The Person - Birth - Death - Marriages - Parents Relation is displayed as:

<table>
<tr><td colspan="2">be person: object
born in: agent
born at:agent
died in: agent
died at: agent
has father: agent
has mother: agent
has number marriages: agent</td><td>be birth date . object
born in: object</td></tr>
<tr><td colspan="2">person</td><td>birth date</td></tr>
<tr><td>_id #_</td><td>name</td><td>date</td></tr>
<tr><td>#</td><td>given name
surname
nickname</td><td>year
month
day</td></tr>
<tr><td></td><td></td><td></td></tr>
</table>

Person - Birth - Death - Marriage - Parents continued

be birth place: object born at: object	be death date: object died in: object	be death place: object died at: object
birth place	death date	death place
place name	date	place name
country state county town	year month day	country state county town

Person - Birth - Death - Marriage - Parents continued

be number marriages: object has number marriages: object	be father: object has father: object	be mother: object has mother : object
number marriages	father	mother
number	_id # _	_id # _
number	number	number

The constraints are:

(1) identification number (person, father, mother) is unique
(2) name is any one or more of the following:
    (a) given name
    (b) surname
    (c) nickname
    (d) alias
(3) date is any one or more of the following:
    (a) year (same as or earlier than current year)
    (b) month (name of month or unique abbreviation)
    (c) day (day of month)
(4) date of birth is the same as or earlier than date of death
(5) date of birth or death may be qualified in some way
(6) date of death is the same as or later than date of birth
(7) place is any one or more of the following:
    (a) country (may be non-ambiguous abbreviation)
    (b) state (two character abbreviation)
    (c) county
    (d) town
    (e) village
    (f) city
    (g) township
    (h) other political subdivision

The eight predicates which form the Person - Birth - Death - Marriages - Parents Relation are (1) Person, (2) Date of Birth, (3) Place of Birth, (4) Date of Death, (5) Place of Death, (6) Number Marriages, (7) Father and (8) Mother 'Be' predicates for this relation are (1) be person: object, (2) be birth date: object, (3) be birth place: object, (4) be death date: object, (5) be death place: object, (6) be number marriages: object, (7) be father: object and (8) be mother: object. Other predicate: case pairs are (1) born in: agent, (2) born in: object, (3) born at: agent, (4) born at: object, (5) died in: agent, (6) died in: object, (7) died at: agent, (8) died at: object, (9) has number marriages: agent, (10) has number marriages: object, (11) has father: agent, (12) has father: object, (13) has mother: agent and (14) has mother: object.

The entity types are (1) person, (2) birth date, (3) birth place, (4) death date, (5) death place, (6) number marriages, (7) father and (8) mother The characteristics of person are (1) ID# (identification number) and (2) name. The identification number is an essential characteristic. It is a unique number in the domain of numbers. The domain of name includes given names, surnames, nicknames, and aliases. The domain for date (birth date, death date or marriage date) is year, month and day. The names of countries, states, counties, towns, villages, cities, townships or other political subdivisions form the domain of place names (place of birth, place of death or place of marriage). The characteristic of Number Marriages is number which belongs to the domain of numbers. The entity types father and mother are each characterized by a unique ID#, an essential characteristic from the domain of numbers.

The interpretation of the relation is 'The PERSON with NAME and unique ID# was born in (on) DATE at PLACE NAME and died in (on) DATE at PLACE NAME. PERSON had father with unique ID# and mother with unique ID# PERSON was married NUMBER times'

The Marriage - Date of Marriage - Place of Marriage Relation is displayed on the following page.

Marriage - Date of marriage - Place of marriage

be marriage: object occurred on: agent occurred at: agent					be marriage date: object occurred on: object	be marriage place: object occurred at: object
marriage					marriage date	marriage place
mid#	hid #	wid #	hm#	wm#	date	place name
#	#	#	#	#	year/month/day	country/state/county/town

The constraints for the MARRIAGE - DATE OF MARRIAGE - PLACE OF MARRIAGE Relation are:

(1) marriage identification number (mid #) is unique
(2) husband identification number (hid #) is unique
(3) wife identification number (wid #) is unique
(4) husband's marriages (hm#) are numbered sequentially
(5) wife's marriages (wm#) are numbered sequentially
(6) date is any one or more of the following:
    (a) year (same as or earlier than current year)
    (b) month (name of month or unique abbreviation)
    (c) day (day of month)
(7) date of marriage is the same as or earlier than date of death of husband or date of death of wife
(8) date of marriage may be qualified in some way
(9) place of marriage is any one or more of the following:
    (a) country (may be non-ambiguous abbreviation)
    (b) state (two character abbreviation)
    (c) county
    (d) town
    (e) village
    (f) city
    (g) township
    (h) other political subdivision

The three relations which are combined to form the Marriage - Date of Marriage - Place of Marriage Relation are (1) Marriage, (2) Date of Marriage and (3) Place of Marriage. 'Be' predicates for these relations are (1) be marriage: object, (2) be marriage date: object and (3) be marriage place: object. Additional predicate: case pairs are (1) occurred on: agent, (2) occurred on: object, (3) occurred in: agent, (4) occurred in: object. The entity types are (1) marriage, (2) marriage date and (3) marriage place. The characteristics of marriage are (1) mid # (marriage identification number), (2) hid # (identification number of husband), (3) wid # (identification number of wife), (4) hm# (marriage number for husband), and (5) wm# (marriage number for wife). The marriage numbers for the

husband and wife are included in the relation because they cannot be determined from the marriage date if the marriage date is fuzzy. The marriage identification number (mid #) is an essential characteristic. It is unique in the domain of numbers. The domain of each of the other four characteristics is numbers. The domain of date is year, month and day. The names of countries, states, counties, towns, villages, cities, townships or other political subdivisions form the domain of place name.

The interpretation of the relation is 'The marriage with unique identification number MID # is the HM# marriage for the husband whose identification number is HID # and the WM# marriage for the wife whose identification number is WID # The marriage occurred on DATE at PLACE NAME'

5.5.2 SEMANTIC RELATION DATA MODEL - GENEALOGICAL EXAMPLE APPLICATION STATE

The application state is based on the information from the Ancestor Chart and Family Group Sheets in Appendix A. The specific examples of fuzzy data discussed in Chapter 5 Section 3 are included. To provide a point of reference, the identification numbers are duplicated at the left margin of the continued segments of the Person - Birth - Death - Marriages - Parents Relation. The column headings are also duplicated at the top of the continued segments of the relation.

Person - Birth - Death - Marriages - Parents

be person: object born in: agent born at: agent died in: agent died at: agent has father: agent has mother: agent has number marriages: agent	be birth date . object born in: object

person		birth date
_id #_	name	date
#	given name surname nickname	year month day
1	Jacob Mygatt	-----
2	Sarah Whiting	-----
3	Robert Webster	-----
4	Susannah Treat	-----
5	Joseph Mygatt	-----
6	Sarah Webster	about 1655
7	Joseph Mygatt	1678 Oct 23
9	Susannah Mygatt	1680 Oct 3
10	Mary Mygatt	1682 Dec 4
11	Jacob Mygatt	1684 Dec 9
12	Jacob Mygatt	1686 Nov 9
13	Thomas Mygatt	1688 Sept 11
14	Sarah Mygatt	1691 Mar 9
16	Zebulon Mygatt	1693 Nov 3
18	Dorothy Mygatt	1696 Jan 26
20	Page Batchelder	-----
21	Elizabeth Hill	-----
22	Joshua Copp	-----
23	Sarah Poore	-----
24	Isaiah Batchelder	1749 Mar 2
25	Molly Copp	1759 July 15
26	Jonathan Batchelder	-----
28	Sally Batchelder	1784
29	James Leach	1780

be person: object born in: agent born at: agent died in: agent died at: agent has father: agent has mother: agent has number marriages: agent		be birth date object born in: object
person		birth date
_id #_	name	date
#	given name surname nickname	year month day
30	Samuel Batchelder	-----
31	Susana Batchelder	-----
33	Betsey Batchelder	-----
34	Abigail Batchelder	-----
36	Hannah Batchelder	-----
38	Mary Batchelder	1797
40	Isaiah Batchelder	-----
41	Joshua Batchelder	-----
42	Elmira Batchelder	-----
44	Oliver Smith	1767
45	Chester Smith	1792
46	Betsy	1776
47	Lyman Smith	1798
50	Electa Smith	1799
52	Charles Francis Smith	1803 Mar 20
53	Emily Louise Leach	1808 Mar 30
54	Oliver Smith	1805
57	Clarissa Smith	-----
59	Miriam Smith	-----
60	Ares Smith	1805-1810
62	Freeman Smith	1810
63	Harriet	1815
64	John Smith	1811-1820
65	Amos Smith	1815

be person: object born in: agent born at: agent died in: agent died at: agent has father: agent has mother: agent has number marriages: agent	be birth date object born in: object

person		birth date
_id #_	name	date
#	given name surname nickname	year month day
67	Hiram Smith	-----
68	Betsy Smith	-----
69	Oliver A. Smith	1834
70	Harriet M. Smith	-----
71	Giles L. Smith	-----
72	George L. Smith	-----
73	Betsey Leach	-----
74	James Leach	-----
75	Isaiah Leach	-----
76	Priscilla Leach	-----
77	Joshua Leach	-----
78	Betsey Smith	1827 Jan 28
80	Amos Smith	1829 Aug
81	Lathrop Winfield Smith	1830 May
82	Sarah Smith	1831 Sept 29
84	Aris A. Smith	1833 Nov 15
85	Charles Smith	1836 Mar 18
86	Laura A. Smith	1840 Apr 8
87	Helen Smith	1842 Aug 11
89	Adelaide Matilda Smith	1845 Feb 7
90	Theodore Benedict Rundell	1835 July 24
91	James Oliver Smith	1847 July 22
93	Emily H. Smith	1851 Jan

be person: object born in: agent born at: agent died in: agent died at. agent has father: agent has mother: agent has number marriages: agent		be birth date object born in: object
person		birth date
_id #_	name	date
#	given name surname nickname	year month day
95	--- E. Smith	1839
96	Cordelia Smith	1843
97	John B. Smith	1844
98	William E. Smith	1847
99	Betsy Smith	1848

	be birth place: object born at: object	be death date: object died in: object	be death place: object died at: object
	birth place	death date	death place
	place name	date	place name
	country state county town	year month day	country state county town
1	-----	-----	-----
2	-----	-----	-----
3	-----	-----	-----
4	-----	-----	-----
5	-----	1698?	-----
6	-----	1743	-----
7	-----	1724 Dec 27	CT Hartford Hartford
9	-----	after 1698	-----
10	-----	-----	-----
11	-----	1685 Jan 29	-----
12	-----	1687 Nov	-----
13	-----	1727 May 16	-----
14	-----	-----	-----
16	-----	-----	-----
18	-----	1775	-----
20	-----	-----	-----
21	-----	-----	-----
22	-----	-----	-----
23	-----	-----	-----
24	NH Rockingham Chester	1823	PA Erie Springfield
25	NH Rockingham Hampstead	1830	PA Erie Springfield
26	-----	-----	-----
28	VT Orange Bradford	1848 Mar 3	PA Erie Springfield
29	NH	1814 July 25	Can Ontario Niagara Falls

	be birth place: object born at: object	be death date: object died in: object	be death place: object died at: object
	birth place	death date	death place
	place name	date	place name
	country state county town	year month day	country state county town
30	-----	-----	-----
31	-----	-----	-----
33	-----	-----	-----
34	-----	-----	-----
36	-----	-----	-----
38	-----	1877	-----
40	-----	-----	-----
41	-----	-----	-----
42	-----	-----	-----
44	NH	1838 May 14	PA Erie Springfield
45	MA / NH	after 1880	-----
46	NH	1861 Dec 21	PA Erie Springfield
47	NY	1875 Sept 3	-----
50	NY	1836 Apr 2	-----
52	PA Erie Springfield	1886 Feb 17	PA Erie Springfield
53	VT	1888 Aug 26	PA Crawford Linesville
54	PA	1861 Sept 29	-----
57	-----	prior 1830	-----
59	-----	prior 1830	-----
60	-----	-----	-----
62	PA	-----	-----
63	NY	-----	-----
64	-----	-----	-----
65	PA	-----	-----

	be birth place: object born at: object	be death date: object died in: object	be death place: object died at: object
	birth place	death date	death place
	place name	date	place name
	country state county town	year month day	country state county town
67	PA	-----	-----
68	PA	-----	-----
69	PA	1856 Mar 14	-----
70	PA	-----	-----
71	PA	-----	-----
72	-----	-----	-----
73	-----	-----	-----
74	-----	-----	-----
75	-----	-----	-----
76	-----	-----	-----
77	-----	-----	-----
78	-----	1855 Aug 31	-----
80	-----	1900 Sept 19	-----
81	-----	1831 Nov 28	-----
82	-----	-----	-----
84	-----	1852 Apr 22	-----
85	-----	1923 Feb 11	-----
86	-----	1855 Aug 13	-----
87	-----	-----	-----
89	-----	1876 Dec 12	-----
90	PA Crawford Cussewago	1919 Mar 19	PA Crawford Conneautville
91	-----	-----	-----
93	-----	1852 July 23	-----

Person - Birth - Death - Marriage - Parents (continued)

	be birth place: object born at: object	be death date: object died in: object	be death place: object died at: object
	birth place	death date	death place
	place name	date	place name
	country state county town	year month day	country state county town
95	PA	-----	-----
96	PA	-----	-----
97	PA	-----	-----
98	PA	-----	-----
99	CT	-----	-----

	be number marriages: object has number marriages: object	be father: object has father: object	be mother: object has mother: object
	number marriages	father	mother
	number	_id #_	_id #_
	number	number	number
1	1	-----	-----
2	1	-----	-----
3	1	-----	-----
4	1	-----	-----
5	1	1	2
6	1	3	4
7	1	5	6
9	-----	5	6
10	0	5	6
11	0	5	6
12	0	5	6
13	-----	5	6
14	1	5	6
16	1	5	6
18	1	5	6
20	1	-----	-----
21	1	-----	-----
22	1	-----	-----
23	1	-----	-----
24	1	20	21
25	1	22	23
26	1	24	25
28	1	24	25
29	1	-----	-----

	be number marriages: object has number marriages: object	be father: object has father: object	be mother: object has mother: object
	number marriages	father	mother
	number	_id #_	_id #_
	number	number	number
30	2	24	25
31	1	24	25
33	0	24	25
34	1	24	25
36	1	24	25
38	1	24	25
40	-----	24	25
41	-----	24	25
42	1	24	25
44	2	-----	-----
45	0	44	-----
46	1	-----	-----
47	2	44	46
50	1	44	46
52	1	44	46
53	1	29	28
54	2	44	46
57	1	44	46
59	-----	44	46
60	1	44	46
62	1	44	46
63	1	-----	-----
64	-----	44	46
65	1	44	46

be number marriages: object has number marriages: object	be father: object has father: object	be mother: object has mother: object
number marriages	father	mother
number	_id #_	_id #_
number	number	number

67	-----	47	48
68	-----	47	48
69	-----	47	48
70	-----	47	48
71	-----	47	48
72	-----	47	48
73	-----	29	28
74	-----	29	28
75	-----	29	28
76	-----	29	28
77	-----	29	28
78	1	52	53
80	-----	52	53
81	0	52	53
82	1	52	53
84	-----	52	53
85	-----	52	53
86	-----	52	53
87	1	52	53
89	1	52	53
90	2	-----	-----
91	1	52	53
93	0	52	53

be number marriages: object has number marriages: object	be father: object has father: object	be mother: object has mother: object
number marriages	father	mother
number	_id #_	_id #_
number	number	number
-----	62	63
-----	62	63
-----	62	63
-----	62	63
-----	62	63

MARRIAGE - DATE OF MARRIAGE - PLACE OF MARRIAGE

be marriage: object occurred on: agent occurred at: agent					be marriage date: object occurred on: object	be marriage place: object occurred at: object
marriage					marriage date	marriage place
mid#	hid#	wid#	hm#	wm#	date	place name
#	#	#	#	#	year/month/day	country/state/county/town
1	5	6	1	1	1677 Nov 15	-----
2	24	25	1	1	1778 June 4	NH Grafton Rumney
3	29	28	1	1	1802 Mar 4	VT Orange Bradford
4	52	53	1	1	1825 Nov 24	PA Erie Springfield
5	90	89	1	1	1863 Sept 28	-----
6	90	94	2	2	1880	-----
7	44	46	2	1	-----	-----

5.6 VETTER / MADDISON RELATIONAL MODEL

The first part of this section is a description of the Vetter / Maddison Relational Model for the genealogical example. The second part is an example of an application state for the model based on the data found in Appendix A.

5.6.1 VETTER / MADDISON RELATIONAL MODEL - GENEALOGICAL EXAMPLE DESIGN

There are thirty-one elementary relations describing the environment of interest in the genealogical example. Copies of the work sheets for the four phases of the design process are included in Appendix C. These include the initial specification of the environment, transformation of the specification into elementary relations, digraph, connectivity matrix, list of transitive dependencies and the minimal cover. The list of non-redundant elementary relations is found on the next page.

VETTER / MADDISON RELATIONAL MODEL GENEALOGICAL EXAMPLE

ER1(Person number, Given name)
ER2(Person number, Surname)
ER3(Person number, Nickname)
ER4(Person number, Alias)
ER5(Person number, Birth.year)
ER6(Person number, Birth.month)
ER7(Person number, Birth.day)
ER8(Person number, Birth.country)
ER9(Person number, Birth.state)
ER10(Person number, Birth.county)
ER11(Person number, Birth.OPS)
ER12(Person number, Death.year)
ER13(Person number, Death.month)
ER14(Person number, Death.day)
ER15(Person number, Death.country)
ER16(Person number, Death.state)
ER17(Person number, Death.county)
ER18(Person number, Death.OPS)
ER19(Person number, Father.person number)
ER20(Person number, Mother.person number)
ER21(Person number, Number of marriages)
ER22(Marriage number, Marriage.year)
ER23(Marriage number, Marriage.month)
ER24(Marriage number, Marriage.day)
ER25(Marriage number, Marriage.country)
ER26(Marriage number, Marriage.state)
ER27(Marriage number, Marriage.county)
ER28(Marriage number, Marriage.OPS)
ER29(Marriage number, Husband.person number)
ER30(Marriage number, Wife.person number)
ER31(Husband.person number, Wife.person number)

Note: OPS stands for Other Political Subdivision
(city, town, township, village)

5.6.2 VETTER / MADDISON RELATIONAL MODEL - GENEALOGICAL EXAMPLE APPLICATION STATE

The application state is based on the information from the Ancestor Chart and Family Group Sheets in Appendix A. The specific examples of fuzzy data discussed in Chapter 5 Section 3 are included. The following list is a key to abbreviations used in the example.

ER	elementary relation
F.P#	father person number
H.P#	mother person number
M#	marriage number
M.P#	mother person number
NM	number of marriages
OPS	other political subdivision city town township village
P#	person number
W.P#	wife person number

ER1

P#	GIVEN NAME
1	Jacob
2	Sarah
3	Robert
4	Susannah
5	Joseph
6	Sarah
7	Joseph
8	Elizabeth
9	Susannah
10	Mary
11	Jacob
12	Jacob
13	Thomas
14	Sarah
15	Thomas
16	Zebulon
17	Dorothy
18	Dorothy
19	Jonathan
20	Page
21	Elizabeth
22	Joshua
23	Sarah
24	Isaiah
25	Molly
26	Jonathan
27	Betsy
28	Sally
29	James
30	Samuel
31	Susana
32	John
33	Betsey
34	Abigail
35	Asa
36	Hannah
37	James
38	Mary
39	Rufus
40	Isaiah
41	Joshua
42	Elmira
43	Elam
44	Oliver
45	Chester
46	Betsy
47	Lyman
48	Maria
49	Harriet
50	Electa

ER1

P#	GIVEN NAME
51	William E.
52	Charles Francis
53	Emily Louise
54	Oliver
55	Jane
56	Sarah Ann
57	Clarissa
58	James
59	Miriam
60	Ares
61	John
62	Freeman
63	Harriet
64	John
65	Amos
66	Jane
67	Hiram
68	Betsy
69	Oliver A.
70	Harriet M.
71	Giles L.
72	George L.
73	Betsey
74	James
75	Isaiah
76	Priscilla
77	Joshua
78	Betsey
79	William
80	Amos
81	Lathrop Winfield
82	Sarah
83	John B.
84	Aris A.
85	Charles
86	Laura A.
87	Helen
88	Henry B.
89	Adelaide Matilda
90	Theodore Benedict
91	James Oliver
92	Sarah Isabel
93	Emily H.
94	Flora Augusta
95	----- E.
96	Cordelia
97	John B.
98	William E.
99	Betsy

ER2

P#	SURNAME
1	Mygatt
2	Whiting
3	Webster
4	Treat
5	Mygatt
6	Webster
7	Mygatt
9	Mygatt
10	Mygatt
11	Mygatt
12	Mygatt
13	Mygatt
14	Mygatt
15	King
16	Mygatt
17	Waters
18	Mygatt
19	Steel
20	Batchelder
21	Hill
22	Copp
23	Poore
24	Batchelder
25	Copp
26	Batchelder
28	Batchelder
29	Leach
30	Batchelder
31	Batchelder
32	Lamb
33	Batchelder
34	Batchelder
35	Lamunyon
36	Batchelder
37	Wilson
38	Batchelder
39	Kendall
40	Batchelder
41	Batchelder
42	Batchelder
43	Meekum
44	Smith
45	Smith
47	Smith
50	Smith

ER2

P#	SURNAME
51	Potter
52	Smith
53	Leach
54	Smith
56	Ellis
57	Smith
58	Hunter
59	Smith
60	Smith
61	Gilbert
62	Smith
64	Smith
65	Smith
67	Smith
68	Smith
69	Smith
70	Smith
71	Smith
72	Smith
73	Leach
74	Leach
75	Leach
76	Leach
77	Leach
78	Smith
79	Brockway
80	Smith
81	Smith
82	Smith
83	McIntosh
84	Smith
85	Smith
86	Smith
87	Smith
88	Line
89	Smith
90	Rundel
91	Smith
92	Pettit
93	Smith
94	Nash
95	Smith
96	Smith
97	Smith
98	Smith
99	Smith

ER3

P#	NICKNAME
89	Addie

ER4

P#	ALIAS
60	Aris
94	Mrs. Flora Augusta Nash Linder

ER5

P#	BIRTH.YEAR
6	about 1655
7	1678
9	1680
10	1682
11	1684
12	1686
13	1688
14	1691
16	1693
18	1696
24	1749
25	1759
28	1784
29	1780
38	1797
44	1767
45	1792
46	1776
47	1798
50	1799
52	1803
53	1808
54	1805
60	1805-1810
62	1810
64	1811-1820
65	1815
69	1834
78	1827
80	1829
81	1830
82	1831
84	1833
85	1836
86	1840
87	1842
89	1845
90	1835
91	1847
93	1851
95	1839
96	1843
97	1844
98	1847
99	1848

ER6

P#	BIRTH.MONTH
7	Oct
9	Oct
10	Dec
11	Dec
12	Nov
13	Sep
14	Mar
16	Nov
18	Jan
24	Mar
25	Jul
52	Mar
53	Mar
78	Jan
80	Aug
81	May
82	Sep
84	Nov
85	Mar
86	Apr
87	Aug
89	Feb
90	Jul
91	Jul
93	Jan

ER7

P#	BIRTH.DAY
7	23
9	3
10	4
11	9
12	9
13	11
14	9
16	3
18	26
24	2
25	15
52	20
53	30
78	28
82	29
84	15
85	18
86	8
87	11
89	7
90	24
91	22

ER9

P#	BIRTH.STATE
24	NH
25	NH
28	VT
29	NH
44	NH
45	MA / NH
46	NH
47	NY
50	NY
52	PA
53	VT
54	PA
62	PA
65	PA
67	PA
68	PA
69	PA
70	PA
71	PA
90	PA
95	PA
96	PA
97	PA
98	PA
99	CT

ER10

P#	BIRTH.COUNTY
24	Rockingham
25	Rockingham
28	Orange
52	Erie
90	Crawford

ER11

P#	BIRTH.OPS
24	Chester
25	Hampstead
28	Bradford
52	Springfield
90	Cussewago

ER12

P#	DEATH.YEAR
5	1698?
6	1743
7	1724
9	after 1698
11	1685
12	1687
13	1727
18	1775
24	1823
25	1830
28	1848
29	1814
38	1877
44	1838
45	after 1880
46	1861
47	1875
50	1836
52	1886
53	1888
54	1861
57	prior 1830
59	prior 1830
69	1856
78	1855
80	1900
81	1831
84	1852
85	1923
86	1855
89	1876
90	1919
93	1852

ER13

P#	DEATH.MONTH
7	Dec
11	Jan
12	Nov
13	May
28	Mar
29	Jul
44	May
46	Dec
47	Sep
50	Apr
52	Feb
53	Aug
54	Sep
69	Mar
78	Aug
80	Sep
81	Nov
84	Apr
85	Feb
86	Aug
89	Dec
90	Mar
93	Jul

ER14

P#	DEATH.DAY
7	27
11	29
13	16
28	3
29	25
44	14
46	21
47	3
50	2
52	17
53	26
54	29
69	14
78	31
80	19
81	28
84	22
85	11
86	13
89	12
90	19
93	23

ER15

P#	DEATH.COUNTRY
29	Canada

ER16

P#	DEATH.STATE
7	CT
24	PA
25	PA
28	PA
29	Ontario
44	PA
46	PA
52	PA
53	PA
90	PA

ER17

P#	DEATH.COUNTY
7	Hartford
24	Erie
25	Erie
28	Erie
44	Erie
46	Erie
52	Erie
53	Crawford
90	Crawford

ER18

P#	DEATH.OPS
7	Hartford
24	Springfield
25	Springfield
28	Springfield
29	Niagara Falls
44	Springfield
46	Springfield
52	Springfield
53	Linesville
90	Conneautville

ER19

P#	F.P#
5	1
6	3
7	5
9	5
10	5
11	5
12	5
13	5
14	5
16	5
18	5
24	20
25	22
26	24
28	24
30	24
31	24
33	24
34	24
36	24
38	24
40	24
41	24
42	24
45	44
47	44
50	44
52	44
53	29
54	44
57	44
59	44

ER19

P#	F.P#
60	44
62	44
64	44
65	44
67	47
68	47
69	47
70	47
71	47
72	47
73	29
74	29
75	29
76	29
77	29
78	52
80	52
81	52
82	52
84	52
85	52
86	52
87	52
89	52
91	52
93	52
95	62
96	62
97	62
98	62
99	62

ER20

P#	M.P#
5	2
6	4
7	6
9	6
10	6
11	6
12	6
13	6
14	6
16	6
18	6
24	21
25	23
26	25
28	25
30	25
31	25
33	25
34	25
36	25
38	25
40	25
41	25
42	25
47	46
50	46
52	46
53	28
54	46
57	46
59	46

ER20

P#	M.P#
60	46
62	46
64	46
65	46
67	48
68	48
69	48
70	48
71	48
72	48
73	28
74	28
75	28
76	28
77	28
78	53
80	53
81	53
82	53
84	53
85	53
86	53
87	53
89	53
91	53
93	53
95	63
96	63
97	63
98	63
99	63

ER21	
P#	NM
1	1
2	1
3	1
4	1
5	1
6	1
7	1
8	1
10	0
11	0
12	0
14	1
16	1
18	1
20	1
21	1
22	1
23	1
24	1
25	1
26	1
27	1
28	1
29	1
30	2
31	1
32	1
33	0
34	1
35	1
36	1
37	1
38	1

ER21	
P#	NM
39	1
42	1
43	1
44	2
45	0
46	1
47	2
50	1
51	1
52	1
53	1
54	2
55	1
56	1
57	1
58	1
60	1
61	1
62	1
63	1
65	1
66	1
78	1
81	0
82	1
83	1
87	1
88	1
89	1
90	2
91	1
92	1
93	0

ER22

M#	MARRIAGE.YEAR
1	1677
2	1778
3	1802
4	1825
5	1863
6	1880
8	1715
9	1822

ER23

M#	MARRIAGE.MONTH
1	Nov
2	Jun
3	Mar
4	Nov
5	Sep
8	May
9	Apr

ER24

M#	MARRIAGE.DAY
1	15
2	4
3	4
4	24
5	28
8	5
9	25

ER26

M#	MARRIAGE.STATE
2	NH
3	VT
4	PA

ER27

M#	MARRIAGE.COUNTY
2	Grafton
3	Orange
4	Erie

ER28

M#	MARRIAGE.OPS
2	Rumney
3	Bradford
4	Springfield

ER29

M#	H.P#
1	5
2	24
3	29
4	52
5	90
6	90
7	44
8	19
9	51

ER30

M#	W.P#
1	6
2	25
3	28
4	53
5	89
6	94
7	46
8	18
9	50

ER31

H.P#	W.P#
5	6
19	18
24	25
44	46
51	50
52	53
90	89
90	94

CHAPTER SIX

EVALUATION OF THREE LOGICAL DATA MODELING APPROACHES

6.1 INTRODUCTION

This chapter includes an evaluation of the Associative Data Model, the Semantic Relation Data Model and the Vetter / Maddison Relational Model. Each of these models is described in Chapter Four of this paper The genealogical example described in Chapter Five was modeled by each of the approaches. The resulting designs are described in Chapter Five. The genealogical example includes the types of fuzzy data described in Chapter Two. The examples of fuzziness are summarized as follows for use in the evaluations in this chapter:

TYPE OF FUZZY DATA	EXAMPLE
imprecise	range of years year and descriptor
inapplicable	no middle name unmarried
incomplete	date no surname middle initial
inconsistent	multiple place names
unknown	date spouse name place name

The evaluation criteria are described in Chapter Three. Criteria for the logical data base design are:

(1) convenient
(2) expressive
(3) minimally redundant

Criteria for the design process are:

(1) complete
(2) consistent
(3) efficient
(4) error-free
(5) implementation independent
(6) precise
(7) theoretically sound
(8) useable

When evaluating each of these approaches it is assumed that the designer understands the environment of interest and is able to identify the objects of interest and their relationships. If this is not the case, it is likely that the designer will introduce inconsistencies in the design as successive iterations of the design process are made.

6.2 ASSOCIATIVE DATA MODEL

The design for the genealogical example is described in Chapter Five Section 4.1. The design process for the Associative Data Model is described in Chapter Four Section 2.2. The application state for the genealogical data listed in Appendix A is found in Chapter Five Section 4.2.

6.2.1 ASSOCIATIVE DATA MODEL - DESIGN

The design for the genealogical example based on the Associative Data Model is (1) convenient, (2) expressive and (3) minimally redundant.

CONVENIENT

The key-level overview and assertion templates (original and derived through application of the Associator Reversal Rule) are good visual descriptions of the environment of interest. Complexity of the design can be described by use of cascades and multiple targets. The roles of objects are clearly stated through the use of associators and non-associators. The data element definitions provide necessary detail. Designation of associators and non-associators as optional or mandatory allows the environment of interest to be precisely described.

EXPRESSIVE

The design is very expressive. The interpretation of the design is explicit. The assertion templates, key-level overview and data element definitions provide a complete, straight-forward description of the environment of interest. There are a reasonable number of constructs and

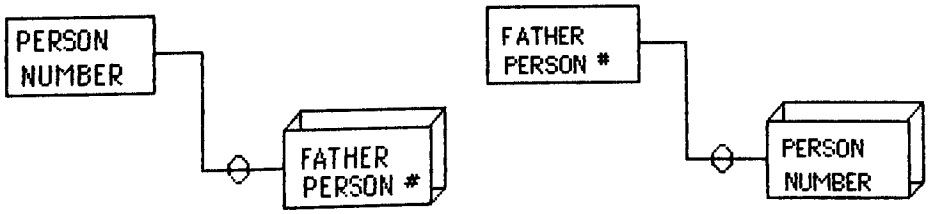
rules governing the use of the constructs. The fuzzy data in the genealogical example are represented in a meaningful way. The following table summarizes the features of the Associative Data Model which were used to model the types of fuzzy data described in Chapter Two.

FUZZY DATA EXAMPLE	MODELED	RELEVANT DESIGN FEATURE
IMPRECISE		
range of years	yes	multiple-valued associator
year and descriptor	yes	optional non-associator
INAPPLICABLE		
no middle initial	yes	optional non-associator
unmarried	yes	optional non-associator data element definition
INCOMPLETE		
date	yes	optional non-associator optional associator
no surname	yes	optional associator
middle initial	yes	data element definition
INCONSISTENT		
multiple place name	yes	multiple-valued associator
UNKNOWN		
date	yes	optional non-associator optional associator
spouse name	yes	optional non-associator
place name	yes	optional associator

Inclusion of imprecise data such as a range of years is by means of the multiple-valued associator YEAR. Each year of the range appears as a value of the associator. The non-asssociator SPECIFIC CHARACTERISTIC OF YEAR accommodates a modifier such as 'range'. Other instances of imprecise data can be modified by 'after', 'prior', and 'about' as necessary.

Marriage data for a person who was never married is inapplicable. The use of 'zero' for a value of MARRIAGE NUMBER OF PERSON distinguishes this case from the case of a person who was married and for whom no marriage data are entered. By definition, a given name may consist of any one or more of a first name, middle name or initials. Use of the optional non-associator NOTES provides a means to indicate when there is no middle name. Incomplete data can be accommodated by appropriate data element definition or through the use of optional associators. The example of a middle initial instead of a complete middle name is a case of the former, no surname of the latter. Multiple place names, an example of inconsistent data, are included through the use of optional multiple-valued associators STATE, COUNTY and OTHER POLITICAL SUBDIVISION. Optional associators and optional non-associators allow the model to represent the environment of interest even though all data are not known at a given moment in time.

This model is very flexible. It can be modified easily to represent slightly different environments. For instance, assume that several persons are cited as the father of an individual. Changing the optional associator FATHER PERSON NUMBER to an optional multiple-valued associator reflects this understanding of the environment. The following relationships would then be included in the assertion templates.



Another example is the use of the multiple-valued non-associator EVENT

In addition to information about birth, marriage and death the environment of interest can be expanded to include such items as schools attended and places lived. Addition of an optional multiple-valued non-associator to each associator or non-associator would make it possible to included the source of the data, a valuable research aid.

MINIMALLY REDUNDANT

The design is minimally redundant. None of the attributes is derived.

6.2.2 ASSOCIATIVE DATA MODEL - DESIGN PROCESS

The design process for the Associative Data Model is (1) complete, (2) consistent, (3) efficient, (4) error-free, (5) implementation independent, (6) precise, (7) theoretically sound and (6) useable.

COMPLETE

This design process is complete. It is necessary to determine whether entities should be represented by associators or non-associators, whether they are optional or mandatory, whether they are single-valued or multiple-valued, and how targets are related to the keys. Particular care needs to be taken in step 2 of the design process (identification of relationships between objects of interest) when modeling fuzzy data. In this step pairs of objects which are related in some way are identified. These relationships are drawn in the assertion templates. Determination of the characteristics of the target allow fuzzy data to be included in an appropriate way. For example, use of the multiple-valued associator STATE facilitates the inclusion of inconsistent data such as a place of birth known to be Massachusetts or New Hampshire. Another example is the selection of the optional indicator which specifies that the data may or may not be present. This is appropriate for unknown or inapplicable data. The optional indicator can be used with a non-associator such as SPECIAL CHARACTERISTIC OF YEAR to describe imprecise data such as a range of years or a year modified by 'about', 'after', 'prior' and so forth. The Associator Reversal Rule is useful in looking at the relationships from more than one perspective. The data element definitions represent an additional level of detail. Redundant data can be identified when the Associator Reversal Rule is applied or when compound templates are

created. There do not appear to be limits on the types of data nor the range of problems which can be modeled by this design process. The key-level overview may indicate inconsistencies in the design or areas needing additional design review.

CONSISTENT

The steps necessary to implement the Associator Reversal Rule and to draw the key level overview make the process consistent from one iteration to the next for these activities. Complexity arises when applying the Associator Reversal Rule to cascades. However, if the definitions of subject and target are remembered, the additional assertion templates can be determined consistently. The information to be included in a data element definition is clearly stated. Therefore, there is consistency in the level of detail in the data element definitions.

EFFICIENT

For the genealogical example the design process was not tedious. Few iterations were required and the time necessary for each iteration was not excessive.

ERROR-FREE

The determination of whether to include an object or a relationship in the design is made by the designer The details of how the object or relationshiop will be included are also determined by the designer The

application of the Associator Reversal Rule forces the designer to consider the impact of his decisions on the design and provides the opportunity for the designer to modify the design to fully describe the environment of interest. The design process does not introduce errors. Errors can, however, be introduced by the designer if the definitions of subject and target are ignored or if the Associator Reversal Rule is not followed.

IMPLEMENTATION INDEPENDENT

The design process is not dependent on the prior selection of implementation hardware or software. The process is independent of any individual application.

PRECISE

The steps used to apply the Associator Reversal Rule and to draw the key-level overview are clear However, it would be helpful to have a list to follow when drawing assertion templates. This list would include (1) consider one subject-target pair at a time, (2) identify target as associator or non-associator, (3) determine if target is single-valued or multiple-valued, (4) determine if target is optional or mandatory, (5) if appropriate, identify cascade associated with target, (6) apply Associator Reversal Rule, (7) combine templates which have the same key and (8) review all steps. The information to be included in the data element definition is clearly stated.

THEORETICALLY SOUND

The design process draws from graph theory and set theory. Multiple values of a data variable, associator or non-associator, are considered sets. An assertion template is a tree structure of a set of domain names. The template describes the corresponence between the subject and its target.

USEABLE

The design process is useable. It is easy to use and not tedious for the genealogical example. It is straight-forward. The process is not too complex and the number of rules is not excessive. The vocabulary is not excessively large. Concepts and terminology are clearly defined. The assertion templates and key-level overview are generated as part of the process. These graphic descriptions are a clear visual representation of the environment of interest.

6.3 SEMANTIC RELATION DATA MODEL

The design for the genealogical example based on the Semantic Relation Data Model is described in Chapter Five Section 5.1. The design process for the Semantic Relation Data Model is described in Chapter Four Section 3.2. The application state for the data displayed in Appendix A is found in Chapter Five Section 5.2.

6.3.1 SEMANTIC RELATION DATA MODEL - DESIGN

The design for the genealogical example based on the Semantic Relation Data Model is not so convenient nor expressive as desired. It is minimally redundant.

CONVENIENT

The design is not convenient. The design is displayed as a table. The tabular display should make the design clear and convenient to use. However, the table for this problem was unwieldy. The number of predicates is so large that the display is fragmented and the sense of the overall design is lost. It is difficult to see the total picture. The predicates are too general to be meaningful to the user. For example, the domain of the entity type PLACE NAME is country, state, county or other political subdivision. This definition of the domain accommodates imprecise, incomplete or inconsistent data. As an alternative, assume that state, county and town have separate domains. If any one of them has the null value, then each of them has the null value and the incomplete data are lost.

EXPRESSIVE

The design is limited in describing the environment of interest. It is not precise. It is confusing. The terminology is not familiar. The handling of fuzzy data is forced. Identification of the characteristics of the entity types was greatly influenced by the fuzzy data being modeled. The restriction of the model that a null value for one of several characteristics of an entity type in a relation forces all characteristics to have the null value resulted in the identification of very general characteristics. For example, NAME was identified as the characteristic instead of first name, middle name and surname. This allowed incomplete data to be included in the design. Definition of the domains was also influenced by the modeling of fuzzy data. Domains had to be defined to include any of the possible values of interest, often a composite of several more specific domains. For instance, the date domain is a composite of the domain of year values, month values and day values. This was necessary to accommodate incomplete data. Constraints are essential when modeling fuzzy data. In this problem no restriction was made to limit the value of YEAR to a single value. In this way a range of years, imprecise data, was accommodated. The constraint regarding qualification of the date in some way also addresses the case of imprecise data in the form 'about 1655' or 'after 1800'.

The following table summarizes the features of the Semantic Relation Data Model which were used to model the types of fuzzy data described in Chapter Two.

FUZZY DATA EXAMPLE	MODELED	RELEVANT DESIGN FEATURE
IMPRECISE		
range of years	yes	constraints
year and descriptor	yes	constraints
INAPPLICABLE		
no middle initial	no	
unmarried	yes	domain
INCOMPLETE		
date	yes	characteristic domain constraints
no surname	yes	characteristic domain constraints
middle initial	yes	characteristic domain constraints
INCONSISTENT		
multiple place name	yes	constraints
UNKNOWN		
date	yes	characteristic domain constraints
spouse name	yes	relation (HAS NUMBER MARRIAGES)
place name	yes	constraints

Selection of appropriate predicates and names of predicates in the predicate: case pairs was awkward. For example, one of several different forms might be used when discussing dates. Examples are 'John was born in 1700', 'John was born May 23, 1700', or 'John was born on May 23rd'. The place of birth may also be expressed in the form 'John was born in New

York' A decision was made to use predicates of the form BORN IN for date and BORN AT for places.

For the design to be meaningful it is necessary to understand multiple predicates, predicate: case pairs and the use of the null value. Familiarity with the Relational Model of Codd increased the time and effort needed to understand multiple predicates in this model.

MINIMALLY REDUNDANT

There is minimal redundancy when multiple predicates are used. Otherwise, the BE PERSON predicate would need to be repeated for each of the following predicates: (1) BORN IN, (2) BORN AT, (3) DIED IN, (4) DIED AT, (5) HAS FATHER, (6) HAS MOTHER and (7) HAS NUMBER MARRIAGES.

6.3.2 SEMANTIC RELATION DATA MODEL - DESIGN PROCESS

The design process for the Semantic Relation Data Model is implementation independent. It is not so complete, convenient, efficient, error-free, theoretically sound nor useable as desired.

COMPLETE

The design process is not so complete as desired. It is not thorough. There are no rules to apply. Details are missing. It was necessary for the designer to identify the steps necessary to create the design. Redundant data are not readily identified in the process. The process is not comprehensive. It is not effective for modeling fuzzy data.

CONSISTENT

Definition by the designer of steps to follow in the design process reduced inconsistencies in the several iterations necessary to complete the design and reduced the ambiguity of the process.

EFFICIENT

The process was not efficient. Time was needed to understand the model and to define and order the steps necessary to create the design. There were many iterations. Familiarity with the Relational Model of Codd increased the time and effort needed to understand multiple predicates in the model.

ERROR-FREE

There are no checks built into the process. Therefore, it is difficult to identify any errors which might be introduced in a given iteration. Identification of predicate: case pairs and entity types depends on the designer's understanding of the environment. It is easier to detect errors introduced when identifing characteristics, domains or constraints. There is no guarantee that redundant data will be identified.

IMPLEMENTATION INDEPENDENT

The process is implementation independent. It is independent of hardware and software considerations. There is a conflict in the use of this design process when some applications have fuzzy data and some do not.

PRECISE

The process was not precise. In the beginning there were no rules to follow. There was ambiguity in how to proceed to create the design.

THEORETICALLY SOUND

The design process does not depend heavily on scientifically acceptable principles. The concept of predicate: case pairs is based on the notion of case grammar. There are no particular procedures for organizing the predicate: case pairs into relations. Two or more relations can be combined to form a more complex relation. When combining relations it is

important to consider the effect of the null value in the new relation. Combining relations can reduce the redundancy in the overall design.

USEABLE

It was difficult to use this design process. The concepts and terminology were not clearly nor concisely defined. It was confusing. The relation formed with multiple predicates to reduce redundancy was unwieldy. The assumptions necessary to accommodate fuzzy data decreased the usefulness of the final design. The process was tedious. The process was forced to accommodate fuzzy data.

6.4 VETTER / MADDISON RELATIONAL MODEL

The design for the genealogical example is described in Chapter Five Section 6.1. The design process for the Vetter / Maddison Relational Model is described in Chapter Four Section 4.2. The application state for the genealogical data listed in Appendix A is found in Chapter Five Section 6.2.

6.4.1 VETTER /MADDISON RELATIONAL MODEL - DESIGN

The design describing the genealogical example in terms of the Vetter / Maddison Relational Model is minimally redundant. It is less convenient and expressive than desired.

CONVENIENT

The design is a list of elementary relations. A list is not graphic. It is difficult to understand the overall environment of interest when reading a list. The number of elementary relations can become very large. This increases the difficulty of understanding the design. The design is not ambiguous. It is necessary to refer to multiple elementary relations for complete information concerning dates and places.

EXPRESSIVE

The design is expressed in terms of elementary relations. Some persons may not be familiar with elementary relations. The definition of domain was forced to accommodate the fuzzy data of the genealogical example. For instance, the domain of year was defined to include a range of years and modifiers. This was necessary to accommodate imprecise data.

The design does not distinguish between the case where a middle initial in the given name is missing and the case where there is no middle initial., inapplicable data. Nor does the design distinguish between the case where the name of a spouse is unknown and the case where the spouse name is missing. All views are not described by the design. For example, the researcher is interested in knowing the identification of all persons with a given surname. The following table summarizes the features of the Vetter / Maddison Relational Model which were used to model the types of fuzzy data described the Chapter Two.

FUZZY DATA EXAMPLE	MODELED	RELEVANT DESIGN FEATURE
IMPRECISE		
range of years	yes	domain definition
year and descriptor	yes	domain definition
INAPPLICABLE		
no middle initial	no	
unmarried	yes	domain definition
INCOMPLETE		
date	yes	elementary relation
no surname	yes	elementary relation
middle initial	yes	domain definition
INCONSISTENT		
multiple place name	yes	domain definition
UNKNOWN		
date	yes	elementary relation
spouse name	no	
place name	yes	elementary relation

MINIMALLY REDUNDANT

The design is minimally redundant. None of the attributes is derived.

6.4.2 VETTER /MADDISON RELATIONAL MODEL - DESIGN PROCESS

The design process for the Vetter / Maddison Relational Model is (1) complete, (2) consistent, (3) error-free, (4) implementation independent and (5) theoretically sound. It is less efficient, less precise, and less useable than desired.

COMPLETE

The design process is completely described. For larger problems there may be a desire to terminate a step before it is complete. This is particularly true when reviewing the transitive dependencies. Use of the digraph helps identify redundant data. The process is comprehensive. It can be used to model a range of data problems. The identification of the domains of the properties of the entity sets in Phase I of the design process was influenced by the fuzzy data to be modeled. For example, the domain for year is a single year (1700), a range of years (1805-1810) or a single year and a modifier (after 1880). The domain for state is not restricted to a single state. Both domains accommodate inconsistent data. Recognizing that zero is part of the domain of number of marriages accommodates the information that a person was never married.

CONSISTENT

There is consistency in using the process for several iterations. The process is not ambiguous.

EFFICIENT

It is important to spend the time necessary, possibly lengthy, to understand the phases in the process and the steps within each phase. To determine all of the minimal covers requires several iterations. Each iteration tends to be lengthy.

ERROR-FREE

The details of implementation of the process are consistent and not contradictory. To keep error to a minimum careful attention should be given to drawing dependencies in the digraph and to using the connectivity matrix.

IMPLEMENTATION INDEPENDENT

The process is not dependent on the prior selection of implementation hardware or software. The process is independent of any individual application.

PRECISE

Some of the phases and steps within the phases are clearly and exactly defined. However, the process is less specific about the details of the steps necessary to work with the connectivity matrix and to derive the transitive closure.

THEORETICALLY SOUND

The design process is theoretically sound. It makes use of concepts from set theory and graph theory.

USEABLE

The connectivity matrix and digraph provide a visual description of the environment. The design process is tedious. There are a number of rules and constructs. The process is complex and confusing. A large vocabulary is necessary. The design process is useable if the designer prefers a relational model, is willing to learn the vocabulary and rules, is willing to become familiar with the constructs and is willing to perform many tedious iterations.

CHAPTER SEVEN

CONCLUSION

This chapter includes comments and recommendations about the conceptualization phase of a design process and the resulting logical data model when it is known that some of the data in the environment of interest are fuzzy.

Increased understanding of the nature of fuzzy data and of the design of a logical data model resulted from the following: (1) investigation of the types of fuzzy data, (2) identification of criteria for evaluating a design process and its logical data model and (3) use of three design processes to model fuzzy data. The following approaches to modeling data are described in Chapter Four: (1) Associative Data Model, (2) Semantic Relation Data Model and (3) Vetter / Maddison Relational Model. In Chapter Five the three logical data models of the genealogical example found in Appendix A are described. The three data modeling processes and the three resulting logical data models are evaluated in Chapter Six.

The original reasons for selecting the three approaches to data modeling discussed in this paper are summarized as follows:

ASSOCIATIVE DATA MODEL

specification of optional variables
specification of domain types

SEMANTIC RELATION DATA MODEL

specification of constraints
interpretation of relations

application of a variety of techniques in the design process

These reasons are not adequate for use in selecting a design process and the associated logical data model. They are not sufficiently comprehensive to evaluate a potential process and its resulting logical data model.

Criteria of a 'good' logical data model and a 'good' design process are defined in Chapter Three. The logical data model should be: (1) convenient, (2) expressive and (3) minimally redundant. Selection of a design process should be made after application of the following criteria: (1) complete, (2) consistent, (3) efficient, (4) error-free, (5) implementation independent, (6) precise, (7) theoretically sound and (8) useable. Identification of a logical data model using the criteria for a 'good' design may dictate the selection of a design process which would not be identified when using the criteria for a 'good' design process. The final logical data model is determined by the design process. The selection of a logical data model is not made independently from the selection of a design process.

The selection of the best combination of logical data model and design process which meets the criteria described in Chapter Three depends on the area of application, problem structure, expertise of the designer and preference of the user. The area of application and the problem structure are especially important when modeling fuzzy data, data which are (1) imprecise, (2) inapplicable, (3) incomplete, (4) inconsistent or (5) not known. For any application it is important to

recognize specific cases of each type of fuzzy data and how they fit into the environment of interest. Different types of fuzzy data may be included in the data model in different ways. The design must be meaningful, clear and easy to use when modeling fuzzy data.

The expertise of the designer affects the size of the vocabulary and the number of rules which are acceptable in a design process. The number of constructs which are acceptable also depends on the expertise of the designer A more experienced designer may accept a larger vocabulary and a greater number of rules and constructs. However, there are limits to the amount of detail and complexity which even the experienced designer will tolerate in a logical data model or design process. The user's preference for the final logical data model may depend on having a graphic description of the problem or a simple, easy to use representaion expressed in familiar terms. The preference of the user for a particular design process may depend on the number of major steps, the number of steps within a given major step, the number of iterations required to complete a design or the amount of time required for each iteration.

The Associative Data Model described in Chapter Four Section 4.2 was effective for modeling the fuzzy data associated with the genealogical example. This design process and the resulting logical data model met the above criteria more closely than the design process and logical data model of either the Semantic Relation Data Model or the Vetter / Maddison Relational Model.

It is recommended that the following steps guide the designer when modeling fuzzy data:

(1) identify examples of each type of fuzzy data in the environment of interest
(2) define subset of environment of interest which includes examples of each type of fuzzy data identified in (1)
(3) use criteria to identify potential logical data model(s)
(4) use criteria to identify potential design process(es)
(5) identify pairs of logical data model and design process for use in modeling subset of environment defined in (2)
(6) use selected process(es) to model subset of the environment defined in (2)
(7) evaluate application of process(es) using design process criteria
(8) evaluate resulting logical data model(s) using design criteria
(9) select 'best' combination of logical data model and design process to model environment of interest
(10) model the entire environment of interest using the 'best' logical data model and design process combination selected in (9)

APPENDIX A

Ancestor Chart

Name _____

Address _____

City, State _____

Date _____

Person No. 1 on this chart is the same person as No. ____ on chart No. ____

CHART NO. _____

(Father of 1)
Charles Francis Smith
Born 20 Mar 1803
At Springfield, PA
Died 17 Feb 1886
At Springfield, PA
Res.

(Mother of 1)
Emily Louise Leach
Born 30 Mar 1808
At VT
Died 26 Aug 1888
At Linesville, PA
Mar. 24 Nov 1825
At Springfield, PA

Theodore Benedict Rundel
Name of Spouse

Adelaide Matilda Smith
Born 7 Feb 1845
At
Died 12 Dec 1876
At
Res.

4 (Father of 2) Oliver Smith
Born 1767
At NH
Died 14 May 1838
At Springfield, PA
Res.

5 (Mother of 2) Betsy
Born 1776
At NH
Died 21 Dec 1861
At Springfield, PA
Mar.
At

6 (Father of 3) James Leach
Born 1780
At NH
Died 25 July 1814
At
Res.

7 (Mother of 3) Sally Batchelder
Born 1784
At Bradford, VT
Died 3 Mar 1848
At Springfield, PA
Mar. 4 Mar 1802
At Bradford, VT

8
Born
At
Died
At
Res.
Cont. on chart # ____

9
Born
At
Died
At
Mar.
At
Cont. on chart # ____

10
Born
At
Died
At
Res.
Cont. on chart # ____

11
Born
At
Died
At
Mar.
At
Cont. on chart # ____

12
Born
At
Died
At
Res.
Cont. on chart # ____

13
Born
At
Died
At
Mar.
At
Cont. on chart # ____

14 Isaiah Batchelder
Born 2 Mar 1749
At Chester, NH
Died 1823
At Springfield, PA
Res.
Cont. on chart # ____

15 Molly Copp
Born 15 July 1759
At Hampstead, NH
Died 1830
At Springfield, PA
Mar. 4 June 1778
At Rumney, NH
Cont. on chart # ____

SHOW

At	Town, State
Born	Date of Birth
Died	Date of Death
Mar.	Date of Marriage
Res.	Town, State of Residence

FAMILY GROUP No. ____ Husband's Full Name Joseph Mygatt

This Information Obtained From:

Husband's Data	Day Month Year	City, Town or Place	County or Province, etc.	State or Country	Add. Info. on Husband
Birth					
Chr'nd					
Mar.	15 Nov 1677				
Death					1698?
Burial					

Places of Residence

Occupation Church Affiliation Military Rec.

Other wives, if any. No. (1) (2) etc. Make separate sheet for each mar.

His Father Jacob Mygatt Mother's Maiden Name Sarah Whiting

Wife's Full Maiden Name Sarah Webster

Wife's Data	Day Month Year	City, Town or Place	County or Province, etc.	State or Country	Add. Info. on Wife
Birth					~ 1655
Chr'nd					
Death	1743				
Burial					

Compiler

Address

City, State

Date

Places of Residence

Occupation if other than Housewife Church Affiliation

Other husbands, if any. No. (1) (2) etc. Make separate sheet for each mar.

Her Father Robert Webster Mother's Maiden Name Susannah Treat

ex	Children's Names in Full (Arrange in order of birth)	Children's Data	Day Month Year	City, Town or Place	County or Province, etc.	State or Country	Add. Info. on Children
1	Joseph	Birth	23 Oct 1678				
	Full Name of Spouse*	Mar.					
	Elizabeth	Death	27 Dec 1724	Hartford		CT	
		Burial					
2	Susannah	Birth	3 Oct 1680				
	Full Name of Spouse*	Mar.					
		Death					after 1698
		Burial					
3	Mary	Birth	4 Dec 1682				
	Full Name of Spouse*	Mar.					um 1729
		Death					
		Burial					
4	Jacob	Birth	9 Dec 1684				
	Full Name of Spouse*	Mar.					
		Death	29 Jan 1685				
		Burial					
5	Jacob	Birth	9 Nov 1686				
	Full Name of Spouse*	Mar.					
		Death	Nov 1687				
		Burial					
6	Thomas	Birth	11 Sept 1688				
	Full Name of Spouse*	Mar.					
		Death	16 May 1727				
		Burial					
7	Sarah	Birth	9 Mar 1691				
	Full Name of Spouse*	Mar.	6 Nov 1721				
	Thomas King	Death					
		Burial					
8	Zebulon	Birth	3 Nov 1693				
	Full Name of Spouse*	Mar.					
	Dorothy Waters	Death					
		Burial					
9	Dorothy	Birth	26 Jan 1696				
	Full Name of Spouse*	Mar.	5 May 1715				
	Jonathan Steel	Death	1775				
		Burial					
10		Birth					
	Full Name of Spouse*	Mar.					
		Death					
		Burial					

*If married more than once No. each mar. (1) (2) etc. and list in "Add. info. on children" column. Use reverse side for additional children, other notes, references or information.

FAMILY GROUP No. _____ Husband's Full Name Isaiah Batchelder

This Information Obtained From:

Husband's Data	Day Month Year	City, Town or Place	County or Province, etc.	State or Country	Add. Info. on Husband
Birth	2 mar 1749	Chester	Rockingham	NH	
Chr'nd					
Mar.	4 June 1778	Rumney	Grafton	NH	
Death	1823	Springfield	Erie	PA	
Burial					

Places of Residence

Occupation Church Affiliation Military Rec.

Other wives, if any. No. (1) (2) etc. Make separate sheet for each mar.

His Father Page Batchelder Mother's Maiden Name Elizabeth Hill

Wife's Full Maiden Name Molly Copp

Wife's Data	Day Month Year	City, Town or Place	County or Province, etc.	State or Country	Add. Info. on Wife
Birth	15 July 1759	Hampstead	Rockingham	NH	
Chr'nd	28 Nov 1762	Hampstead	Rockingham	NH	
Death	1830	Springfield	Erie	PA	
Burial					

Compiler

Address

City, State

Date

Places of Residence

Occupation if other than Housewife Church Affiliation

Other husbands, if any. No. (1) (2) etc. Make separate sheet for each mar.

Her Father Joshua Copp Mother's Maiden Name Sarah Poore

Sex	Children's Names in Full (Arrange in order of birth)	Children's Data	Day Month Year	City, Town or Place	County or Province, etc.	State or Country	Add. Info. on Children
1	Jonathan	Birth					
	Full Name of Spouse* Betsy	Mar.					
		Death					
		Burial					
2	Sally	Birth	1784	Bradford	Orange	VT	
	Full Name of Spouse* James Leach	Mar.	4 mar 1802	Bradford	Orange	VT	
		Death	3 mar 1848	Springfield	Erie	PA	
		Burial					
3	Samuel	Birth					
	Full Name of Spouse* (1) (2)	Mar.					
		Death					
		Burial					
4	Susana	Birth					
	Full Name of Spouse* John Lamb	Mar.					
		Death					
		Burial					
5	Betsey	Birth					
	Full Name of Spouse* (unmarried)	Mar.					
		Death					
		Burial					
6	Abigail	Birth					
	Full Name of Spouse* Asa Lamunyon	Mar.					
		Death					
		Burial					
7	Hannah	Birth					
	Full Name of Spouse* James Wilson	Mar.					
		Death					
		Burial					
8	Mary	Birth	1797				
	Full Name of Spouse* Rufus Kendall	Mar.					
		Death	1877				
		Burial					
9	Isaiah	Birth					
	Full Name of Spouse*	Mar.					
		Death					
		Burial					
10	Joshua	Birth					
	Full Name of Spouse*	Mar.					
		Death					
		Burial					

*If married more than once No. each mar. (1) (2) etc. and list in "Add. info. on children" column. Use reverse side for additional children, other notes, references or information.

(2)

FAMILY GROUP No. ____ Husband's Full Name Isaiah Batchelder

This Information Obtained From:

Husband's Data	Day Month Year	City, Town or Place	County or Province, etc.	State or Country	Add. Info. on Husbar
Birth					
Chr'nd					
Mar.					
Death					
Burial					

Places of Residence

Occupation ____ Church Affiliation ____ Military Rec. ____

Other wives, if any. No. (1) (2) etc.
Make separate sheet for each mar.

His Father ____ Mother's Maiden Name ____

Wife's Full Maiden Name Molly Copp

Wife's Data	Day Month Year	City, Town or Place	County or Province, etc.	State or Country	Add. Info. on Wife
Birth					
Chr'nd					
Death					
Burial					

Compiler

Address

City, State

Date

Places of Residence

Occupation if other than Housewife ____ Church Affiliation ____

Other husbands, if any. No. (1) (2) etc.
Make separate sheet for each mar.

Her Father ____ Mother's Maiden Name ____

Sex	Children's Names in Full (Arrange in order of birth)	Children's Data	Day Month Year	City, Town or Place	County or Province, etc.	State or Country	Add. Info. on Childre
	1 Elmira	Birth					
	Full Name of Spouse* Elam Meekum	Mar.					
		Death					
		Burial					
	2	Birth					
	Full Name of Spouse*	Mar.					
		Death					
		Burial					
	3	Birth					
	Full Name of Spouse*	Mar.					
		Death					
		Burial					
	4	Birth					
	Full Name of Spouse*	Mar.					
		Death					
		Burial					
	5	Birth					
	Full Name of Spouse*	Mar.					
		Death					
		Burial					
	6	Birth					
	Full Name of Spouse*	Mar.					
		Death					
		Burial					
	7	Birth					
	Full Name of Spouse*	Mar.					
		Death					
		Burial					
	8	Birth					
	Full Name of Spouse*	Mar.					
		Death					
		Burial					
	9	Birth					
	Full Name of Spouse*	Mar.					
		Death					
		Burial					
	10	Birth					
	Full Name of Spouse*	Mar.					
		Death					
		Burial					

FAMILY GROUP No. ____ Husband's Full Name James Leach

This Information Obtained From:

Husband's Data	Day Month Year	City, Town or Place	County or Province, etc.	State or Country	Add. Info. on Husband
Birth	1780			NH	
Chr'nd					
Mar.	4 Mar 1802	Bradford	Orange	VT	
Death (Battle of Lundy's Lane)	25 July 1814	Niagara Falls	Ontario	Canada	
Burial					

Places of Residence

Occupation ____ Church Affiliation ____ Military Rec. War of 1812

Other wives, if any. No. (1) (2) etc. Make separate sheet for each mar.

His Father ____ Mother's Maiden Name ____

Wife's Full Maiden Name Sally Batchelder

Wife's Data	Day Month Year	City, Town or Place	County or Province, etc.	State or Country	Add. Info. on Wife
Birth	1784	Bradford	Orange	VT	
Chr'nd					
Death	3 Mar 1848	Springfield	Erie	PA	
Burial					

Compiler

Address

City, State

Date

Places of Residence

Occupation if other than Housewife ____ Church Affiliation ____

Other husbands, if any. No. (1) (2) etc. Make separate sheet for each mar.

Her Father Isaiah Batchelder Mother's Maiden Name Molly Copp

Sex	Children's Names in Full (Arrange in order of birth)	Children's Data	Day Month Year	City, Town or Place	County or Province, etc.	State or Country	Add. Info. on Children
1	Betsey	Birth					
	Full Name of Spouse*	Mar.					
		Death					
		Burial					
2	James	Birth					
	Full Name of Spouse*	Mar.					
		Death					
		Burial					
3	Isaiah	Birth					
	Full Name of Spouse*	Mar.					
		Death					
		Burial					
4	Emily Louise	Birth	30 Mar 1808			VT	
	Full Name of Spouse*	Mar.	24 Nov 1825	Springfield	Erie	PA	
	Charles Francis Smith	Death	26 Aug 1888	Linesville	Crawford	PA	
		Burial					
5	Priscilla	Birth					
	Full Name of Spouse*	Mar.					
		Death					
		Burial					
6	Joshua	Birth					
	Full Name of Spouse*	Mar.					
		Death					
		Burial					
7		Birth					
	Full Name of Spouse*	Mar.					
		Death					
		Burial					
8		Birth					
	Full Name of Spouse*	Mar.					
		Death					
		Burial					
9		Birth					
	Full Name of Spouse*	Mar.					
		Death					
		Burial					
10		Birth					
	Full Name of Spouse*	Mar.					
		Death					
		Burial					

*If married more than once No. each mar. (1) (2) etc. and list in "Add. info. on children" column. Use reverse side for additional children, other notes, references or information.

FAMILY GROUP No. ____ Husband's Full Name Oliver Smith

This Information Obtained From:

Husband's Data	Day Month Year	City, Town or Place	County or Province, etc.	State or Country	Add. Info. on Husband
Birth	1767			NH	
Chr'nd					
Mar.					
Death	14 May 1838	Springfield	Erie	PA	
Burial					

Places of Residence

Occupation | Church Affiliation | Military Rec.

Other wives, if any. No. (1) (2) etc. Make separate sheet for each mar.

His Father | Mother's Maiden Name

Wife's Full Maiden Name (1)

Wife's Data	Day Month Year	City, Town or Place	County or Province, etc.	State or Country	Add. Info. on Wife
Birth					
Chr'nd					
Death					
Burial					

Compiler

Address

City, State

Date

Places of Residence

Occupation if other than Housewife | Church Affiliation

Other husbands, if any. No. (1) (2) etc. Make separate sheet for each mar.

Her Father | Mother's Maiden Name

Sex	Children's Names in Full (Arrange in order of birth)	Children's Data	Day Month Year	City, Town or Place	County or Province, etc.	State or Country	Add. Info. on Children
1	Chester	Birth	1792			MA *	1850 Census
	Full Name of Spouse: (unmarried)	Mar.			*	NH	1880 Census
		Death					after 1880
		Burial					
2		Birth					
	Full Name of Spouse:	Mar.					
		Death					
		Burial					
3		Birth					
	Full Name of Spouse:	Mar.					
		Death					
		Burial					
4		Birth					
	Full Name of Spouse:	Mar.					
		Death					
		Burial					
5		Birth					
	Full Name of Spouse:	Mar.					
		Death					
		Burial					
6		Birth					
	Full Name of Spouse:	Mar.					
		Death					
		Burial					
7		Birth					
	Full Name of Spouse:	Mar.					
		Death					
		Burial					
8		Birth					
	Full Name of Spouse:	Mar.					
		Death					
		Burial					
9		Birth					
	Full Name of Spouse:	Mar.					
		Death					
		Burial					
10		Birth					
	Full Name of Spouse:	Mar.					
		Death					
		Burial					

FAMILY GROUP No. ____ Husband's Full Name Oliver Smith

This Information Obtained From:

Husband's Data	Day Month Year	City, Town or Place	County or Province, etc.	State or Country	Add. Info. on Husband
Birth	1767			NH	
Chr'nd					
Mar.					
Death	14 May 1838	Springfield	Erie	PA	
Burial					

Places of Residence

Occupation ____ Church Affiliation ____ Military Rec. ____

Other wives, if any. No. (1) (2) etc. Make separate sheet for each mar.

His Father ____ Mother's Maiden Name ____

Wife's Full Maiden Name (2) Betsy

Wife's Data	Day Month Year	City, Town or Place	County or Province, etc.	State or Country	Add. Info. on Wife
Birth	1776			NH	
Chr'nd					
Death	21 Dec 1861	Springfield	Erie	PA	
Burial					

Compiler

Address

City, State

Date

Places of Residence

Occupation if other than Housewife ____ Church Affiliation ____

Other husbands, if any. No. (1) (2) etc. Make separate sheet for each mar.

Her Father ____ Mother's Maiden Name ____

Sex	Children's Names in Full (Arrange in order of birth)	Children's Data	Day Month Year	City, Town or Place	County or Province, etc.	State or Country	Add. Info. on Children
1	Lyman	Birth	1798			NY	
	Full Name of Spouse: (1) Maria (2) Harriet	Mar.					
		Death	3 Sept 1875				
		Burial					
2	Electa	Birth	1799			NY	
	Full Name of Spouse: William E. Potter	Mar.	25 Apr 1822				
		Death	2 Apr 1836				
		Burial					
3	Charles Francis	Birth	20 Mar 1803	Springfield	Erie	PA	
	Full Name of Spouse: Emily Louise Leach	Mar.	24 Nov 1825	Springfield	Erie	PA	
		Death	17 Feb 1886	Springfield	Erie	PA	
		Burial					
4	Oliver	Birth	1805			PA	
	Full Name of Spouse: (1) Jane (2) Sarah Ann Ellis	Mar.					
		Death	29 Sept 1861				
		Burial					
5	Clarissa	Birth					
	Full Name of Spouse: James Hunter	Mar.					
		Death					prior 1830
		Burial					
6	Miriam	Birth					
	Full Name of Spouse:	Mar.					
		Death					prior 1830
		Burial					
7	Ares (Aris)	Birth	1805-1810				
	Full Name of Spouse: John Gilbert	Mar.					
		Death					
		Burial					
8	Freeman	Birth	1810			PA	
	Full Name of Spouse: Harriet	Mar.					
		Death					
		Burial					
9	John	Birth	1811-1820				
	Full Name of Spouse:	Mar.					
		Death					
		Burial					
10	Amos	Birth	1815			PA	
	Full Name of Spouse: Jane	Mar.					
		Death					
		Burial					

FAMILY GROUP No. ____ Husband's Full Name Lyman Smith

This Information Obtained From:

Husband's Data	Day Month Year	City, Town or Place	County or Province, etc.	State or Country	Add. Info. on Husba
Birth	1798			NY	
Chr'nd					
Mar.					
Death	2 Oct 1875				
Burial					

Places of Residence

Occupation ____ Church Affiliation ____ Military Rec. ____

Other wives, if any. No. (1) (2) etc. Make separate sheet for each mar.

His Father Oliver Smith Mother's Maiden Name Betsy

Wife's Full Maiden Name (1) Maria

Wife's Data	Day Month Year	City, Town or Place	County or Province, etc.	State or Country	Add. Info. on Wife
Birth	1802				
Chr'nd					
Death	3 Aug 1840				
Burial					

Compiler

Address

City, State

Date

Places of Residence

Occupation if other than Housewife ____ Church Affiliation ____

Other husbands, if any. No. (1) (2) etc. Make separate sheet for each mar.

Her Father ____ Mother's Maiden Name ____

Sex	Children's Names in Full (Arrange in order of birth)	Children's Data	Day Month Year	City, Town or Place	County or Province, etc.	State or Country	Add. Info. on Childre
1	Hiram	Birth				PA	
	Full Name of Spouse*	Mar.					
		Death					
		Burial					
2	Betsy	Birth				PA	
	Full Name of Spouse*	Mar.					
		Death					
		Burial					
3	Oliver A.	Birth	1834			PA	
	Full Name of Spouse*	Mar.					
		Death	14 Mar 1856				
		Burial					
4	Harriet M.	Birth				PA	
	Full Name of Spouse*	Mar.					
		Death					
		Burial					
5	Giles L.	Birth				PA	
	Full Name of Spouse*	Mar.					
		Death					
		Burial					
6	George L.	Birth					
	Full Name of Spouse*	Mar.					
		Death					
		Burial					
7		Birth					
	Full Name of Spouse*	Mar.					
		Death					
		Burial					
8		Birth					
	Full Name of Spouse*	Mar.					
		Death					
		Burial					
9		Birth					
	Full Name of Spouse*	Mar.					
		Death					
		Burial					
10		Birth					
	Full Name of Spouse*	Mar.					
		Death					
		Burial					

*If married more than once No. each mar. (1) (2) etc. and list in "Add. info. on children" column. Use reverse side for additional children, other notes, references or information.

FAMILY GROUP No. ____

This Information Obtained From:

Husband's Full Name Charles Francis Smith

Husband's Data	Day Month Year	City, Town or Place	County or Province, etc.	State or Country	Add. Info. on Husband
Birth	20 Mar 1803	Springfield	Erie	PA	
Chr'nd					
Mar.	24 Nov 1825	Springfield	Erie	PA	
Death	17 Feb 1886	Springfield	Erie	PA	
Burial					

Places of Residence

Occupation ____ Church Affiliation ____ Military Rec. ____

Other wives, if any. No. (1) (2) etc. Make separate sheet for each mar.

His Father Oliver Smith Mother's Maiden Name Betsy

Wife's Full Maiden Name Emily Louise Leach

Wife's Data	Day Month Year	City, Town or Place	County or Province, etc.	State or Country	Add. Info. on Wife
Birth	30 Mar 1808			VT	
Chr'nd					
Death	26 Aug 1888	Linesville	Crawford	PA	
Burial					

Compiler

Address

City, State

Date

Places of Residence

Occupation if other than Housewife ____ Church Affiliation ____

Other husbands, if any. No. (1) (2) etc. Make separate sheet for each mar.

Her Father James Leach Mother's Maiden Name Sally Batchelder

Sex	Children's Names in Full (Arrange in order of birth)	Children's Data	Day Month Year	City, Town or Place	County or Province, etc.	State or Country	Add. Info. on Children
1	Betsey Full Name of Spouse* William Brockway	Birth	28 Jan 1827				
		Mar.					
		Death	31 Aug 1855				
		Burial					
2	Amos Full Name of Spouse*	Birth	Aug 1829				
		Mar.					
		Death	19 Sept 1900				
		Burial					
3	Lathrop Winfield Full Name of Spouse*	Birth	May 1830				
		Mar.					
		Death	28 Nov 1831				
		Burial					
4	Sarah Full Name of Spouse* John B. McIntosh	Birth	29 Sept 1831				
		Mar.					
		Death					
		Burial					
5	Aris A. Full Name of Spouse*	Birth	15 Nov 1833				
		Mar.					
		Death	22 Apr 1852				
		Burial					
6	Charles Full Name of Spouse*	Birth	18 Mar 1836				
		Mar.					
		Death	11 Feb 1923				
		Burial					
7	Laura A. Full Name of Spouse*	Birth	8 Apr 1840				
		Mar.					
		Death	13 Aug 1855				
		Burial					
8	Helen Full Name of Spouse* Henry B. Line	Birth	11 Aug 1842				
		Mar.					
		Death					
		Burial					
9	Adelaide Matilda Full Name of Spouse* Theodore Benedict Rundel	Birth	7 Feb 1845				
		Mar.	28 Sept 1863				
		Death	12 Dec 1876				
		Burial					
10	James Oliver Full Name of Spouse* Sarah Isabel Pettit	Birth	22 July 1847				
		Mar.	21 Apr 1873				
		Death					
		Burial					

FAMILY GROUP No.

Husband's Full Name Charles Francis Smith

This Information Obtained From:

Husband's Data	Day Month Year	City, Town or Place	County or Province, etc.	State or Country	Add. Info. on Husban
Birth					
Chr'nd					
Mar.					
Death					
Burial					

Places of Residence

Occupation | Church Affiliation | Military Rec.

Other wives, if any. No. (1) (2) etc.
Make separate sheet for each mar.

His Father | Mother's Maiden Name

Wife's Full Maiden Name Emily Louise Leach

Wife's Data	Day Month Year	City, Town or Place	County or Province, etc.	State or Country	Add. Info. on Wife
Birth					
Chr'nd					
Death					
Burial					

Compiler

Address

City, State

Date

Places of Residence

Occupation if other than Housewife | Church Affiliation

Other husbands, if any. No. (1) (2) etc.
Make separate sheet for each mar.

Her Father | Mother's Maiden Name

Sex	Children's Names in Full (Arrange in order of birth)	Children's Data	Day Month Year	City, Town or Place	County or Province, etc.	State or Country	Add. Info. on Children
M	1 Emily H.	Birth	Jan 1851				
	Full Name of Spouse*	Mar.					
		Death	23 July 1852				
		Burial					
	2	Birth					
	Full Name of Spouse*	Mar.					
		Death					
		Burial					
	3	Birth					
	Full Name of Spouse*	Mar.					
		Death					
		Burial					
	4	Birth					
	Full Name of Spouse*	Mar.					
		Death					
		Burial					
	5	Birth					
	Full Name of Spouse*	Mar.					
		Death					
		Burial					
	6	Birth					
	Full Name of Spouse*	Mar.					
		Death					
		Burial					
	7	Birth					
	Full Name of Spouse*	Mar.					
		Death					
		Burial					
	8	Birth					
	Full Name of Spouse*	Mar.					
		Death					
		Burial					
	9	Birth					
	Full Name of Spouse*	Mar.					
		Death					
		Burial					
	10	Birth					
	Full Name of Spouse*	Mar.					
		Death					
		Burial					

FAMILY GROUP No. ____ Husband's Full Name Freeman Smith

This Information Obtained From:

Husband's Data	Day Month Year	City, Town or Place	County or Province, etc.	State or Country	Add. Info. on Husban
Birth	1810			PA	
Chr'nd					
Mar.					
Death					
Burial					

Places of Residence

Occupation ____ Church Affiliation ____ Military Rec. ____

Other wives, if any. No. (1) (2) etc. Make separate sheet for each mar.

His Father Oliver Smith Mother's Maiden Name Betsy

Wife's Full Maiden Name Harriet

Wife's Data	Day Month Year	City, Town or Place	County or Province, etc.	State or Country	Add. Info. on Wife
Birth	1815			NY	
Chr'nd					
Death					
Burial					

Compiler

Address

City, State

Date

Places of Residence

Occupation if other than Housewife ____ Church Affiliation ____

Other husbands, if any. No. (1) (2) etc. Make separate sheet for each mar.

Her Father ____ Mother's Maiden Name ____

Sex	Children's Names in Full (Arrange in order of birth)	Children's Data	Day Month Year	City, Town or Place	County or Province, etc.	State or Country	Add. Info. on Children
M	1 E. Full Name of Spouse*	Birth	1839			PA	
		Mar.					
		Death					
		Burial					
F	2 Cordelia Full Name of Spouse*	Birth	1843			PA	
		Mar.					
		Death					
		Burial					
M	3 John B. Full Name of Spouse*	Birth	1844			PA	
		Mar.					
		Death					
		Burial					
M	4 William E. Full Name of Spouse*	Birth	1847			PA	
		Mar.					
		Death					
		Burial					
F	5 Betsy Full Name of Spouse*	Birth	1848			CT	
		Mar.					
		Death					
		Burial					
	6 Full Name of Spouse*	Birth					
		Mar.					
		Death					
		Burial					
	7 Full Name of Spouse*	Birth					
		Mar.					
		Death					
		Burial					
	8 Full Name of Spouse*	Birth					
		Mar.					
		Death					
		Burial					
	9 Full Name of Spouse*	Birth					
		Mar.					
		Death					
		Burial					
	10 Full Name of Spouse*	Birth					
		Mar.					
		Death					
		Burial					

FAMILY GROUP No. ____ Husband's Full Name Theodore Benedict Rundel

This Information Obtained From:

Husband's Data	Day Month Year	City, Town or Place	County or Province, etc.	State or Country	Add. Info. on Husbar
Birth	24 July 1835	Cussewago	Crawford	PA	
Chr'nd					
Mar.	1880				
Death	19 Mar 1919	Conneautville	Crawford	PA	
Burial					

Places of Residence

Occupation | Church Affiliation | Military Rec.

Other wives, if any. No. (1) (2) etc. Make separate sheet for each mar.

His Father | Mother's Maiden Name

Wife's Full Maiden Name (2) Mrs. Flora Augusta Nash Linder

Wife's Data	Day Month Year	City, Town or Place	County or Province, etc.	State or Country	Add. Info. on Wife
Birth					
Chr'nd					
Death					
Burial					

Compiler

Address

City, State

Date

Places of Residence of Vandalia, Michigan

Occupation if other than Housewife | Church Affiliation

Other husbands, if any. No. (1) (2) etc. Make separate sheet for each mar.

Her Father | Mother's Maiden Name

Sex	Children's Names in Full (Arrange in order of birth)	Children's Data	Day Month Year	City, Town or Place	County or Province, etc.	State or Country	Add. Info. on Children
1	Full Name of Spouse*	Birth					
		Mar.					
		Death					
		Burial					
2	Full Name of Spouse*	Birth					
		Mar.					
		Death					
		Burial					
3	Full Name of Spouse*	Birth					
		Mar.					
		Death					
		Burial					
4	Full Name of Spouse*	Birth					
		Mar.					
		Death					
		Burial					
5	Full Name of Spouse*	Birth					
		Mar.					
		Death					
		Burial					
6	Full Name of Spouse*	Birth					
		Mar.					
		Death					
		Burial					
7	Full Name of Spouse*	Birth					
		Mar.					
		Death					
		Burial					
8	Full Name of Spouse*	Birth					
		Mar.					
		Death					
		Burial					
9	Full Name of Spouse*	Birth					
		Mar.					
		Death					
		Burial					
10	Full Name of Spouse*	Birth					
		Mar.					
		Death					
		Burial					

APPENDIX B

ALTERNATE SPELLINGS

BATCHELDER

Bachaler
Bachalor
Bachelder
Bacheleder
Bacheldor
Bacheldore
Bacheler
Bachelere
Bacheller
Bachellor
Bachelor
Bachilder
Bacheler
Bachillor
Bachilor
Bachlicor
Bachlor
Bacholter
Bachylere
Bactherer
Barchaldor
Bashelor
Batchalder
Batchalor
Batchelar
Batcheldor
Batcheler
Batcheller
Batchelor
Batchellow
Batchelter
Batcherder
Batchilder
Batchilor
Batchlar
Batchldor
Batchlor
Batcholder
Batcholdor
Battchelor
Bocldr

COPP

Cope
Copps

LEACH

Leah
Leatch
Leceh
Lech
Leech
Leetch
Lirtch
Lish

COUNTY HISTORY

STATE	COUNTY	CREATED FROM	DATE
Pennsylvania	Allegheny	Westmoreland & Washington	1788
Pennsylvania	Crawford	Allegheny	1800
Pennsylvania	Erie	Allegheny	1800

APPENDIX C

VETTER / MADDISON RELATIONAL MODEL - PHASE I

STEP 1. IDENTIFY ENTITY SETS

person
marriage

STEP 2. IDENTIFY PROPERTIES OF ENTITY SETS

PERSON

- id #
- name
- nickname
- alias
- birth date
- birth place
- death date
- death place
- father
- mother
- married

MARRIAGE

- id #
- marriage date
- marriage place
- person # (husband)
- person # (wife)

STEP 3. DOMAINS OF PROPERTIES OF ENTITY SETS

- id numbers
- given name
- surname
- nickname
- alias
- year
- month
- day
- country
- state
- county
- other political subdivision (OPS)
- number

STEP 4. KEY DOMAINS OF ENTITY SETS

PERSON

id number (P#)

MARRIAGE

id number (M#)

STEP 5. IDENTIFY ENTITY ATTRIBUTES

PERSON

P-name (P#, given name, surname)
P-nickname (P#, nickname)
P-alias (P#, alias)
P-birth (P#, year, month, day)
P-death (P#, year, month, day)
P-father (P#, F.P#)
P-mother (P#, M.P#)
P-number of marriage (P#, NM)

MARRIAGE

M-marriage date (M#, year, month, day)
M-marriage place (M#, country, state, county, OPS)
M-husband (M#, H.P#)
M-wife (M#, W.P#)

STEP 6. IDENTIFY RELATIONSHIP SETS

spouse-of (H.P#, W.P#)

STEP 7. IDENTIFY RELATIONSHIP ATTRIBUTES

none

VETTER / MADDISON RELATIONAL MODEL - PHASE II

TRANSFORM DESCRIPTION OF ENVIRONMENT OF INTEREST INTO ELEMENTARY RELATIONS

CK	candidate key
ER	elementary relation
FD	functional dependence (--->)
FFD	full functional dependence (===>)
PK	primary key
R	relation

M#	marriage number
NM	number of marriages
OPS	other political subdivision
P#	person id #
F.P#	father id #
H.P#	husband id #
M.P#	mother id #
W.P#	wife id #

R1. P-NAME (P#, Given name, Surname)

CK & PK	P#
FD	P#--->Given name
	P#--->Surname
ER1	(P#, Given name)
ER2	(P#, Surname)

R2. P-NICKNAME (P#, Nickname)

CK & PK	P#, Nickname
FD	
ER3	(P#, Nickname)

R3. P-ALIAS (P#, Alias)

CK & PK	P#, Alias
FD	
ER4	(P#, Alias)

R4. P-BIRTH DATE (P#, Year, Month, Day)

CK & PK	P#
FD	P#--->Year
	P#--->Month
	P#--->Day
ER5	(P#, Year)
ER6	(P#, Month)
ER7	(P#, Day)

R5. P-BIRTH PLACE (P#, Country, State, County, OPS)

CK & PK	P#
FD	P#--->Country
	P#--->State
	P#--->County
	P#--->OPS
ER8	(P#, Country)
ER9	(P#, State)
ER10	(P#, County)
ER11	(P#, OPS)

R6. P-DEATH DATE (P#, Year, Month, Day)

CK & PK	P#
FD	P#--->Year
	P#--->Month
	P#--->Day
ER12	(P#, Year)
ER13	(P#, Month)
ER14	(P#, Day)

R7 P-DEATH PLACE (P#, Country, State, County, OPS)

CK & PK	P#
FD	P#--->Country
	P#--->State
	P#--->County
	P#--->OPS
ER15	(P#, Country)
ER16	(P#, State)
ER17	(P#, County)
ER18	(P#, OPS)

R8. P-FATHER (P#, F.P#)

CK & PK	P#, F.P#
FD	
ER19	(P#, F.P#)

R9. P-MOTHER (P#, M.P#)

CK & PK	P#, M.P#
FD	
ER20	(P#, M.P#)

R10. P-NUMBER MARRIAGES (P#, NM)

CK & PK	P#
FD	P#--->NM
ER21	(P#, NM)

R11. M-MARRIAGE DATE (M#, Year, Month, Day)

CK & PK	M#
FD	M#--->Year
	M#--->Month
	M#--->Day
ER22	(M#, Year)
ER23	(M#, Month)
ER24	(M#, Day)

R12. M-MARRIAGE PLACE (M#, COUNTRY, STATE, COUNTY, OPS)

CK & PK	M#
FD	M#--->Country
	M#--->State
	M#--->County
	M#--->OPS
ER25	(M#, Country)
ER26	(M#, State)
ER27	(M#, County)
ER28	(M#, OPS)

R13. M-HUSBAND (M#, H.P#)

CK & PK	M#, H.P#
FD	
ER29	(M#, H.P#)

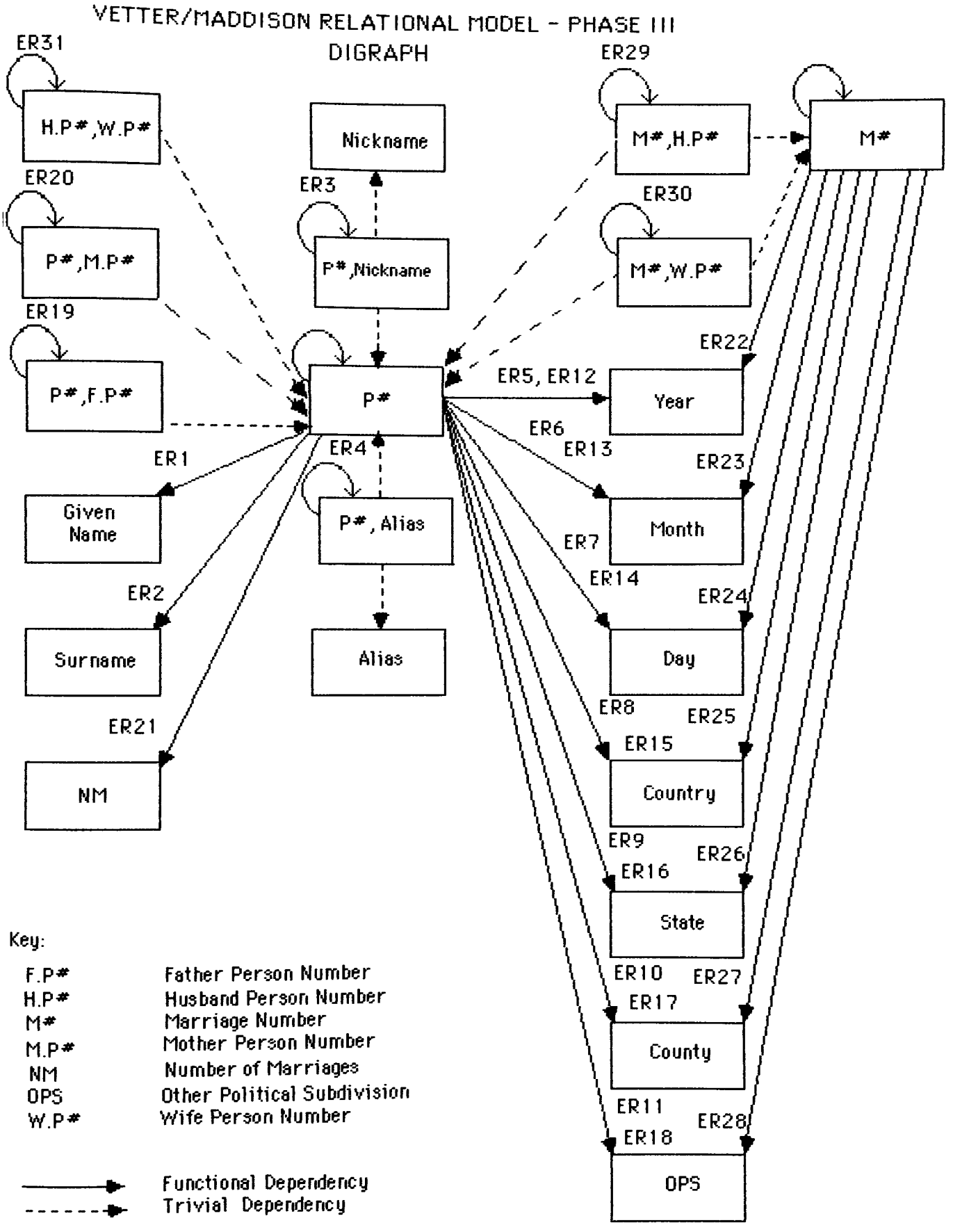
R14. M-WIFE (M#, W.P#)

CK & PK	M#, W.P#
FD	
ER30	(M#, W.P#)

R15. SPOUSE OF (H.P#, W.P#)

CK & PK	H.P#, W.P#
FD	
ER31	(H.P#, W.P#)

Note: This list of Elementary Relations is used in Phase III.


VETTER/MADDISON RELATIONAL MODEL - PHASE III
DIGRAPH
ER31
H.P#,W.P#
ER20
P#,M.P#
ER19
P#,F.P#
ER1
Given Name
ER2
Surname
ER21
NM
Nickname
ER3
P#,Nickname
P#
ER4
P#, Alias
Alias
ER29
M#,H.P#
ER30
M#,W.P#
M#
ER22
ER5, ER12
Year
ER6
ER13
ER23
Month
ER7
ER14
ER24
Day
ER8
ER25
ER15
Country
ER9
ER26
ER16
State
ER10
ER27
ER17
County
ER11
ER28
ER18
OPS
Key:
F.P# Father Person Number
H.P# Husband Person Number
M# Marriage Number
M.P# Mother Person Number
NM Number of Marriages
OPS Other Political Subdivision
W.P# Wife Person Number
Functional Dependency
Trivial Dependency

	P#	GIVEN NAME	SURNAME	NICKNAME	P#,NICKNAME	ALIAS	P#,ALIAS	P#,F.P#	P#,M.P#	BIRTH YEAR	BIRTH MONTH	BIRTH DAY	BIRTH COUNTRY	BIRTH STATE	BIRTH COUNTY	BIRTH OPS
P#	1	1	1							1	1	1	1	1	1	1
Given Name																
Surname																
Nickname				1												
P#,Nickname	1			1	1											
Alias						1										
P#,Alias	1					1	1									
P#,F.P#	1							1								
P#,M.P#	1								1							
Birth Year																
Birth Month																
Birth Day																
Birth Country																
Birth State																
Birth County																
Birth OPS																
# Marriages																
M#																
M#,H.P#	1															
M#,W.P#	1															
H.P#,W.P#	1															
Marriage Year																
Marriage Month																
Marriage Day																
M Country																
M State																
M County																
M OPS																
Death Year																
Death Month																
Death Day																
Death Country																
Death State																
Death County																
Death OPS																

	M#	H.P#	W.P#	MARRIAGE YEAR	MARRIAGE MONTH	MARRIAGE DAY	M COUNTRY	M STATE	M COUNTY	M OPS	DEATH YEAR	DEATH MONTH	DEATH DAY	DEATH COUNTRY	DEATH STATE	DEATH COUNTY	DEATH OPS
P#		1	1														
Given Name																	
Surname																	
Nickname																	
P#,Nickname																	
Alias																	
P#,Alias																	
P#,F.P#																	
P#,M.P#																	
Birth Year																	
Birth Month																	
Birth Day																	
Birth Country																	
Birth State																	
Birth County																	
Birth OPS																	
# Marriages																	
M#																	
M#,H.P#	1																
M#,W.P#	1																
H.P#,W.P#	1																
Marriage Year																	
Marriage Month																	
Marriage Day																	
M Country																	
M State																	
M County																	
M OPS																	
Death Year																	
Death Month																	
Death Day																	
Death Country																	
Death State																	
Death County																	
Death OPS																	

							M												
						M	A									D			
						A	R	M								E		D	
	#					R	R	A						D		A	D	E	
						R	I	R					D	E		T	E	A	
	M				H	I	A	R	M				E	A	D	H	A	T	D
	A					A	G	I			M		A	T	E		T	H	E
	R		M	M	P	G	E	A	C	M			T	H	A	C	H		A
	R		#	#	#	E		G	O		C		H		T	O		C	T
	I		,	,	,		M	E	U	S	O	M		M	H	U	S	O	H
	A		H	W	W	Y	O		N	T	U		Y	O		N	T	U	
	G					E	N	D	T	A	N	O	E	N	D	T	A	N	O
	E	M	P	P	P	A	T	A	R	T	T	P	A	T	A	R	T	T	P
	S	#	#	#	#	R	H	Y	Y	E	Y	S	R	H	Y	Y	E	Y	S
**********	**********	**********	**********	**********	**********	**********	**********	**********	**********	**********	**********	**********	**********	**********	**********	**********	**********	**********	**********
P#	1												1	1	1	1	1	1	1
Given Name																			
Surname																			
Nickname																			
P#,Nickname																			
Alias																			
P#,Alias																			
P#,F.P#																			
P#,M.P#																			
Birth.Year																			
Birth.Month																			
Birth.Day																			
Birth.Country																			
Birth.State																			
Birth.County																			
Birth.OPS																			
# Marriages																			
M#		1				1	1	1	1	1	1	1							
M#,H.P#		1	1																
M#,W.P#		1		1															
H.P#,W.P#					1														
Marriage.Year																			
Marriage.Month																			
Marriage.Day																			
M.Country																			
M.State																			
M.County																			
M.OPS																			
Death.Year																			
Death.Month																			
Death.Day																			
Death.Country																			
Death.State																			
Death.County																			
Death.OPS																			

VETTER / MADDISON RELATIONAL MODEL - PHASE III

TRANSITIVE DEPENDENCIES

DEPENDENCY	COMPOSITION OF	MEANINGFUL
P#,F.P#--->Given name	P#,F.P#--->P#--->Given name	No
P#,F.P#--->Surname	P#,F.P#--->P#--->Surname	No
P#,F.P#--->Birth year	P#,F.P#--->P#--->Birth.year	No
P#,F.P#--->Birth month	P#,F.P#--->P#--->Birth.month	No
P#,F.P#--->Birth day	P#,F.P#--->P#--->Birth.day	No
P#,F.P#--->\# marriages	P#,F.P#--->P#--->\# marriages	No
P#,F.P#--->Death year	P#,F.P#--->P#--->Death.year	No
P#,F.P#--->Death month	P#,F.P#--->P#--->Death.month	No
P#,F.P#--->Death day	P#,F.P#--->P#--->Death.day	No
P#,F.P#--->Birth country	P#,F.P#--->P#--->Birth.country	No
P#,F.P#--->Birth state	P#,F.P#--->P#--->Birth.state	No
P#,F.P#--->Birth county	P#,F.P#--->P#--->Birth.county	No
P#,F.P#--->Birth OPS	P#,F.P#--->P#--->Birth.OPS	No
P#,F.P#--->Death country	P#,F.P#--->P#--->Death.country	No
P#,F.P#--->Death state	P#,F.P#--->P#--->Death.state	No
P#,F.P#--->Death county	P#,F.P#--->P#--->Death.county	No
P#,F.P#--->Death OPS	P#,F P#--->P#--->Death.OPS	No
P#,M.P#--->Given name	P#,M.P#--->P#--->Given name	No
P#,M.P#--->Surname	P#,M.P#--->P#--->Surname	No
P#,M.P#--->Birth year	P#,M.P#--->P#--->Birth.year	No
P#,M.P#--->Birth month	P#,M.P#--->P#--->Birth.month	No
P#,M.P#--->Birth day	P#,M.P#--->P#--->Birth.day	No
P#,M.P#--->\# marriages	P#,M.P#--->P#--->\# marriages	No
P#,M.P#--->Death year	P#,M.P#--->P#--->Death.year	No
P#,M.P#--->Death month	P#,M.P#--->P#--->Death.month	No
P#,M.P#--->Death day	P#,M.P#--->P#--->Death.day	No
P#,M.P#--->Birth country	P#,M.P#--->P#--->Birth.country	No
P#,M.P#--->Birth state	P#,M.P#--->P#--->Birth.state	No
P#,M.P#--->Birth county	P#,M.P#--->P#--->Birth.county	No
P#,M.P#--->Birth OPS	P#,M.P#--->P#--->Birth.OPS	No
P#,M.P#--->Death country	P#,M.P#--->P#--->Death.country	No
P#,M.P#--->Death state	P#,M.P#--->P#--->Death.state	No
P#,M.P#--->Death county	P#,M.P#--->P#--->Death.county	No
P#,M.P#--->Death OPS	P#,M.P#--->P#--->Death.OPS	No

VETTER / MADDISON RELATIONAL MODEL - PHASE III

TRANSITIVE DEPENDENCIES

DEPENDENCY	COMPOSITION OF	MEANINGFUL
P#,Nickname--->Given name	P#,Nickname--->P#--->Given name	No
P#,Nickname--->Surname	P#,Nickname--->P#--->Surname	No
P#,Nickname--->Birth year	P#,Nickname--->P#--->Birth.year	No
P#,Nickname--->Birth month	P#,Nickname--->P#--->Birth.month	No
P#,Nickname--->Birth day	P#,Nickname--->P#--->Birth.day	No
P#,Nickname--->* marriages	P#,Nickname--->P#--->* marriages	No
P#,Nickname--->Death year	P#,Nickname--->P#--->Death.year	No
P#,Nickname--->Death month	P#,Nickname--->P#--->Death.month	No
P#,Nickname--->Death day	P#,Nickname--->P#--->Death.day	No
P#,Nickname--->Birth country	P#,Nickname--->P#--->Birth.country	No
P#,Nickname--->Birth state	P#,Nickname--->P#--->Birth.state	No
P#,Nickname--->Birth county	P#,Nickname--->P#--->Birth.county	No
P#,Nickname--->Birth OPS	P#,Nickname--->P#--->Birth.OPS	No
P#,Nickname--->Death country	P#,Nickname--->P#--->Death.country	No
P#,Nickname--->Death state	P#,Nickname--->P#--->Death.state	No
P#,Nickname--->Death county	P#,Nickname--->P#--->Death.county	No
P#,Nickname--->Death OPS	P#,Nickname--->P#--->Death.OPS	No
P#,Alias--->Given name	P#,Alias--->P#--->Given name	No
P#,Alias--->Surname	P#,Alias--->P#--->Surname	No
P#,Alias--->Birth year	P#,Alias--->P#--->Birth.year	No
P#,Alias--->Birth month	P#,Alias--->P#--->Birth.month	No
P#,Alias--->Birth day	P#,Alias--->P#--->Birth.day	No
P#,Alias--->* marriages	P#,Alias--->P#--->* marriages	No
P#,Alias--->Death year	P#,Alias--->P#--->Death.year	No
P#,Alias--->Death month	P#,Alias--->P#--->Death.month	No
P#,Alias--->Death day	P#,Alias--->P#--->Death.day	No
P#,Alias--->Birth country	P#,Alias--->P#--->Birth.country	No
P#,Alias--->Birth state	P#,Alias--->P#--->Birth.state	No
P#,Alias--->Birth county	P#,Alias--->P#--->Birth.county	No
P#,Alias--->Birth OPS	P#,Alias--->P#--->Birth.OPS	No
P#,Alias--->Death country	P#,Alias--->P#--->Death.country	No
P#,Alias--->Death state	P#,Alias--->P#--->Death.state	No
P#,Alias--->Death county	P#,Alias--->P#--->Death.county	No
P#,Alias--->Death OPS	P#,Alias--->P#--->Death.OPS	No

VETTER / MADDISON RELATIONAL MODEL - PHASE III

TRANSITIVE DEPENDENCIES

DEPENDENCY	COMPOSITION OF	MEANINGFUL
H.P#,W.P#--->Given name	H.P#,W.P#--->P#--->Given name	No
H.P#,W.P#--->Surname	H.P#,W.P#--->P#--->Surname	No
H.P#,W.P#--->Birth year	H.P#,W.P#--->P#--->Birth.year	No
H.P#,W.P#--->Birth month	H.P#,W.P#--->P#--->Birth.month	No
H.P#,W.P#--->Birth day	H.P#,W.P#--->P#--->Birth.day	No
H.P#,W.P#--->* marriages	H.P#,W.P#--->P#--->* marriages	No
H.P#,W.P#--->Death year	H.P#,W.P#--->P#--->Death.year	No
H.P#,W.P#--->Death month	H.P#,W.P#--->P#--->Death.month	No
H.P#,W.P#--->Death day	H.P#,W.P#--->P#--->Death.day	No
H.P#,W.P#--->Birth country	H.P#,W.P#--->P#--->Birth.country	No
H.P#,W.P#--->Birth state	H.P#,W.P#--->P#--->Birth.state	No
H.P#,W.P#--->Birth county	H.P#,W.P#--->P#--->Birth.county	No
H.P#,W.P#--->Birth OPS	H.P#,W.P#--->P#--->Birth.OPS	No
H.P#,W.P#--->Death country	H.P#,W.P#--->P#--->Death.country	No
H.P#,W.P#--->Death state	H.P#,W.P#--->P#--->Death.state	No
H.P#,W.P#--->Death county	H.P#,W.P#--->P#--->Death.county	No
H.P#,W.P#--->Death OPS	H.P#,W.P#--->P#--->Death.OPS	No

VETTER / MADDISON RELATIONAL MODEL - PHASE III

TRANSITIVE DEPENDENCIES

DEPENDENCY	COMPOSITION OF	MEANINGFUL
M#,H.P#--->Given name	M#,H.P#--->P#--->Given name	No
M#,H.P#--->Surname	M#,H.P#--->P#--->Surname	No
M#,H.P#--->Birth year	M#,H.P#--->P#--->Birth.year	No
M#,H.P#--->Birth month	M#,H.P#--->P#--->Birth.month	No
M#,H.P#--->Birth day	M#,H.P#--->P#--->Birth.day	No
M#,H.P#--->* marriages	M#,H.P#--->P#--->* marriages	No
M#,H.P#--->Death year	M#,H.P#--->P#--->Death.year	No
M#,H.P#--->Death month	M#,H.P#--->P#--->Death.month	No
M#,H.P#--->Death day	M#,H.P#--->P#--->Death.day	No
M#,H.P#--->Marriage year	M#,H.P#--->M#--->Marriage.year	No
M#,H.P#--->Marriage month	M#,H.P#--->M#--->Marriage.month	No
M#,H.P#--->Marriage day	M#,H.P#--->M#--->Marriage.day	No
M#,H.P#--->Birth country	M#,H.P#--->P#--->Birth.country	No
M#,H.P#--->Birth state	M#,H.P#--->P#--->Birth.state	No
M#,H.P#--->Birth county	M#,H.P#--->P#--->Birth.county	No
M#,H.P#--->Birth OPS	M#,H.P#--->P#--->Birth.OPS	No
M#,H.P#--->Death country	M#,H.P#--->P#--->Death.country	No
M#,H.P#--->Death state	M#,H.P#--->P#--->Death.state	No
M#,H.P#--->Death county	M#,H.P#--->P#--->Death.county	No
M#,H.P#--->Death OPS	M#,H.P#--->P#--->Death.OPS	No
M#,H.P#--->Marriage country	M#,H.P#--->M#--->Marriage.country	No
M#,H.P#--->Marriage state	M#,H.P#--->M#--->Marriage.state	No
M#,H.P#--->Marriage county	M#,H.P#--->M#--->Marriage.county	No
M#,H.P#--->Marriage OPS	M#,H.P#--->M#--->Marriage.OPS	No
M#,W.P#--->Given name	M#,W.P#--->P#--->Given name	No
M#,W.P#--->Surname	M#,W.P#--->P#--->Surname	No
M#,W.P#--->Birth year	M#,W.P#--->P#--->Birth.year	No
M#,W.P#--->Birth month	M#,W.P#--->P#--->Birth.month	No
M#,W.P#--->Birth day	M#,W.P#--->P#--->Birth.day	No
M#,W.P#--->* marriages	M#,W.P#--->P#--->* marriages	No
M#,W.P#--->Death year	M#,W.P#--->P#--->Death.year	No
M#,W.P#--->Death month	M#,W.P#--->P#--->Death.month	No
M#,W.P#--->Death day	M#,W.P#--->P#--->Death.day	No
M#,W.P#--->Marriage year	M#,W.P#--->M#--->Marriage.year	No
M#,W.P#--->Marriage month	M#,W.P#--->M#--->Marriage.month	No
M#,W.P#--->Marriage day	M#,W.P#--->M#--->Marriage.day	No
M#,W.P#--->Birth country	M#,W.P#--->P#--->Birth.country	No
M#,W.P#--->Birth state	M#,W.P#--->P#--->Birth.state	No
M#,W.P#--->Birth county	M#,W.P#--->P#--->Birth.county	No
M#,W.P#--->Birth OPS	M#,W.P#--->P#--->Birth.OPS	No
M#,W.P#--->Death country	M#,W.P#--->P#--->Death.country	No
M#,W.P#--->Death state	M#,W.P#--->P#--->Death.state	No
M#,W.P#--->Death county	M#,W.P#--->P#--->Death.county	No
M#,W.P#--->Death OPS	M#,W.P#--->P#--->Death.OPS	No
M#,W.P#--->Marriage country	M#,W.P#--->M#--->Marriage.country	No
M#,W.P#--->Marriage state	M#,W.P#--->M#--->Marriage.state	No
M#,W.P#--->Marriage county	M#,W.P#--->M#--->Marriage.county	No
M#,W.P#--->Marriage OPS	M#,W.P#--->M#--->Marriage.OPS	No

VETTER / MADDISON RELATIONAL MODEL GENEALOGICAL EXAMPLE

PHASE III AND PHASE IV

ER1(Person number, Given name)
ER2(Person number, Surname)
ER3(Person number, Nickname)
ER4(Person number, Alias)
ER5(Person number, Birth.year)
ER6(Person number, Birth.month)
ER7(Person number, Birth.day)
ER8(Person number, Birth.country)
ER9(Person number, Birth.state)
ER10(Person number, Birth.county)
ER11(Person number, Birth.OPS)
ER12(Person number, Death.year)
ER13(Person number, Death.month)
ER14(Person number, Death.day)
ER15(Person number, Death.country)
ER16(Person number, Death.state)
ER17(Person number, Death.county)
ER18(Person number, Death.OPS)
ER19(Person number, Father.person number)
ER20(Person number, Mother.person number)
ER21(Person number, Number of marriages)
ER22(Marriage number, Marriage.year)
ER23(Marriage number, Marriage.month)
ER24(Marriage number, Marriage.day)
ER25(Marriage number, Marriage.country)
ER26(Marriage number, Marriage.state)
ER27(Marriage number, Marriage.county)
ER28(Marriage number, Marriage.OPS)
ER29(Marriage number, Husband.person number)
ER30(Marriage number, Wife.person number)
ER31(Husband.person number, Wife.person number)

Note: OPS stands for Other Political Subdivision
(city, town, township, village)

At the conclusion of Phase III there were no redundant elementary relations. Therefore, the list of elementary relations for Phase IV is the same as that for Phase III.

BIBLIOGRAPHY

INTRODUCTION

Buchmann, A. P., and Dale, A. G., "Evaluation Criteria For Logical Database Design Methodologies", Tutorial: Data Base Management in the 1980's, Larson, James A., and Freeman, Harvey A., Computer Society Press, 1981

FUZZY DATA

Babad, Yair M., and Hoffer, Jeffrey A., "Even No Data Has Value", Communications of the ACM, Vol. 27 Number 8, August 1984

Bouille, F., "Fuzzy Data Processing With The Hypergraph-based Data Structure", Proceedings of the International Conference of Cybernetics and Society, 3-7 Nov 1978, IEEE, NY, 1978

Gaines, B. R., "Logical Foundations for Database Systems", Fuzzy Reasoning and Its Applications, ed. E. H. Mandani and B. R. Gaines, Academic Press, Inc., London, 1981

Gray, M. A., "Implementing Unknown and Imprecise Values in Databases", Databases: Proceedings of First British National Conference on Databases held at Jesus College, Cambridge, 13-14 July 1981, S. M. Deen and P. Hammersley (ed.), John Wiley & Sons, NY, 1981

Kambayashi, Y., Tanaka, K., Yajima, S., "Problems of Relational Database Design", Lecture Notes in Computer Science, G. Goos and J. Hartmanis (ed.), Springer-Verlag, Germany, 1982

Vassiliou, Y., "Null Values in Data Base Management, A Denotational Semantics Approach", Proceedings ACM SIGMOD 1979, International Conference on Management of Data, Boston, MA, May 30 - June 1, 1979, Philip A. Berstein, (ed.), ACM 1979

Vassiliou, Y., "Functional Dependencies and Incomplete Information", Proceedings of Sixth International Conference on VLDS, October 1980

Wiederhold, Gio, Database Design, Second Edition, McGraw-Hill Book Co., 1983

Atre, S., Data Base: Structured Techniques For Design, Performance, & Management, John Wiley & Sons, NY 1980

Biller, Horst and Neuhold, Erich J., "Concepts For The Conceptual Scheme", Architecture and Models in Data Base Management Systems, G. M. Nijssen (ed.), North Holland Publishing Company, Amsterdam, The Netherlands, 1977

Borkin, Sheldon A., Data Models: A Semantic Approach for Database Systems, MIT Press, Cambridge, MA, 1980

Bracchi, Giampio, "Methodologies and Tools For Logical Database Design", Data Base Management: Theory and Applications, Proceedings of the NATO Advanced Study Institute held at Estoril, Portugal, June 1-14, 1981, Clyde W. Holsapple and Andrew B. Whinston (ed.), D. Reidel Publishing Company, Dordrecht, Holland, 1983

Curtice, Robert M., and Jones, Paul E., Logical Data Base Design, Van Nostrand Reinhold Company Inc., NY, 1982

Date, C. J., An Introduction to Database Systems, Third Edition, Addison-Wesley Publishing Company, USA, 1981

Deen, S. M., "The State of the Art in Database Research", Databases: Proceedings of First British National Conference on Databases held at Jesus College, Cambridge, 13-14 July 1981, S. M. Deen and P Hammersley (ed.), John Wiley & Sons, NY, 1981

Falkenberg, Eckhard, "Some Aspects of Conceptual Data Modeling", Data Base Management: Theory and Applications, Proceedings of the NATO Advanced Study Institute held at Estoril, Portugal, June 1-14, 1981, Clyde W. Holsapple and Andrew B. Whinston (ed.), D. Reidel Publishing Company, Dordrecht, Holland, 1983

Hubbard, George U., Computer Assisted Data Base Design, Van Nostrand Reinhold Company, NY, 1981

Kahn, Beverly K., "LDDM - A Structured Logical Database Design Methodology", Lecture Notes in Computer Science, G. Goos and J. Hartmanis (ed.), Springer-Verlag, Germany, 1982

Martin, James, An End-User's Guide to Data Base, Prentice-Hall, Inc., Englewood Cliffs, NJ, 1981a

Martin, James, Principles of Data-Base Management, Prentice-Hall, Inc., Englewood Cliffs, NJ, 1981b

Nijssen. G. N., "Current Issues in Conceptual Schema Concepts", Architecture and Models in Data Base Management Systems, G. N. Nijssen, North-Holland, 1977

Olle, T W., "Data Modelling Using Entity Types, Attributes and Relationships", Data Base Management: Theory and Applications, Proceedings of the NATO Advanced Study Institute held at Estoril, Portugal, June 1-14, 1981, Clyde W. Holsapple and Andrew B. Whinston (ed.), D. Reidel Publishing Company, Dordrecht, Holland, 1983

Robinson, Hugh, Database Analysis and Design, Chartwell-Bratt, England, 1981

Smith, John Miles and Smith, Diane C. F., "Principles of Database Conceptual Design", Lecture Notes in Computer Science, G. Goos and J. Hartmanis (ed.), Springer-Verlag, Germany, 1982

Ullman, Jeffrey D., Principles of Database Systems, Computer Science Press Inc., MD, 1980

Vetter, Max and Maddison, R. N., Database Design Methodology, Prentice-Hall International, Englewood, NJ, 1981

DESIGN PROCESSES AND LOGICAL DATA MODELS

Borkin, Sheldon A., Data Models: A Semantic Approach for Database Systems, MIT Press, Cambridge, MA, 1980

Curtice, Robert M. and Jones, Paul E., Logical Data Base Design, Van Nostrand Reinhold Company Inc., NY, 1982

Vetter, Max and Maddison, R. N., Database Design Methodology, Prentice-Hall International, Englewood, NJ, 1981

GENEALOGICAL EXAMPLE

Genealogical Computing, Vol. 2, No. 1, July 1982, published by Data Associates, Inc., Fairfax, VA