COMPUTERS IN ENGINEERING 1993

CAD/CAM/CAE
COMPUTERS IN ENGINEERING EDUCATION
COMPUTATIONAL FLUID MECHANICS
COMPUTER GRAPHICS
COMPUTATIONAL HEAT TRANSFER
DESIGN AUTOMATION
EXPERT SYSTEMS
FINITE ELEMENT ANALYSIS
PROCESS CONTROL
ROBOTICS

93-AS

Computers in Engineering — 1993
ASME 1993

APPLICATION OF SEMANTIC DATA MODELING TO SCIENCE AND ENGINEERING

Naphtali Rishe
School of Computer Science
Florida International University
Miami, Florida

Qiang Li
Department of Computer Engineering
Santa Clara University
Santa Clara, California

ABSTRACT. This paper describes the application of semantic modeling using the Semantic Binary Database Model, SBM, to the design of databases having embedded measurement and scientific data, as well as complex interrelations between various objects. A case study of an environmental engineering database is considered.

1. INTRODUCTION

The semantic database models offer a simple, natural, implementation-independent, flexible, and non-redundant specification of information and its semantic aspects. Since the original idea of [Abrial-74], many semantic data models have been studied in the Computer Science literature. Many semantic models have been surveyed in [Hull&King-87] and [Peckham&Maryanski-88]. Although somewhat differing in their terminology and their selection of tools used to describe the semantics of the real world, the various semantic models are roughly equivalent. This paper's methodology uses the Semantic Binary Model (SBM) ([Rishe-92-DDS], [Rishe-89-SD]) a descendant of the model of [Abrial-74]. Models similar to SBM have been studied in: [Bracchi&*al.*-76], [Nijssen-77], [Nijssen-81], [Breutman&*al.*-79], [Senko-78], [Mark-83], [Mark-87], [Mark-89], [Meersman&Assche-83], [Vermeir-83], and others. SBM does not have as rich an arsenal of tools for semantic description as can be found in some other semantic models, e.g. the IFO model [Abiteboul&Hull-84], SDM [Hammer&McLeod-81], the Functional Model ([Kerschberg&Pacheco-76], [Shipman-81], [Chan&*al.*-82]), SEMBASE [King-84], NIAM ([Nijssen-81], [Verheijen&VanBekkum-82], [Leung&Nijssen-87]), Taxis [Nixon&*al.*-87], SIM [Jagannathan&*al.*-88], SAM [Su&Lo-80], OSAM* [Su-88-OS], GEM [Tsur&Zaniolo-84], GENESIS [Batory&*al.*-88], ER [Chen-76], Extended ER [Batini&*al.*-92]. Nevertheless, the SBM has a small set of sufficient simple tools by which all of the semantic descriptors of the other models can be constructed.

Non-procedural languages for semantic databases have been studied in [Meersman-81], [Shipman-81], [Senko-78], [Rishe-86-PS], [Rishe-91-PC], and others. Extended Pascal data manipulation language for semantic databases is defined in [Rishe-88-TM] and [Rishe-88-DDF]. Data definition languages integrated with data manipulation languages for semantic databases have been studied in [Roussopoulos&Mark-85]. Some implementation issues and architectural support have been discussed in [Li&Rishe-93-TT].

This research has been supported in part by grants from the U.S. Department of the Interior and Florida High Technology and Industry Council.

2. THE SEMANTIC BINARY MODEL

This section describes the Semantic Binary Model. A more detailed description can be found in [Rishe-92-DDS]. The semantic binary database model represents information of an application's world as a collection of elementary facts of two types: unary facts categorizing objects of the real world and binary facts establishing relationships of various kinds between pairs of objects. A definition of the model's concepts follows.

Object — any item in the real world. It can be either a concrete object or an abstract object as follows. *Value*, or *Concrete Object* — a printable object, such as a number, a character string, or a date. *Abstract Object* — a non-value object in the real world. An abstract object can be, for example, a tangible item (such as a person, a table, a country), or an event (such as an offering of a course by an instructor), or an idea (such as a course). Abstract objects cannot be represented directly in the computer.

Category (also called *Entity Type* or *Entity Set* in some semantic models) — any concept of the application's real world which is used for classification of objects. Two categories are *disjoint* if no object may simultaneously be a member of both categories. A category is a *subcategory* of another category if at every point in time every object of the former category should also belong to the latter. *Binary Relationship* — any concept of the application's real world which is a binary property of objects, that is, the meaning of a relationship or connection between two objects.

Notation: "xRy" means that object x is related by the relationship R to object y. Binary relationships are classified as *many-to-one* (*m:1*, *functional*), *one-to-many* (*1:m*), *many-to-many* (*m:m*), and *one-to-one* (*1:1*). The descriptor *proper* may be used: e.g. *proper* m:1, means that the relation is m:1 and not 1:1.

A category C is the *domain* of R if it satisfies the following two conditions: (a) whenever xRy then x belongs to C (at every point in time for every pair of objects); and (b) no proper subcategory of C satisfies (a). A category C is the *range* of R if: (a) whenever xRy then y belongs to C (at every point in time for every pair of objects); and (b) no proper subcategory of C satisfies (a). A relationship R whose domain is C is *total* if *at all times* for every object x in C there exists an object y such that xRy. (At different times different objects y may be related to a given object x.)

A *non-binary relationship* is regarded in the Binary Model as group of several simple relationships, specifically:

a. An abstract category of events. Each event symbolizes the existence of a relationship between a group of objects.

b. Functional binary relationships, whose domain is the category (a). Each of those functional binary relationships corresponds to a role played by some objects in the non-binary relationship.

Thus, the fact that objects $x_1, \ldots, x_n$ participate in an n-ary relationship R in roles $R_1, \ldots, R_n$, is represented by: an object e in the category R', and binary relationships $eR_1x_1, \ldots, eR_nx_n$.

3. SCHEMA DESIGN GOALS

A schema is said to be of *high quality* if it satisfies the following criteria (described in greater detail in Chapter 1 of [Rishe-92-DDS] and Chapter 2 of [Batini&*al.*-92]): the schema is a *natural* description of the real world; contains very little or *no redundancy*; does *not impose implementational restrictions*; covers as many *integrity constraints* as possible; the schema is *flexible* to design changes; and other minor criteria of [Rishe-92-DDS]. The most important issue of the database design is the design of a high-quality schema within the restrictions of the available DBMS and database model. A low-quality schema increases the chances of corruption of the data, makes it very hard to use and maintain the database, and makes it very hard, if not impossible, to adjust the database to the changing concepts of the application's real world. It is easy to design a high quality schema in semantic models, particularly the Semantic Binary Model. The task is much harder in the Relational Model. Moreover, it is usually impossible to describe an application world by a schema in the Relational Model with the same high quality as with which that application can be described in the Semantic Binary Model.

4. APPLICATION

This section describes a portion of the semantic schema that has been developed for the Hydrology Division of the Everglades National Park. (Actually this application is a self-contained sub-application of a larger database covering various activities of the Park and consisting of more than 1000 categories, relations, and attributes.)

The specification of every concept herein consists of:

- the **concept's name**, which should be clear and meaningful to the database users;
- **technical characteristics** of the concept; and
- a **comment** defining the meaning of the concept.

A correct definition in the comment is important. Its purposes are:

- to verify that the systems analysts correctly understand the meanings of the application's concepts;
- to concisely convey the meanings of the application to the programming personnel who will work on the application in the future;
- to provide online comments on all database entities to the future users of the database on the Client's side;
- to provide an information reference manual for use by the Client's personnel and for training of new employees, whether they will be using the database or not;
- to facilitate decision making at the Client's managerial and executive levels by providing a graphic overview and a comprehensive directory of the information owned by the Client (as a supplement to the other decision support resources: a directory of the personnel employed, a directory of financial and tangible assets owned, and the database itself);
- to specify informally integrity constraints beyond those shown in the graphical schema.

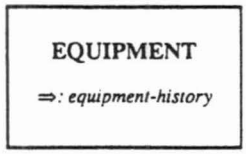
For every numeric attribute, a range of its possible values is given. For example, 23.5..100.7 means that the values may not be less than 23.5 or greater than 100.7, and that the precision is one digit after the decimal point. It is desirable that a range be as narrow as possible while still allowing for all the possible values that may be meaningful in the database at any time in the future. The range specification is used to check the input in order to eliminate data entry typos.

The schema is partitioned into several sub-schemas each of which is small enough to be displayed on one page.

The interconnections between the sub-schemas are shown by having some categories appear in more than one sub-schema. For every category, one sub-schema is the **home sub-schema** of that category and contains all of its attributes. If that category appears also in other sub-schemas, it has no attributes there, but, instead, a reference to its home sub-schema. The category appears in sub-schemas other than its home sub-schema in order to show relations with other categories of those sub-schemas.

Example 4-1.

The Equipment History sub-schema covers information relevant to all the equipment of the Park. It is the home sub-schema of the category *EQUIPMENT*. The Hydrology Stations Equipment sub-schema has additional information about the equipment installed at hydrological stations. The category *EQUIPMENT* appears here as a pointer to the Equipment History sub-schema.



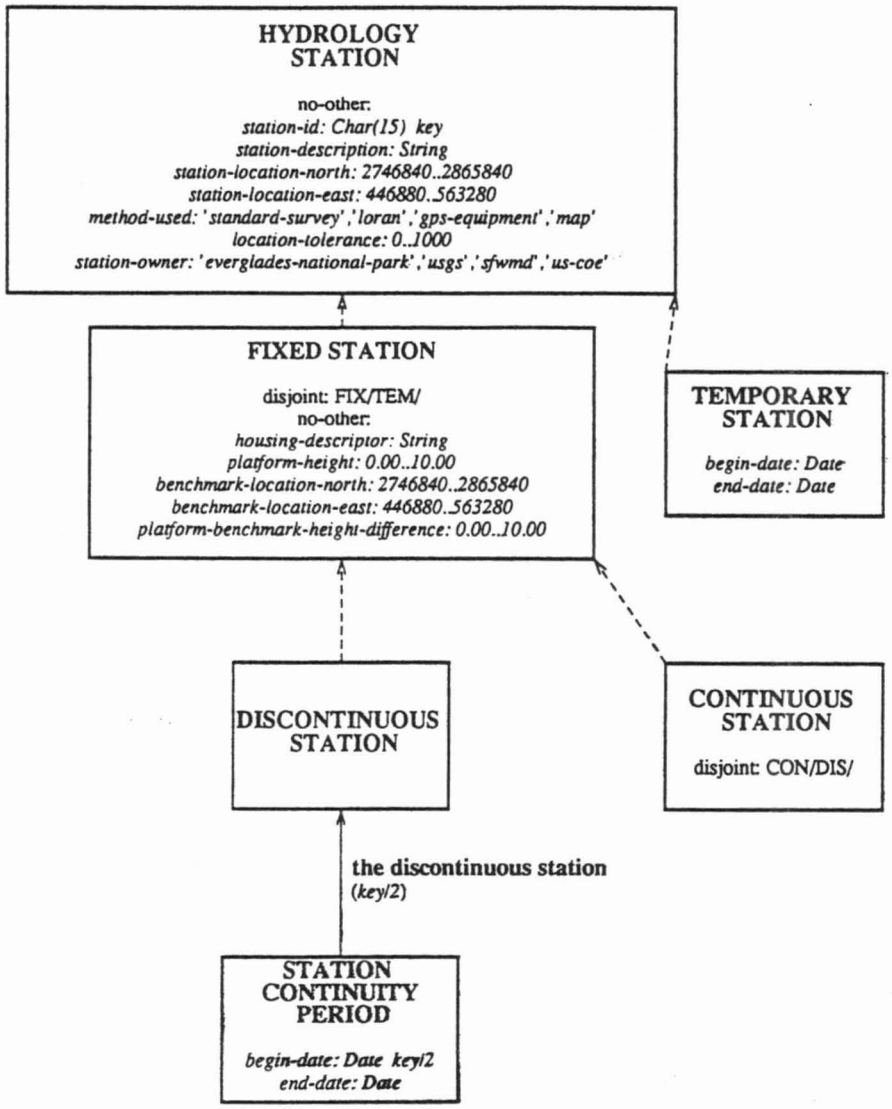


Figure 4-1. Semantic sub-schema for Hydrology Stations.

4.1. Semantic Analysis

4.1.1. Hydrology stations

HYDROLOGY-STATION — category. (A catalog of hydrology stations which reside within the Park.)

FIXED-STATION — subcategory of *HYDROLOGY-STATION*. (A hydrology station which is housed in a permanent structure.)

TEMPORARY-STATION — subcategory of *HYDROLOGY-STATION*. (A hydrology station which only exists for a period of time and is not housed in a permanent structure.)

CONTINUOUS-STATION — subcategory of *FIXED-STATION*. (A fixed hydrology station which collects data continuously.)

DISCONTINUOUS-STATION — subcategory of *FIXED-STATION*. (A fixed hydrology station which collects data only for specific intervals of time.)

STATION-CONTINUITY-PERIOD — category. (A catalog of periods during which a discontinuous station is active and various data are collected.)

the-discontinuous-station — relation from *STATION-CONTINUITY-PERIOD* to *DISCONTINUOUS-STATION* (*key/2*). (The discontinuous station which was active for periods of time collecting data.)

station-id — attribute of *HYDROLOGY-STATION*, range: *Char(15)* (*key*).

station-description — attribute of *HYDROLOGY-STATION*, range: *String* (*m:1*). (English name or designation of the station.)

station-location-north — attribute of *HYDROLOGY-STATION*, range: *2746840..2865840* (*m:1*). (UTM north coordinate of a hydrology station.)

station-location-east — attribute of *HYDROLOGY-STATION*, range: *446880..563280* (*m:1*). (UTM east coordinate of a hydrology station.)

method-used — attribute of *HYDROLOGY-STATION*, range: *'standard-survey','loran','gps-equipment','map'* (*m:1*). (The method used to derive the location coordinates of a station.)

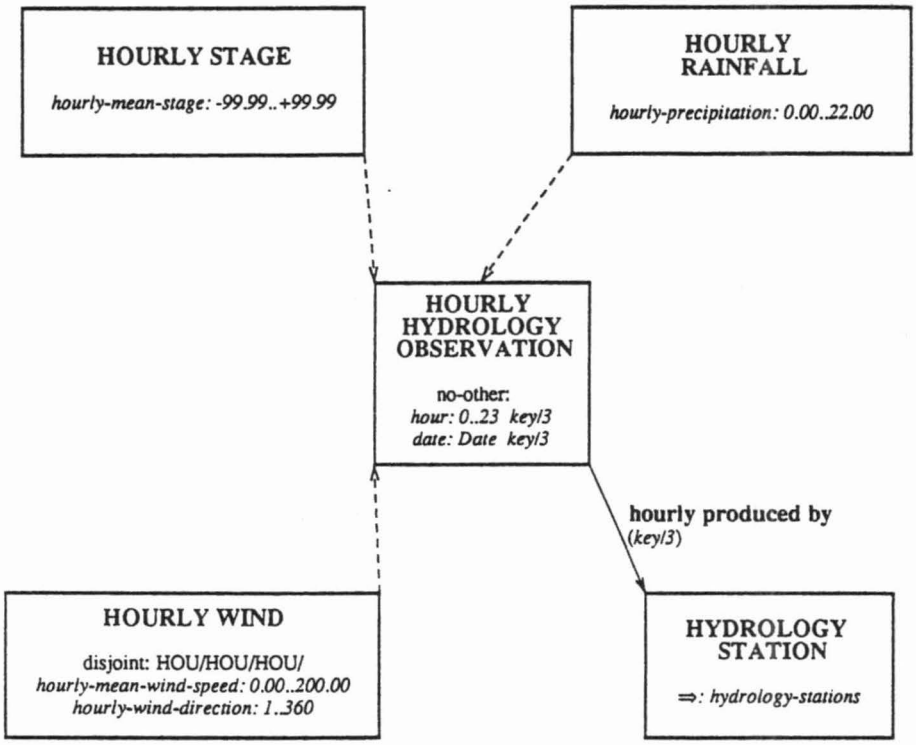


Figure 4-2. Semantic sub-schema for hourly stage and rainfall observations.

location-tolerance — attribute of *HYDROLOGY-STATION*, range: *0..1000* (*m:1*). (Tolerance of the location of a station, in feet. A value X assigned to this attribute means that the tolerance is +/-X feet.)

station-owner — attribute of *HYDROLOGY-STATION*, range: *'everglades-national-park','usgs','sfwmd','us-coe'* (*m:1*). (The agency which owns the station.)

housing-descriptor — attribute of *FIXED-STATION*, range: *String* (*m:1*). (Description of the housing of a fixed station.)

platform-height — attribute of *FIXED-STATION*, range: *0.00..10.00* (*m:1*). (The height of the station platform from the water surface, in feet.)

benchmark-location-north — attribute of *FIXED-STATION*, range: *2746840..2865840* (*m:1*). (UTM north coordinate of the benchmark which corresponds to a fixed station.)

benchmark-location-east — attribute of *FIXED-STATION*, range: *446880..563280* (*m:1*). (UTM east coordinate of the benchmark which corresponds to a fixed station.)

platform-benchmark-height-difference — attribute of *FIXED-STATION*, range: *0.00..10.00* (*m:1*). (The difference between the height of the station platform and the height of its corresponding benchmark, in feet.)

begin-date — attribute of *STATION-CONTINUITY-PERIOD*, range: *Date* (*key/2*). (The date during which a discontinuous station was activated and started the generation of data for some parameters.)

end-date — attribute of *STATION-CONTINUITY-PERIOD*, range: *Date* (*m:1*). (The date during which a period of activation for some discontinuous station ended.)

begin-date — attribute of *TEMPORARY-STATION*, range: *Date* (*m:1*). (The starting date of the life of a temporary station.)

end-date — attribute of *TEMPORARY-STATION*, range: *Date* (*m:1*). (The ending date of the life of a temporary station.)

4.1.2. Hourly hydrology observations

HOURLY-HYDROLOGY-OBSERVATION — category. (A catalog of hourly hydrology observations which originate from stations within the Park.)

HOURLY-STAGE — subcategory of *HOURLY-HYDROLOGY-OBSERVATION*. (Hourly mean stage measurements.)

HOURLY-RAINFALL — subcategory of *HOURLY-HYDROLOGY-OBSERVATION*. (Hourly total rainfall measurements.)

HOURLY-WIND — subcategory of *HOURLY-HYDROLOGY-OBSERVATION*. (Hourly wind speed and wind direction measurements.)

HYDROLOGY-STATION — category. (See subschema hydrology-stations.)

hourly-produced-by — relation from *HOURLY-HYDROLOGY-OBSERVATION* to *HYDROLOGY-STATION* (*key/3*). (The station which generates hourly either stage, or rainfall, or wind speed and direction measurements. Hourly stage and rainfall measurements can only be generated by continuous stations owned by the Everglades Research Center. Hourly wind speed and direction measurements can only be generated by stations owned by any of the following agencies: Everglades Research Center, SFWMD, US Army Corps of Engineers, NOAA.)

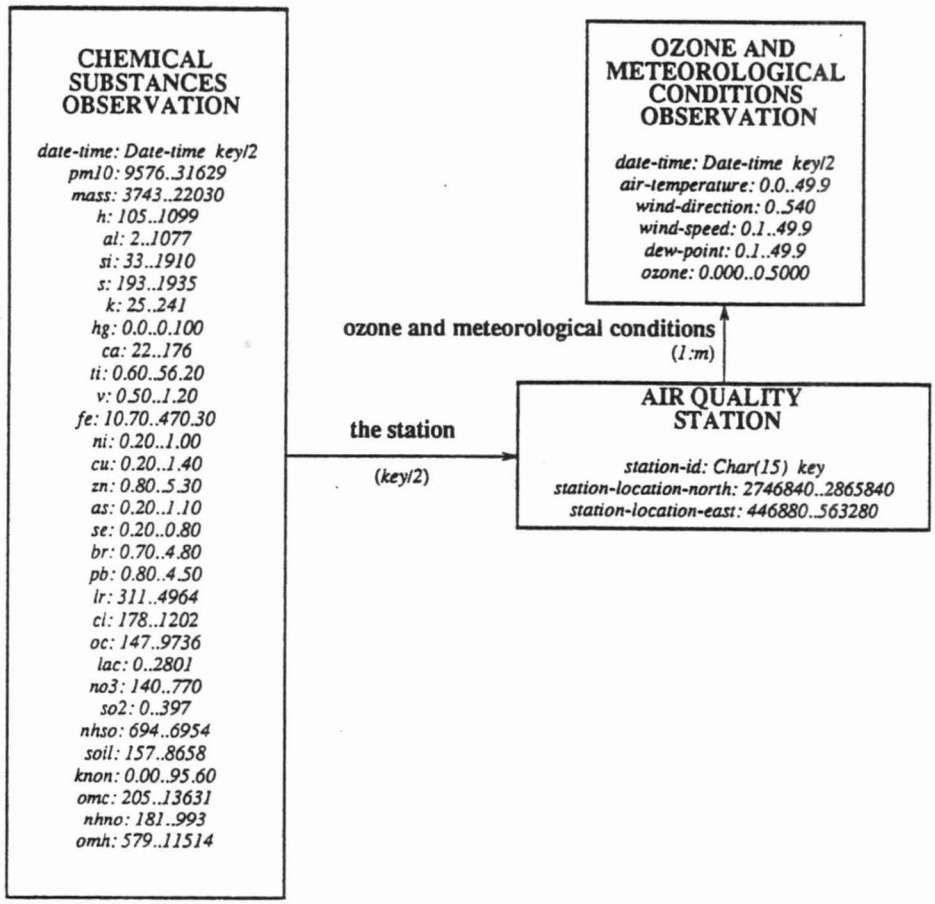


Figure 4-3. Semantic subschema for air quality.

hour — attribute of *HOURLY-HYDROLOGY-OBSERVATION*, range: *0..23* (*key/3*). (The hour during which the hydrology observation was made.)

date — attribute of *HOURLY-HYDROLOGY-OBSERVATION*, range: *Date* (*key/3*). (The date on which the hourly hydrology observation was made.)

hourly-mean-stage — attribute of *HOURLY-STAGE*, range: *-99.99..+99.99* (*m:1*). (The hourly mean stage quantity measured in ft/100. That is, the value 1.23 means 0.0123 feet. This field is left blank (null) when data is not available.)

hourly-precipitation — attribute of *HOURLY-RAINFALL*, range: *0.00..22.00* (*m:1*). (The hourly total precipitation quantity measured in inches. This field is left blank (null) when data is not available.)

hourly-mean-wind-speed — attribute of *HOURLY-WIND*, range: *0.00..200.00* (*m:1*). (The mean wind speed for the hour, measured in miles per hour. This field is left blank (null) when data is not available.)

hourly-wind-direction — attribute of *HOURLY-WIND*, range: *1..360* (*m:1*). (The wind direction for the hour, measured in degrees. This field is left blank (null) when data is not available.)

4.1.3. Air quality

AIR-QUALITY-STATION — category. (A catalog of air-monitoring stations within the Everglades National Park.)

CHEMICAL-SUBSTANCES-OBSERVATION — category. (A record of observations of selected chemicals found in the atmosphere.)

OZONE-AND-METEOROLOGICAL-CONDITIONS-OBSERVATION — category. (A record of observations of meteorological conditions and of ozone quantities found in the atmosphere.)

the-station — relation from *CHEMICAL-SUBSTANCES-OBSERVATION* to *AIR-QUALITY-STATION* (*key/2*). (The station where the chemical substances observation took place.)

ozone-and-meteorological-conditions — relation from *AIR-QUALITY-STATION* to *OZONE-AND-METEOROLOGICAL-CONDITIONS-OBSERVATION* (*1:m*). (Meteorological conditions and ozone measurements.)

station-id — attribute of *AIR-QUALITY-STATION*, range: *Char(15)* (*key*). (A string of characters which uniquely identifies each air quality station within the Park.)

station-location-north — attribute of *AIR-QUALITY-STATION*, range: *2746840..2865840* (*m:1*). (The UTM-north coordinate of the station.)

station-location-east — attribute of *AIR-QUALITY-STATION*, range: *446880..563280* (*m:1*). (The UTM-east coordinate of the station.)

date-time — attribute of *OZONE-AND-METEOROLOGICAL-CONDITIONS-OBSERVATION*, range: *Date-time* (*key/2*). (The date and time during which the observation took place.)

air-temperature — attribute of *OZONE-AND-METEOROLOGICAL-CONDITIONS-OBSERVATION*, range: *0.0..49.9* (*m:1*). (The air temperature, in degrees Celsius.)

wind-direction — attribute of *OZONE-AND-METEOROLOGICAL-CONDITIONS-OBSERVATION*, range: *0..540* (*m:1*). (The wind direction, in degrees compass.)

wind-speed — attribute of *OZONE-AND-METEOROLOGICAL-CONDITIONS-OBSERVATION*, range: *0.1..49.9* (*m:1*). (The wind speed in meters per second.)

dew-point — attribute of *OZONE-AND-METEOROLOGICAL-CONDITIONS-OBSERVATION*, range: *0.1..49.9* (*m:1*). (A general estimate of general humidity, in degrees Celsius.)

ozone — attribute of *OZONE-AND-METEOROLOGICAL-CONDITIONS-OBSERVATION*, range: *0.000..0.5000* (*m:1*). (The ozone quantity measured, in PPM units.)

date-time — attribute of *CHEMICAL-SUBSTANCES-OBSERVATION*, range: *Date-time* (*key/2*). (The date and time during which the observation took place.)

pm10 — attribute of *CHEMICAL-SUBSTANCES-OBSERVATION*, range: *9576..31629* (*m:1*). (The total mass of the particles smaller than 10 micrometers, in nanograms/cubic meter.)

mass — attribute of *CHEMICAL-SUBSTANCES-OBSERVATION*, range: *3743..22030* (*m:1*). (The total mass of the particles greater than 10 micrometers, in nanograms/cubic meter.)

h — attribute of *CHEMICAL-SUBSTANCES-OBSERVATION*, range: *105..1099* (*m:1*). (Concentration of hydrogen, in nanograms/cubic meter, for particles smaller than 2.5 micrometers.)

al — attribute of *CHEMICAL-SUBSTANCES-OBSERVATION*, range: *2..1077* (*m:1*). (Concentration of aluminum, in nanograms/cubic meter, for particles smaller than 2.5 micrometers.)

si — attribute of *CHEMICAL-SUBSTANCES-OBSERVATION*, range: *33..1910* (*m:1*). (Concentration of silicon, in nanograms/cubic meter, for particles smaller than 2.5 micrometers.)

s — attribute of *CHEMICAL-SUBSTANCES-OBSERVATION*, range: *193..1935* (*m:1*). (Concentration of sulfur, in nanograms/cubic meter, for particles smaller than 2.5 micrometers.)

k — attribute of *CHEMICAL-SUBSTANCES-OBSERVATION*, range: *25..241* (*m:1*). (Concentration of potassium, in nanograms/cubic meter, for particles smaller than 2.5 micrometers.)

hg — attribute of *CHEMICAL-SUBSTANCES-OBSERVATION*, range: *0.0..0.100* (*m:1*). (Concentration of mercury, in nanograms/cubic meter, for particles smaller than 2.5 micrometers.)

ca — attribute of *CHEMICAL-SUBSTANCES-OBSERVATION*, range: *22..176* (*m:1*). (Concentration of calcium, in nanograms/cubic meter, for particles smaller than 2.5 micrometers.)

ti — attribute of *CHEMICAL-SUBSTANCES-OBSERVATION*, range: *0.60..56.20* (*m:1*). (Concentration of titanium, in nanograms/cubic meter, for particles smaller than 2.5 micrometers.)

v — attribute of *CHEMICAL-SUBSTANCES-OBSERVATION*, range: *0.50..1.20* (*m:1*). (Concentration of vanadium, in nanograms/cubic meter, for particles smaller than 2.5 micrometers.)

fe — attribute of *CHEMICAL-SUBSTANCES-OBSERVATION*, range: *10.70..470.30* (*m:1*). (Concentration of iron, in nanograms/cubic meter, for particles smaller than 2.5 micrometers.)

ni — attribute of *CHEMICAL-SUBSTANCES-OBSERVATION*, range: *0.20..1.00* (*m:1*). (Concentration of nickel, in nanograms/cubic meter, for particles smaller than 2.5 micrometers.)

cu — attribute of *CHEMICAL-SUBSTANCES-OBSERVATION*, range: *0.20..1.40* (*m:1*). (Concentration of copper, in nanograms/cubic meter, for particles smaller than 2.5 micrometers.)

zn — attribute of *CHEMICAL-SUBSTANCES-OBSERVATION*, range: *0.80..5.30* (*m:1*). (Concentration of Zinc, in nanograms/cubic meter, for particles smaller than 2.5 micrometers.)

as — attribute of *CHEMICAL-SUBSTANCES-OBSERVATION*, range: *0.20..1.10* (*m:1*). (Concentration of arsenic, in nanograms/cubic meter, for particles smaller than 2.5 micrometers.)

se — attribute of *CHEMICAL-SUBSTANCES-OBSERVATION*, range: *0.20..0.80* (*m:1*). (Concentration of selenium, in nanograms/cubic meter, for particles smaller than 2.5 micrometers.)

br — attribute of *CHEMICAL-SUBSTANCES-OBSERVATION*, range: *0.70..4.80* (*m:1*). (Concentration of bromium, in nanograms/cubic meter, for particles smaller than 2.5 micrometers.)

pb — attribute of *CHEMICAL-SUBSTANCES-OBSERVATION*, range: *0.80..4.50* (*m:1*). (Concentration of lead, in nanograms/cubic meter, for particles smaller than 2.5 micrometers.)

lr — attribute of *CHEMICAL-SUBSTANCES-OBSERVATION*, range: *311..4964* (*m:1*). (The coefficient of absorption or equivalent soot in nanograms/cubic meter.)

cl — attribute of *CHEMICAL-SUBSTANCES-OBSERVATION*, range: *178..1202* (*m:1*). (Concentration of Chlorine, in nanograms/cubic meter, for particles smaller than 2.5 micrometers.)

oc — attribute of *CHEMICAL-SUBSTANCES-OBSERVATION*, range: *147..9736* (*m:1*). (Concentration of organic carbon, in nanograms/cubic meter, for particles smaller than 2.5 micrometers.)

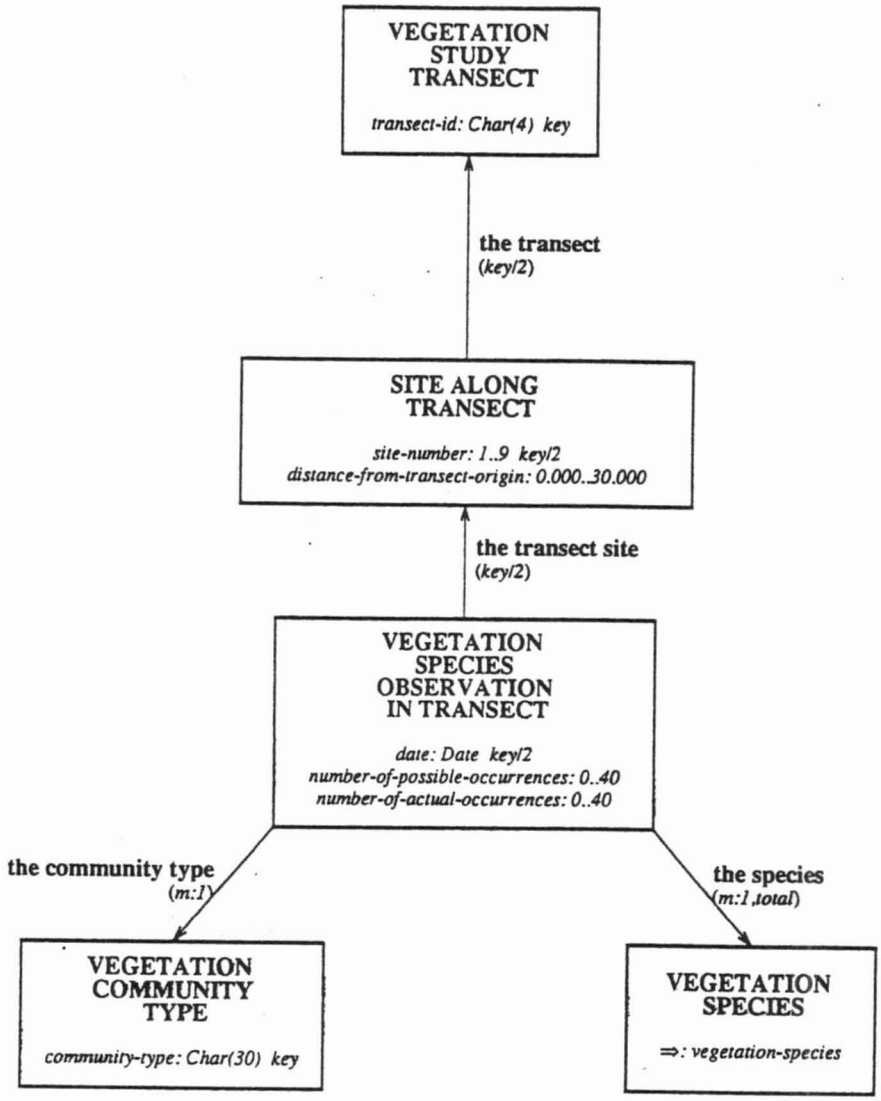


Figure 4-4. Semantic subschema for the effects of water quality on vegetation.

lac — attribute of *CHEMICAL-SUBSTANCES-OBSERVATION*, range: *0..2801* (*m:1*). (Concentration of light absorbing carbon, in nanograms/cubic meter, for particles smaller than 2.5 micrometers.)

no3 — attribute of *CHEMICAL-SUBSTANCES-OBSERVATION*, range: *140..770* (*m:1*). (Concentration of nitrate ion, in nanograms/cubic meter, for particles smaller than 2.5 micrometers.)

so2 — attribute of *CHEMICAL-SUBSTANCES-OBSERVATION*, range: *0..397* (*m:1*). (Concentration of sulfur dioxide gas, in nanograms/cubic meter, for particles smaller than 2.5 micrometers.)

nhso — attribute of *CHEMICAL-SUBSTANCES-OBSERVATION*, range: *694..6954* (*m:1*). (Concentration of ammonium bisulfate, in nanograms/cubic meter, for particles smaller than 2.5 micrometers. Measured as 3.94 * s.)

soil — attribute of *CHEMICAL-SUBSTANCES-OBSERVATION*, range: *157..8658* (*m:1*). (The sum of al, si, ca, ti, and fe, estimated oxides, soil potassium and measured components. Measured in nanograms/cubic meter, for particles smaller than 2.5 micrometers.)

knon — attribute of *CHEMICAL-SUBSTANCES-OBSERVATION*, range: *0.00..95.60* (*m:1*). (Concentration of nonsoil potassium, in nanograms/cubic meter, for particles smaller than 2.5 micrometers. Estimated from fe and k.)

omc — attribute of *CHEMICAL-SUBSTANCES-OBSERVATION*, range: *205..13631* (*m:1*). (Organic mass from carbon, in nanograms/cubic meter, for particles smaller than 2.5 micrometers.)

nhno — attribute of *CHEMICAL-SUBSTANCES-OBSERVATION*, range: *181..993* (*m:1*). (Concentration of ammonium nitrate, in nanograms/cubic meter, for particles smaller than 2.5 micrometers. Measured as 1.29 * no3.)

omh — attribute of *CHEMICAL-SUBSTANCES-OBSERVATION*, range: *579..11514* (*m:1*). (Organic mass from hydrogen and sulfur, in nanograms/cubic meter, for particles smaller than 2.5 micrometers. Measured as s*(11*(h-0.156*s)).)

The objects of the category *CHEMICAL-SUBSTANCES-OBSERVATION* are identified by: date-time the-station.

The objects of the category *AIR-QUALITY-STATION* are identified by: station-id.

The objects of the category *OZONE-AND-METEOROLOGICAL-CONDITIONS-OBSERVATION* are identified by: date-time ozone-and-meteorological-conditions.

4.1.4. Water quality effects on vegetation

VEGETATION-SPECIES-OBSERVATION-IN-TRANSECT — category. (A record of vegetation species observations along different transects.)

VEGETATION-STUDY-TRANSECT — category. (A catalog of the various transects used for the vegetation observations.)

SITE-ALONG-TRANSECT — category. (A catalog of sites along different transects.)

VEGETATION-COMMUNITY-TYPE — category. (A catalog of vegetation community types.)

VEGETATION-SPECIES — category. (See subschema vegetation-species.)

the-species — relation from *VEGETATION-SPECIES-OBSERVATION-IN-TRANSECT* to *VEGETATION-SPECIES* (*m:1,total*). (The vegetation species observed.)

the-transect-site — relation from *VEGETATION-SPECIES-OBSERVATION-IN-TRANSECT* to *SITE-ALONG-TRANSECT* (*key/2*). (The site along a transect where the observation took place.)

the-transect — relation from *SITE-ALONG-TRANSECT* to *VEGETATION-STUDY-TRANSECT* (*key/2*). (The transect upon which the site lies.)

the-community-type — relation from *VEGETATION-SPECIES-OBSERVATION-IN-TRANSECT* to *VEGETATION-COMMUNITY-TYPE* (*m:1*). (The community type of the vegetation species observed.)

community-type — attribute of *VEGETATION-COMMUNITY-TYPE*, range: *Char(30)* (*key*). (A description of the vegetation community type.)

transect-id — attribute of *VEGETATION-STUDY-TRANSECT*, range: *Char(4)* (*key*). (A code that uniquely identifies each transect.)

site-number — attribute of *SITE-ALONG-TRANSECT*, range: *1..9* (*key/2*). (A code which uniquely identifies each site along a transect.)

distance-from-transect-origin — attribute of *SITE-ALONG-TRANSECT*, range: *0.000..30.000* (*m:1*). (The distance of the site from the origin of the transect, in kilometers.)

date — attribute of *VEGETATION-SPECIES-OBSERVATION-IN-TRANSECT*, range: *Date* (*key/2*). (The date on which the observation took place.)

number-of-possible-occurrences — attribute of *VEGETATION-SPECIES-OBSERVATION-IN-TRANSECT*, range: *0..40* (*m:1*). (The number of possible occurrences of the vegetation species under observation.)

number-of-actual-occurrences — attribute of *VEGETATION-SPECIES-OBSERVATION-IN-TRANSECT*, range: *0..40* (*m:1*). (The number of actual occurrences of the vegetation species under observation.)

The objects of the category *VEGETATION-SPECIES-OBSERVATION-IN-TRANSECT* are identified by: date the-transect-site.

The objects of the category *VEGETATION-STUDY-TRANSECT* are identified by: transect-id.

The objects of the category *SITE-ALONG-TRANSECT* are identified by: site-number the-transect.

The objects of the category *VEGETATION-COMMUNITY-TYPE* are identified by: community-type.

4.1.5. Exotic plants control

EXOTIC-PLANT-SITE — category. (A catalog of sites to be monitored to control exotic vegetation.)

SEED-TRAP — category. (A catalog of traps for seeds in sites.)

SEED-IN-TRAP-OBSERVATION — category. (A record of monitoring of a trap to find the number of seeds of a particular species.)

VEGETATION-LAYER — category. (A catalog of layers observed to classify the colonizing vegetation into layers according to height.)

WELL — category. (A catalog of wells. This helps in calculating hydrology data. There is one well in each site, 1..63, plus one central well, #64)

WETLAND-TYPE — category. (A catalog of Federal wetland types.)

WELL-OBSERVATION — category. (The monitoring of a well to calculate water levels.)

SUBPLOT-OBSERVATION-FOR-SEEDLINGS — category. (A record of observations of seedlings in the subplots of sites. A 1 sq.meter area in the south east corner of a site is taken as subplot.)

VEGETATION-SPECIES-OBSERVATION — category. (A record of observation of the species found in the layer.)

VEGETATION-SPECIES — category. (See subschema vegetation-species.)

observes-well — relation from *WELL-OBSERVATION* to *WELL* (*key/2*). (The well that was being monitored.)

observes-trap — relation from *SEED-IN-TRAP-OBSERVATION* to *SEED-TRAP* (*key/2*). (The trap that was monitored.)

site — relation from *VEGETATION-SPECIES-OBSERVATION* to *EXOTIC-PLANT-SITE* (*key/4*). (One species can be observed more than once in a site.)

observed-in-site — relation from *SUBPLOT-OBSERVATION-FOR-SEEDLINGS* to *EXOTIC-PLANT-SITE* (*key/2*). (Each site has one subplot.)

wetland-type — relation from *VEGETATION-SPECIES* to

WETLAND-TYPE (*m:1*). (The Federal wetland type in which the species is found.)

the-species-observed — relation from *VEGETATION-SPECIES-OBSERVATION* to *VEGETATION-SPECIES* (*key/4*). (One species can be observed many times.)

in-layer — relation from *VEGETATION-SPECIES-OBSERVATION* to *VEGETATION-LAYER* (*key/4*). (More than one species can be observed in one vegetation layer.)

well-number — attribute of *WELL*, range: *1..64* (*key*). (The identification number for the well.)

The objects of the category *WELL* are identified by: well-number.

site-number — attribute of *EXOTIC-PLANT-SITE*, range: *1..63* (*key*). (The unique identification number of the site.)

The objects of the category *EXOTIC-PLANT-SITE* are identified by: site-number.

the-date — attribute of *VEGETATION-SPECIES-OBSERVATION*, range: *Date* (*key/4*). (The date on which the site was monitored.)

cover-class — attribute of *VEGETATION-SPECIES-OBSERVATION*, range: *'1','2','3','4','5','8','9'* (*m:1*). (The cover class in which vegetation species is found. 1-less than 5% covered, 2-5 to 25% covered, 3-25 to 50% covered, 4-50 to 75% covered, 5-greater than 75 % covered, 8-few with small cover, 9-solitary with small cover.)

life-form — attribute of *VEGETATION-SPECIES-OBSERVATION*, range: *'H','S','T','LI'* (*m:1*). (The growth habit of the plant. H-Herb, S-Shrub, T-tree, LI-liana.)

The objects of the category *VEGETATION-SPECIES-OBSERVATION* are identified by: the-species-observed the-date in-layer site.

trap-number — attribute of *SEED-TRAP*, range: *1..81* (*key*). (The unique identification number of the trap. Trap number is same as the site number if it is within the site. Else it is greater than 63.)

trap-location-east — attribute of *SEED-TRAP*, range: *446880..563280* (*m:1*). (The UTM-east coordinate of the station.)

trap-location-north — attribute of *SEED-TRAP*, range: *2746840..2865840* (*m:1*). (The UTM-north coordinate of the trap.)

The objects of the category *SEED-TRAP* are identified by: trap-number.

the-date — attribute of *SEED-IN-TRAP-OBSERVATION*, range: *Date* (*key/2*). (The date on which the trap was monitored)

number-of-seeds — attribute of *SEED-IN-TRAP-OBSERVATION*, range: *0..9999* (*m:1*). (Number of seeds found in the trap during the observation.)

The objects of the category *SEED-IN-TRAP-OBSERVATION* are identified by: observes-trap the-date.

shelf-elev-ft — attribute of *WELL-OBSERVATION*, range: *4..6* (*m:1*). (The shelf elevation when the well is monitored: the top of the well above the mean sea level, in feet.)

down-to-water-ft — attribute of *WELL-OBSERVATION*, range: *-5.0..+5.0* (*m:1*). (Height from top to meniscus. Measured in feet.)

date-time — attribute of *WELL-OBSERVATION*, range: *Date* (*key/2*). (The date and time when the well was monitored.)

The objects of the category *WELL-OBSERVATION* are identified by: observes-well date-time.

layer-number-in-a-site — attribute of *VEGETATION-LAYER*, range: *1..6* (*m:1*). (The vegetation layer number in a site. 1-submerged, 2-liana, 3-0 to 1m, 4-1 to 2m, 5-2 to 5m, 6-greater than 5m.)

wetland-code — attribute of *WETLAND-TYPE*, range: *Char(6)* (*key*). (The unique code for the particular wetland type.)

wetland-description — attribute of *WETLAND-TYPE*, range: *String* (*m:1*). (Description of the particular wetland type.)

The objects of the category *WETLAND-TYPE* are identified by: wetland-code.

the-date — attribute of *SUBPLOT-OBSERVATION-FOR-SEEDLINGS*, range: *Date* (*key/2*). (The date in which the subplot was observed.)

number-of-seedlings — attribute of *SUBPLOT-OBSERVATION-FOR-SEEDLINGS*, range: *0..9999* (*m:1*). (Number of seedlings observed in the subplot observation.)

The objects of the category *SUBPLOT-OBSERVATION-FOR-SEEDLINGS* are identified by: the-date observed-in-site.

Acknowledgment

The authors greatly appreciate the help of Michael Alexopoulos and Ravichandra Kallem.

References

[Abiteboul&Hull-84] S. Abiteboul and R. Hull. "IFO: A Formal Semantic Database Model," *Proceedings of ACM SIGACT-SIGMOD Symposium on Principles of Database Systems*, 1984.

[Abrial-74] J.R. Abrial, "Data Semantics," in J.W. Klimbie and K.L. Koffeman (eds.), *Data Base Management*. North Holland, 1974.

[Batini&*al.*-92] Batini, C., Ceri, S., Navathe, S.B., *Conceptual Database Design: an Entity Relationship Approach*, Benjamin Cummings, 1992.

[Batory&*al.*-88] D.S. Batory, T.Y. Leung, and T.E. Wise. "Implementation Concepts for an Extensible Data Model and Data Language." *ACM Transactions on Database Systems*, vol. 13, no. 3, September 1988, pp. 231-262.

[Bracchi&*al.*-76] G. Bracchi, P. Paolini, G. Pelagatti. "Binary Logical Associations in Data Modelings." In G.M. Nijssen (ed.), *Modeling in Data Base Management Systems*. IFIP Working Conference on Modeling in DBMSs, 1976.

[Breutman&*al.*-79] B. Breutman, E. Falkenberg, and R. Mauer. "CSL: A Conceptual Schema Language." *Proceedings IFIP TC 2 WG 2.6 Working Conference*, March 1979.