USER PERSPECTIVE AND DESIGN EMPHASIS:
EXPERIENCE WITH A NETWORK MODEL OF THE FOOD PROCESSING SYSTEM (ALINET)*

Elizabeth R. Ducot
Alexander H. Levis
Laboratory for Information and Decision Systems
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

ABSTRACT
An approach that stimulates both qualitative and quantitative network study in the same modeling environment is presented. Insights are supported by the experience with ALINET, a network model designed for the analysis of energy use in the food processing and distribution sector and for the evaluation of the potential effectiveness of energy conserving technologies. The model specifications and the design of the computer software are described in order to provide the background for some observations on the relationship between the issues to be addressed by the model, the needs of the users and the evolution of the operational modeling system. Examples are drawn from the procedures by which portions of the network are specified and constructed.

1. INTRODUCTION
The controversy surrounding the use of computer models to support human design and analysis activities continues, across many disciplines and application areas. With the increasing complexity of the problems being addressed and the systems being studied, the disillusioned voices form a greater part of the controversy. These complaints, common in varying degrees to all modeling approaches, are indeed legitimate; touching on the inadequacy of tools, data and methodology [Greenberger (1976)], the incomprehensibility of models as implemented [Rosenberg (1980)], and the poor communication between the model builder, the model user and the client or decision maker [Greenberger (1976)], [Lipinski, et.al. (1978)].

Motivated by these concerns, progress is being made through the development of theory, algorithms and languages; the design of procedures for implementation and documentation and the understanding of considerations for system use, validation and interpretation of results. Yet it is significant to note that the increasingly important role of network modeling in the design and analysis of a variety of systems is due only in part to these advances. The explanatory power of the network construct is fundamental, derived not only from the ease with which many systems can be conceived of as networks, but also from the effectiveness of the visual representation in communicating the significant features of the analysis [Whitehouse (1975)].

How best to exploit these advantages in order to create a supportive design and analysis environment is therefore an area of ongoing interest. In Section 2 of the paper that follows, some of the aspects of existing applications are explored in order to distinguish the primary modeling forms used and to identify the support requirements for qualitative versus quantitative analysis. The attributes of problems that require a cooperative modeling effort, the blending of modeling forms (designated "mixed networks") and a framework for describing these modeling systems are suggested. Section 3 describes the experience with a particular mixed network modeling application in the development of a model of the food processing and distribution sector (ALINET); while Section 4 indicates selected features of the system components necessitated by the modeling environment.

* This work was supported by the Office of Industrial Programs, U. S. Department of Energy, under Contract No. DE-AC-03-77CS40021, awarded to Systems Control, Inc., with the Laboratory for Information and Decision Systems of MIT as subcontractor.
2. NETWORK MODELING

Models in general can be distinguished according to form (schematic, physical, symbolic and role playing), purpose (descriptive, prescriptive and normative) and use [Greenberger (1976)]. Added to this may be a designation of the generality of the model, and its treatment of time or uncertainty [Murdick and Ross (1976)]. Each of these characteristics denotes a major area of development, requiring the support of a number of modeling methodologies and techniques. Examples of network models, drawn from such diverse areas as medical diagnosis, computer communications, semantics, chemistry, production planning, electric power, water and transportation, can be found that exhibit almost any combination of these attributes. The unifying characteristic of computer-supported network modeling activities, however, is the fact that schematic and symbolic forms are integrated within a single approach. Indeed, much of the power attributed to this general modeling area is derived from the way in which the aspects of this blended form are exploited.

In order to identify the desirable features of a modeling and simulation environment, it may be useful to begin with a subjective decomposition of network modeling applications according to which of the two aspects of the blended form (schematic or symbolic) is dominant. While this is admittedly a fuzzy distinction, it provides a vehicle for articulating user requirements. For purposes of discussion, the designations may be defined as follows:

- **Iconic dominance** indicates the emphasis on the visual aspects of the graph. The underlying systems described may be physical or abstract. The primary interaction between the user and the network representation consists of modifying the output structure based on subjective considerations. The symbolic components of the models may be extremely complex; for example, the automatic routing software for developing layouts of dense, double sided printed circuit boards [Hosking, et.al. (1978)]. Yet, despite the network analysis conducted during the design process, the fundamental value of the model is measured by the effectiveness of the schematic it constructs.

- **Symbolic dominance** indicates the emphasis on the outcome of the symbolic (computer supported) portions of the model, conveyed in terms of the values of system variables, measures of system performance etc. The network representation may be viewed in this case as a construct for improving communication between the user and the model; e.g. for input specification and output display. The extensiveness efforts to simplify the user/modeling system interface (i.e., the development of interactive graphics for direct graphical input and display) should not obscure the primary modeling emphasis. The majority of the network models of large scale systems fall into this category; exemplified by water, power, transportation and communication systems. There is an underlying assumption inherent in many of the systems studies: While detailed data required to specify the network parameters and to formulate the scenarios for study may be difficult to obtain, the specification of the network is a straightforward (albeit not necessarily simple) process.

From the preceding description, one can perceive a fundamental difference in the direction of the modeling flow for models in the two basic categories. The interconnection of generic modeling elements is depicted in Figure 1:

![Figure 1. Generic Modeling Elements](attachment:image.png)

The primary user interaction point is assumed to be attached to a system data base (D). For iconic dominant models, the primary flows are described by: 1) user requests translated, through the execution of the modeling software, into desired network (D-S-I), 2) network structures, viewed by user (I-D) and modified (D-I), and 3) results saved for future reference (I-D). The symbolic dominant models, on the other hand, primarily support user requests to 1) describe structure and combine it with appropriate system data (D-I), 2) execute required studies (I-S), and 3) display (S-I) and store (S-D) results.

However, regardless of the differences identified, an integrated support system for either approach ideally provides for 1) simple user/system interface languages, 2) network display and interaction tools, 3) error checking and iteration mechanisms, 4) data processing and analysis.
routines, 5) multiple access points for user intervention and 6) provisions for nested study execution, evaluation and restart.

Two questions naturally arise: Does a "mixed network" designation make sense? And, if so, are additional support features required? In defining a mixed network model it is important that it not

Figure 2. Mixed Network Modeling
merely imply a question of indifference in the choice between an iconic dominant or symbolic dominant designations; e.g., for models that may be viewed either way. Rather, what is implied is the potential for the same network structure to serve two distinct functions in different user/modeling situations. In order to characterize such an instance, one needs only imagine a potential project in which the output of a structural modeling exercise is used to motivate data gathering and analysis activities and, ultimately, to drive a complex simulation model.

Figure 2 depicts the user activities within a mixed network modeling environment drawn to indicate the separation of the qualitative (iconic dominant) aspects from the quantitative (symbolic) aspects of the problem and the sequence of steps outlined above. A desirable modeling environment clearly includes the features of an Integrated Modeling System identified above, however there are several implications underlying the flow indicated in Figure 2: 1) the existence of two distinct symbolic components (indicated by blocks labeled “Network Analysis/Synthesis” and “Study”), 2) a difference in access frequency for Data and Schematic components to support distinct user groups and 3) an increased likelihood of a discontinuous study, thus suggesting greater need for on-line documentation and summary of previous results.

After some consideration, two interrelated attributes can be conjectured that lead to mixed network formulation:

1) Technical considerations that require the extension of the network boundaries across organizational lines. This problem has been described in terms of a conflict between mission and organization [Churchman (1968)], and is expected to appear with increasing frequency in instances requiring interdisciplinary expertise.

2) A non-traditional problem orientation which makes the network difficult to represent directly. Gaps in the knowledge of the network structure and a lack of relevant data have become increasingly evident in dealing with current problems, for example, energy conservation.

In each case, a schematic of the system under study, reflecting various levels of understanding, becomes a tool for 1) integrating individual perceptions of the appropriate network elements and relationships and 2) eliciting a form of collective intuition to improve the consistency of the results.

3. ON A NETWORK MODEL OF THE FOOD PROCESSING SYSTEM

The attributes of mixed network modeling identified above were characteristic of the modeling environment surrounding the development of ALINET (Allimentary Industry Network). Inherent conflicts between the model requirements, derived from the problem statement, and the operational conditions of the real system are apparent from the description that follows.

The modeling effort was motivated by the need to provide the capability of analyzing energy use in the energy intensive food processing and distribution sector and to evaluate energy conserving technologies. This implied a modeling framework that had to be both comprehensive in order to reflect the structural changes that continue to occur in this sector and, at the same time, sufficiently detailed to permit the study of innovations in specific processes. In addition, explicit characterization of the transformation that the farm commodities undergo during conversion to food products was required in order to account for factors of processing and distribution. Measures of productivity were needed in order to provide insight into the effectiveness of the new and alternative technologies.

These characteristics were translated, late in 1977, into a model in which the processing of the commodities (e.g., wheat, soybeans or milk into bread, oil or cheese) is represented by a sequence of transformations. Each of the processes changes the state of the material, whether in form, in location in ownership or in value. In order to carry out these transformations, inputs are required and, specifically, energy in various forms. Nodes designate the states, while the links depict the transformations (e.g., processes such as milling, baking, canning, etc., each characterized by a production function). The material flow representation, therefore, is an acyclic directed graph -- which allows for the explicit consideration of time and the sequential determination of the flows in the network. Associated with the material flow there is information flow, and specifically, information about prices and quantities of commodities and inputs. In order to avoid problems inherent in flow graphs with loops, the information is placed on a tableau and partitioned according to stages in the material flow graph. The resulting information matrix evolves with time within the annual cycle and is available to the whole system at each time instant.

The modeling of a specific sector of the food processing industry, whether at the national or regional level (suitable for government policy studies), or at the facility level (more appropriate for company studies), consists of two parts: First, the determination of the structure, i.e., the interconnection of nodes (states) and links (processes) that represent the segment of the industry being modeled, and second, the establishment of the numerical values of the attributes that characterize the material flow and the various processes.

Once a specific model has been developed and verified, many different analyses can be carried out: from simple simulations of the material flow and accounting of energy use, to the assessment of the impact of alternative technologies. The structure of the model is also consistent with the formulation of operations research type of problems such as optimum resource allocation or distribution.
However, the operational conditions of the real system, for providing much needed insight and information to support this approach, were far from ideal. The overall system, a complex and loosely organized network of processors and distributors, was, for the most part, one finely tuned to prevailing economic conditions of the recent past. Requirements for labor and raw materials were balanced against the need for capital resources for plants and equipment; using standard procedures, experience, and many traditional plant optimization and forecasting models to generate local decisions.

Common practices and regulations governing the transportation and distribution of products and supplies were based on considerations that largely predated the issues of energy cost or conservation. Moreover, technological innovation, allowing not only for higher production but also for the creation of increasingly specialized food products, to a substantial degree relied on the extensive substitution of non-renewable resources for labor. Clearly, the abundance of cheap energy on which these technologies and practices are based can no longer be taken for granted. Yet, so imbedded were the energy related assumptions in the operating procedures of this industry that little relevant data, experience or expertise specific to energy use had been developed. Without it, representatives of this industry recognized the considerable difficulties to be faced in assessing their energy requirements, whether in response to the rapid changes in the economics of energy or to the policies directed toward conservation and the development of energy conserving technologies. In addition, it was clear that changes in existing technologies often require substitution between two fundamental types of energy consuming processes -- those related to distribution and storage (changes in location and time) and those related to the engineering processes (changes in form). Thus the effects of proposed innovation are likely to extend well beyond the boundaries of a facility or segment of the industry. Moreover, despite the fact that a single corporation may control a significant portion of the flow of a specific product from farm to consumer, it was recognized that the perception of commodity flows tends to be local within organizations; expressed in terms of individual inputs and outputs for specific processes.

On the one hand, the need for a comprehensive, yet detailed, approach was apparent; on the other hand, the problems inherent in taking a non-traditional energy-related view of the system (e.g., organizational contraints and insufficent data) were also evident. Two workshops, innumerable technical exchanges and a cooperative pilot project with the Pillsbury Company were conducted in order to stimulate the collective insight and resources of users, sponsors and developers. As a result, a data base was created that provided not only opportunity for model development but also, of necessity, for experimentation in the design of a network modeling and simulation support system for mixed network studies.

4. SYSTEM STRUCTURE AND USE

The three functional entities; the Data System, the Flow Graph Generator and the Simulation System; correspond to the network building blocks described in Section 2. The potential for independent access to the blocks, interconnected as shown in Figure 3, is significant and is, in some sense, contrary to many of the current trends for integration evident in both modeling language development and network application [Pritsker (1978)], [Roberts (1978)].

The requirements for "stand alone" access, in the context of an integrated modeling system, are derived from the division between qualitative and quantitative modeling support.

Since the Simulation System is not a shared resource in the mixed network modeling environment, its function is clear and therefore a brief description of its structure is sufficient. (A detailed discussion can be found in [Levis (1979)].) The representation of the network structure, acquired through interactions with the Data System and the Flow Graph Generator, provides a framework within which a variety of studies or analyses of the material flow under different conditions can be performed. The execution and post processing steps are straightforward, with iteration points indicated by instruction blocks. The steps for input and study definition, while equally simple, are presented to indicate the types of user support provided. According to user instructions, the data extraction routine produces a sequence of data base requests for the network of interest. The process and material flow data appropriate to a specified time period are returned by the Data System and integrated with default values for all system inputs to create a baseline simulation file. In the absence of specific change instructions, the baseline is used to initialize the information tableau. Two forms of modifications to the data, are then accepted: 1) changes to single data values via entries in a temporary update file, and 2) global changes through the introduction of the user specified process models. The existence of a temporary update capability was considered essential in that it permits a variety of "what if" questions to be posed without affecting the permanent data base.

The Data System and Flow Graph Generator constitute the shared resources of the system. A brief description of these entities, as implemented in ALINET, may be appropriate before indicating the general features of the user environment provided. The Data System controls and maintains two distinct ALINET data bases. In addition to the data bases, it also contains software for data validation and element creation as well as routines that process users' requests and produce reports. (Thus the Data System can also be used independently as a data base of the food processing industry.) The primary component of the system is the Element Data Base, containing
information about the processes, commodities and constraints operating within the food processing and distribution sector. Subordinate to it is the facility data base, containing specific information (i.e., capacity, shipping, product types, etc.) for individual food processing plants within the U. S. An interconnection between the Facility Data Base and the system Element Data Base is provided via a set of algorithms for

**Figure 3. Structure of ALINET**
performing simple verification and data generation tasks [Levis et al. (1979)]. Each node in the
Element Data Base is characterized by an open ended string of attributes that includes a coded
node name that reflects the commodity in question (e.g., wheat), the level of geographical detail
(national, state, or facility specific), and an index number. Other node attributes are used to
specify the particular type of the commodity (e.g., durum wheat), the form the commodity has at
that node (e.g., durum wheat flour), a qualifier (e.g., packaged durum wheat flour), the location
(e.g., packaged durum wheat flour at the mill), and the end use or market destination (e.g., for
domestic use).

A similar procedure is used to specify a link, where in addition to the other attributes characterizing the process represented by the link (process identifier, process coefficient, process
capacity, distance, allocation coefficients, input vector), the corresponding origin and destination
nodes are also specified. A distinction is made between the link and the process it represents in
that the process entry in the data base contains the generic characteristics of the process (e.g.,
transportation by truck), while a link corresponds to a specific occurrence of that process (e.g.,
transportation by truck of packaged durum flour from mill at a given location to a pasta plant at
another given location). Each process, node or
link attribute can serve as a basis for user
interrogation of the system. A hierarchical data
model is used in which information on lower tiers
represents the seasonal variations appropriate to
a specific element. This approach, focused on
individual elements, provides the desired
flexibility for incorporating new technologies and
modifying the characteristics of existing ones.

In order to simulate the flows of materials
through the system, an identifiable set of nodes
and links must be associated to create an ordered
network. Depending on user instructions, many
different partial flow graphs at different levels of
detail may be drawn from the same basic data
base. Loops, implying an error in the data
specification, can be detected as the graph is
constructed. The result of the graph generation
procedure, stored in the Network Data Base, is the
organization of the selected data elements into a
sequence of levels, each containing a set of
nodes, which are related to the stages of
processing within the physical system. The paths
through the network, along which material
flows, correspond to the sequence of links
interconnecting these nodes. The Network Data
Base is also specified directly by the Simulation
System. In addition it drives the software that
creates the visual displays of the graph
automatically in a way that permits both
verification and structural insight.

In order to describe the user support features
implied by this implementation, the interactive
modeling team may be considered as two distinct
groups: one responsible primarily for the
determination of the network structure, (User 1)
and the other (User 2) for development of
numerical parameters and execution of the study.

A conversational data handling language is
important for both user groups. Since the
desirability of particular language forms is to
some extent a matter of personal preference, no
further detail is presented here. It is
sufficient to note that a commercial data
management package, selected to accommodate other
user requirements [MRI Systems Corp. (1974)] was
used. The desirability of atomic data elements
corns both groups in different ways. User 1
requires the capability of manipulating individual
elements to build local portions of the graph.
Although he relies on visual representation of the
working graphs to suggest what to do next, he
should not be required to relate his
inputs/changes directly to the structure. This
independence is possible because, unlike some
activity network based applications [Elmaghraby
(1977)] there is, in general, no precedence
ambiguity for the system; thus no dummy elements
or other structural references are required. The
independence of elements is also of interest to
User 2, primarily in specifying scenarios and
changes in parameters. Operators for creating
sets (according to element attributes) are
important primarily to assist User 1 in
qualitative analysis; however these same operators
are fundamental to the graph construction process
as well. Operators for performing simple
arithmetic calculations and data analysis, on the
other hand, support the needs of User 2 for
parameter development. More extensive modeling
support is certainly available [Lipinski, et al.
(1978)] and could be included in a general
interactive modeling environment. However for
this application more modest tools were
sufficient.

The need for additional supporting data,
accessible by the system, is problem specific.
However in general, one may include 1) information
about the organization and operation of the system
(i.e., facilities, their products, technologies,
locations, etc.; primarily of interest to User 1)
and 2) information about the problem (i.e., energy
data, process descriptions, etc.; primarily of
interest to User 2).

The Flow Graph Generator as indicated above must
support (and capitalize upon) the independent
element assumption. In addition due to the
structural interpretation placed on the graph and
focus on individual technologies, it must maintain
element integrity, thus no reliance on structural
equivalence appropriate in many other applications
[Elmaghraby (1977)] can be made. Aggregation,
although automated should be under user control
according to attributes of interest. Likewise,
the partition of graphs into subgraphs should be
based on qualitative (attributes, stages of
production, etc.) rather than analytical
considerations.

A small portion of a state level aggregated flow
graph -- a hand drafted copy of a computer
generated graph -- serves as a basis for the
remaining discussion. The overall graph from
which this piece was abstracted depicted the flow
of durum grain (used for the production of pasta) in
Minnesota. The structure of the graph in
Figure 4, indicating the left to right flow of grain from collection to the mills, can be described in terms of a sequence of vertical lines (e.g., levels), each containing a set of nodes. The levels correspond to physical stages in the processing of durum wheat. Major stages, in this case collection (the left hand edge of Figure 4) and preprocessing (the right) are identified.
explicitly and correspond to partitions on the information tableau.

Direct user specification of the structure, in terms of lists of nodes on each level and sets of links emanating from each node, although effective in other application [Collins and Defanti (1980)], is clearly impractical in the setting of iterative network definition for large systems.

The algorithm employed by the Flow Graph Generator translates the notion of sequential flows in the physical system to the simple graphical requirement that all transitions advance from left to right. The additional restriction that a node appear no more than once in a given representation allows a unique graph to be generated from a series of starting points or source nodes, the "one step forward" information contained in the link definitions which specify origin and destination and the left to right progression. The procedure is basically the same as that for most current graph manipulation languages [Delgrande (1980)]. For given values in the data base, the set of elements in the unique graph is designated the reachable set.

The element set selected for a particular graph is a subset of the reachable set, modified by simple focus instructions specified by the user. Focus can be defined, for example, in terms of major processing stages of interest, level of detail (i.e., national, state, local facility) on a stage by stage basis, or by the specification of geographic boundaries. The segment thus specified is designated a subsector, with source or sink nodes created to represent the interconnection points with other subsectors. In this case, a request to display the flow of Minnesota durum grain from collection to preprocessing initiated the element selection process by placing a single source node, WMN6 (grain at a country elevator in Minnesota), on a consideration list. Each node is processed, its outgoing links are examined and any destination nodes, not yet appearing on the list, are added. Processing continues until the data are exhausted or boundaries are encountered. The user boundaries create sink nodes from which outgoing links are not considered; in the case nodes representing grain in New York and Louisiana mills (WNY23 and WLA23 respectively) are "bound," both by the fact that they are out-of-state with respect to Minnesota and by their location on the preprocessing level. Node WMN23 (grain at the mill in Minnesota) is restricted only by its association, by attribute, as a preprocessing node. If the user focus were shifted to include processing stages in Minnesota, the graph would have been expanded automatically to include the milling elements emanating from node WMN23. Nodes WNY23 and WLA23 would remain as sink nodes. Additional source nodes are included by examining incoming links, and the identification of the element set, (links and nodes), is complete.

Loop detection and the organization of the element set into the sequence of levels that define the graph are accomplished in a single simple step. Beginning with the source nodes, outgoing links are traversed creating requests for level assignments. A node is not "assigned" until the number of requests received corresponds to the number of incoming links; outgoing links are not traversed until their origin node has been assigned to a level. In this way the left to right progression is enforced; any request for a previously assigned node constitutes a cycle. The primary benefit of this approach for large networks accrues from the fact that no form of incidence matrices or other complete representation of the graph are required; processing is one level at a time. Structural symmetry is maintained despite frequent changes. For example, the disaggregation of the "vessel, rail" link into a sequence of operations for the shipment of grain to the New York mill effectively "stretches" the graph, shifting the preprocessed level to the right to incorporate additional elements.

5. CONCLUSIONS

Some of the advantages of network modeling have been explored in terms of the distinction between the iconic and symbolic functions of the network representation. Derived from the nature of the formulation of many complex problems, certain problem attributes, that lead to the requirements of cooperative modeling and "mixed network" support, are suggested. The experience gained in the development of ALINET, from inception, to design and implementation and to its application to two examples of wheat processing, has formed the basis for the mixed network modeling approach presented. On the one hand, a methodology has been developed that utilizes qualitative as well as quantitative information, on the other, a computer simulation model has been implemented that permits different parts of the system to be modeled and analyzed at different levels of detail. The direct interaction with potential users and the cooperative effort with the Pillsbury Company provided the joint modeling effort necessary for network definition. The interest expressed and the advice and constructive criticism received from potential users in government and industry may indicate 1) that ALINET could become a useful tool in the analysis of energy conservation issues in the food processing industry, and 2) that the overall modeling and simulation approach may be applicable to other classes of mixed network problems.

REFERENCES


