On the Market Value of Information Commodities. III. Demand Price

Abbe Mowshowitz*

Rotterdam School of Management, Erasmus University-Rotterdam, 300 DR Rotterdam, The Netherlands

The derived-demand price of an information commodity depends on the commodity's cost impact on a user's production process. We model an arbitrary production process as a collection of interrelated tasks which can be represented in the form of a production digraph. The nodes of the digraph represent the subtasks in the production process. A directed edge from node a to node b signifies that the subtask corresponding to node b, in order to perform its function, requires input from the subtask corresponding to node a. Appropriate weights, representing costs, are assigned to nodes and arcs. The production digraph allows for investigating cost reduction possibilities. To reduce production costs we have to lower subtask-costs, or lower the total by restructuring. The trick is to determine which subtasks to choose as reduction candidates, or to discover how to restructure the production process so as to reduce the costs. The digraph model is used, in particular, to estimate the maximum price a user should be willing to pay for an information commodity. If a producer could, by introducing some information commodity, lower production costs by x dollars, the user may be willing to pay as much as x dollars for the commodity.

Introduction

This is the last in a series of three articles on the market value of information commodities. In the two earlier articles (Mowshowitz, 1992a, b), we addressed the problem of defining information and assessing its economic value. We showed, first of all, that our proper object of study is not information, per se, but information commodities. The latter is a product (e.g., a software package or a book) or a service (e.g., an online database or a consultant) that provides information, and, in addition, can be owned and has a determinate market value. With this definition, we reduced the assessment problem to the study of a particular class of commodities.

The market for information commodities, like other markets, may be analyzed in terms of supply and de-

mand. Thus, we distinguish the economic perspective of the producer from that of the user. A previous article (Mowshowitz, 1992b) was concerned with the producer's perspective, while this one focuses on the user's perspective. To simplify our task, we restrict the analysis to information commodities employed in the making of other products or services.

For the producer of an information commodity, the critical factor in assessing value is the cost of production. Therefore, we modeled information commodities to allow for a systematic analysis of cost. The model shows how value is added to an information commodity on its way to market.

For the user, the critical factor in assessing value is the impact of the information commodity on the process employed to make the products or services being offered for sale.

We are concerned with the information commodity as an element of production. Moreover, we assume that the potential buyer of an information commodity intends to use it to make a product or service that is to be offered for sale. Thus, assessing the maximum price a user should pay for an information commodity can be reduced to determining how much the commodity in question will reduce production costs, or, more generally, increase profits from the sale of the product or service.¹ To simplify our task, we restrict the analysis of

^{*}Author's permanent affiliation: Department of Computer Science, The City College of the City University of New York.

Received March 7, 1991; revised May 20, 1991; accepted May 29, 1991.

¹Here, production refers to the activities of a firm or subdivision thereof. We are not looking at the production of an industry or a sector of the economy. In other words, we are studying production at the micro rather than macro level. Thus, our analysis is quite different from the work of Braunstein et al. (1980), Hayes and Erickson (1982), and Braunstein (1985, 1987). Whereas studies of production at the macro level are properly concerned with production functions, investigation at the micro level calls for analysis of the structure of production processes. For further discussion of production function models, see Hayes and Borko (1982).

The distinction between macro and micro level analysis holds in particular for cost-benefit studies. For example, Mason and Sassone (1978) present a cost-benefit model for information services. Their model quantifies costs and benefits in terms of properties of information services in general. Our model examines the effect of information commodities (such as the information services studied by Mason and Sassone) on the specific production processes in which those commodities might be used.

^{© 1992} by John Wiley & Sons, Inc.

demand pricing to the effects of an information commodity on the cost of production.

Since price determination depends on the role of the information commodity in a production process, we need to model such a process in a way that shows the effects of an information commodity on the cost of production. We model an arbitrary production process as a collection of inter-related tasks which can be represented in the form of a weighted digraph, called here a "production digraph."

Production Digraphs

The concept of a production digraph was introduced informally in Mowshowitz (1992b). Now we give a precise definition. P is a production digraph if P = P(V, E, a, z, c, w) where

- (1) V is a set of vertices or nodes;
- (2) E is a set of arcs or directed edges joining distinct pairs of nodes;
- (3) a is a unique node (the source) of indegree zero;
- (4) z is a unique node (the *sink*) of outdegree zero;
- (5) c is a function mapping V to the non-negative reals;
- (6) w is a function mapping E to the non-negative reals;

and P satisfies conditions (1) and (2) shown below.²

The nodes of the production digraph represent the subtasks in the production process. A directed edge from node a to node b signifies that the subtask corresponding to node b, in order to perform its function, requires input from the subtask corresponding to node a. Since we want to model a production process as an isolated system, we identify initiating and terminating subtasks. These subtasks are represented in the digraph by the source and sink, respectively. The function c assigns a weight to each node. This weight is the cost of the processing performed by the subtask on its inputs. Finally, the function w assigns a weight to each directed edge in the digraph. For an edge ab, directed from a to b, w(ab) is that portion of the cost of the output of a that is allocated to b.

Defining I(v) as the *input* of v (i.e., the sum of the weights of the edges directed to node v), and O(v) as the *output* of v (i.e., the sum of the weights of the edges directed from node v), we can state condition (1)

$$O(v) = c(v) + I(v) \text{ for all nodes } v \text{ in } P.$$
(1)

This condition may be interpreted to mean that the output of a node equals the input plus the value added by (or the processing cost of) the subtask corresponding to the node. Note that this condition implies O(a) = c(a), since a has no incoming edges; and I(z) = -c(z), since z has no outgoing edges.

P is acyclic

This condition is a simplifying assumption.

A basic property of production digraphs is given in the following proposition.

Proposition 1

 $I(z) = \sum c(i)$, where the sum is over *i* unequal *z*. **Proof.** For any edge (i, z), we have the relation

$$w(i,z) = c(i) + I(i) - \sum_{j \neq z} w(i,j)$$

Summing over *i* gives

(*)
$$\sum_{i} w(i,z) = \sum_{i} c(i) + \sum_{i} I(i) - \sum_{i} \sum_{j \neq z} w(i,j)$$

The last term may be expressed in the form

$$\left\lfloor \sum_{i} w(i, j) - w(i, z) \right\rfloor = \sum_{i} I(i) + c(z)$$

Substituting this in (*) yields the desired result.

According to Proposition 1, the entire cost of production, i.e., the sum of the costs of the subtasks in the production process, is given by the input to the sink in the production digraph. Note that the edges incident to z form a cutset that divides P into two subdigraphs, one containing z alone, the other containing all the remaining nodes. Using the proposition, we can show the following.

Proposition 2

Let T be a cutset of edges separating production digraph P into disjoint subdigraphs Q and R containing source a and sink z, respectively. Then

$$\sum_{i\in Q} c(i) = \sum_{e\in T} w(e)$$

Proof. Consider the digraph Q' formed from Q by adding a new node z' and terminating all the edges in T on z'. If we let

$$c(z') = -\sum_{e \in T} w(e)$$

then Q' is a production digraph. Now, the conclusion of Proposition 2 is an immediate consequence of Proposition 1, since

$$\sum_{e\in T} w(e) = \sum_{i} w(i, z').$$

Proposition 2 allows for an incremental approach to cost determination. First, the production digraph can be built in successive stages. Suppose we have a production digraph Q that represents part of the production process of a given product or service. For

²For definitions of graph theoretic terms, see Harary (1969).

concreteness, consider the production of an information commodity such as the word processing package discussed in Mowshowitz (1992b). Suppose that Q captures the respective contributions of kernel, storage, and processing; and that R is another production digraph with nodes corresponding to elements of distribution and presentation. The two production digraphs can be combined to form a new production digraph Pby redirecting the edges incident to the sink in Q to the nodes incident from the source in R.

Secondly, given production digraphs Q and R, we can replace (expand) a node v of Q by R to obtain a new production digraph P. This is done by the following procedure:

- (1) Remove the node v of Q and redirect all the edges incident on v to the source of R.
- (2) Redirect the edges incident from v so they are incident from the sink of R.
- (3) Adjust the costs and weights of the nodes and edges, respectively, of R to be consistent with condition (1).

There are two major steps in our approach to determining the cost of making a product or service: (a) Construction of the production digraph. (b) Determination of the costs of the subtasks (i.e., node costs), and establishment of the cost relations between subtasks (i.e., edge weightings). After presenting an example, we will examine each of these steps in turn. Then, we will discuss the use of the production digraph as a means of determining how to reduce production costs, and show how to analyze the impact of an information commodity on the cost of a production process.³

Example of a Production Digraph

To illustrate the use of production digraphs, we will exhibit one representing the creation of a research report. This particular task may be viewed as a project or development task, since the final product is a unique text—the processes of reproducing and distributing the report are not included in the example. The task of developing a research report is of interest because many different reports may be created using the same network of interacting subtasks. Thus, a production digraph representation lends itself to analysis of the cost-efficiency of the network.

Suppose the digraph shown below in Figure 1 represents the way a certain consulting group prepares a research report for a client. Nodes a and z represent the

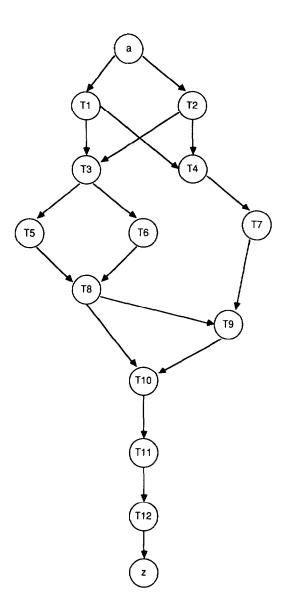


FIG. 1. Production digraph for research report.

source and sink, respectively, of the digraph; nodes T1–T12 correspond to the following 12 subtasks:

- T1 :scanning indexes;
- T2 :discussions with colleagues;
- T3 :preparing literature list;
- T4 :identifying subject experts to interview;
- T5 :obtaining books and journal articles;
- T6 :obtaining unpublished documents;
- T7 :making appointments with subject experts;
- T8 :reviewing literature;
- T9 :conducting interviews;
- T10:drafting the report;
- T11:obtaining reviews of the first draft;
- T12:preparing final draft.

Although the arcs in the digraph of Figure 1 could possibly be interpreted as precedence relations in a PERT or CPM network, we interpret them as dependence relations. This means, for example, that preparing a literature list (T3) requires input from the

³The model presented here captures the effects of information commodities on production processes. Unlike the approach taken by Yovits and Foulk (1985), and Yovits, Foulk, and Rose (1981a, b, c), the model does not purport to determine directly the value of information to a decision maker.

discussions with colleagues (T2) and from the scanning of indexes (T1). Unlike the arcs in PERT or CPM networks, an arc xy in a production digraph does not imply that subtask x must be completed before task y. So, some of the index scanning (T1) or the discussions with colleagues (T2) may occur while the preparation of the literature list is underway. The concern is with dependence that reflects cost relations, e.g., the total output from T3 includes the cost of the inputs from T1 and T2, as well as the cost incurred in processing at T3.

According to Proposition 1, the total cost of producing the research report is the weight of the arc from T12 to the sink. This, in turn, is equal to the cost, c(T12), of preparing the final draft plus the cost of the input to T12 from T11. Since T12 receives input from T11 alone, one is led to consider the possibility of lowering O(T12)by restructuring the input to T12. Perhaps, for example, the time required to obtain reviews could be reduced if T11 were split into several, parallel subtasks. By the same reasoning, O(T10) might be reduced by splitting T10 into several parallel subtasks. This variation is shown in Figure 2. The subdigraph on the nodes a, T1,...,T9 is the same as in Figure 1.

The production digraph leads one also to consider reducing costs by lowering the cost of processing at a node. For example, C(T12) could perhaps be lowered if the work of preparing the final draft could be divided into two or more parallel subtasks, each performed by lower cost personnel. Similar questions may be raised about all the nodes in the production digraph.

By dividing the cost of subtask output into processing and input components, the production digraph calls

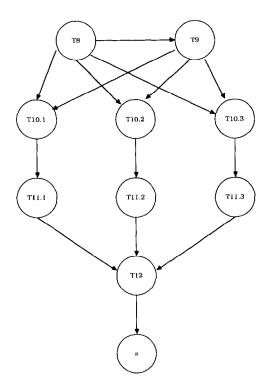


FIG. 2. Variant of research report digraph.

attention to the possibility of reducing overall costs by lowering input and/or processing costs. Moreover, by accumulating costs from source to sink, the digraph representation provides a systematic way of investigating these cost-reduction possibilities. We will show this more formally in the section on cost reduction.

Our main objective in introducing the digraph model is to provide a means for investigating the effects of an information commodity on production tasks. In particular, we want to use the production digraph as a means of determining the maximum price a user would be willing to pay for an information commodity.

Note that the information commodities we are concerned with are products or services designed for use in the production of other products or services. If a producer could, by introducing some information commodity, lower production costs by x dollars, he may be willing to pay as much as x dollars for the commodity. Although not a least upper bound, this quantity x could provide a useful approximation of the demand price of an information commodity.

Thus far, we have shown how the production digraph in Figure 1 could be used to examine, systematically, ways of lowering the costs of production. Now, we will show how the digraph could be used to investigate the ways in which an information commodity might be introduced so as to lower production costs.

Consider the question of introducing word processing software. To make for a sharp comparison, suppose that in our research report example, documents are typed on a conventional typewriter. To analyze the cost impact of word processing in our example, we must first determine which subtasks might be altered by the use of word processing. Note that a variety of document preparation activities could be affected by the introduction of word processing.

Subtasks T12, T10, T8, and T3 all involve a significant amount of document preparation, and thus are likely to be strongly affected by a shift to word processing; but other subtasks might be affected as well. For example, some of the reviews in subtask T11 could be received as work processor files through an electronic mail system. A word processor would make it possible to edit and/or combine various reviews and to prepare hard copy for circulation within the group. Parts of subtask T2 might also be affected by word processing: if the discussions were written up on a word processor, they could more easily be adapted for use in the report.

The production digraph provides a framework for analyzing the cost effects of information commodities such as word processing software. Clearly, it does not resolve cost accounting issues, such as how to deal with fixed as opposed to variable or direct vs. indirect costs. We assume that some cost accounting method is in use in a given production environment, so that a production manager could analyze the specific cost impact of an information commodity on a subtask. Let us try to analyze the effect of a shift from typewriter to word processor in subtask T12 (preparation of final draft). For the moment, suppose that T12 is the only subtask involved. Four cost factors have to be examined in this case:

- (1) purchase of word processing software:
- (2) purchase of a computer;
- (3) training of typist to use the new equipment and software;
- (4) word processor operations.

The first three are fixed costs which could be amortized over appropriate time periods and allocated as variable costs; the third is strictly a variable cost. The sum of these costs would then have to be compared with the current unit cost of preparing the final draft on a typewriter.

If T12 were the only subtask potentially affected by the replacement of the typewriter by a word processor, the production digraph would not be terribly useful. However, as noted earlier, several subtasks are likely to be affected. The digraph representation facilitates identification of the subtasks, as we have seen; but more importantly it provides a means of analyzing possible changes in the interrelations between the affected subtasks. That is to say, the digraph model is a vehicle for analyzing the impact of structural changes in the production process. What is more, the same production digraph may be used to investigate the impact on production of many different types of information commodities.

Constructing Production Digraphs

There are four major steps in the construction of a digraph model of a production process:

- (1) analyzing the process into its constituent subtasks;
- (2) determining the interdependencies among the subtasks identified in (1);
- (3) determining the unit processing costs incurred by the subtasks;
- (4) allocation of each subtask's output cost among the subtasks dependent upon it.

Steps (1) and (2) yield the digraph structure; steps (3) and (4) assign weights (or costs) to the nodes and edges of the digraph.

There is no algorithm for carrying out the steps outlined above. If the production task consists of a number of clearly recognizable subtasks, Steps 1 and 2 may be quite straightforward. On the other hand, if task subdivisions are not clearly defined, considerable empirical effort may be required to generate the digraph structure. In this case, one might use structured problem solving as a framework for the empirical investigation.

One must keep in mind that the production digraph is intended to clarify cost relations. This means, in particular, that the nodes of the graph may not correspond to organizational divisions or to job definitions. Thus, the empirical investigation is a crucial part of the digraph construction process.

Once the digraph structure has been determined, the node and arc weights must be ascertained (Steps 3 and 4). Determination of node weights (Step 3) involves the following: (1) Determining precisely what categories of costs are to be considered; and (2) Obtaining the relevant cost data. These requirements are nontrivial. As discussed earlier, it may be desirable in certain cases to take account of some fixed as well as variable costs. Since accounting principles and practices vary, careful attention must be given to proper conceptualization of cost factors. Moreover, the data needed to assign values to the various costs identified for a subtask may not be readily available.

Allocating input plus processing cost, I(x) + c(x), of a node x to the nodes incident from it calls for analysis of the outputs of x. Analysis here means estimating the portion of a node's output that goes to each of the nodes incident from it. This must be done on a case by case basis.⁴

Cost Reduction

In our discussion of the production digraph corresponding to the development of a research report, we indicated how the digraph might be used to investigate cost reduction possibilities. Now we will show this more formally.

Suppose P = P(V, E, a, z, c, w) is a production digraph. Recall that the total cost of production is I(z), the sum of the weights on the arcs incident to the sink z. Moreover, from Proposition 1 we know that I(z)equals the sum over all the nodes *i* (other than z) of c(i). So, to reduce production costs we have to lower some of the c(i), or to lower the total by restructuring. The trick is to determine which subtasks to choose as reduction candidates, or to discover how to restructure the production process so as to reduce the costs. Proposition 2 suggests a systematic procedure.

Let T = (Q, R) be a cutset of edges in the digraph Pwhose removal partitions P into subdigraphs Q and Rcontaining a and z, respectively; and let w(T) denote the sum of the weights of the edges in T. According to Proposition 2, w(T) is equal to the sum of the costs of all the nodes in the subdigraph Q. Now, let N(T, Q) denote the set of nodes of Q that are incident to arcs in T. Finally, for S and V, let $N_{in}(S)$ be the set of all nodes of P that are adjacent to nodes in S; and let $A_{in}(S)$ be the set of arcs of P that are incident to nodes of S.

⁴See Bellin (1991) for a detailed discussion of node and arc weighting in the context of software development.

First, construct a sequence of cutsets as follows

$$T_{1} = A_{in}(\{z\})$$

For $k > 1$, $T_{k} = A_{in}[N(T_{k-1}, Q_{k-1})]$ (*)

Clearly, this gives a finite sequence whose last element is the set of arcs incident from the source. Note also that this sequence of cutsets determines a corresponding sequence of subdigraphs Q_k beginning with $P - \{z\}$ and ending with the trivial subdigraph consisting of the source alone. This sequence of subdigraphs in reverse order, represents a sequence of "subproduction" processes that terminates in the given production process.

Second, construct a sequence of subsets V_k of nodes corresponding to the cutsets of (*): for k > 0, $V_k = N(T_k, Q_k)$ The first element, V_1 , consists of the nodes incident to the sink; the last element (say V_n) is the set consisting of the source alone. These subsets of nodes represent clusters of subtasks at successive levels in the production process. Each cluster is a candidate for cost reduction.

Third, for each V_k , from k = 1 to *n*, determine the balance of the set $\{O(x)|x \text{ is in } V_k\}$ of output costs. This analysis could be done by computing the deviation of each pair-difference, O(x) - O(y), from the mean of all the pair-differences. A "high" deviation suggests considering one of the following actions:

- merging one subtask with another to achieve economy of scope;
- (2) subdividing one subtask to achieve the cost savings of specialization and division of labor.

If any rebalancing is done, outputs must be reallocated, and the production digraph modified to reflect the changes.

Production Effects of Information Commodities

Finally, we return to our point of departure in this study, namely, the impact of information commodities on production. We will show how the production digraph can be used to analyze this impact—especially how it can be used to estimate the demand price of an information commodity.

An information commodity such as a computer, a piece of software, or a database may affect production in several ways.

Type 1. Processing within a subtask may be altered without establishing new connections or eliminating existing connections to other subtasks;

Type 2. Connections between subtasks may be modified;

Type 3. A subtask may be divided into two independent subtasks;

Type 4. Two subtasks may be combined to form one subtask.

We will discuss these effects of information commodities in relation to the production digraph model. The simple production process represented by the digraph in Figure 3 will be used to illustrate the effects.

Type 1 changes in a node x are reflected only in modifications of the cost, c(x), of node x. This implies no change in either the set of arcs incident to x or the set incident from x. However, since the cost is altered, the output O(x) must be redistributed among the arcs incident from x to other nodes, and all the other node and edge weights must be readjusted in accordance with the definition of a production digraph.

Suppose the cost of the subtask corresponding to node b is reduced from 15 to 10. This reduction could result, for example, from replacing some software package by a more efficient one. Now, the cost of the output from node b is reduced from 25 to 20. Let us say that the weight of arc bd is lowered to 16 and the weight of arc be drops to 4. Then the weight of arc dz becomes 26 and ez becomes 39. The resulting weighted digraph satisfies the definition of a production digraph, and its total cost is 65, 5 less than the original (Fig. 3).

Type 2 changes involve the elimination or addition of arcs in the production digraph. If an arc incident from a node x is eliminated, then the entire output O(x)must be reallocated to the remaining arcs issuing from node x, and the weighting function w must be adjusted. If a new arc is added from node x to node y, again the output O(x) of x must be reallocated, and the function w adjusted.

Consider the removal of arc be from the digraph of Figure 3. This means that the subtask corresponding to node e no longer receives (nor needs) output from b. The output cost of node b is thus reduced by 5, which implies that the cost of input plus the cost of processing at b is reduced by 5. This could result from simply cutting the cost of input to b from 10 to 5, in which case the cost of the source a would be lowered to 25. The only remaining change to be made in the digraph is on the output side of node b, i.e., the weight of arc ez must be reduced from 40 to 35.

Type 3 changes are more complicated. These involve the elimination of an existing node together with all the arcs incident to it and from it, and the introduction of two new nodes together with arcs linking them to other nodes in the digraph. This is a major change requiring

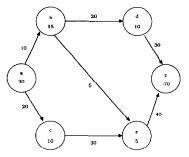


FIG. 3. A Simple production process.

an updating of the set V of nodes and the set E of arcs as well as modification to both of the weighting functions c and w.

Suppose the subtask corresponding to node b in the digraph is split so as to disaggregate its outputs. The resulting production digraph is shown in Figure 4.

The total cost of the production process represented in Figure 4 is the same as that of Figure 3. However, if sum of the costs of nodes x and y (Fig. 4) were less than that of node b (Fig. 3), then the total cost of the modified process would also be lower.

Type 4 changes involve similar adjustments. When two nodes x and y are merged into x', the arcs incident to either x or y will be incident to the new node x'; however, not all of the arcs incident from either x or y will remain in the new digraph. In particular, arcs xz and yz would be merged, and arcs xy and yx would disappear. The merger of nodes x and y would presumably result in a cost c(x') < c(x) + c(y). The output O(x')must be reallocated to the arcs issuing from x' and the weighting functions c and w as well as the node set V and arc set E must be adjusted so as to yield a new production digraph.

An estimated upper bound for the demand price of an information commodity may be obtained by determining the change induced by the commodity in the total cost of production. This change is the difference, I(z') - I(z), where z and z' are the sinks of the old and new production digraphs, respectively.

Further research is needed to classify information commodities according to the types of changes they induce in production digraphs.

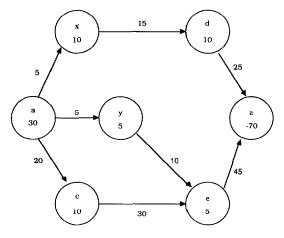


FIG. 4. Subtask splitting.

References

- Bellin, D. (1991). Information commodities and the production process: how the introduction of CASE tools affects the software development life cycle. unpublished PhD Dissertation, The City University of New York.
- Braunstein, Y. M. (1985). Information as a factor of production: substitutability and productivity. *The Information Society*, *3*, 261–273.
- Braunstein, Y. M. (1987). Information inputs and outputs in firms and quasi-firms. In M. Jussawala (Ed.), The Cost of Thinking: Information Economics of Ten Pacific Countries (pp. 71-80). Norwood, NJ: Ablex.
- Braunstein, Y. M., Baumol, W. J., & Mansfield, E. (1980). The economics of R&D. TIMS Studies in the Management Sciences, 15, 19-32.
- Harary, F. (1969). Graph theory. Reading, MA: Addison-Wesley.
- Hayes, R. M., & Borko, H. (1983). Mathematical models of information system use. *Information Processing & Management*, 19, 173-186.
- Hayes, R. M., & Erickson, T. (1982). Added value as a function of purchases of information services. *The Information Society*, 1, 307-338.
- Mason, R. M., & Sassone, P. G. (1978). A lower bound cost benefit model for information services. *Information Processing & Man*agement, 14, 71-83.
- Mowshowitz, A. (1992a). On the market value of information commodities: I. The nature of information and information commodities. Journal of the American Society for Information Science, 43, 225-232.
- Mowshowitz, A. (1992b). On the market value of information commodities: II. Supply price. Journal of the American Society for Information Sciences, 43, 233–241.
- Yovits, M. C., & Foulk, C. R. (1985). Experiments and analysis of information use and value in a decision-making context. *Journal* of the American Society for Information Processing, 36, 63-81.
- Yovits, M. C., Foulk, C. R., & Rose, L. L. (1981a). Information flow and analysis: theory, simulation, and experiments. I. Basic theoretical and conceptual development. *Journal of the American Soci*ety for Information Science, 31, 187–202.
- Yovits, M. C., Foulk, C. R., & Rose, L. L. (1981b). Information flow and analysis: theory, simulation, and experiments. II. Simulation, examples, and results. *Journal of the American Society for Information Science*, 31, 203-210.
- Yovits, M. C., Foulk, C. R., & Rose, L. L. (1981c). Information flow and analysis: theory, simulation, and experiments. III. Preliminary Experiments and Analysis. *Journal of the American Society for Information Science*, 31, 243–248.