On the Market Value of Information Commodities. II. Supply Price

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

Information commodities make use of storage, processing, and communication capabilities in varying degrees to acquire market value. We have identified five major value-adding dimensions of information commodities: (1) kernel; (2) storage; (3) processing; (4) distribution; and (5) presentation. The kernel is the (organized) information provided by an information commodity. The storage dimension of an information commodity encompasses both the medium used to store information and the method used to gain access to it. Computer-based, as opposed to traditional, commodities can process and reconfigure the information stored within them. Processing adds value by making it possible to transform the kernel. Distribution adds value by making the kernel of an information commodity accessible to the user. Presentation, the final link in the value-adding chain, involves the display of information in forms suitable to the user. The five value-added components provide the basis for an analysis of the cost of producing an information commodity. Major steps in the analysis include: (1) Identification of subtasks in the production process. (2) Determining interdependencies between subtasks. (3) Construction of digraph representing interacting subtasks. (4) Determining the costs of the subtasks and their respective allocations. (5) Using the digraph representation to answer questions about production costs.

Introduction

This is the second in a series of three articles on the market value of information commodities. In the previous article (Mowshowitz, 1992a), we introduced an abstract definition of information commodity, based on our definitions of information and commodity, respectively. Information is defined as ability rather than as stuff, i.e., information is the ability of a goal seeking system to decide or control. Anything that can be

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

owned and valued in the marketplace can be a commodity. So, an information commodity is something that furnishes the ability to decide or control, and can be owned and valued in the marketplace. In this article, we will interpret this abstract definition by analyzing the components of an information commodity in terms of their respective contributions to its market value.

For most of human history, the principal carrier of information has been the human being. Before the advent of writing, the record of human experience rested entirely with living human beings—the survival of that record depended on human memory and oral transmission from person to person. The social uses of information were confined in space and time. All the knowledge and skill of the community was carried by the individual members; knowledge was passed from generation to generation in the form of oral tradition. With writing came the possibility of embedding information in artifacts, and thus began an evolutionary process of separating information from human beings.

The ability to store information outside the human being provided the technological foundation for a collective memory transcending space and time. As Innis (1951) has argued so eloquently, media of communication made it possible to administer large territories, and to achieve cultural continuity over time. Media can be differentiated, according to Innis, on the basis of characteristics such as portability and durability. Among the earliest media, stone, being durable but not very portable, lent itself to applications emphasizing control over time, e.g., religion. The more portable clay tablet was put to use in trade and political administration.

Stone and clay gave way to a succession of media including papyrus, parchment, and paper. Each of these media has left a powerful and distinctive mark on society and culture. But that influence derives entirely from their rather limited ability to *store* information. Modern information technology represents an order of magnitude increase in the ability to embed information in artifacts. Computers, for example, can transform in-

formation as well as store it. Written documents store information in a fixed format and, unlike computers, have no ability to execute instructions. All the processing to extract information from a document must be done by the human user. Transmitting the information in a document usually means the transfer from one person to another of the document itself or a copy thereof.

Information artifacts can thus be classified in terms of three fundamental characteristics: (1) storage capabilities, (2) processing capabilities, and (3) communication capabilities. An important attribute of the second characteristic is the degree to which human intervention is required in processing information. Storage, processing, and communication are highly complex characteristics and figure prominently in the production and marketing of information commodities. Here we merely want to point out the key elements involved in the separation of information from the human being.

Modern information technology has ushered in a new era of information handling in artifacts, but it is not without antecedents. The history of calculating devices reveals the evolution of information processing capability, or, the development of artifacts capable of responding to procedural statements. Early mechanical adding machines, for example, had very limited storage capacity, but were able to perform arithmetical operations automatically or semiautomatically. Typically, such devices could accumulate a number consisting of a few decimal digits by adding a second number entered by a human operator to the one currently stored in the machine.

Mechanical storage devices also have a considerable history. The Jacquard loom, perfected at the end of the 18th century, contained a digital storage device based on punched cards. The presence or absence of a hole in a given position on a card determined whether or not a particular warp thread would be drawn. This made it possible to store an entire pattern on a sequence of punched cards and to realize the pattern in a weave.

Until the invention of the telegraph, the history of advances in communication is closely linked to transportation. Modern telecommunications technology allows us to move information from place to place without having to move ourselves or to mechanically transport a physical object such as a book. Moreover, with electronic or optical transmission, information can be exchanged by artifacts without any human mediation.

Composition of Information Commodities

Information commodities make use of storage, processing and communication capabilities in a variety of ways to add market value to information.[1] We have

identified five major value-adding dimensions of information commodities. These are:

(1) kernel
(2) storage
(3) processing
(4) distribution
(5) presentation

An information commodity may thus be viewed as a point in a five dimensional space. That is to say, these dimensions will be reflected in varying degrees in different information commodities. One information commodity may derive its market value from the kernel, while another may be valuable because of innovations in, say, storage or presentation.

These dimensions may be interpreted as the means for providing access to information.[2] In this sense, the dimensions correspond to a series of Chinese boxes. The innermost, kernel, is contained within storage, which is contained within processing, which is, in turn, contained within distribution, which is, finally, contained within presentation. These relations, illustrated in Figure 1, signify logical rather than physical containment.

Kernel

The *kernel* of an information commodity is the information, i.e., the ability to decide or control, furnished by the commodity. This includes the organization and structure of the information—the kernel is an organized system of declarative and procedural statements.[3] Indeed, the distinction between information and its structure is somewhat artificial, since information is always structured to some degree.

Consider an online, legal database such as Lexis or Westlaw. This is an information commodity whose kernel consists of a collection of texts reporting judicial decisions, together with an indexing scheme that allows for orderly updating and systematic access to the reports. The contribution of the indexing scheme to the market value of these legal databases is underscored by recent litigation involving the vendors of Lexis and Westlaw, Mead Data, and West Publishing, respectively. West Publishing filed suit against Mead Data on the grounds that the latter was using an indexing scheme developed by the former. The dispute was settled out of court by Mead's agreement to pay West an undisclosed amount for the use of the scheme.

[1]For further discussion of the value of information, see Griffiths (1982), Duncan (1985), and Repo (1986,1989).

[2]Generally speaking, this view of an information commodity has much in common with Taylor's view of an information system as "a series of formal processes by which the potential usefulness of specific input messages being processed is enhanced" (Taylor, 1982, p. 341).

[3]The interpretation of information (i.e., the ability to decide or control) as a system of interacting subsystems—belief and command—is discussed in Mowshowitz (1992a).
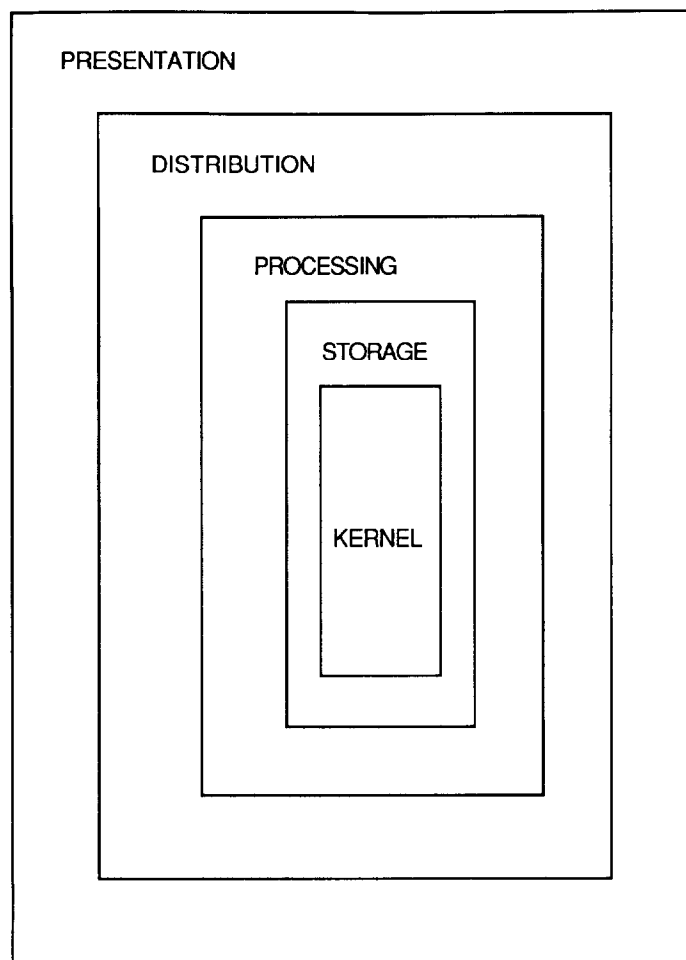
FIG. 1.   Chinese box model of information commodity.

A book is a traditional kind of information commodity. The kernel of a book consists of the information represented by the text, together with the structure implicit in the organization of the text. Aids to the reader, such as tables of contents, summaries, indexes, glossaries, etc. correspond to the retrieval algorithms associated with online databases. But unlike retrieval algorithms, these aids are "executed" by the user, not the artifact itself. That is to say, books have no internal processing capability.

A software package differs from the previous two examples because its kernel consists primarily of procedural statements. Declarative statements are important in a software package, but the procedural component is its dominant feature. Typically the kernel is a collection of computer programs, i.e., a collection of ordered sets of programming language instructions, that can be executed by a computer.

The kernel of an information commodity is but one determinant of market value. Its specific contribution to the price of such a commodity derives partly from its cost of development, and partly from its perceived economic value to potential users. The relative weights attached to these factors may vary widely. At one

extreme is a 300-page book that sells for $20, little more than the unit costs of printing, binding, and distribution. The development costs associated with, say, research, writing, and editing, contribute practically nothing in this case. At the other extreme is a specialized market report of the same length that sells for $20,000. Here, development cost is a major factor in the price asked by the producer.

A complete characterization of our notion of kernel is not possible, since it depends on the organization of knowledge itself.[4] We must content ourselves with a sketch of the idea. Thus, we will illustrate the ways in which information may be structured and organized, thereby enhancing its usefulness to potential buyers.

Indexing is an important feature of the kernel of an information commodity. This holds for research reports and textbooks as well as catalogues. The index is an aid to the user. It may be viewed as a metalanguage in terms of which questions about the body of information (or object language) can be formulated. Moreover, an

_____

[4]Kochen (1984) reviews recent developments in the conceptualization of information.

index or indexing scheme defines a mapping of queries to the body of information. So, an appropriately formulated question can be expected to yield answers drawn from the body of information.

The implementation of the mapping is especially important for the marketing of information commodities. Here is where the computer plays a vital role. The processing power of the computer makes it possible to implement such mappings in "real time." Computer-based searching algorithms, for example, are essential components of online databases. We will return to this issue when we discuss the processing part of the model.

Arrangement is another important feature of the organization of information. This refers to the ways in which the parts of a body of information are linked to each other. The chapters of a book follow each other in sequence. Reference may be made in one part to another part of the book or to another body of information by means of a citation. Subprograms of a structured computer program are also linked to each other by means of a reference scheme. Each subprogram constitutes an identifiable block of instructions which may be invoked by another (or itself, in some cases).

Different applications call for different types of linkage. Textbooks for college freshman typically make heavy use of internal references, but do not usually cite many external sources. On the other hand, a scientific research paper would normally have a number of references to source material, but relatively few internal references.

The structure of an individual item is another important feature of the organization of information. For a file (e.g., products, customers, subscribers, employees, etc.), the structure chosen to represent an item has implications for data entry, storage, retrieval, and reporting. For some applications a tree structure is appropriate; for others, a relational structure may be preferred. The idea of a *data structure* is also applicable to more traditional forms of information, such as reports. Sentences, charts, tables, diagrams, etc. are data structures commonly used in reports. Accuracy of data items is another important factor in assessing the kernel's contribution to market value.[5]

Storage

The *storage* dimension of an information commodity encompasses both the medium used to store information and the method used to gain access to the medium. Historically, storage has been the most important aspect of information artifacts (e.g., books and periodicals). It warrants attention as a dimension of modern information commodities because of the role it

plays in making such commodities available to users, as well as its basic function in the operation of automatic computing devices.

The information associated with a legal database, for example, is stored in the secondary memory of a computer system. In this case, storage is integrally connected with the processing system, although it is conceivable that derivative products might be offered on CD/ROMs or other storage devices. Information is stored in a book as strings of symbols inscribed on paper. The information (represented as statements) in a software package is usually stored in machine sensible form on a medium that can be handled by a computing device. Examples of media include the high speed Random Access Memory of a computer, magnetic disks, optical disks, and magnetic tape. Software for personal computers, text files, and data files are now routinely exchanged by means of floppy disks. Perhaps this medium's importance will diminish as network storage facilities, in conjunction with electronic information transfer, become more commonplace.

The main attributes of storage are: capacity, speed of access, reusability, reliability, portability, and longevity. Storage shares with other areas of information technology a pattern of rapid advances. Of particular interest here is the progress in floppy, magnetic disk drives. More and more information is being packed on smaller and smaller surfaces, and accessed at increasingly higher speeds. In addition, the evolution of high-capacity optical storage systems are of major importance in the creation of information commodities. Both the floppy and the CD/ROM are critical value-adding components of an increasingly large class of products.

Processing

Traditional information commodities such as books are essentially passive: they do not have the ability to reorganize or re-present the information they contain. The human user (reader, in the case of a book) has to do whatever *processing* there is to be done. A computer system, on the other hand, replete with CPU, operating system, and application programs can process and reconfigure the information stored within it.

Most information commodities fall somewhere in between books and computer systems on the processing dimension. An online database, for example, makes use of a computer system but the system is not a part of the commodity itself. In this case, the computer makes it possible to execute the storage and retrieval algorithms designed to update the files and to search for specific items of information. A software package consisting of computer programs stored on floppy disks does not itself have the ability to execute instructions, but it is designed to be used in a computerized processing environment. Both the online database and the software package presuppose the availability of computing

[5]Bradford and Kelejian (1978) estimate the value added to an information system by improvements in the quality of information used to construct crop forecasts.

power to actualize the procedural statements of their respective kernels.

The value added by processing depends on what elements are included in the information commodity, and on whether those elements contribute directly or indirectly to that processing power. Direct contributions include both hardware (e.g., specialized devices) and software (e.g., program code of the commodity). Indirect contributions are linked to the processing environment (e.g., operating system, translators, and other utilities) in which the commodity is designed to be used. A software package, for example, may be designed for use with a particular operating system such as UNIX, or it may be intended for use on a certain type of computer. Thus, the peculiar advantages or disadvantages of the processing environment are conferred on the commodity.

Distribution

Like processing, the value added by means of the *distribution* of information in a commodity may be direct or indirect. Data communication systems are to distribution what computer systems are to processing. The specialized communication hardware and software of the commodity are direct, value-adding elements of distribution. Value is added indirectly when the use of an information commodity presupposes a particular distribution environment.

For example, the kernel together with the storage and processing power of the computer are not enough to make an online, legal database available to practicing attorneys in different parts of the country. The package is incomplete without a communications network allowing users access to the files from remote locations. Specialized hardware and software that users must purchase with the package are direct value-added elements; the network itself adds value indirectly to the commodity.

Traditional information commodities such as books are distributed by means of mechanical transport systems. Such systems are inherently slower than electronic or optical transmission of information. Distribution of software may be accomplished by physically transporting the medium on which it is stored, or by transmitting the information from one computer to another via a network or direct link.

An information commodity such as a stock market quotation service would have relatively little value to a trader if it were delivered in the form of printed lists through the postal service. Timeliness of delivery is crucial to this and many other commodities (e.g., online information systems). So, elements of distribution, although not necessarily integrated into a given commodity, may be essential for its use.

Information commodities can be classified according to their requirements for timeliness of delivery. At one extreme are products and services (such as archival storage and retrieval, and batch computing systems) that must be updated and/or accessed relatively infrequently. At the other extreme are commodities (such as online quotation, real-time control, and time-sharing systems) containing information that must be updated and/or accessed from moment to moment. Computer conferencing is an example of a service that falls somewhere in between these extremes.

Cost is a basic attribute of all the value-adding components of information commodities, but it plays a particularly important role in the distribution environment. Economies of scale in distribution can be achieved by load and resource sharing. Public switched networks, for example, provide telecommunication services to a community of users over shared lines. Since the investment in infrastructure is lower in this case than it would be for pairwise connections, lower rates can be charged. Similarly, the sharing of resources, such as software and data, among the users of a network reduces the cost per user of maintaining those resources.

Economies of scale also figure in the mass production of processors and storage media. We raise the issue here because the way economy of scale is achieved in distribution, i.e., thorough common infrastructure, allows for the development of a rich variety of information commodities. That is to say, the distribution environment, by linking actors in the information marketplace, is itself a product development platform.

Presentation

The final link in the value-adding chain between producer and consumer is *presentation*. In addition to providing a structured kernel, storage, processing, and distribution, the producer must arrange to have the information appropriately presented to the user. Items retrieved from an online database must be identified and presented in a way that enables the user to read and interpret the information as responses to queries. This usually means the display of text on the screen of a monitor or in printed form. If the user is a computing device or computer application, the information must be formatted according to a protocol or convention shared by both sender and receiver.

Presentation requirements vary with the intended application. If the user is a person, various aids (e.g., menus and command options) are usually incorporated into the package. Typically, the package is installed on a computer, and messages that describe its functions and report on its behavior are displayed on a monitor. If the user is a machine or an application program, then the package must structure its output so as to be compatible with the format adopted by the user.

In the case of a book, the obvious considerations include page size, paper quality, type face, style and size of the font, page layout, etc. But factors external to the

book may also affect its presentation. For example, if a very small type font is used to compress the text (as in the two-volume edition of the Oxford English Dictionary), many readers would require a magnifying glass, or, perhaps, special lighting.

This example suggests that two classes of factors affect presentation. Both of these classes are defined in relation to an object or device (printed paper, in the case of a book) on which the information of the commodity is displayed. *Internal* factors refer to characteristics of the object or device and the ways in which it displays information; *external* factors have to do with the environment in which the object or device is positioned for use.

Computer-based, information commodities are more complex than traditional ones, but both can be analyzed in essentially the same way. Consider the presentation of software or database packages. Typically, these packages are presented to the user on a display screen attached to a microprocessor; instructions from the user are entered at a keyboard or related device linked to the processor. Internal factors in this case include properties of the display (e.g., size, resolution, chromaticity, etc.) and keyboard (e.g., functions available, size and layout of keys, action of keys, etc.). Also included are features of the display of information on the screen. Mechanisms designed to call attention to particular items on the screen, spatial arrangement of the information displayed, division of the screen into multiple workareas, etc. are cases in point.

Factors external to the display and keyboard are also important in this example. Lighting and physical placement of the display and keyboard are important ergonomic concerns and thus affect the presentation of the information commodities.

All of the abovementioned characteristics of presentation signal opportunities for adding value to an information commodity. A software package may be superior to competing products by virtue of its exceptional screen layout; a database system may outshine its rivals because it accepts commands from the user in the form of simple key combinations. Certainly, this is not news to the designers of such products.

On the Pricing of Information Commodities

The Chinese box model—by identifying opportunities for adding value—provides a framework for determining the cost of producing an information commodity. That is to say, the value-adding dimensions may be used to analyze systematically the subtasks involved in the production of an information commodity. These subtasks may then be represented by the nodes of a weighted digraph. The directed edges in the digraph represent cost relations between the nodes. Assuming the weights and cost relations can be obtained, this digraph representation allows for determining the

overall production cost quite easily. The properties of such digraphs are investigated in Mowshowitz (1992b).

The first step in assessing the costs of producing an information commodity is identifying the subtasks in the production process. By elaborating the elements of each of the five boxes in our model, we can obtain a first approximation to these subtasks. Consider, for example, a word processing package designed for use on a microcomputer.

Kernel of Word Processing

Typically, the kernel contains both procedural and declarative elements. The procedural component includes the text processing and text formatting algorithms of the system, as well as ancillary algorithms for interaction with peripheral devices such as printers, and for communication with other programs such as spreadsheet applications. A common example of declarative information in a word processing kernel is a dictionary used as the database for a spelling checker.

Storage Component of Word Processing

Most word processing packages are stored on floppy disks for distribution. They are also commonly stored on hard disks when installed on a microcomputer equipped with a hard disk drive.

Processing Component of Word Processing

The processing component of word processing software usually consists of the computer programs that collectively implement the algorithms of the kernel. Part of the value added by the implementation derives from the range of computers and operating systems for which the package is designed. Some word processors may be embedded in larger packages that include, for example, spreadsheet and database applications. In these cases, the value added by processing should take account of the additional functionality of these related applications.

Distribution Component of Word Processing

The typical word processing package is distributed in the form of a floppy disk containing the object code and associated data of the programs defined in the processing component. Mechanical transport (rather than electronic communication) is the principal method for bringing these packages to the users.

Presentation Component of Word Processing

The functions of word processing packages are usually invoked by selecting items on a menu or by issuing specific commands. One package may be more "menu-

driven" than another, but all make use of a display screen and a keyboard. Variable features of such packages include the capability of using color and high resolution display screens, auxiliary input devices, special printers, and split screen programs. These features help to distinguish the presentation environment.

The value added to a word processing package by each of the five components described above is summarized in the following table. Moreover, the table shows the relationship between value added and development tasks. As illustrated in the table, it is useful to distinguish between direct and indirect contributions to value for each of the five components.

The items listed in the right-hand column of Table 1 are the major subtasks in the development of a word processing system. We focus here on development rather than volume production of the package, because the marginal production costs of a disk containing software are not much more than the disk itself. However, our method of cost analysis, making use of a digraph representation of a complex task, is applicable to both development and production.

Determination of the major components in the development of an information commodity (such as a word processing package) is the first step in identifying the subtasks used in constructing the digraph representation. The word processing example is intended to show the role of the value-adding framework in cost analysis. In general, cost analysis involves the following procedure:

(1) Identify major development steps corresponding to value adding dimensions.
(2) Refine the major steps to obtain the separable subtasks in the development process.
(3) Establish the interdependencies between the subtasks, i.e., determine which subtasks feed information to which other subtasks.
(4) Construct a production digraph representing the interacting subtasks obtained in (1) and (2).
(5) Determine the costs of the subtasks, and, for each one, determine the allocation of outputs to the subtasks dependent upon it.
(6) Use the digraph representation to answer questions about overall development costs, and to explore the cost-implications of changes in the development process.

If the major development tasks identified in Table I were independent of each other, the production digraph would have the form shown in Figure 2.

We will show in the sequel (Mowshowitz, 1992a) that for any production digraph, the total cost of the process represented therein is the sum of the weights on the edges coming into the node labelled "Sink." The digraph pictured in Figure 2 is not very interesting be-

TABLE I. Analysis of word processing package.

| Value Added	Major Development Subtask
Kernel	Algorithm design
Direct	
text manipulation	
control structure	
writing aids	
Indirect	
complementary packages	
Storage	Media specification
Direct	
floppy disk	
Indirect	
type of disk drive	
Processing	Program design and coding of kernel algorithms
Direct	
effectiveness and efficiency of software for kernel	
Indirect	
operating system	
Distribution	Packaging and shipping
Direct	
disk packaging	
Indirect	
shipping service	
Presentation	Design and implementation of display
Direct	
display features of package	
Indirect	
user interface	
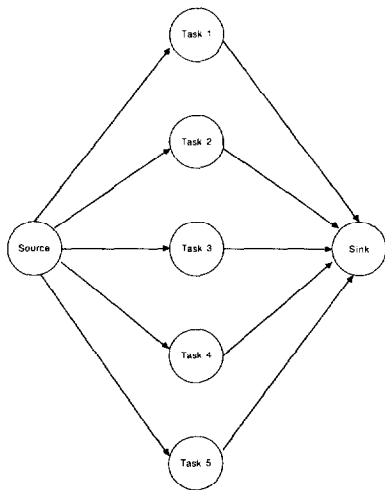
FIG. 2. Production digraph with independent subtasks.

cause the cost relations among the subtasks would be immediately evident from a verbal description. However, the relations among development subtasks are usually more complex than those depicted in Figure 2. This holds in particular for our example of the development of a word processing package. Figure 3 provides a more realistic picture, and gives more insight into the usefulness of the digraph representation.

Figure 3 shows that Task 2 (media specification), Task 3 (program design and coding of kernel algorithms), and Task 5 (design and implementation of display) are dependent on Task 1 (algorithm design); Tasks 2 and 5 are also dependent on Task 3; and Task 4 (packaging and shipping) is dependent on Tasks 2, 3, and 5. The dependence relationships are not necessarily temporal, e.g., Task 3 (program design and coding) may be in progress before Task 1 (algorithm design) is completed. In general, task b is dependent on task a, if b requires some output from a.

Although the digraph of Figure 3 is an improvement over the representation given in Figure 2, further refinement is needed to obtain a practicable model of the development of word processing software. The tasks identified in the digraph must be further differentiated. For example, both Task 1 (algorithm design) and Task 3

(program design and coding) encompass a number of distinct, production subtasks.

The next step in constructing the digraph model is the assignment of weights to the nodes and the arcs. Nodes are assigned weights reflecting the costs of processing performed by the subtasks they represent. Note that the output of a given node is the sum of the processing cost at the node and the inputs to the node. The weight assigned to an edge ab (directed from node a to node b) is the cost of the input passed from a to b. The total cost of producing the commodity is the sum of the inputs to the sink. In our example, this is the sum of the weights attached to the edges directed from Tasks 2, 3, and 5 to Task 4.

Knowledge of the total cost of production or development would enable the producer to determine a break-even price for the commodity. This provides an approximate lower bound for the supply price.

The production digraph model can also be used to assess the cost implications of introducing an information commodity into a production process. Because of the general applicability of the production digraph, the theory of these digraphs is developed in relation to the demand for information commodities. (Mowshowitz, 1992b).
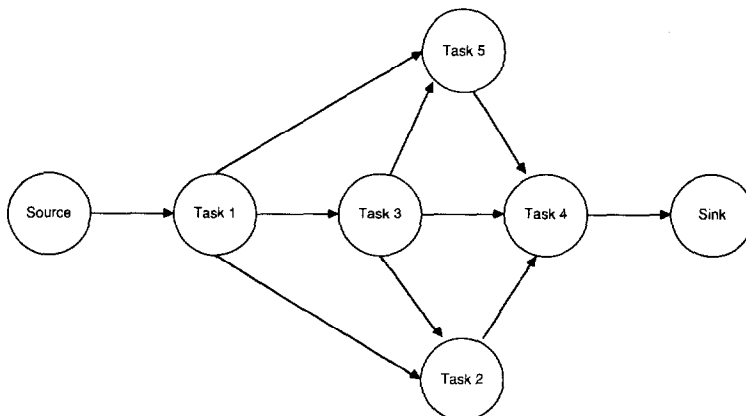
Acknowledgments

FIG. 3. Production digraph: Second approximation.

References

Bradford, D. F. & Kelejian, H. H. (1978). The value of information for crop forecasting with Bayesian speculators: theory and empirical results. *Bell Journal of Economics, 9*, 123–144.

Duncan, J.W. (1985). *The worth and value of information.* Paper presented at American Federation of Information Societies Annual Meeting.

Griffiths, J.M. (1982). The value of information and related systems, products, and services. In M. E. Williams (Ed.), *Annual review of information science and technology,* volume 17 (pp. 270–285). White Plains, NY: Knowledge Industry Publications.

Innis, H. A. (1951). *The bias of communication.* Toronto: University of Toronto Press.

Kochen, M. (1984). Information science research: the search for the nature of information. *Journal of the American Society for Information Science, 35*, 194–199.

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: III. Demand price. *Journal of the American Society for Information Science, 43*, 242–248.

Repo, A. J. (1986). The dual approach to the value of information: an appraisal of use and exchange values. *Information Processing & Management, 22*, 373–383.

Repo, A. J. (1989). The value of information: approaches in economics, accounting, and management science. *Journal of the American Society for Information Science, 40*, 68–85.

Taylor, R. S. (1982). Value-added processes in the information life cycle. *Journal of the American Society for Information Science, 32*, 341–346.