Chapter 1.
Information

Figure 1.1: redraw

1.1. What is Information?

The value of information seems simple. If I'm lost and I don't know the way to get someplace, I can look at a map or ask directions: (usually) my confusion is ended and my path home is clear. What then is information? A good staring point is that "'information" helps people make better choices and be more effective in their actions than they would without it. [1] Suppose that a child hears a key turning in the door lock and this sound is always heard just before the child's mother returns home from work. The sound of the key in the lock provides information that enables a prediction. Upon hearing the key turn in the lock, the child may get up and be ready to greet her mother (or may hurry to hide some candy under the rug). Further, when visiting a friend, the child may generalize the sound of the turning key to the new situation and expect that the same noise at the friend's house signals the return of the friend's mother.

Information allows people to make effective predictions about the future. In this sense, information helps people to interact more effectively with our world, thus influencing the way we think and act. This definition focuses on the utility of information for a particular person — it has meaning because it helps a person construct their actions. We can also talk about information as constituting a property of an object. One might hear that a book or a Web site is "full of information." This means that the content has the potential to be meaningful for the reader.

1.1.1. Structure and Context

Information necessarily has a limited range of applicability. Structure helps to specify that range. Lists are organized by dimensions such as chronology, space, luminosity, or alphabetical order, and that structure can give us clues as to the meaning they are intended to impart.

Natural languages, those languages used by human beings, combine symbols in a complex structure. Compare the sentence "The robber was described as a tall man with a black mustache weighing 150 pounds" with "The robber was described as a tall man weighing 150 pounds with a black mustache". In the first sentence — common sense aside — the structure of the sentence leads us to believe that the robber's mustache weighed 150 pounds. The second sentence, composed of the exact same words, conveys a more plausible, though perhaps less amusing, meaning (7.2.3). Structure, especially that of a natural language, is referred to as "syntax". Syntax is often contrasted with "semantics," or meaning.

Understanding the *Dilbert* cartoon in Figure 1.2 seems easy, but to even begin to understand this

[1] While this definition is a convenient place for us to get started, there is a more extended discussion about the definition of information later (1.4.0). Note that (1.4.0) indicates there is related material in Section 1.4.0.

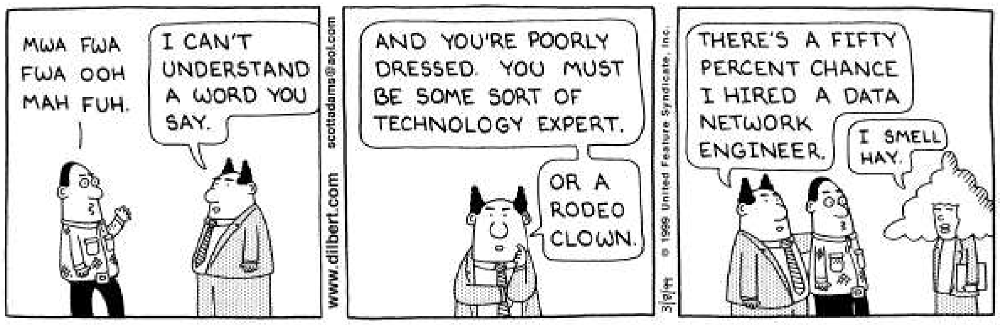

Figure 1.2: Note how much context must be inferred to understand the pictures. (redraw)(check permission).

cartoon, the reader must know that the speakers are in an office, that the man in the suit is the boss, that networking is a somewhat mysterious art to the layperson, and that rodeos are held in places where there is hay. In short, without the proper context, a reader would have trouble even understanding this cartoon, much less finding it funny. Pragmatics.

Individual symbols (e.g., words) have very little meaning in themselves, but do in their relationship to other things. A smile may indicate a greeting in one context and irony in another. When a situation is described by a sign or a set of symbols, many of the subtleties of its context are lost. Consider how a personal photo often seems hollow in comparison to the original event or meaningless to a stranger; that is because much of the context is missing. However, by its very nature, context is difficult to represent accurately. People are continually evaluating their environment (Figure 1.3).

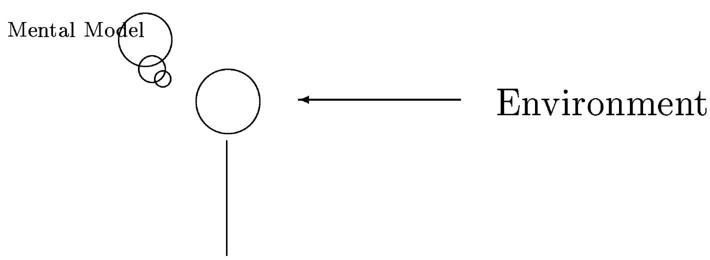

Figure 1.3: People actively interpret what is going on in the world based on their experiences.

Simplest type of information is facts about the world which can confirmed and explained. By comparison, facts and constructs about the social interaction are generally much more subjective. We tend to have beliefs about such things. Social science is very different from natural science.

1.1.2. Information in Social Interaction

When two people interact, each forms a model of the other and what the other means by their statements. One person may come to understand the other person's sense of humor, and when they are joking and when they are not. Indeed, this model formation is a property of all mutually interactive systems including people interacting with computers. Human beings have developed language as a system to transmit information from one person to another. The key turning in the lock to announce the mother coming home is a good predictor of her arrival, but it is not intended to announce the arrival; that is, she does not intend to use the sound of the key as a signal of her arrival. However, the mother might call the child's name or say "I'm home." That is an intentional, direct transmission of information through language. For language to communicate information successfully, there must be agreement about the meaning of symbols. Although this agreement is largely defined by the arbitrary assignations of meaning to words or phrases, as may be defined in a dictionary, agreement is also based on shared experiences, including cultural expectations and education. A message such as "This is the end of the line" may mean one thing to the sender and something entirely different to the recipient (Figure 1.4).

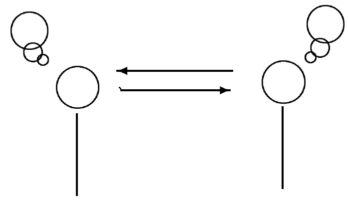

Figure 1.4: Communication is an exchange of information between people, but it does not result in both participants having exactly the same interpretation. One part of the interaction is consideration of the motives of the other participant in the interaction.

There are several senses of meaning. Meaning could be what a speaker intends. Or, it could be what the listener undersands, It could be some absolute interpretation.

People interact with information all the time. Most of this information is routine but some types of information are critical. People often work hard to get the information they need. This can be interacting with information systems but it can also be interacting with other people. Furthermore, individuals and organizations spend a great deal of time organizing information.

1.2. Capturing, Representing, and Processing Information

Because information helps people to make predictions about their environment, it is natural to develop systems for the organization, access and use of information. Using representations to building tools to change environments.

1.2.1. Representations and Processes

Aspects of the world are captured in representations or abstractions. The terms "representation" and "model" are often used interchangeably. We make the distinction here that a representation is static while a model is dynamic. A representation captures some aspect of the world as an "abstraction". Figure 1.5 is a schematic for the use of a representation. Every representation is accompanied by techniques for using or reconstructing its content, and the "model" is the combination of the representation with those routines.

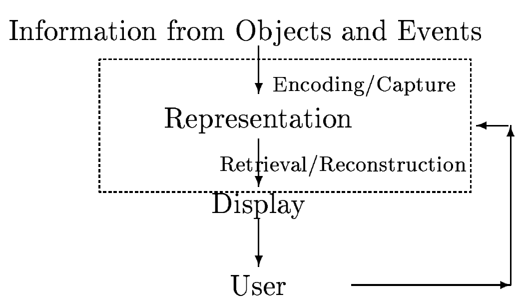

Figure 1.5: The components of a very simple information system. Information processing occurs both during encoding and during retrieval. The box shows the "model" which is the combination of the representation and associated information processing.

Attributes

Types of Representations

There are many kinds of representations with different levels of semantics, or meaning. Some of the simplest representations are categories. But, representations can also be mathematical functions. The points in the graph on the left of Figure 1.6 are well described by a straight line. Indeed, the distinction between categorical and numerical representations is the basis of a lot of discussion. Some people argue that categorical representations are essential for symbol processing but numerical processing could approximate categories and might account for some characteristics of human information processing.

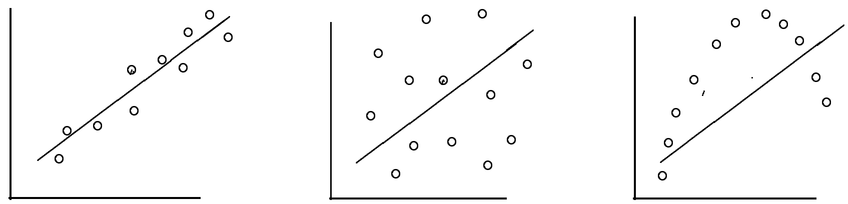

Figure 1.6: A straight line can be an effective representation to describe a set of points (left). However, the line is less satisfactory if the points are scattered (center) or if a curved line may be a better representation (right). A representation that fits curved lines will be more complex.

We can identify several types of representations. Categories, entities, concepts, classes, and nouns are all similar, though not identical ways of describing objects. What's more, some models are purely numeric. Other representations are quantitative models. Still others, emphasize processes. Even sequences of DNA can serve as a representation.

Representing Processes

When designing information systems, we often use models that attempt to duplicate a natural process (Figure 1.7). A computer program or simulation may use the model to generate a result for a given set of conditions. Interactor's models are the mental representations of processes. The actual results are a "goal" or a "target". Nonetheless, these models do their best to reflect natural situations. model may be abstracted and used to generate new responses for situations beyond those from which it was originally developed. As with other representations, many of these environments and systems are too complex to model accurately, and many details are lost. It is often helpful to develop several related models, each of which specializes in different aspects of the same task. When we examine models used in education, we will consider the student model, the domain model, and the task model. There may be multiple levels of goals or targets for each model, and these may even conflict with each other. Indeed, even static representations have implicit processes. The ability of a model to match complex activities depends on how well the representation and processing match the task. A process can store (or encode) information if it causes actions to be more effective. Furthermore, there is necessarily a limited context in which a model applies.

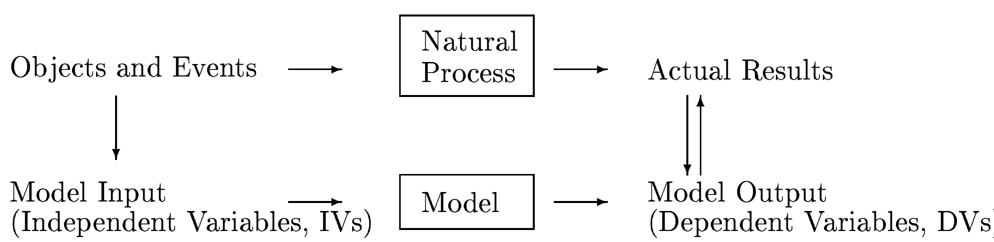

Figure 1.7: A simple model tries to emulate a natural process as closely as possible. Given the same input (the IVs), the output of a model (i.e., the DVs) should match the output of the natural process it is emulating.

These processes are specified by workflow models. These models often lead to the idea of causal associations. Indeed this is the basis of science. Mathematics explores abstract procedures and computer programming implements them. A system integrates individual components and coordinates the interaction among these components.

Adaptation and Learning

In some system information is used for automated process monitoring and adjustment – that is for adaptation. The information may signal when to move to a new state or the basic model has simple feedback with a controller (Figure 1.8). It is common for such systems attempt to attempt maintain current conditions (i.e., homeostasis). Control point. A typical example is a thermostat. Cybernetics and mechanisms of control [9]. Feedback to go to the next stage in a process. By some viewpoints, control is a type of information. It is is information which requires action and is enforced by consequences. Tradtional view of feedback controls a single parameter but with computing systems, the contoller itself can be modified.

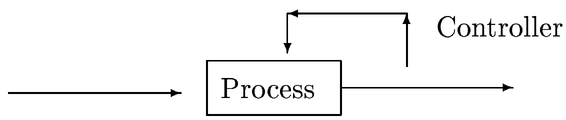

Figure 1.8: Feedback uses the output of a process to determine how to control that process.

Beyond simple adaptation, learning is a change in a representation. This may be through a conscious process of knowledge discovery. Or, it more spontaneous. That is, the information which has been received is captured by the representation. Learning requires an adaptive representation. Some models are adaptive – they adjust representations and processes to match changing conditions (Figure 1.9). Human learning is particularly complex and people seem to have multi-layered representations.

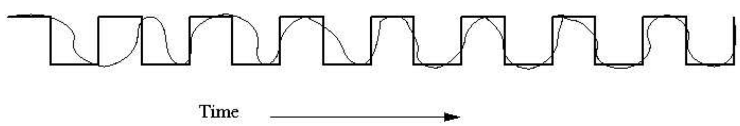

Figure 1.9: Imagine a system that produces a regular pattern (dark line) across time and a model that is trying to adapt to it (light line). The model will be successful at matching the target only if it has an adequate representation and learning algorithms. In this example, by the end of the sequence, the model seems to have learned the phase but has trouble matching the sharp angles of the original pattern.

Using Representations

Computation or information processing operates on representations. It can include the creation of representations and extraction of information from them. Indeed, the representations and algorithms for operating on them go hand-in-hand. They are both required for the task to be accomplished.

Inferences must be drawn when synthesizing those fragments. Logical and statistical approaches. Even if we have an effective representation, we need to be able to operate on it. Information processing versus computing. An information system may "know" something (i.e., it is represented in the system) but not be able to retrieve it; this would be a "retrieval failure".

Ideally information can be generalized to situations beyond which it is collected. We can use such information to navigate novel situations. The representations and information processing techniques can be symbolic or numeric; that is, the representations can be comprised of symbols (e.g., language) or images (e.g., photos). When there are unstructured tasks with many possible outcomes, the synthesis of facts (i.e., inference) of information is often needed.

Rules of thumb. Heuristics. Logic is one approach to inference; logic (A.6.0) is a set of procedures for generating inferences from categories and attributes; one type of basic logical inference is the syllogism (Figure 1.10). This type of reasoning is valid only if the category assignments are valid. A second type of inference is based on statistics or probability. We many infer that the sun will rise tomorrow based since it has risen every other day of our lives. Inferences made by humans are susceptible to cognitive distortions such attentional bias and poor ability to calculate probabilities.

Assertion	if all animals breath and
Assertion	if Lassie is an animal
Inference	then Lassie breathes.

Figure 1.10: When information is stored as attributes those attributes can often be combined to form additional propositions. Inference may be logical or statistical. A syllogism is an example of deductive logical inference.

Recognition is similar to inference. It helps to answer questions such as – "Is this person Bob Allen?" or "Is this car a Buick?". Recognition may occur as part of information capture or during retrieval. The identification of phonemes is a part of speech recognition systems (12.3.3). Recognition may simply mean identification or categorization.

Creating an instance from a model. Rendering. This is most often useful for simulation.

1.2.2. Interactivity and Coordination: Communication, Natural Language, and Social Processes

Creating and sharing information. Interaction suggests a complex, dynamic process between two systems. We may have human-computer interaction or human-human interaction. While have already discussed the nature of information transfer in social interaction. Language allows people to communicate complex information. At the level of culture and society, language and writing are related to the development of complex social organizations that are composed of semi-autonomous units coordinated for a goal. Symbols may be combined into larger meaningful units. Sometimes, the composition of symbols simply modifies the original and sometimes it creates a much richer meaning such as discourse, explanations, and narrative. Conversation as a model for interactivity.

When a person interacts with an information system by making a query or by following a hyperlink, they retrieve resources that are presented by the system, often in combination with other resources. The process of interaction is sometimes described as a negotiation because it is often involves having the participants find a common ground for coordination – whether that is between two people or between a person and an automated information system. Interactors model.

Information is interwoven with social interaction; thus, the development of information systems affects the development of complex social organization (Figure 1.13). The importance of information in modern society is evident in the many information-intensive social institutions such as schools and libraries as well as the prominence of information services in government. Science is a set of procedures for generating knowledge has shaped our technological society. Similarly, the ease of collection and dissemination of large amounts of data greatly increases social complexity. Social needs help determine the ways information systems are developed so that there is a socio-technical interaction.

Freedom of information is interdependent with market economics and democracy. Indeed, information is almost always developed to facilitate meeting human needs. In a social situation, essential information about the situation is often obscured – either by sheer complexity or by intentional manipulation of appearances. This tendency can be corrected by transparency.

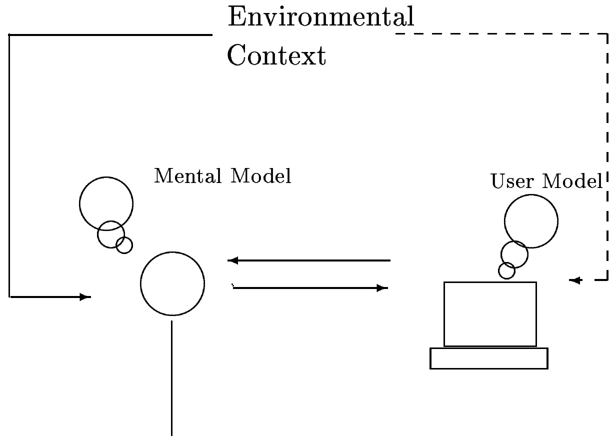

Figure 1.11: In a conversation, each participant between a user and an information system. The user and information system may tailor their responses to each other. Moreover, they may be affected by the environmental context. We would say the user has a mental model of the computer and that the computer has a user model.

Speaker A	**Speaker B**
How do I get to Berlin?	
	You should take the 10 o'clock train.
Do I need a passport?	
	No, not anymore.

Figure 1.12: Interaction with an information system may be viewed as a type of dialog. Here we could imagine that Speaker B was natural language interface. Beyond verbal interaction, all types of interfaces need to have properties of a dialog.

In addition to the areas on which they focus, information systems have indirect effects on the user and the user's organization or social context. The way that information is presented and used can have lasting impact on both the actions of a societal group and its consciousness. Rhymes and fables can provide simple and easily remembered lessons that become part of society's collective cultural heritage, while on a smaller scale, a system of transmitted information can help define a corporate or organizational structure.

Information technologies and systems affects human culture. Writing and civilization developed simultaneously. Papyrus allowed early libraries to develop, the transistor radio encouraged the development of rock-and-roll, and the computer and computing networks, among many other effects, have created video games and made long-distance communication easy. Stored information helps people to make good predictions.

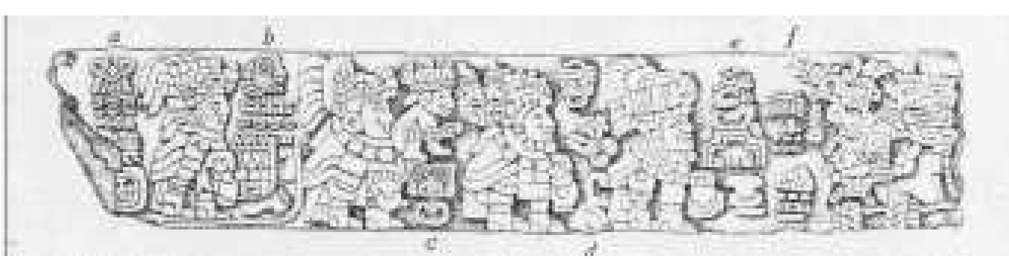

Figure 1.13: Written language is often associated with the development of complex social structures. In this case, written Mayan, accompanied urban development in Central America. (check permission)

Situatedness of social interaction.

Structure. Discourse. Tasks, work, activities. Explanation, narrative, argumentation, and Indeed, narrative seems to be a way that people analyze the world.

The focus in on what is conceptually meaningful rather than what is easily structured and manipulated.

Genres.

Self presentation constrained by norms

Interaction.

Norms and maxims.

Artifacts, personal memories, and social memories.

Social networks.

1.2.3. Systems

Systems are made of interlocking processes which pass information among them. When a system – such as a social organization – has grown up by itself it is often not very well structured. In order to model them we need a set of interlocking models. Indeed, such models may be purely descriptive or they may be used to help re-engineer the system.

While systems are decomposable into sub-systems others are not easily decomposed. Handling complexlity. Complex systems often have non-linear behavior. Feedback. That is, small changes is the inputs can have dramatic changes in the outputs. Complex systems can be stable but it is often difficult to predict what might make them become unstable.

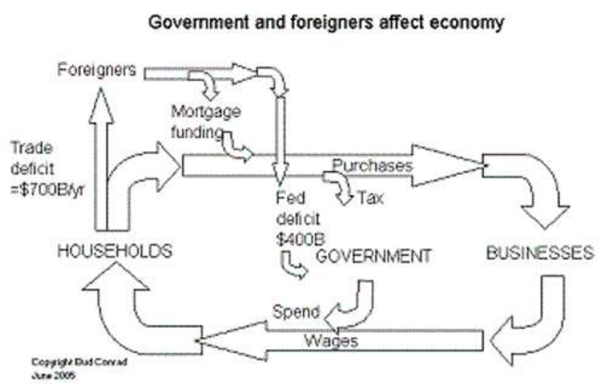

Figure 1.14: A small example of complex system. Complex systems have many feedback loops and often have non-obvious behavior. (check permission)(redraw)

Some systems develop a stable equilibrium without external control. These are self-organizing systems. Other systems are chaotic. Emergent systems have high-level regularities which emerge from seemingly random patterns at lower levels. There are many examples. Culture and society are complex systems. Biology, economics, and even society itself are examples; typically, these structure incentives.

But, many of these are not enterly stable and, further, the environment in which they exist may be changing. In other words, the strucutres are poorly defined and the data flow among them is uneven.

1.2.4. Formal Models for Representations

Categories. State machines. Numerical representations.

Simulations as models.

1.2.5. Design

Handling complexity with design. Designing individual features versus seeing the whole.

1.3. Information Resources, Services, and Systems

When information is captured and managed, it becomes an information resource. When the environment is stable, and the effective representations employed, information can help people adapt to new situations. When we save representations, they can become information resources. Because it is so useful and because it can be applied across situations, information is often collected and organized into information resources. Moreover, information systems are developed to manage and present the information resources. What we know about the world. Ordinary interaction with the world.

1.3.1. Variety, Management, and Use of Information Resources

Information resources differ most fundamentally in terms of their representations but they also differ in the ways the information is captured and the ways it used. Information resources include works such as documents, aggregations such as databases, distributed content such as neural networks, and model-based simulatons. We also include collections and ad hoc sets such as web pages and working materials of project teams. Whether they are work or aggregates, their organization or structure is an essential aspect of the resource. Such organizational structures may range from hierarchical or data-flow models. By adding secondary representations such as summaries, links, and descriptions. The needs of users for information can be supported either by interfaces that allow the users to explore the available information, or by interaction with other people who may be able to help them.

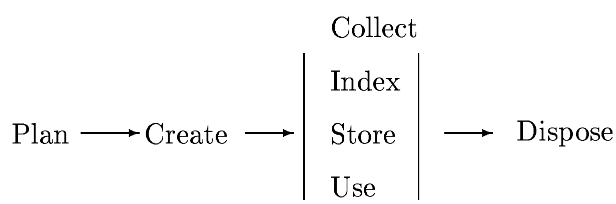

Figure 1.15: Information resources have a lifecycle which starts with planning and ends with disposal of the resource. In some cases, an old version is replaced by newer versions.

Information is integral to the completion of tasks. In some cases, the information the person uses consists of formal resources (e.g., journal articles, books) and in other cases, the information is derived from databases and email messages). An information system can adapt to a user's interests, abilities, and roles. That is, it develops user models and task-models. Beyond supporting access to specific resources. Some users are engaged in complex intellectual activities. Thus, the information environments should support complex activities such as critical thinking, design, and science.

There are several shematics to describe how information resources are used. Here we consider three of those models. The first (Figure 1.17) is most typical of scholarly research in which a lot of background information is collected. While these two models, focus in interaction with existing resources, information resources also include materials which are generated in the process of completing a task (Figure 1.18). Such materials may be informal such as notes and annotations or relatively formal such as working papers. Understanding the tasks required of the user helps us to define information systems. This is very similar to the schematic for social interaction shown earlier and extended in Figure 1.11. Some decisions, may involve extensive research ("look") while for others the person already knows the necessary information and doesn't need to acquire more. This model also describes collaborative interaction with information resources.

Increasing opportunities for managing stored information. Informatics.

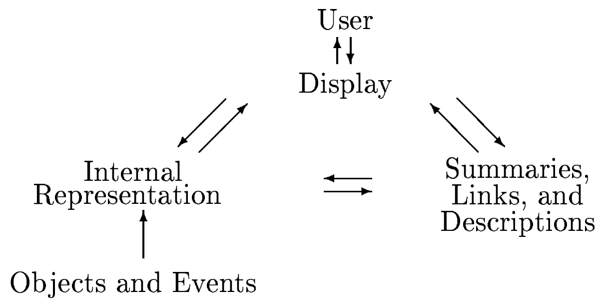

Figure 1.16: Related information resources are often collected and aids like summaries, links, and descriptions are useful in retrieval of resources from those collections. Multiple representations.

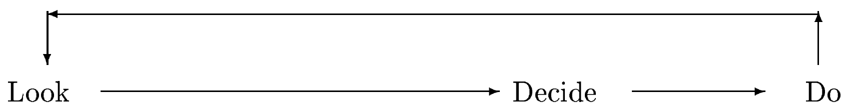

Figure 1.17: A very simple model for using information to complete tasks: $Look \rightarrow Decide \rightarrow Do$.

Figure 1.18: In still other cases, a significant information artifact is developed and the Developer may refer frequently to it.

Information systems are socio-technical systems. That is, the policies, the technology, and their acceptance co-evolve.

1.3.2. Description and Access

Developing approaches for finding information.

1.3.3. Managing Content and Information Policy

Information systems manage sets of information resources. As Figure 1.15 shows resources evolve, and in fact, entire areas evolve. So the information systems cannot be static. The policies for their management involves tradeoffs. Collections of objects have common descriptive themes. A simple collection may aggregate objects such as the sea shells you picked up at the beach. However, more formal collections try to select objects; thus, we might try to include a range of different shell types. A collection of information resources should not be static. Collections of information resources are often managed by libraries (8.3.0), and the items in a collection are often selected with some common theme. A collection could consist of books by a certain author, or it may attempt to provide all the important viewpoints on a given issue. With regard to information, this means that collections attempt to provide a variety of perspectives on certain topics. A bibliography is a list of records that survey the issues surrounding a given topic, whether or not the collection actually has those records. Bibliographies are usually generated in response to specific user requests. They may be either "descriptive" or "systematic". Tagging the internal content of documents with XML.

Policy sets objectives which are used as the basis for rules. The intention of policy is to create more effective systems. We might have policies for the management of information systems and their content. Policy is also important in human social systems.

Beyond ensuring simple accuracy, an information system should useful content; thus, the information manager must understand the needs and capabilities of the user population. Content in an information system is often associated with specific tasks; when those tasks are complete, the content can be discarded. Even material that is not task-oriented becomes dated and should be removed. The lifecycle for content is illustrated in Figure 1.15. The management of information content necessarily involves trade-offs and judgment about selection, access, organization, and preservation. Tradeoff between best practices and triage.

1.3.4. Variety of Information Systems

Information systems manage information collection, storage, and delivery. There are many types of information systems; search engines, collaborative systems, and databases. Information systems can be as diverse as airline reservation systems, digital libraries, Web servers, or virtual-reality environments. These types of systems add value to raw information in several ways: they make information easier to use, reduce noise, increase information product quality, increase adaptability, and save time and money [8]. We consider information systems more broadly than just those which serve businesses. Services science.

Information systems help information to show its greatest utility; they help people use information to direct their actions toward a desired outcome. To this end, it is helpful to consider one approximation of how information helps people to complete tasks. Given a goal, a person must *look* for information relevant to that goal and use that information to *decide* on a course of action, which they then carry out, or *do*. This may be summarized as *Look*→*Decide*→*Do*. For the completion of complex tasks, this sequence is often repeated as a cycle at various stages. More complex tasks interweave information and action at several levels (8.6.3).

Information Retrieval Systems	Information systems manage the capture, storage, and retrieval from sets of representations. They provide the broader context for the management of information resources, document-management systems, search engines, and libraries. Indeed, the human cognitive system is a type of information retrieval system
Simulations, Decision Support, and Interactive Media	Simulations goes beyond simple information retieval to synthesize projections. Games and entertainment.
Artifact Preparation Tools and Communication Management Systems	Information systems may also serve as a platform for mediating human interaction. That is they may be a community resource such as Wikis and an information commons.

Figure 1.19: A rough categorization of information systems.

1.3.5. Developing and Managing Information Systems

We have defined information systems broadly to include many approaches for providing information to people. An information system is more than a simple technology; it also consists of content and users. Designing, constructing or using an effective information system thus involves much more than simply writing a good piece of code or using the fastest computer; it requires all of these pieces — the task, the content, the users, and the system — to function as a coordinated whole. The following sections outline various issues for information storage and processing, use and impact, and content and system management.

Developing a design.

When information systems are introduced or changed, they often disrupt existing practice.

Making interaction easy. Indexing and user interfaces. Information resources should be easy to access and to use. Tools for access.

Services.

The management of information systems and content should reflect the types of users and their needs. In addition to managing the information system should be distinguished from managing the content. To develop a useful information system, it is usually necessary combine human procedures and technology upgrades: detailed analysis of how the system will be used; a clear vision of the lifecycle of the entire system, as well as the content; audits and best practices guidelines; the incorporation of new technology; and wise information management (8.7.2). Systems themselves also have a lifecycle (Figure 1.20).

Best practices.

Figure 1.20: A simple model for the lifecycle of an information systems. We will consider more complex development processes in later chapters. Note the similarity to other types of information systems.

The development and maintenance of a set of information resources and of the underlying information system are often costly. A business model is needed to offset these costs. Among the most common business models are support for organizational activities (as in knowledge management systems), direct payment (akin to a newspaper subscription), and public service (as in public libraries).

Information systems include the social and organizational support necessary for managing the resources. For an information system to be of value, the information that it contains must be accurate. Decisions need to be made about what should be included, and also to whom the information should be available. Even when the system has been deployed, content often needs to be refreshed. Poor information management may result in information disasters [5]. Figure 1.21 gives an example of the possible serious consequences of inadequate information.

1.3.6. Information Technology

Advances in technology also affect a system's viability. System development has become easier, and the systems can be more widely deployed. Certain elements of technology have also become much less expensive: Figure 1.22 shows the change in disk prices as a function of time. Several other aspects of computers, such as the CPU, network speed, and display size are also changing rapidly. Indeed, this change has become predictable: the consistency of the change in CPU cycles was originally noted by the engineer Gordon Moore, so it is known as Moore's Law. Of course, this is pattern is driven both by technological developments but also by demand and economies of scale.

Algorithms, Operating system. Cloud computing.

Sensons report the world without human observation Figure 1.23. Indeed, the great proliferation of sensors has contributed to the development of data libraries.

Operating systems, network infrastructure, and careful systemic evaluations, or audits, can provide that security. Also, standards complement infrastructure development — the improved infrastructure provided by hardware and networking is paralleled by greater standardization of content so it can be

As NATO and the United States continue to deal with diplomatic fallout from Friday's Chinese embassy bombing in Belgrade, a senior U.S. intelligence official told Salon News that the CIA team in charge of choosing Yugoslav targets does not include any agents or experts with recent on-the-ground experience in Belgrade.

Speaking on condition of anonymity Tuesday, the official said that no CIA officer with an up-to-date, walking familiarity with the Yugoslav capital was on the targeting team when China's embassy was mistakenly bombed Friday, killing three occupants and injuring 20 more. Nor, apparently, does the CIA have clandestine spotters in Belgrade helping verify targets picked from maps and satellite photos.

The issue has taken on added gravity because the CIA has admitted it used a partially updated 4-year-old street map and "educated guesses" to select the target, which was thought to be a Yugoslav arms agency. In this case, the maps did not show that China had vacated its old property and built a new embassy elsewhere in 1996, even though American officials, from the U.S. ambassador to the semi-public chief of the CIA mission, frequented the embassy for events. The U.S. embassy in Belgrade was closed and its staff evacuated March 24.

Figure 1.21: One example of the problems caused by faulty information (from [1]). (check permission)

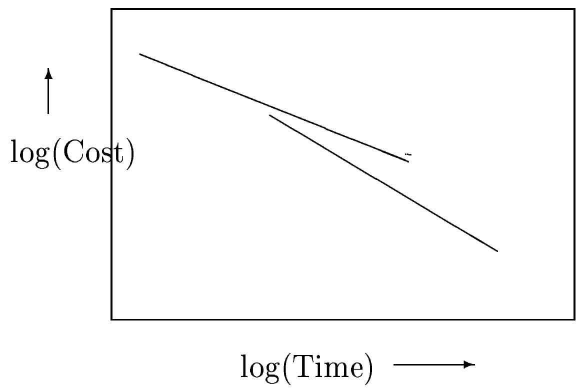

Figure 1.22: The cost per megabyte of storage keeps dropping (adapted from [6]).

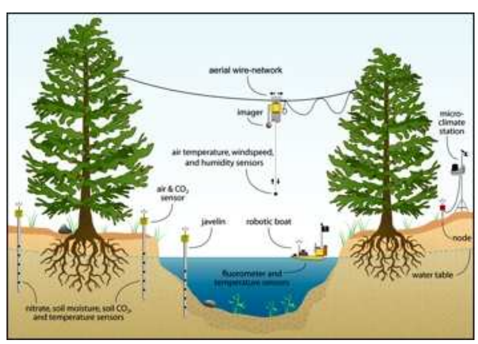

Figure 1.23: Sensors collect information from their enviorment. Here, a network of sensors monitors conditions in a forest and the ecology is them able to be modeled. (check permission)

exchanged more readily. Not only are individual services able to operate, but they interoperate with other services more easily. Indeed, much of the success of the Web may be attributed to developing standards for exchanging content such as HTML, HTTP, and MIME.

We are in the midst of a “digital convergence”. There are many types of information can now be represented digitally, and in many cases, the semantics of this information can be converted from one medium to another. These changes alter the nature of processing information – digital content is easy to copy, easy to convert, and easy to process.

Senors primarily collect information (or data). Basic sensors don’t process that information but pass it to a central processor to make inferences about it. The data from somewhat more complex sensor networks may be processed with sensor fusion in several layers.

1.4. The Language of Information

Many of these terms, however, are confusing and often inconsistent. Here, we discuss some of these confusions but we will encounter other ambiguous terms throughout the text.

1.4.1. What is Information

Information is clearly not something physical. Thus, we must define information in terms of its effects. There are several issues for defining information. Information creates a change in a representation.

A related definition of information is that it “reduces uncertainty”. This is derived from Shannon [7] though we note several issues and implications. The first, is that how we measure uncertainly. Shannon defined it in terms of bits (A.1.1), but that is often difficult to apply in complex situations involving humans. A second difficulty is that although uncertainly may be reduced, that may not be mean the information is accurate. If somebody told me that today is Tuesday when it is actually Wednesday, I may believe the incorrect information. There are several ways this problem could be addressed. We could claim that information helps a person to predict the future. Or, perhaps that information helps a person to make good decisions. Another important distinction is that information is a construct at the system level. Thus information as a non-physical interaction between two systems which results in a change of behavior of the receiving system.

A distinction is sometimes made between subjective information versus stored information. Information is sometimes thought of as a thing – that which is stored in an information resources such a book or journal. Such instantiation allows information to be managed. It can be difficult to determine whether a change of behavior will actually occur so we may examine representations to assess whether changes in behavior are likely. In some cases, such as in computer programs, it it relatively easy to see what the effects will be by examination of the representations, while in other cases, such as when trying to understand a person, it it very difficult, if not impossible, to examine cognitive representations.

Information is the content which fills models and representations whereas learning is a change in the structure of the representation. There are many other definitions of information [4]. In several ways, the representations are at least as important as the information which constitutes that information.

Ambiguity.

Representation and description.

Cultural and social information.

Information is essential for interaction.

1.4.2. Data, Information, and Knowledge

It is often claimed that there is a hierarchy for Data to Information to Knowledge, specifically, information and knowledge are said to include more context than data. Data might include just a table of values whereas information and knowledge might be include a paragraph of text or knowledge in a person's head. Another dimension concerns the generality of the knowledge. Data or information is often about individual entities whereas knowledge more often concerns categories and classes. If that user realizes that they have a need for that piece of content, then at that point it ceases to be data and becomes information. "Knowledge" is a term that is generally used to mean that an individual possesses the minimum amount of information necessary to accomplish a certain task, such as having the know-how to fix a flat bicycle tire [3]. Traditionally, a continuum is established from "data" to "information" to "knowledge". This is generally associated with the amount of context. However, the terms are used so inconsistently that the continuum is at best, only an approximation. There are certainly distinctions to be made about the context and richness of information but these three terms are used so loosely that they have lost most of their meaning (Figure 1.24).

Make observations and collect the results of those observations

Organize and summarize those results

Integrate with other information and beliefs

Apply them to making a decision

Figure 1.24: The information value chain describes important levels of activities for processing stored information. Actually a great many services can be applied to information resources.

Attitudes as cognitive constructs. Complexity. Computing.

1.4.3. Entertainment and Information

Entertainment shares many aspects of information including being stored by information systems. While entertainment often emphasizes emotional reactions, quite a bit of entertainment is also informative. We can learn about how people act under pressure from movies and novels. Moreover some information seeking, such as reading the newspaper, can be a type of entertainment and playing some games can be educational. On the other hand, it seems like some entertainment is about mood optimization. A horror movie seems to provide an emotional jag rather than reducing uncertainty about the future. In any event, the technologies for managing entertainment and information overlap so much that we consider them together here. Affect.

Figure 1.25: The Fox and the Grapes from Aesop's Fables. Entertainment can also provide information (from [2]).

Emotional content needs to be reconciled with information. Emotional material also changes the representation but in a transient way.

Many other difficult definitions. Intelligence. Effectiveness for functioning. Common-sense.

1.4.4. Beliefs and Objectivity

We generally think of a fact as a statement which conforms to observation of the external world. In most cases, this is straightforward common sense. However, it can be complicated either when the observation of the external world is ambiguous or when the statements are about complex processes. There is a long philosophical debate about the nature of the physical world and our perceptions of it. One approach, which is the "logical positivist" view, claims that the world exists very much as we perceive it. A second approach, which is called "phenomenology," suggests that the world is largely a construction of our minds. This view is reflected in Descartes' famous saying, *Cogito ergo sum* ("I think therefore I am").

In general, daily life requires that we accept certain facts about our world — objects exist, we can affect the world around us, and we can gather data about that world. Truth, then, is congruence with the world. However, there are are many ways that a statement can be confirmed in the context of world. In most cases, the statements are made in the context of a broad conceptual system such as a scientific or cultural analysis.

Rather than attempting to determine an absolute truth about a situation "objectivity" is better thought of the practice or policy of avoiding overt personal references or personal interpretation to whatever extent that is possible. Rather, it objectivity can be thought of as following logical inferences and also of providing as many significant perspectives from many people in the population.

Relativism argues that these is no objectivity.

Rather than being objective in some absolute sense, journalists, for instance, attempt to maintain a neutral point of view with regard to their reporting of observable facts. Even if objectivity is not possible, a journalist may try to be an "honest broker" of information. That is, to fairly represent opinions relevant to minorities in society.

Another position is that and that news organizations should state their positions directly. Note that this is different from always presenting exactly two alternatives. Even if there is no absolute objectivity that does not imply that all alternatives are equally plausible. How to make an objective presentation of information. Objectives weighted by population, by expertize, or by outliers.

Exercises

Short Definitions:

Abstraction
Compositionality
Context
Data
Digital convergence
Entertainment
Feedback
Information
Information resource
Information value chain
Lifecycle (content)
Lifecycle (system)
Moore's Law
Objectivity
Positivism
Recognition
Relativism
Representation
Retrieval failure
Semantics
Structure
System
Top-down processing

Review Questions:

1. Give five examples of structured information. (1.1.1)
2. Explain how information relates to information systems. (1.2.0)
3. What makes a representation effective? (1.2.1)
4. Describe a typical representation for (1.2.1): a) music, b) images, c) video?
5. How is recognition different from classification? (1.2.1, 2.1.2)
6. Distinguish between lifecycle of the content and the system lifecycle. (1.3.3)
7. Name five information sources in your school or university. How are they managed? (1.3.6)
8. Give some examples of policies and procedures for the management of information systems (1.3.6)

9. Give an example of digital convergence. (1.3.6)
10. Distinguish between information management and information system management. (1.3.6)
11. Give an example of how an information system has transformed an organization. (1.3.6, 6.6.0)
12. Distinguish between brute facts and social facts. (1.4.4)

Short-Essays and Hand-Worked Problems:

1. Describe several different senses in which a video program (e.g., a television sitcom) has "structure". (1.1.1)
2. Identify three information systems and describe their representation, information processing techniques, content, and typical use. (1.2.1)
3. In what sense are the following examples of representations: a book, a library collection, a library index, a video game. (1.2.1)
4. Describe what types of inference techniques can be applied to the following types of representations: a data model, formal logic, numerical descriptions. (1.2.1, A.6.0)
5. Describe a situation in which a friend needed to find information and the strategies they adopted to do that. (1.3.1)
6. Give an example of failed information systems introduction in an organization. (1.2.2)
7. Many communication and information media have been developed. It is sometimes proposed that new media do not replace existing media types, rather the old media forms continue to exist. Do you agree? How will television and radio in the future be fundamentally affected by the Internet? (1.2.2)
8. There is a saying that "content is king" which suggests that in a highly networked world, content is more unique than the communication medium. Do you agree? Give an example. ((sec:content))
9. Estimate how many Web servers there are in the state or country where you live. Then, estimate the average number of pages on each server and the average number of bytes per page. Finally, approximate the total number of bytes available from the servers in your state or country. (1.3.6)
10. The Web is growing rapidly. Estimate how large your answer to the previous question will be in two more years. Explain how you derived this estimate. (1.3.6)
11. Write an equation for Moore's Law. (1.3.6)
12. As of this writing (2009) a fast CPU is about 3 Giga-Hertz. If Moore's Law continues to apply, how fast will processors be in 2015? (1.3.6)
13. Do you agree that video games are an information genre? (1.4.0, 12.6.0)
14. Does fiction convey information? Does a sculpture convey information? Does an antelope convey information? Explain. (1.0.0, 1.4.1)
15. Flowers have evolved with distinctive shapes and colors to attract certain insects. Would you say the flowers have learned to convey information? (1.0.0, 1.4.1)
16. Is gossip a type of information? (1.0.0, 1.4.1)
17. If you were viewing a basketball game, would you say that information is inherent in the basketball game or is it there only for the players and the viewers? (1.0.0, 1.4.1)
18. There are many difficulties with definitions of information (1.0.0, 1.4.1). Give your definition of "information" and discuss the following puzzles:
    a) Do nerve impulses in your brain carry information?
    b) Where is the information in an organization?
    c) Distinguish between "information acquisition" and "learning".
    d) Distinguish between "information" and "entertainment".

Going Beyond:

1. Is it possible to estimate the total amount of information there is in the world? (1.0.0)
2. Can we have information without structure? (1.0.0)
3. What are some techniques authors might use for making their writing cross boundaries? (1.1.1)
4. Since the somebody searching for information is necessarily in a different context from the person who created the information, how is the searcher ever be sure the context is correct? Is there a tendency to uncritically accept such information. (1.1.1, 6.9.0)
5. How important are language and words for defining expectations and social interactions. (1.1.1, 7.2.1)
6. Develop a model that could learn square-waves like the example in Figure 1.9. (1.2.1)
7. What is the connection between the development of the printing press and the rise of science in Renaissance Europe? (1.2.2)
8. Do you agree with the statement that "Whatever a person believes is true for that person". (6.9.1, 5.4.0)
9. List the information resources in your immediate environment. (1.3.6)

10. Keep an information diary for two hours while at your school. Describe what information resources you access. (1.3.6)
11. Does a computer have intentionality? (1.0.0, 1.4.1)
12. What are some of the difficulties with the concept of "objectivity". (1.4.4)
13. What does it mean to understand what somebody else is saying? (1.4.1)
14. Are there always "two sides to every question"? (1.4.4)
15. What is common sense? (1.4.0)

Teaching Notes

Objectives and Skills: Introduce the concept of information as a fundamental construct. Representations. Modeling. Definitions.

Instructor Strategies: Chapters 1 and 2 might be combined in one lesson especially for students already familiar with XML.

Supplemental Readings

- American Library Association *Information Power.* Chicago, 1998.
- Biermann, A. *Great Ideas in Computer Science: A Gentle Introduction* 2^{nd} edition, MIT Press, Cambridge MA, XXXX
- Ciborra, C. *The Labyrinth of Information: Challenging the Wisdom of Systems*, Oxford, 2002.
- Brown, J.S. and Duguid, P. *Social Life of Information.* Harvard Business School Press, Cambridge MA, 2000.
- Case, D.O. *Looking for Information: A Survey of Research on Information Seeking, Needs, and Behavior.* Academic Press, San Diego, 2002.
- Day, R. *The Modern Invention of Information: Discourse, History, Power.* Southern Illinois University Press, Carbondale IL, 2001.
- Dodsworth, C. *Digital Illusion: Entertaining the Future with High Technology.* Addison-Wesley, Boston, 1997.
- Eisenstein, E. *The Printing Revolution in Early Modern Europe.* Cambridge University Press, New York, 1983.
- Grafton, A. *The Footnote: A Curious History*, Harvard University Press, Cambridge MA, 1999.
- Latour, B. *Science in Action: How to Follow Scientists and Engineers through Society.* Harvard University Press, Cambridge MA, 1990.
- Pool, J. *Beyond Engineering: How Society Shapes Technology.* Oxford University Press, New York, 1997.
- Spinello, R.A. *Case Studies in Information and Computer Ethics.* Prentice Hall, Upper Saddle River NJ, 1996.
- von Beralanffy, L. *General Systems Theory: Foundations, Development, Applications.* George Braziller Inc., New York, NY, 1969.
- Whyte, J. *Crimes Against Logic.* McGraw Hill, New York, 2005.

References

1. (belgrade china embassy).
2. Aesop.
3. Bates, M. Fundamental forms of information. 1033–1045.
4. Capurro, R., and Hjorland, B. The concept of information. *ARIST 37* (2003), 343–411.
5. Horton, F. W., and Lewis, D., Eds. *Great Information Disasters.* Aslib, London, 1991.
6. Law, M. (title).
7. Shannon, C., and Weaver, W. *The Mathematical Theory of Communication.* University of Illinois Press, Champaign IL, 1963.
8. Taylor, R. S. *Value-Added Processes in Information Systems.* Ablex, Norwood NJ, 1986.
9. Weiner, N. *Cybernetics: The Control and Communication in the Animal and the Machine.* MIT Press, Cambridge UK, 1948.