The cost of communication in economic organization: II

Hajime Oniki

1 Introduction

Professor Kenneth Arrow, in his presidential address at the 1973 annual meeting of the American Economic Association, discussed problems in the economics of uncertainty and information. In dealing with the efficiency of the price system, he stated:

In equilibrium, at least, the [market] system as a whole gives the impression of great economy in the handling of information, presumably because transmission of prices is in some significant sense much cheaper than transmission of the whole set of production possibilities and utility functions ... But what was left obscure is a more definite measure of information and its costs, in terms of which it would be possible to assert the superiority of the price system over a centralized alternative ...

if we are going to take informational economy seriously, we have to add to our usual economic calculations an appropriate measure of the costs of information gathering and transmission. (Arrow 1974a, pp. 4-5)

This chapter responds partly to the point raised by Professor Arrow; this is an attempt to find a measure of the cost of internal communication in economic systems.

For the convenience of the reader, we shall first give a brief and informal explanation of this work. We consider the problem of comparing two alternative economic systems, the centralized system and the (decentralized) market mechanism. In this work, each system is composed of a *center* and *agents*, the latter being interpreted to be productive firms. In the centralized system, the center may be considered as the planning board of a socialist state, whereas in the market mechanism it is an *auctioneer*, who runs the system by executing the law of supply and demand.

Our strategy is to have each system "solve" an allocation problem and to calculate the cost of internal communication arising from this. For simplicity, we choose a textbook problem of cost minimization.

Suppose that each agent is given a marginal cost schedule to produce an output commodity and that the total output is required to be at a given level. In the centralized economy, this level may be an output target set by the socialist government. In the market system, it may represent a level of inelastic demand for the output commodity. We know, of course, that optimality (or equilibrium) requires equality of marginal costs among all agents.

One can observe that the optimal level of output for each agent is determined from the data of this

allocation problem, that is, from the marginal cost schedules and the required level of total output. In other words, the optimal level is a *function* of the given data; the former may be *computed* once the latter is given. We can state, in more general terms, that desired decisions are obtained from data describing a given environment. In this work, we regard each of the two systems, the centralized economy and the market mechanism, as a computer," that is, an information processing system that can generate what is desired from what is given; this work deals with the *cost* of running such a computer.

From our standpoint, the two systems differ only in the way they handle economic data; otherwise, they are treated equally. In particular, the data given to each system at the beginning of computation are identical. Specifically, we assume that initially, before any computation starts, each agent knows his marginal cost schedule and the center knows the required level of total output. This condition, common to the two systems as remarked above, is called the initial dispersion of information. Further, we require that the computation output from each system be identical; the solution to the given allocation problem obtained by the centralized system must be the same as that obtained by the market mechanism, although the *internal algorithm* in each system may not. Thus, our work is like comparing two (real) computer systems that produce an identical line printer output from an identical card deck.

Computation of desired allocation proceeds as explained below. In the centralized system, each agent transmits to the center the entire marginal cost schedule, and the center, with all the data at hand, calculates the optimal allocation by itself. The decentralized system, on the other hand, simulates a competitive market. The center announces a (tentative) price of the output commodity, and each agent tells the center the quantity of output to be supplied at this price. The center then calculates the total excess demand for the output commodity and revises the price according to the law of supply and demand. The process is continued until an optimal allocation is obtained.

One sees that various kinds of informational activities need to be performed even in the simple process described above; they include observing, gathering, storing, sending, and transforming economic data. (See, e.g., Marschak 1968 for a comprehensive study of informational activities in economic systems.) In this work, however, we deal with internal communication only; the reason for this is that it is the easiest to investigate.

In order to calculate the cost of communication, we must choose a measurement unit. Our strategy for doing this is best explained by comparing communication, which is to move information from one place to another, to transportation, which is to move, say, passengers from one place to another. The amount of passenger transportation is measured by the unit of, say, passenger-kilometer, and the cost of transportation depends on this and the choice of actual means of transportation, which may be characterized by mode (air, auto), route, speed, and so on. For any given means, the greater the amount of transportation expressed by the unit of passenger-kilometer, the higher the cost. Thus, we may state that the amount of transportation is a measure of transportation cost independent of its means.

In this work, we seek to formulate a model in terms of which the amount of information to be transmitted may be obtained independent of the means of information transmission. In communication, we have no common sense unit for measuring information like the unit of passenger-kilometer in transportation. However, information theory fills the gap; in fact, as explained in Section A in the appendix, it provides a universal unit of measuring information, which is independent of the actual content of information and of the channel through which information is transmitted (see also Hess 1983, Chapter 11).

In information theory, "having a piece of information" means that a particular object is designated in a collection of possible objects. For example, the statement "the air temperature is now increasing" may mean that the object *increase* has been selected from the collection of three objects, *increase*, *decrease*, and *unchanged*. It is essential that this set be specified completely beforehand and its meaning be understood both by the sender and the receiver of information. Roughly speaking, the amount of information is measured by the *degree of difficulty* to identify an object in the set of possible objects. It depends on the number of objects contained in the set and the probability distribution according to which the object in the set takes place. It does not depend on the interpretation attached to them. As shown in Section A (Appendix), the amount of information is expressed by the expected number of letters (bits) needed to code the objects, which can be

approximated by the entropy of the probability distribution.¹

In Section 2 of this chapter, we use the entropy function to calculate the amount of information to be transmitted in the centralized system and the price mechanism when the optimal allocation is computed. In fact, the author's earlier work (1974) did this by using an elementary combinatorial method. It was found that, in terms of the cost of communication, the price mechanism was more economical than the centralization of information if the required accuracy of resource allocation was not very low; the ratio of the communication cost of the centralized system to that of the price mechanism increases as the required accuracy tends to be high.

This contribution deals with the same problem as summarized above. We shall, however, present an improved formulation of the problem so that the results to be shown will be more accurate and the analysis needed to get them will be much simpler than in the earlier work.

The main difference between the earlier and the present works lies in their formulation. In the earlier work, all the data of the model were discretized, and the entropy function was calculated by combinatorial enumeration. This made the model elementary and easy to understand, but it also made the calculation very cumbersome. In this essay, we shall construct our model by using analytical tools. This will simplify our task greatly, but there is a price for this; the distance between the reality and the model is greater in this work than in the earlier work.

There seems to exist an intrinsic difficulty in formulating models to measure informational costs.

¹ The entropy function is also called the Shannon measure (Shannon 1948), which is a "direct" measure of information in the sense that it focuses on the actual transmission of information and uses bits to measure it. As a measure for comparing resource allocation systems, it is in contrast with the one using the "size" of a space from which information arises (like the dimension of a message

As we know, analytical models are simpler than combinatorial ones. For example, differential equations are easier to solve than difference equations, and normal distributions have nice properties not shared by binomials. However, to employ analytical construction, we have to introduce a continuum like the real variables, but this is not directly compatible with the objective of calculating informational costs. The reason is that the cost of identifying an object in a continuous set is infinite (e.g., a real number can only be represented by an infinite decimal sequence). This means that if we use analytical tools with the entropy function to express the informational cost, we must introduce "approximation" into our model and somehow bridge the gap between the continuous and the discrete spheres. In this chapter, this is done by Proposition A6 (Appendix).

2 Complete centralization and the price mechanism

2.1 Assumptions common to the two systems

In this section, we construct a model to compare the cost of internal communication in the centralized system and in the price mechanism. First, let us present assumptions that are common to both of the two systems.

We consider a simple problem of resource allocation to be solved by a center" and agents (producers), each being indexed by *i* (*i* = 1, ..., *I*). It is assumed that there is only one commodity, to be denoted by x. The price of the commodity will be expressed by t. Let $T = \{t \mid 0 < t < +\infty\} = (0, +\infty)$ and $X = \{x \mid 0 < x < +\infty\} = (0, +\infty)$ be the price and the quantity spaces, respectively.

The environment to be given to agent *i* is a supply function (an inverse marginal cost schedule) g^{i} (t). More precisely, define

$$G = \{g: T \rightarrow X \mid g(0) = 0, g \text{ is nondecreasing and left continuous}\}$$

to be the set of supply functions, G being common to all agents.

It is assumed that the supply function arises from G subject to what is called the first-passage time distribution of Brownian motions, of which a brief summary is given in Section B (Appendix). In this chapter, we consider the case in which this distribution is identical and independent over all agents. Given $g^i \in G$, let g^i (t) denote the quantity supplied at price $t \in T$. By Proposition B3(i) (Appendix), the random variable g^i (t) has the density function $f(\cdot, t)$ of the one-sided stable distribution with parameter t, which is also explained in Section B (Appendix). Let z > 0 be a constant denoting the aggregate demand for the commodity. The objective of the center and the agents is to find a $t^* \in T$ such that the equilibrium condition Σg^i (t^*) = z be satisfied at least approximately.

We shall assume the following information structure. At the beginning of each period, a state of

space) (see Hurwicz 1977; Mount and Reiter 1974).

the world $(g^l, ..., g^l, z)$ obtains. Agent *i* knows g^i (and g^i only), whereas the center knows *z* (and *z* only). The equilibrium price t^* is theoretically determined once the state of the world is given. As explained in Section 1, however, our problem here is to consider adjustment processes (algorithms) that specify the detailed steps leading from the given data describing the state of the world to the equilibrium price and the equilibrium level of output for each agent. In any such process, information about the state of the world, that is, information about $(g^l, ..., g^l, z)$, must be exchanged between the agents and the center.

The cost of communication to execute a process may be calculated by examining how information is transmitted at each step of the process. As summarized in Section A (Appendix), the cost of communication is determined by the amount of information transmitted, which may be expressed by the expected number of letters needed to code the information to be transmitted.

In order to calculate the amount of information transmitted, we need to impose further assumptions on our model. First, we specify a degree of accuracy required for optimization. Let Δ_t > 0 and $\Delta_x > 0$ denote, respectively, the length of an interval on the price axis and the length of an interval on the quantity axis. In the following, when we consider decision making with approximation, we shall work only with *integer* prices specified by the interval Δ_t , (i.e., multiples of Δ_t); a price that is not equal to a multiple of Δ_t , will be represented by an integer nearest to it. On the quantity axis, we shall work only with intervals of length Δ_x (i.e., intervals with endpoints that are equal to multiples of Δ_x ; any quantity of output commodity will be represented by an interval containing it. (We assume that "ties" are resolved in some way. It will be seen later that the way they are done does not matter.) When we consider a model in terms of integer prices and quantity intervals, we say that it is in the approximation mode, and when we consider a model without approximation, we say that it is in the theoretical mode. Decision making in the approximation mode accompanies allocation errors. The length of the intervals Δ_t and Δ_x determines the degree of accuracy for optimization, but it is not necessarily equal to the allocation error (the latter will be considered later). Observe that the cost of in formational activities can be considered only in the approximation mode (as long as we use the Shannon measure). We expect that as the accuracy requirements Δ_t and Δ_x tend to be small, the amount of information transmitted will be increased.

Second, we choose a number $M \in T$ so large that the probability for the equilibrium price t^* to lie outside the interval (0, M) may be ignored. It is assumed that all of the adjustments to be considered below are carried out within this interval. Furthermore, for analytical simplicity, we assume that the equations $\Delta_t = 2^{-m}M$ and $\Delta_x = 2^{-n}N$ hold for some positive integers m > 0 and n > 0 and a number N > 0. (See Figure 1.)

Below, we calculate the cost of communication in the centralized system and that in the price mechanism. That is, we calculate the expected number of letters (bits) to be transmitted between the center and the agents for computing the desired allocation to some level of accuracy. For each of the two systems, we explain about the data in the approximation mode to be transmitted between the center and the agents, the maximum allocation error that may arise from the adjustment process using the data, and the cost of transmitting the data.

2.2 The centralized system

We first deal with the centralized system. Consider an agent who is given a particular supply function g(t). In order for the center to calculate the equilibrium price in the approximation mode, the center needs to obtain the values of the function g(t) at $t = r \Delta_t (r = 1, 2, ..., 2^m)$. There may be several alternative ways to do this. In this chapter, we assume, for the sake of making our calculation of the communication cost simple, that the agent transmits to the center the quantity interval containing the *increment* of the supply function, that is, it sends the data approximating $\{g(r \Delta_t) - g((r - 1) \Delta_t)\}$ for $r = 1, 2, ..., 2^m$ successively in this order. (See Figure 2.) The center, having received these data from each agent, obtains information to approximate the aggregate supply curve. It can then compute the optimal solution of the given allocation problem in the approximation mode. The error arising from this depends on the parameters *m* and *n* and also on the sample supply functions.

The maximum error that arises from the center's computation of optimum in the approximation mode is determined as follows. Since the agent transmits a quantity interval containing the "true" increment of the quantity supplied at an integer price, the maximum error arising from this is equal to the length of this interval, that is, Δ_x . In the worst case, this error is accumulated for each increment of the supply function and for each agent; accordingly, the maximum error in estimating the aggregate quantity supplied is equal to $2^m \Delta_x I$, where 2^m is the number of the integer prices and I is the number of the agents. Suppose that we can neglect the probability that the equilibrium price t^* lies outside the interval (0, M) and, in addition, the probability that the aggregate quantity supplied at t = M exceeds N. Then we can state that if $N = 2^n \Delta_x$ is much greater than the maximum error $2^m \Delta_x I$, that is, if n is much greater than $m \log_2 I$, the maximum error relative to the aggregate quantity supplied is small. Figures 3(a)-(c) illustrate such a case, whereas Figures 4(a)-(c) illustrate a case in which the relative error is not small. The reader who is not quite satisfied with this result is reminded that the primary objective of this work is to construct a model for comparing economic systems with respect to the cost of communication in a simple setting, not to construct a model in which the error arising from adjustments in the approximation mode is expressed in a simple form.²

Let us now turn to calculating the cost of transmitting the data approximating $\{g(r \Delta_t) - g((r-1) \Delta_t)\}$, for $r = 1, 2, ..., 2^m$, from the agent to the center. Proposition B3(ii) (Appendix) states that these data are distributed independently and each of them has the density function $f(., \Delta_t)$ of the one-sided stable distribution with parameter Δ_t . In other words, the agent is given 2^m random variables distributed independently and identically. (See Figure 2.)

Consider one of these random variables. Since the agent reports to the center a quantity interval of length Δ_x to which the true value of the random variable belongs, the set of messages from which the agent chooses is $S = \{S_1, S_2, ...\}$, where *Su* denotes the interval ((*u* - 1) Δ_x , $u \Delta_x$). It is now evident

² An obvious alternative way for the agent to transmit the supply function to the center in the approximation mode is to send $g(r \Delta_i)$, $r = 1, 2, ..., 2^m$, i.e., to send the values of the quantity supplied at the integer prices rather than their increments. With this assumption, the maximum allocation error would be expressed in a simpler term, but the communication cost would be obtained in a more

that we can approximate $p_u = \text{Prob} \{s_u\}$ by $f(u \Delta_x, \Delta_t) \Delta_x$.

Proposition A7 (Appendix) allows us to state that the amount of information needed to send a message from the message set S defined above is approximately equal to $\{H(f(., \Delta_t)) - \log \Delta_x\}$, where $H(\cdot)$ is the entropy function defined by Definition A2 (Appendix). Since there are 2^m such random variables and the number of the agents is *I*, we may conclude that the communication cost in the centralized system, say C_{I_i} is expressed as

$$C_{I} = I \times 2^{m} \{H(f(., \Delta_{t})) - \log \Delta_{x}\}$$
(1)

By making use of Proposition B2 (Appendix), we obtain³

$$C_{I} = I \times 2^{m} \{ (-2m+n)\log 2 + (2\log M - \log N) + K \} , \qquad (2)$$

where the constant K > 0 is given in Proposition B2.

For simplicity, let us assume that the parameters *M*, *N*, *m*, and *n* are chosen in such a way that $M^2 = N$ and 2m = n.⁴ We then get

$$C_1 = I \times 2^m K \tag{3}$$

2.3 The price mechanism

Next, we consider the price mechanism. The center here is an auctioneer and follows the ordinary scheme of price adjustments (i.e., the law of supply and demand) in order to get to the equilibrium price. In the theoretical mode, the center announces a tentative price, gets the quantity supplied at that price from each agent, and calculates the excess demand. If it is positive, the center will raise the price, whereas if the excess demand is negative, the tentative price will be lowered. In any case, at the following step, the center announces a revised price to the agents and repeats the procedure described above. The adjustment will be terminated when the excess demand becomes zero.

In our model, the adjustment is considered in the approximation mode so that we work only with integer prices and quantity intervals. We assume that, on the price axis, the center adjusts tentative prices in the following way: Let *t_j* denote the integer price chosen by the center at the *j*th step of the

³ $C_1 = I \times 2^m \{2 \log \Delta^t + K - \log \Delta_x\}$

 $= I \times 2^{m} \{ 2 \log(2^{-m}M) + K \log (2^{-n}N) \}$

complicated formula, than with the assumption adopted in the text.

 $⁼ I \times 2^{m} \{(-2m+n) \log 2 + 2 \log M - \log N + K\}.$

⁴ Note that this assumption is consistent with the condition (stated on p. 198) for the relative allocation error to be small.

adjustment procedure. We postulate that the initial price is chosen so that $t_j = 2^{-1}M$, and, for each j > 1, $t_{j+1} = t_j + 2^{-(j-1)}M$ if the excess demand at is positive and $t_{j+1} = t_j - 2^{-(j+1)}M$ if the excess demand at t_j is negative (or zero).

Since, in our model, each agent reports the quantity supplied at a given price by specifying a quantity interval of length Δ_x , the center obtains the aggregate quantity supplied in the form of an *aggregate interval* of length $I \Delta_x$, where *I* is the number of the agents. We postulate, for the sake of definiteness, that the center calculates the excess demand at each adjustment step by subtracting the number equal to the midpoint of this aggregate interval from the number equal to the (given) demand for the output commodity.

We assume that, in each case, the center performs exactly *m* adjustment steps. Since there are 2^m integer prices, this implies that the center obtains information approximating the position of the aggregate supply curve *near the equilibrium price* to the same degree of accuracy as in the centralized system. The allocation error arising from this procedure, which is performed in the approximation mode, depends on the parameters *m* and *n* and also on the sample supply functions.⁵

Let us now calculate the cost of communication of this procedure. For simplicity, we ignore the cost attached to the transmission of prices by the center to the agents; we concentrate on the cost attached to the transmission of the quantity by an agent to the center. At the *j*th step, agent *i* transmits a message designating the interval containing the value $g^i(t_j)$ of supply function g^i . It is noted that the position of the interval containing $g^i(t_j - 2^{-j}M)$ in the quantity axis is known both to the center and agent *i* at the *j*th step, and as stated by Proposition B3(ii) (Appendix), the random variable $g^i(t_j)$ given $g^i(t_j - 2^{-j}M)$ has the probability distribution with the density function $f(.,2^{-j}M)$. The amount of information that agent *i* has to transmit at the *j*th step with the accuracy corresponding to the quantity interval Δ_x is therefore equal to $\{Hf(.,2^{-j}M)\} - \log \Delta_x\}$. Since this expression is independent of *i* and there are *I* agents, we may write down the cost of communication in the price system, say C_2 , as

$$C_{2} = I \sum_{j=1}^{m} \{ (H f(., 2^{j} M)) - \log \Delta_{x} \} ,$$

which may be simplified by using Proposition B2 (Appendix) as⁶

$$C_2 = Im \left\{ (-m + n - 1) \log 2 + (2\log M - \log N) + K \right\} .$$
(5)

If, as in the preceding section, we assume that $M^2 = N$ and 2m = n, then we get

⁵ Since each agent transmits a quantity interval exactly *m* times, the maximum allocation error for this case is equal to $m \Delta_x I$. We have assumed that the center revises output prices according to the rule called *bisectioning* or *dichotomization*. This rule does not minimize the expected number of adjustment steps needed to reach equilibrium unless the equilibrium price *is* distributed uniformly. Our rule has been chosen here for the sake of simplicity.

⁶ $C_2 = I \Sigma \{2 \log (2^{-j}M) + K - \log (2^{-n}N)\}$

 $⁼ I\{2 (\log 2) \Sigma (-j) + m(2 \log M - \log N + n \log 2) + K\}$

 $⁼ I\{1/2[-2(\log 2)m(m+1)] + m(2\log M - \log N + n\log 2) + K\}$

 $⁼ Im\{(-m+n-1)\log 2 + (2\log M - \log N) + K$

$$C_2 = Im \{ (m-1) \log 2 + K \} .$$
(6)

2.4 Summary

To summarize, we state that

$$C_1/C_2 = O(2^m/m^2). (7)$$

In other words, the cost of communication in the centralized system grows linearly in the number of approximation intervals and the cost in the price mechanism is proportional to the square of the logarithm of the number.

It is noted that the results obtained above basically reflect the difference in the number of messages transmitted for adjustments in the two systems. As seen easily, in the centralized system, the agent needs to send to the center 2^m messages to report the quantity supplied at prices $t=r \Delta_t$ $(r=1,...,2^m)$. On the other hand, in the market mechanism, the agent needs to send only *m* messages, since the center revises tentative prices exactly *m* times.⁷ Our estimation of the relative cost of communication in the two systems [i.e., equation (7)] could be obtained by counting the number of messages transmitted in each system (i.e., without relying on information theory). This is because we have used the supply function as the objective of communication both for the centralized system and for the market mechanism. If we formulated a model of centralization in which the production function, not the supply function, were transmitted, then we would not be able to compare communication costs merely by counting the number of messages.

3 Conclusion

From what is stated at the end of the preceding section, the reader may observe that the main reason that, in our model, the once-and-for-all centralization of information is shown to be less efficient than the price mechanism is that the former "wastes" a great deal of information. To solve the allocation problem, some, but *not all*, of the information about the environment of each agent is needed. Which portion of the information about the environment of an agent is relevant to solve the problem depends

$$C'_1 = \sum^{m} \{H(f(\cdot, 2^{-j}M) - \log \Delta x) \cdot 2^{j}\}$$

 $^{^{7}}$ In the text, the market mechanism is formulated in such a way that at each step of adjustments the agent sends a message designating the quantity supplied at a price selected according to the bisectioning rule, whereas in the centralized system no bisectioning rule is used; prices are scanned *linearly* from left to right. If in the centralized system the agent sends messages by selecting prices according to the bisectioning rule, then the cost of communication will be equal to

on the environment of every other agent; it cannot be determined before the problem is actually solved. Thus, if the information about the environment has to be sent to the center before any computation begins there, each agent must send *all* information about the environment (to the required degree of accuracy); for, otherwise, the center might not be able to compute the optimum because of a lack of relevant information. The advantage of the price mechanism over the centralization lies in the fact that at each step of successive adjustments it allows the agent to send only that information relevant to solve the problem at that step, thus making it possible to avoid sending information that will never be used.

The conclusion stated above has been obtained from our model, which is constructed on a number of simplifying assumptions. Whether it holds in a more general setting is, of course, an open question. If, however, we limit our consideration to the case in which the centralized system "simulates" a decentralized algorithm (as in the Lange—Lerner socialist state performing a marketlike adjustment procedure within a computer), then we can assert that the centralization of information is inferior to the decentralized adjustment with respect to the cost of communication for the reason that the former wastes information the latter does not.

This advantage of *decentralizing* information has long been recognized by Lange and Taylor (1938), Hayek (1945), and others in relation to the classical controversy on socialist planning and more recently by Hurwicz (1960, 1971, 1973, 1977), Marschak and Radner (1972), Marschak (1959), Mount and Reiter (1974), Reiter (1977), and others in relation to the problem of designing economic mechanisms satisfying certain performance requirements. The results presented in Section 2 may be regarded as a quantitative confirmation, in terms of the entropy measure of information, of the advantage of decentralizing information for a simple problem of resource allocation.⁸ A question that arises naturally from the foregoing discussion is whether it is possible to explain the existence of centralized information in the real world. A large amount of information is collected and

compiled by economic (and other) organizations (e.g., by modern corporations and government institutions) to facilitate their operations. Centralized information in such organizations is stored partly in human beings and partly in various informational devices. One may think of libraries, dictionaries, and data bases as examples of nonhuman means to store centralized information. It is a fact that usually only a fraction of such centralized information is used; the remaining part will never be used and will thus be wasted. There must be factors that make the centralization of information so economical that the disadvantage of wasting information may be offset.

Arrow (1974a, b) has pointed out that one of the major advantages of centralizing information lies in the economies of scale in "handling" information. For example, a particular piece of printed information, once produced, may be copied with a negligible cost and may be distributed widely. It is

⁸ We point out that this essay may be regarded as a formulation of what Simon (1978) called bounded rationality in resource allocation. In particular, decision making in the approximation mode introduced in this contribution is a way to express the fact that in reality optimization cannot be carried out with perfect accuracy because of the presence of adjustment costs. Further, our work to compare two economic systems with respect to the cost of communication may be considered as a step toward formulating Simon's procedural rationality.

well known that this property of information likely prevents it from being traded as an ordinary commodity since it is difficult to protect the ownership of information, but it seems less well known that the same property may explain the existence of systems in which information *is* centralized. To summarize, we state that centralization of information brings about both advantages and disadvantages. The former comes from scale economies of information handling and the latter from the waste of centralized information. Thus, an *optimal* degree of centralization (or decentralization) may be obtained by weighing these two factors. Whether it is possible to formulate a model explaining this point is an open question.

Appendix

Α

Theorems from information theory

Here we assemble from information theory definitions and theorems used in the previous sections. We first define the set of messages (possible objects) and the set of letters (symbols). Let $S = \{S_I, S_{2,...}\}$ be a countable set, of elements S_I (i = 1, 2,...), denoting the collection of messages to be transmitted. It is assumed that the message S_I occurs with probability p_1 , where $p_1 \ge 0$, $\sum p_1 = 1$, and $P = (p_1, p_2,...)$. Let $A = \{a_{1L},..., a_L\}$ be a finite set, denoting the collection of letters (*alphabet*) to be used for coding messages, where the positive integer L denotes the number of letters. Given a positive integer K > 0, let $b_1 b_{2k}...b_k = (b_1,..., b_k)$ be a sequence of K elements of A, where $b_k \in A$ (k = 1,..., K). Let A^+ be the set of all finite sequences of elements of A:

$$A^+ = \{b_1, ..., b_k : K \in N, b_k \in A(k = 1, ..., K)\}$$
,

where *N* is the set of positive integers.

Definition Al. A *coding h* of the messages *S* by alphabet *A* is a one-to-one function *h*: $S \rightarrow A^+$ from *S* into A^+ :

$$S \rightarrow h(s) = b_1 ... b_k.$$

The value h(s) is the *code* of *s*, and the *length* |h(s)| of code $h(s)=b_1...b_k$ is the number *K* of letters composing the code. The *average length* |h| of coding *h* of *S* is given by the weighted sum of all codes:

$$|h| = \Sigma P_i / h(s_i) /.$$

Definition A2. The entropy of a probability distribution $P = (p_1, p_2,...)$ is

$$H(P) = -\Sigma P_i \log P_i$$

Theorem A3 (Shannon 1948; Huffman 1952).

 $H(P) \ge \inf |h| < H(P) + 1,$

where the infimum is taken over all codings and the base of the logarithm in the entropy function is equal to *L*, the number of letters available for coding.

Given a positive integer J>0, let $T_J = S \times S \times ... \times S$ (*J* times) be the *J*-fold Cartesian product of *S*. To each element $t = (t_1, ..., t_J)$ T_I , $t_I \in S$ (j = 1, ..., J), assign the product probability $q = q(t) = q_1 ... q_J$ so that $q_J = p_i$ if and only if $t_i = s_i$ (j = 1, ..., J). Let $h_J : T_J \rightarrow A^+$ be a coding of the messages T_J by *A*.

Theorem A4 (Shannon 1948).

$$inf | h_J | \rightarrow H(P)$$
 as $J \rightarrow +\infty$,

where the infimurn is taken over all codings of T_J by A, and the base of

the logarithm in the entropy function is equal to *L*.

Remark A5: The two theorems stated above suggest that the expected number of letters needed to send a message arising from the set S according to a given probability distribution *P* is approximated by the entropy H(P) of *P*. If a message is to be coded each time it arises, the average length of a code is equal to a number between H(P) and H(P) + 1. If one can code a sequence of messages, then the average length of a code per message can be made arbitrarily close to H(P). (Note, however, that the cost and the delay of coding or decoding is not considered in this essay.)

Messages are always transmitted through a channel using some coding. If the channel uses coding from a finite set *A* (e.g., binary coding using bits 0 and 1; in this case L = 2), H(P) measures the average length of letters used to transmit a message arising from S. Note that the entropy is invariant in the number of letters *L* up to multiplication by a positive constant.

The *capacity* of a channel is defined by the number of letters that it can transmit per unit of time. Therefore, if the capacity is denoted by *C*, then the time needed to transmit information in the amount H(P) is approximately equal to H(P)/C. Since the actual means of transmitting infortion is determined by the choice of a channel, we may conclude that the entropy is a measure of communication cost independent of its means.

Furthermore, it is known that what is stated above holds for wider classes of channels: channels using a finite set of letters but with erroneous transmission (noisy channels), channels using "Continuous

means" like electric or radio waves, and so on. For example, a noisy channel transmits, say, bits with possible errors: 0 as 1 or 1 as 0. If the probability law governing errors in transmission is known, then it is possible to define the capacity from that probability law *only*, so that the relation among the capacity, the amount of information, and the transmission time is exactly the same as in noiseless channels. Hence, we may state that the entropy is a measure of communication costs for a very wide class of channels.

Definition A6. Let $f: R \to R_+$ be a probability density function defined on the reals R, where R_+ denotes the nonnegative reals. The *entropy* H(f) of f is $H(f) = -\int_{\mathbb{R}} f(s) \log f(x) \, dx.$

Proposition A7. Let $f: R \to R_+$ be a probability density function and let a positive number $\Delta > 0$ be given. Let

$$P_{i} = \int_{i\Delta} f(x)dx, I = ..., -1, 0, 1, 2, ...,$$
$$P = (..., p_{-1}, p_{0}, p_{1}, p_{2}...)$$

Then

$$H(P) + \log \Delta \rightarrow H(f)$$
, as $\Delta \rightarrow 0$.

[That is, H(P) may be approximated by $H(f) - \log \Delta$ if Δ is small. Note that H(f) is independent of Δ ; *P* is called a *quantization of f with the interval length* Δ .]

Proof: Let x_i satisfy $i \Delta \leq x_i \leq (i+1) \Delta$ and $p_i = f(x_i) \cdot \Delta$. (We know from the mean value theorem that such x_i always exists.) We have

$$H(P) = -\Sigma f(x_i) \cdot \varDelta \cdot \log [f(x_i) \cdot \varDelta]$$

$$= -\Sigma [\log f(x_i) + \log \varDelta] \cdot \int_{i\varDelta} f(x) dx$$

$$= -\int_{-\infty}^{\infty} \log [f(v(x))] f(x) dx - \log \varDelta$$

where $v(x) = x_i$ if and only if $i \Delta \leq x < (i+1) \Delta$. Since $v(x) \to x$ as $\Delta \to 0$, we have $H(P) \to H(f) - \log \Delta$ as $\Delta \to 0$.

B The one-sided stable distribution and the first-passage time distribution of Brownian motions

Given a positive number t > 0, let $f: R \rightarrow R+$ be defined by

$$f(x,t) = \begin{cases} (2\pi)^{-1/2} tx^{-3/2} \exp(-t^2/2x), & \text{if } x \leq 0. \\ 0, & \text{if } x \leq 0. \end{cases}$$

The function $f(\cdot, t)$ is the density of the *one-sided stable probability distribution* with parameter t.

Proposition B1 (Feller 1971, pp. 52, 173-5).

(i)
$$\int_{0}^{y} f(x, t) f(y - x, s) dx = f(y, t + s),$$

where t > 0, s > 0, and y > 0. That is, if X_1 and X_2 are independent random variables with density functions $f(\cdot, t)$ and $f(\cdot, s)$, respectively, then $Y = X_1 + X_2$ has the density function $f(\cdot, t + s)$.

(ii) Let X_i (i = 1, ..., I) be independent random variables with the common density function f(•, t), where t >0. Let I >0 be a positive integer. Then the random variable $Y = \sum_{i=1}^{I} X_i$ has the density function f(•, $I^{1/2}t$).

(iii) Let *X* be a random variable with the density function $f(\cdot, t)$, where t > 0. Let I > 0 be a positive number. Then the random variable Y = IX has the density function $f(\cdot, I^{1/2}t)$.

Proposition B2

 $H(f(\cdot, t)) = 2 \log t + K,$

where

 $K = -\log 2 + 2[\log \pi + 3\Gamma' (1/2)\pi^{-1/2} + 1] = 3.3245,$

and Γ' (1/2) is the derivative of the gamma function $\Gamma(\cdot)$ at 1/2.

Proof: Let
$$f(x, t) = ax^{-3/2}e^{-b/x}$$
, where $a = t/\sqrt{2}\pi$ and $b = 1/2t^2$.

$$H(f) = \int_{0}^{\infty} f(x,t) \log f(x,t) dx$$

$$= -\log a + a \int_{0}^{\infty} (\{3/2 \log x + bx - 1\}x - 3/2e^{-b/x})dx$$

$$= -\log a + \frac{3a}{2\sqrt{b}} \int_{0}^{\infty} (\log b \cdot y^{-1/2} \cdot e^{-y} - \log y \cdot y^{-1/2} \cdot e^{-y})dy$$

$$+ ab \int_{0}^{\infty} b^{-3/2}y^{1/2}e^{-y} dy (\operatorname{put} y = b/x)$$

$$= -\log a + \frac{3a}{2\sqrt{b}} [\log b \cdot \Gamma(1/2) - \pi \Gamma'(1/2)] + abb^{-3/2}\Gamma(3/2)$$

$$= -\log t + \log\sqrt{2}\pi + 2\sqrt{\pi} (2\log t) - \left[\log 2\sqrt{\pi} - \Gamma'(1/2) + 1/2\right]$$

$$= 2\log t + K,$$

where $\Gamma(\cdot)$ is the gamma function, and

 $K = 1/2 \ (1-2 \log 2 + \log \pi - \sqrt{\pi} \Gamma' 1/2) = 3.32448,$ Using $\Gamma' (1/2) = 3.480226$, which is computed numerically.

Let $T = X = R_+ = (0, +\infty)$ be the set of positive reals. Let *G* be the set of left continuous nondecreasing functions *g*: $T \rightarrow + X$ from *T* into *X* such that g(0) = 0. Let *F* (*x*, *t*) be the (cumulative) distribution function of the one-sided stable distribution with the density function f(x, t).

Proposition B3 (Karlin 1966, pp. 276-80). There exists a probability distribution over *G* satisfying the following properties:

- (i) For any t > 0, Prob $\{g(t) \le y = F(y, t), y > 0$.
- (ii) For any t > s > 0, the random variable g(t) g(s) is independent of g(r)($0 < r \leq s$). Furthermore,

Prob $\{g(t) - g(s) \leq y/g(s)\} = F(y, t-s).$

Remark: The distribution over *G* introduced in the preceding proposition is called the *first'*-*passage time distribution* (the *FPT distribution*) of Brownian motions (or the inverse-Gaussian process), since the random variable g(t) is the time that the standard Brownian motion passes the given point t > 0 for the first time. It is a stochastic process whose sample function is monotone and has finite-dimensional density functions in a closed form. In this chapter, we have exploited the *fact* that the entropy function of this distribution has a simple form, as shown in Proposition B3, as well as properties similar to those of Brownian motions, as shown in Proposition B1.

REFERENCES

- Arrow, K. J. (1974a), "Limited knowledge and economic analysis," *American Economic Review*, 64(1): 1—10.
- Arrow, K. J. (1974b), The limits of organization, New York: W. W. Norton.
- Arrow, K. J. and L. Hurwicz (1977), *Studies in resource allocation processes*, Cambridge: Cambridge University Press.
- Feller, W. (1971), *An introduction to probability theory and its applications*, Vol. 2, 2nd ed., New York: Wiley.
- Hayek, F. A. von (1945), "The use of knowledge in society," *American Economic Review*, 35(4): 519–30.
- Hess, J. D. (1983), The economics of organization, Amsterdam: North-Holland.
- Huffman, D. A. (1952), "A method for the construction of minimum redundancy codes," *Proc. IRE*, 40(10): 1098—1101.
- Hurwicz, L. (1960), "Optimality and informational efficiency in resource allocation processes," In Arrow, Karlin, and Suppes (Eds.), *Mathematical methods in social sciences*, Stanford: Stanford University Press, pp. 27—46.
- Hurwicz, L. (1971), "Centralization and decentralization in economic processes," in A. Eckstein (Ed.), *Comparison of economic systems: Theoretical and methodological approaches*. Berkeley: University of California Press.
- Hurwicz, L. (1973), "The design of resource allocation mechanisms," *American Economic Review: Papers and Proceedings*, 58(2): 1-30. Reprinted in Arrow and Hurwicz (1977), pp. 3—37.
- Hurwicz, L. (1977), "On the dimensional requirements of Pareto-satisfactory processes," in K. I. Arrow and L. Hurwicz, *Studies in resource allocation processes*, Cambridge: Cambridge University Press, pp. 413—24.
- Johnson, N. L. and S. Katz (1970), *Continuous univariate distributions—1: Distributions in statistics*, New York: Wiley.
- Karlin, S. (1966), A first course in stochasiic processes, New York: Academic Press.
- Lange, 0. and F. M. Taylor (1938), *On the economic theory of socialism*, University of Minnesota Press. Reprinted by McGraw-Hill, 1964.
- Marsehak, J. (1968), "Economics of inquiring, communicating, deciding," American Economic Review: Papers and Proceedings, 58(2): 1—18.
- Marschak, J. and R. Radner (1972), Economic theory of teams, New Haven: Yale University Press.
- Marschak, T. (1959), "Centralization and decentralization in economic organizations," *Econometrica*, 27, 399–430.
- Mount, K. R. and S. Reiter, "The informational size of message spaces," *Journal of Economic Theory*, 8:161-92.

- Oniki, H. (1974), "The cost of communication in economic organization," *Quarterly Journal of Economics*, 88(4): 529-50
- Reiter, S. (1977), "The (new)² welfare economics," *American Economic Review: Papers and Proceedings*, 67(2): 226—34.
- Shannon, C. E. (1948), "The mathematical theory of communication," *Bell System Technical Journal*, 27: 379-423, 623—56; reprinted in C. E. Shannon and W. Weaver, *The mathematical theory of communication*, University of Illinois Press, 1949.
- Simon, H. A. (1978), "Rationality as process and as product of thought," *American Economic Review*, 68(2): 1—16.