INFORMATION STORAGE AND ACCESS IN DECISIONMAKING ORGANIZATIONS

by

GHASSAN J. BEJJANI

Ingénieur des Arts et Manufactures

(Ecole Centrale de Paris, 1983)

SUBMITTED TO THE DEPARTMENT OF
CIVIL ENGINEERING
IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE
IN
TECHNOLOGY AND POLICY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

January 1985

Signature of Author

Department of Civil Engineering
January 25, 1985

Certified by

Alexander H. Levis
Thesis Supervisor

Accepted by

Richard de Neufville
Chairman, Technology and Policy Program

Accepted by

Francois M. M. Morel
Chairman, Department Committee on Graduate Students
Civil Engineering Department

INFORMATION STORAGE AND ACCESS IN DECISIONMAKING ORGANIZATIONS

BY

GHASSAN J. BEJJANI

Submitted to the Department of Civil Engineering

on January 29, 1985

in partial fulfillment of the requirements for the degree of

Master of Science in Technology and Policy

ABSTRACT

Information storage and access in decisionmaking organizations is modeled using a Petri Net representation. A centralized and a decentralized database usage configuration are analyzed and their impacts on the decisionmakers' workload assessed. Organizational protocols are defined and their criteria of acceptability presented. Protocols' key variables, minimum allowable input interarrival time and response time, are determined for two organizational structures: parallel and hierarchical. A numerical example illustrates the theoretical results in the case of two decisionmaking organizations. It suggests the quantitative use of timeliness as a third organizational attribute -- the first two being workload and peformance. It also demonstrates the importance of updating coordination in evaluating the organization's performance.

Thesis Supervisor: Dr. Alexander H. Levis

Title: Senior Research Scientist

ACKNOWLEDGEMENT

Writing this thesis was a very enriching experience for me. I am deeply grateful to Dr. Alexander H. Levis for that, and for fulfilling his promise to make a researcher out of me.

Lisa M. Babine proves once again her typing skills, both in quality and speed, that were an essential factor in the timely completion of this work.

I wish also to thank Art Giordani for matching Ms. Babine's pace with his beautiful graphs and figures.

Last but not least, I would like to thank my friends Tania, Katia, Kathryn, Anne-Claire, Roland, Rafi, Bechara, Nasri, Joseph, Philippe, Philip, Francois, Jean-Jacques, and Nabil for their support and understanding, although it was at times provided from miles and miles away...

This research was conducted at the MIT Laboratory for Information and Decision Systems with support provided in part by the Office of Naval Research under Contracts N00014-83-K-0185 (NR 274-349) and N00014-84-K-0519 and in part by the U.S. Army Research Institute for the Behavioral and Social Sciences under Contract No. MDA903-83-C-0196.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

To my parents
To my sister and brothers
To the rest of my family

To Darcy and Tony

CHAPTER I

INTRODUCTION

1.1 INTRODUCTION

During the past decade, Information Theory has been increasingly adapted to the analysis and evaluation of organizations. First developed by Shannon (Shannon and Weaver, 1949), Information Theory matured into a mathematical theory in its own right, and was applied to the study of various communications systems (Gallager, 1968). It was then used as a basic tool for modeling human decisionmaking (see Sheridan and Ferrel, 1974). The apparition of the Partition Law of Information (Conant, 1976) provided a physical interpretation of the mathematical expressions derived by using the n-dimensional version of the Theory.

The basic information theoretic model of the decisionmaker was introduced by Boettcher (1981). Quantitative means for measuring human decisionmakers' workload and organizations' performance were designed, under speficic assumptions. Subsequent research effort was oriented towards relaxing some of those assumptions, and resolving more complex issues related to a realistic use of the decisionmaking model.

Significant work has been conducted regarding a more flexible interpretation of a primary working assumption that the model is memoryless. This thesis follows from the awareness that decisionmakers do possess and use memory, and from the fact that information storage and acccess devices are actually put to service in most modern organizations. It approaches the problem of databases in acyclical organizations along two directions: the information theoretic aspects of the use of those devices are addressed by computing modified activity terms for the decisionmakers; non-information theoretic notions appear as well in a closer consideration of time in the normal functioning of the organization. The two directions of analysis are presented separately in two distinct theoretical developments. They are then brought together in the example of the last chapter, where numerical results are derived to illustrate the theory. In the coming sections, a more complete description of the problems tackled and the results obtained will be made.

1.2 PRESENTATION OF THE PROBLEMS

The first part of the thesis is concerned with the definition of the problems to be addressed and the set-up of the analytical framework within which all the issues are to be approached. The broader problem was the initiation of a systematic understanding of the use of databases by decisionmaking organizations. It was soon realized that Information Theory, the primary analytical tool in this work, erected considerable obstacles to that endeavor. Since it did not provide any means of calculating the entropies of variables whose successive values depend on the chronological order in which they appear, it imposed a restriction on the categories of databases to be examined. Mainly, this drawback excluded considering recursive databases. The model finally adopted was that of an information storage and access device receiving data from an element external to the organization, and delivering it to the decisionmakers upon their request.

Once the logical features of the database were established, three major issues were identified: the actual modeling of the database, a tool for assessing its impacts on previous decisionmaking models, and the integration of the data storage and retrieval processes into the standard operations of the organization. The analytical framework developed as a response to the needs created by the above three issues was three-fold as well: it consisted of (a) the Petri Net representation technique, introduced by Petri (Peterson, 1981) and adopted by Tabak and Levis (1984), (b) the Information Theoretic modeling of two traditional organizational attributes: human workload and performance level (Levis, 1983) and (c) organizational protocols. This framework was used in Chapter 3 to represent adequately the database model in a Petri Net diagram, and to derive the activity rate terms of decisionmakers using database networks. In Chapter 4, the third instrument of that framework, protocols, was defined and its key variables determined for four database-organization structures. The conclusions drawn in these two chapters and the following one, Chapter 5, which illustrates numerically the theoretical results, are summarized in the next section.

1.3 SUMMARY OF THE ANALYSIS OF DATABASES

Starting from a general concept of memory, this thesis has attempted to present an information theoretic model of information access and storage in decisionmaking organizations. This model was then used to derive theoretical results under specific assumptions. These results are discussed in this section, along with the extent of their applicability.

The actual modeling of the logical features of the database used in this work was presented in Chapter 3, using the Petri Net theory to represent the access and retrieval procedures at each stage of the decisionmaking process. These same procedures were used to perform the updating of the information contained by the databases. These were presented in two alternative usage configurations: centralized, and decentralized.

The computation, in Chapter 3, of the activity rates for the n^{th} decisionmaker in the organization revealed an interesting property of databases: whether these are used in a centralized or decentralized network, the formal workload expressions they induce are almost equivalent. Minor differences exist, however, that pertain to the interrelationships between the data delivered to different decisionmakers. It was concluded that, in order to establish a theoretical differentiation between the two database configurations, one should resort to the analysis of non-information theoretic aspects of the problem. The more salient of these aspects were time-related features, as first shown in Section 3.6. Furthermore, the introduction of databases in a system appeared to enhance the sensitivity of the orderly functioning of the organization to time factors. These considerations were already of concern in the first memoryless model, but their impacts are compounded in the present study. Indeed, an efficient use of variable databases requires a strict updating schedule, that assures that all decisionmakers are provided with data having similar levels of accuracy and relevance. For these reasons, the concept of protocols was introduced in Chapter 4. Protocols were defined to be the description of the chronological order in which all elementary tasks in the organization have to be carried out, as given by the Petri Net representation, in addition to determining the processing times of all transitions. The key variables in protocols, the minimum allowable input interarrival time (IT) and the response time (RT), were determined according to a general proposition introduced in the same chapter. It essentially stated that IT was to be greater than any processing time and any amount of time spent by any token in any given place.

Protocols provided a quantitative criterion against which different organizational (parallel and hierarchical) structures or different database (centralized and decentralized) configurations could be compared. By using the proposition mentioned above, IT could be calculated in each case, based on the transition processing times present in the system. Under the assumption that all SA and RS algorithms in both organizations had the same processing time τ, it was found that organization P (parallel) allowed for a much smaller IT(τ) than organization H (hierarchical) did ($11\tau/3$ or $10\tau/3$, depending on the database configuration used). It was also proven that because of the more complex interrelationships between the different decisionmakers, organization H exhibited a greater RT (8 or 7τ) than organization P did ($19\tau/3$ or $17\tau/3$). These results are summarized in Table 4.1. Since the consideration of time characteristics led to clear-cut differences between the two organizational structures and between the two database configurations, it was suggested that a third organization evaluation criterion, timeliness, be established.

The use of that criterion in tactical configurations was illustrated in the numerical example of Chapter 5. Two specific organizations were used to provide quantitative interpretations of the theoretical results reached in the preceding chapters. It appeared that inadequate timeliness could expose a military organization to great risks in the event it is attacked by a fast-moving enemy, and that therefore the nature of the input had a definite influence on the evaluation of this organization. Workload and performance levels were actually computed, and represented on (P-W) loci to point to the tradeoffs between these two attributes. The relative advantages and drawbacks of each organization appeared to be mission-dependent, and not inherent to the organization's structure. Another interesting result was that the use of activity rates as a workload measure enhanced the tradeoff between timeliness and workload. This followed from the concept that, by increasing the organization's IT, the designer is in a sense giving the decisionmakers more time to accomplish their tasks, decreasing their workload by the same token. Conversely, given an organization, its IT is fixed, and one cannot envision admitting inputs with a smaller interarrival time than IT, at the cost of increasing the decisionmakers' workload; what actually takes place in that event is that the decisionmakers ignore part of or all the inputs, and that their workload does not vary.

Databases have been proven to induce a rise in the total activity of the system. This result was not, however, aimed at the questioning of the use of such devices; they are indispensable to today's organizations and information theoretic analysis should be rather used to compare alternative database usage networks. As far as the second traditional attribute of decisionmaking models, performance level, is concerned, the example in Chapter 5 produced yet another means of comparing database configurations and, more generally the Decision Support Systems (DSS) that manage these databases. It was shown that loss in updating coordination among various databases or within a centralized database led to considerable degradation of the average organization's performance (in that particular case, performance deterioration ranged from 29 to 68%).

In summary, a global overview of this thesis reveals that the two research directions introduced in section 1.1 have been investigated sufficiently to produce a substantial basis for a comprehensive study of organizations. Main additions to previous work have been made; they are (a) the extension of the use of Petri Nets to the representation of information storage and retrieval, (b) the computation of the decisionmakers' workload in a database equipped organization, and (c) the introduction of organizational protocols as a framework within which the actual functioning of the organization can be examined and its timeliness assessed. While many parts of this work have unveiled nearly as many issues as they have resolved, it is hoped that it will contribute to the continuous advancement of techniques for the design and analysis of organizations.

CHAPTER II

INFORMATION THEORY AND THE DECISIONMAKING MODEL

2.0 INTRODUCTION

Information theory has been used recently in the analytical modeling of decisionmaking process (Sheridan and Ferrell, 1974; Boettcher, 1981; Levis and Boettcher, 1983). It has provided a theoretical support for the development of increasingly complex models of the decisionmaker, and a quantitative means of assessing his workload and performance. In this chapter, relevant definitions and laws of Information Theory are given. The basic information theoretic model of the decisionmaker is then presented, in the version appropriate for organizations. Finally, an information theoretic interpretation of the concepts of bounded rationality and satisficing behavior is developed. In each section, the reader is referred to specific relevant previous work for more information about the topic discussed.

2.1 INFORMATION THEORY: DEFINITIONS AND LAWS

2.1.1 Entropy and Transmission

Information theory was developed by Shannon (Shannon and Weaver, 1949) for the most part, to be applied to communications systems. Since then, further developments of the theory have made it a valid mathematical theory in its own right (Gallager, 1968, Conant, 1976). It has been applied to many fields outside communications, notably in some simple models of the decisionmaker (Sheridan and Ferrell, 1974).

Two primary quantities are defined in Information Theory: entropy, and transmission. The entropy of a variable x, which is an element of an alphabet X and occurs with a probability p(x), is noted H(x), and defined as follows:

$$H(x) \equiv - \sum_{x} p(x) \log p(x) \tag{2.1}$$

It is measured by bits if the logarithm base is two. Entropy is also known as the average information or uncertainty in x, where information does not refer to the content of the variable x, but rather to the average amount by which knowledge of x reduces the uncertainty about it. H(x) is independent of the real nature of the variable x.

The other quantity of interest in information theory, transmission, is also called average mutual information: given two variables x and y, elements of the alphabets X and Y, given p(x), p(y) and p(x/y), the transmission between x and y, T(x:y), is defined to be:

$$T(x:y) \equiv H(x) - H_y(x) \equiv H(y) - H_x(y) \tag{2.2}$$

$H_x(y)$ is the conditional uncertainty in the variable x, given the full knowledge of y. It is computed as follows:

$$H_y(x) = - \sum_{y} p(y) \sum_{x} p(x|y) \log p(x|y) \tag{2.3}$$

Transmission measures the relatedness or constraint holding between two variables. It can be interpreted as the amount by which knowledge of y reduces the uncertainty in x or vice-versa, as it is a symmetric quantity in x and y. Entropy and transmission are particularly useful for measuring the uncertainty and relatedness of quantities which are not numerically defined.

McGill (1954) extended this two-variable input-output theory to N dimensions. Noticing that Eq. (2.2) can be rewritten as:

$$T(x:y) = H(x) + H(y) - H(x,y) \tag{2.4}$$

McGill's extension to N dimensions:

$$T(x_1:x_2:\ldots:x_N) = \sum_{i=1}^{N} H(x_i) - H(x_1,x_2,\ldots,x_N) \tag{2.5}$$

seems natural. N-dimensional mutual information measures the total constraint holding between all N variables of a system (between pairs, triples, etc.). A very nice feature of this measure is, as Ashby (1965) and Conant (1976) have pointed out, that it may be expressed as the sum of simpler quantities. For example, a system may be decomposed into subsystems; then, the N-dimensional mutual information of the overall system is the sum of two quantities: the transmissions of the individual subsystems, and the transmission between subsystems. With N=4, this might be expressed as follows:

$$T(x_1:x_2:x_3:x_4) = T(x_1:x_2) + T(x_3:x_4) + T(x_1,x_2:x_3,x_4) \tag{2.6}$$

Note: A very useful Information Theoretic identity is the following:

$$H(x,y) = H(x) + H_x(y) \tag{2.7}$$

It is easily generalized to the N-dimensional case:

$$H(x_1,x_2,\ldots.,x_N) = H(x_1) + H_{x_1}(x_2) + H_{x_1,x_2}(x_3) + H_{x_1,x_2,\ldots,x_{N-1}}(x_N) \tag{2.8}$$

2.1.2 Rates

Major Information Theory's competitors, like variance analysis and auto-and cross-correlation techniques, share an important deficiency in dealing with dynamic systems: they do not provide any means of taking into account constraints exercised by the system's past history on its present values. This deficiency is nonetheless somewhat overcome in Information Theory by the definition of rates:

Entropy rate:

$$\bar{H}(x) = \lim_{m \to \infty} \frac{1}{m} H(x(t),x(t+1),\ldots,x(t+m-1)) \tag{2.9}$$

It is approximately the average uncertainty of x per step. A similar definition stands for the conditional entropy rate, $\bar{H}_x(y)$, which is the uncertainty of y per step, conditional on complete knowledge of x -- past, present and future. The transmission rate can then be written:

$$\bar{T}(x:y) = \bar{H}(x) - \bar{H}_y(x) \tag{2.10}$$

$$\bar{T}(x:y) = \bar{H}(y) - \bar{H}_x(y) \tag{2.11}$$

$$\bar{T}(x:y) = \bar{H}(x) + \bar{H}(y) - \bar{H}(x,y) \tag{2.12}$$

All information theoretic identities remain valid when rates are used, provided every entropy and transmission is replaced by the corresponding rate (by overlining).

Rates will be used in this thesis, because the general concept of memory involves dependence between successive iterations of the system, and the possibility to use this dependence as a first step towards learning (Chapter 3).

2.1.3 Partition Law of Information

In his work on the application of Information Theory to the analysis of information processing systems, Conant defined the total information theoretic activity of a system to be the sum of the entropies of all the individual variables in the system, without considering constraints between variables. He then decomposed this total activity into quantities that can be interpreted to correspond to what actually takes place in such a system. This decomposition is called the Partition Law of Information (PLI) (Conant, 1976). It is defined for a system with N-1 internal variables, w_1 through w_{N-1}, an input variable, x, and an output variable, y, also called w_N. It states:

$$\sum_{i=1}^{N} H(w_i) = T(x:y) + T_y(x:w_1,w_2,\ldots w_{N-1}) + T(w_1:w_2:\ldots:w_{N-1}:y)$$

$$+ H_x(w_1,w_2,\ldots,w_{N-1},y) \qquad (2.13)$$

As noted above, the left-hand side of Eq. (2.13) refers to the total activity of the system, also designated G. Each of the quantities on the right-hand side has its own interpretation, as discussed below.

The first term, T(x:y), is called the throughput of the system and designated G_t. It measures the amount by which the output of the system is related to its input. The second term:

$$T_y(x:w_1,\ldots,w_{N-1}) = T(x:w_1,\ldots,w_{N-1},y) - T(x:y) \qquad (2.14)$$

is called the blockage of the system and designated G_b. As the above expansion demonstrates, blockage may be thought of as the amount of information in the input that is not included in the output. The third term, $T(w_1:w_2:\ldots:w_{N-1}:y)$, is called the coordination of the system and designated G_c. It is just the amount by which all of the internal variables in the system constrain each other. It reflects all system variables' interactions and can be interpreted as the coordination required among the system variables to accomplish the processing of the input to obtain the output. Usually, it is more helpful to break the system into subsystems and to compute the coordination activity as shown in Eq. (2.6). The fourth term, $H_x(w_1,w_2,\ldots,w_{N-1},y)$, is called the noise of the system, G_n. It represents the uncertainty that remains in the system variables when the input is completely known. It may be also thought of as internally generated information, i.e., information supplied by the system to supplement the input and facilitate the decisionmaker's task. Finally, the Partition Law of Information may be abbreviated as:

$$G = G_t + G_b + G_n + G_c \tag{2.15}$$

Notes:

1. By replacing entropies and transmission in the PLI by the corresponding rates, one gets another identity, called the Partition Law of Information Rates (PLIR). It states:

$$\sum_{i=1}^{N} \bar{H}(w_i) = \bar{T}(x:y) + \bar{T}_y(x:w_1,w_2,\ldots,w_{N-1}) + \bar{T}(w_1:w_2:\ldots:w_{N-1}:y) + \bar{H}_x(w_1,w_2,\ldots,w_{N-1},y) \tag{2.16}$$

or again:

$$F = F_t + F_b + F_n + F_c \tag{2.17}$$

The PLIR will be used in the next two chapters to derive the activity rate expressions for the decisionmaking models considered.

2. Throughput and blockage rates represent two possible dispositions of the input to a particular system. A third possibility is that the input may not even cross the boundary into the system. Such a phenomenon is termed rejection, denoted F_r, and can be expressed as:

$$F_r = \bar{H}_{w,y}(x) \tag{2.18}$$

F_r is a passive form of blockage, that does not appear in the PLIR, but in the auxiliary expression:

$$\bar{H}(x) = F_t + F_b + F_r \tag{2.19}$$

2.2 THE DECISIONMAKING MODEL

Decisionmaking systems do more than transmit information from input to output. Internal decisionmaking takes place, in a way that has been modeled by many researchers (Sheridan and Ferrell, 1974; Simon and March, 1958; Simon, 1947; March, 1978). March and Simon have hypothesized that the decisionmaking process of the satisficing decisionmaker is a two-stage process of "discovery and selection". This concept has been applied to an information theoretic model of the decisionmaker by Boettcher and Levis (Boettcher, 1981; Boettcher and Levis, 1982, 1983; Levis and Boettcher, 1983). In this section, a Petri Net representation is used as first developed by Petri (Peterson, 1981) and adapted by Tabak and Levis (1984).

Figure 2.1 shows the two-stage model of the nth member of an organization. His input x^n is a component of a single vector source distributed by a set of partitioning matrices among all the decisionmakers (Stabile and Levis, 1984). The decisionmaker processes this input in the situation assessment (SA^n) stage to determine or select a particular value of the variable z^n that denotes the situation. He may communicate his assessment of the situation to other members (z^{no}) and he may receive their assessments in return (z^{on}). This supplementary information may be used to modify his assessment, i.e., it may lead to a different value of z^n. Possible alternatives of action are evaluated in the response selection (RS^n) stage. The outcome of this process is the selection of a local action or decision y^n that may be communicated to other team members or may form all or part of the organization's response. A command input from other decisionmakers, v^{on}, may affect the selection process. (It should be noted here that y^{no} corresponds to v^{no} and y^{on} is nothing else than v^{on}).

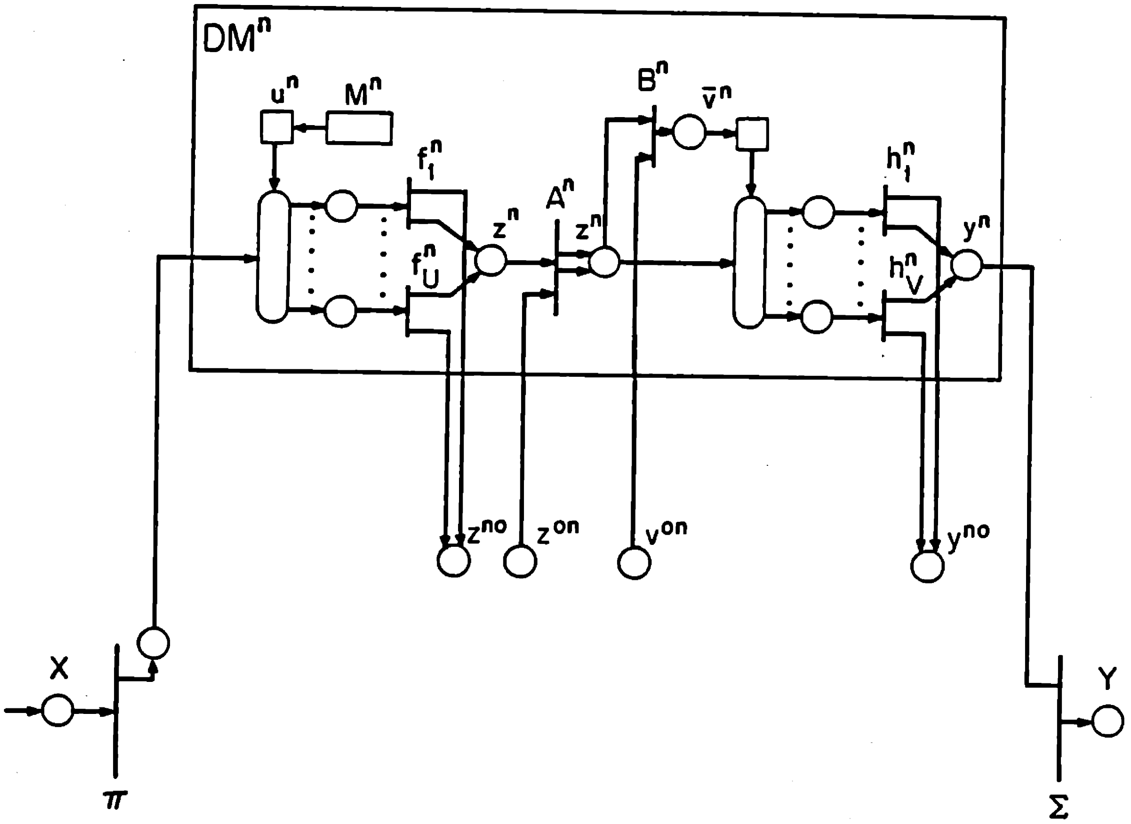


Figure 2.1 Petri Net Representation of the n^{th} Decisionmaker of the Organization

The situation assessment stage consists of U algorithms (f_i^n,i=1,...,U). The value taken by the variable u determines the number of the algorithm to be used, and is chosen according to the probability p(u). Similarly, the choice of an algorithm in the RS stage is determined by the variable v, with the probability p(v).

A fundamental assumption that is made throughout the study of this decisionmaking model is that it is memoryless. This means that no learning is allowed to occur between two iterations. All the algorithms are initialized when used. The decisionmakers are assumed to be well-trained, performing their tasks in a steady-state configuration.

The PLI has been applied to the study of the decisionmaker's activity in the model above, by Boettcher and Levis (Boettcher, 1981; Levis and Boettcher, 1983). The different activity expressions outlined in the previous section have been derived, and will be presented later in this thesis during the study of a database-equipped organization (section 3.5).

At this point, it is useful to summarize the main assumptions under which the results of the studies so far have been derived:

- the sets of algorithm variables are mutually disjoint.

- the algorithms are deterministic (the stochastic case has been resolved in G. Chyen's work (Chyen, 1984)).

- the algorithms have no rejection.

- the situation assessment strategy is satistically independent of the input ($p(u^n|x^n) = p(u^n)$). The case of $p(u|x)$ has been analyzed by Chyen (1984).

- the response selection strategy depends on $\bar{z}^n$ according to the law $p(v^n|\bar{z}^n)$. The response selection can be influenced by the rest of the organization, through v^{on}, after v^n has been determined.

Under these assumptions, if $p(x)$, $f_i(x)$, $h_j(z)$, and the internal coordination terms of each algorithm are known, then the activity is a function only of the internal decision strategy: $(p(u), p(v|z))$.

2.3 CONSTRAINTS AND EVALUATION CRITERIA

2.3.1 Bounded Rationality

One of the basic premises of the information theoretic analysis of human decisionmakers is that they exhibit bounded rationality. This concept, first developed in detail by Simon and March (Simon and March, 1958), basically states that human decisionmakers have a limited capacity as information processors and problem solvers. The notion of bounded rationality is modeled here as a constraint on the total decisionmaker's workload, or activity. This activity threshold has also been introduced by empirical work on the human reaction time to external stimuli (Sheridan and Ferrell, 1974).

The quantitative interpretation of the bounded rationality principle is used under the following form, for decisionmaker n:

$$G^n = G_t^n + G_b^n + G_n^n + G_c^n \leq F_o^n \tau \quad , \quad \forall n \tag{2.20}$$

or, using activity rates:

$$F^n = F_t^n + F_b^n + F_c^n + F_n^n \leq F_o^n \tag{2.21}$$

where τ is the mean input interarrival time and F_o^n is the maximum rate of information processing that characterizes decisionmaker n. This constraint implies that the decisionmaker must process his inputs at a rate that is at least equal to the rate with which they arrive.

2.3.2 Performance Evaluation

Since the overall task of the organization is to process inputs so adequate responses are selected , it is safe to assume that all decisions made are not perfect. A logical way of evaluating the organization's performance is then to compare the actual response y, to the desired response, y'. A mapping of the inputs into the desired responses, $y' = L(x)$ is originally established by the organization designer. A performance measure of the organization is then the cost function associated to having y as an output instead of y', given x. This cost function is usually closely related to the probability $p(y \neq y')$, using weights to account for particularly costly errors (for instance, it could be more costly to shoot down a friendly plane than to let an enemy go away). The performance J is then a function of the decision strategy: $J(p(u), p(v|z))$, that is minimized by the designer according to one of two goals:

- optimal organizational design: minimize J. Theories of Rational Behavior assume such a design to be feasible.

- satisficing organizational design: verify:

$$J \leq \bar{J} \tag{2.22}$$

One corollary of the principle of bounded rationality is that decisionmakers exhibit a satisficing, and not optimizing, behavior (Simon and March 1958; Simon, 1979).

Another very important performance measure of a system introduces the relevance of time. Indeed, delays are very costly in strategic configurations; obtaining the optimal decision too late is worse than obtaining a suboptimal decision on time. This notion is discussed in more detail in Chapter 4, after the introduction in Chapter 3 of databases and of their impacts on the model presented in the previous section.

CHAPTER III

MODELS OF INFORMATION STORAGE AND ACCESS

3.1 THE INTRODUCTION OF THE CONCEPT OF MEMORY

In the previous chapter, the basic decisionmaking model was presented as being memoryless. Decisionmakers were assumed to perform in a steady-state configuration, carrying out a task for which they were well trained. However, the reality is that decisionmakers, whether they are humans or machines, do have the capability of learning: humans can learn from their mistakes and be continuously trained, machines can be programmed, to adapt to new situations or to perform better iterative tasks. Furthermore, the results of learning must be remembered for experience to build-up and improvements to take place. In the human mind, this kind of remembering is assured by what is called the long-term memory (Bailey, 1982).

On the other hand, memory is an indispensable tool for today's organizations. They have to process increasingly complex and burdensome amounts of information, and their performance is seriously impeded when the necessary data storage and access devices are not available. In fact, information handling has been deemed so cumbersome and time-consuming that it is increasingly managed by computer systems. Since the first apparition of the concept of Decision Aids a little more than a decade ago, many generations of Management Information Systems have purported to help human decisionmakers, by performing the mechanical, fastidious part of their tasks. These devices have evolved from mere calculators to well-integrated Decision Support Systems (DSS) (Keen, 1981). Their designers are primarily concerned with the way they fit into the organization (Keen and Scott-Morton, 1978; Huber, 1981), or adapt to the particular need or managerial style of the decisionmaker (De Waele, 1978; Huber, 1983).

In parallel with the interest in DSS developed by psychologists and computer scientists, the concept of memory has been addressed recently from the point of view of Information Theory. Sen and Drenick (1981) obtained interesting results by allowing the outputs of their model to depend on present and past inputs. However, the actual modeling of information storage devices was initiated in Hall's work (1982). She modeled buffer storage explicitly, which provides a way of processing together strings of successive inputs when they are statistically related. This is particularly useful in tactical configurations, where it is of equal importance to monitor both changes in the environment and the actual nature of this environment. Hall also lay the foundation for the study of databases used during the internal decisionmaking process. These databases contain the main information about the history of the system and some relevant elements of its environment, or its context, that are not carried by its formal inputs. Databases can be updated regularly, or when the context changes, to keep providing the decisionmaker with up-to-date information.

The database is one of the three main parts of a Decision Support System. The other two are an information management program, and a machine-user interface (computer terminal) (Sprague, 1980; Sprague and Carlson, 1982). The primary goal of this thesis is to address the database and decisionmaker machine interface issues from an information theoretic point of view. The database's storage and access procedures, and their impact on the decisionmaker's workload and performance levels, will be described in this chapter and the next one. In the next sections, the general model and the basic assumptions it relies upon are presented; relevant remarks about the properties of the different database structures are made in the last section of the chapter.

3.2 THE GENERAL DATABASE MODEL

The database model developed in this chapter and used throughout the remainder of the thesis answers the traditional definition of an information storage device: it can receive information from an external source that it stores adequately, and it delivers this information, or part of it, whenever accessed by its users. The Petri Net model adopted here consists of two stages (see Fig. 3.1). The first stage, transition C, receives an input from the decisionmaker who requests access to the data. This input represents the situation in which the user is. Transition C determines then the nature of the information needed to cope with that situation, and sends a query to the next stage, D. Transition D performs the actual search, and delivers the data to the decisionmaker, at a predetermined stage of his internal decisionmaking process. The usage of such a database is made more explicit in the coming sections, in the specific context of the Petri Net decisionmaking model presented in the previous chapter.

A database is the memory of the system. In the model discussed here, this memory can be either fixed or variable. The data can be updated regularly, to keep up with the evolution of the environment. It is assumed that this updating is not performed by the decisionmakers themselves (i.e., this memory is not recursive), but rather by an element external to the organization considered. This technical assumption is made to respect the steady-state configuration that still is the basic premise to the use of Information Theory. Indeed, in a recursive memory case, the entropy of the data variables cannot be computed.

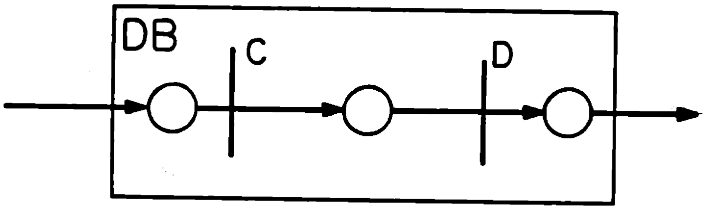


Figure 3.1 Petri Net Representation of the General Database Model

The updating process can take the form of an access to the database following the procedure described above. An updating signal, with a new set of data, is received by transition C, which in turn passes on the information, along with the necessary commands, to the data storage and access transition, D. This process will not be represented in the models used in this thesis, both because it adds on complexity to the figures and because its exact definition is out of the scope of this work. These models will be basically related to two configurations of database usage: one centralized database, or overall decentralized databases.

3.3 CENTRALIZED DATABASE

A centralized database is a database shared by all members of the organization. It is physically located in one place, and individual terminals allow the decisionmakers to access it independently. In the Petri Net representation, a centralized database is modeled as one unit, comprising several transition C - transition D sequences. There are two such databases, one for the SA stage, called DBSA, and one for the RS stage, DBRS. The inputs to transition C^n in DBSA are the input to the n^{th} decisionmaker, x^n, and the variable u^n indicating the SA algorithm he is about to use (see Fig. 3.2). Transition C^n emits then a message towards transition D_1, that carries a query for the information needed for DM^n to process x^n through the u^{th} SA algorithm. D^n in turn delivers the requested data, d^n_{SA}, to the decisionmaker, who receives it as an input to the algorithm he is using. The usage of DBRS follows a similar rationale, applied to the RS stage.

The model described above shows the main characteristics of the functioning of centralized databases as decision aids: although all the data is stored in the same memory, different decisionmakers can access different parts of it at different times. Furthermore, the data they receive depends not only on what has been stored in the database by the external updating element, but also on the decisionmaker's precise needs, based on his input and the algorithm he uses. It is interesting to note that the decision strategy in the RS stage is not influenced by the data input, and remains a function of $\bar{z}^n$ only.

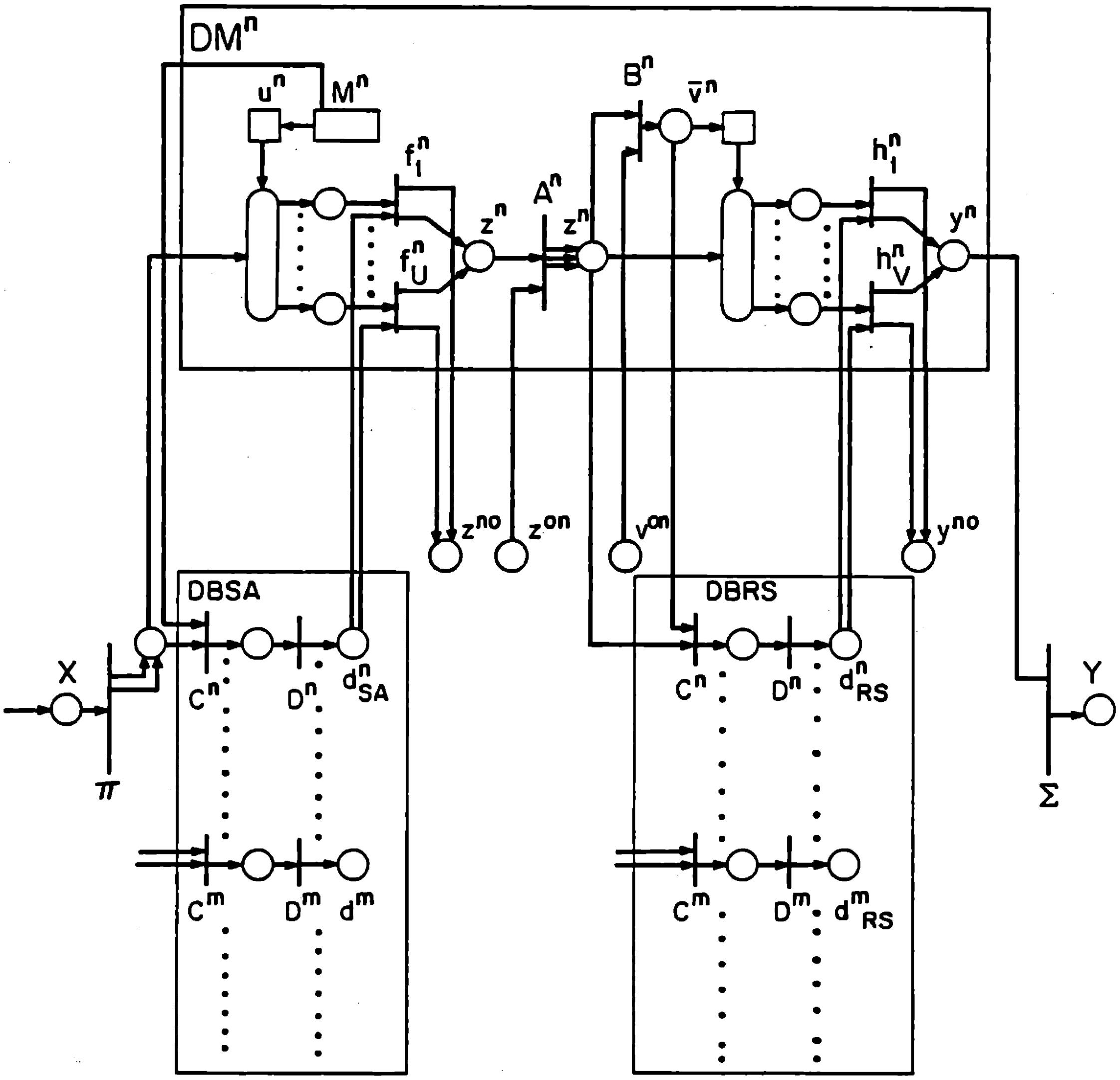


Figure 3.2 Petri Net Representation of DM^n Using Centralized Databases

The use of databases has a significant impact on the decisionmaker's workload, as can be seen in the following developments. Activity rate terms are derived by applying the PLIR to the decisionmaking model used here. For a more complete description of the calculations, the reader is referred to Appendix A. The modifications brought to the basic model presented in the previous chapter are due to the presence of two supplementary variables, d^n_{SA} and d^n_{RS}, and to their relationships with the existing structure. For simplicity's sake, the superscript n will be omitted in the following equations whenever confusion may not arise.

Throughput Rate:

$$F_t = \bar{T}(x,d_{SA},z^{on},v^{on},d_{RS}:u,z^{no},\bar{z},\bar{v},y^{no},y) \quad (3.1)$$

Blockage Rate:

$$F_b = \bar{H}(x,d_{SA},z^{on},v^{on},d_{RS}) - F_t \quad (3.2)$$

Noise Rate:

$$F_n = \bar{H}(u) + \bar{H}_{\bar{z}}(v) \quad (3.3)$$

Given the structure of the decisionmaking process modeled in this thesis, the only decisions that do not deterministically depend on the input to DM^n are the choise of an SA algorithm, and the adoption of a decision strategy for the RS stage. One must also note that the data is not internally generated information, but rather input provided by an external source. All this is adequately rendered by Eq. (3.3).

Coordination Rate:

For computational purposes, the global system is divided into the following four subsystems:

$$S^{I} = \{u, W^{1}, W^{2}, \ldots, W^{U}, z\} \tag{3.4}$$

$$S^{A} = \{W^{A}\} \tag{3.5}$$

$$S^{B} = \{W^{B}\} \tag{3.6}$$

$$S^{II} = \{\bar{v}, W^{U+1}, \ldots, W^{U+v}, y\} \tag{3.7}$$

The coordination rate can then be calculated using identity (2.6):

$$F_{c} = F_{c}^{I} + F_{c}^{A} + F_{c}^{B} + F_{c}^{II} + \bar{T}\,(S^{I}:S^{A}:S^{B}:S^{II}) \tag{3.8}$$

The developments in Appendix A yield:

$$F_{c}^{I} = \sum_{i=1}^{U} \left(\bar{g}_{c}^{i}(p(x,d_{SA})) + \frac{\alpha_{i}}{\tau_{SA}}\,H(p_{i})\right) + \bar{H}(z) \tag{3.9}$$

$$F_{c}^{A} = \bar{g}_{c}^{A}\,(p(z,z^{on})) \tag{3.10}$$

$$F_{c}^{B} = \bar{g}_{c}^{B}\,(p(\bar{z},v^{on})) \tag{3.11}$$

$$F_{x}^{II} = \sum_{j=1}^{V} \left(\bar{g}_{c}^{U+j}\,(p(\bar{z},d_{RS})) + \frac{\alpha_{U+j}}{\tau_{RS}}\,H(p_{j})\right) + \bar{H}(y) \tag{3.12}$$

and finally:

$$\bar{T}(S^{I}:S^{A}:S^{B}:S^{II}) = \bar{H}(z) + \bar{H}(\bar{z}) + \bar{H}(\bar{z},\bar{v}) + \bar{T}_{z}(x,d_{SA}:z^{on})$$

$$+ \bar{T}_{\bar{z}}(x,d_{SA},z^{on}:v^{on}) + \bar{T}_{\bar{z},\bar{v}}(x,d_{SA},z^{on},v^{on}:d_{RS}) \tag{3.13}$$

The final form of F_c is therefore:

$$F_c = \sum_{i=1}^{U} (\bar{g}_c^{i}\ (p(x,d_{SA})) + \frac{\alpha_i}{\tau_{SA}} H(p_i)) + \bar{H}(z)$$

$$+ \bar{g}_c^{A}\ (p(z,z^{on})) + \bar{g}_c^{B}\ (p(\bar{z},v^{on}))$$

$$+ \sum_{j=1}^{V} (\bar{g}_c^{U+j}\ (p(\bar{z},d_{RS})) + \frac{\alpha_{U+j}}{\tau_{RS}}\ H(p_j)) + \bar{H}(y)$$

$$+ \bar{H}(z) + \bar{H}(\bar{z}) + \bar{H}(\bar{z},\bar{v}) + \bar{T}_{z}(x,d_{SA}:z^{on}) + \bar{T}_{\bar{z}}(x,d_{SA},z^{on}:v^{on})$$

$$+ \bar{T}_{\bar{z},\bar{v}}(x,d_{SA},z^{on},v^{on}:d_{RS}) \tag{3.14}$$

<u>Definitions and comments related to the coordination rate term, F_c:</u>

$$p_i = p(u=i) \tag{3.15}$$

$$p_j = p(v=j) \qquad (3.16)$$

$$H(p) = (p\log_2 p + (1-p)\log_2 (1-p)) \qquad (3.17)$$

α_i is the number of variables of the algorithm i that are reinitialized at each iteration. The symbol τ_{SA} designates the mean interarrival time of the input to the SA stage. τ_{RS} has an equivalent meaning with respect to the RS stage. The mean input interarrival time can be used in the equations if the interarrival time is not constant, by regulating the source (Hall, 1982). The functions $\bar{g}_c^i$, $\bar{g}_c^{U+j}$, $\bar{g}_c^A$ and $\bar{g}_c^B$ are the individual coordination rate functions of the SA, A, B, and RS algorithms, and are of the following form:

$$\bar{g}_c^i = \sum_{j=1}^{\alpha_i} \bar{H}_u (w_j^i) - \bar{H}_u(W^i) \qquad (3.18)$$

The terms $\bar{H}(z)$, $\bar{H}(\bar{z})$, $\bar{H}(\bar{z},\bar{v})$ in the term $\bar{T}(S^I:S^A:S^B:S^{II})$ can be interpreted to represent the direct coordination rate between subsystems, through the fact that one subsystem's output is another's input. However, indirect coordination between the four subsystems is accounted for by the transmission rate terms. $\bar{T}_z(x,d_{SA}:z^{on})$ represents the coordination rate between S^I and S^A that is due to the relationship between x and d_{SA}, and z^{on}. Indeed, if the inputs to DM^n and those to the rest of the organization (RO) are related, or if d_{SA}^n and d_{SA}^m, $m \neq n$, are not totally independent, due to the structure of the storage or the updating process in the centralized database, then z^{on} can bring to S^A information about the inputs to the system that is not contained in z. Similar interpretations stand for the other two transmission rate terms. There is something more to say, however, about the term $\bar{T}_{\bar{z},\bar{v}}(x,z^{on},d_{SA},v^{on}:d_{RS})$: it brings the question of the relationship between d_{SA} and $_{dRS}$. Indeed, the data stored in DBSA by the external operator can be dependent on data stored in DBRS, or vice-versa, depending on the tactical configuration. An example of this mutual constraint can be the following: let DBRS contain the inventory levels for slow and fast missiles when the decisionmaker commands an anti-aircraft battery, receiving information about the speed and trajectory of the target as an input, and firing slow or fast missiles as an output. Let DBSA contain the threshold of speed according to which the decisionmaker decides whether the target is slow or fast. This threshold can be reevaluated (increased, for instance) if the relative level of fast missiles inventory varies (decreases).

3.4 DECENTRALIZED DATABASES:

Although a centralized database often is an adequate information storage device for the organization, a decentralized database network is sometimes preferred, for reasons discussed in the following section. Decentralized databases are individual storage units, accessed exclusively by one decisionmaker, and holding and delivering information relevant to this decisionmaker's task only. Such a structure is represented in Figure 3.3. One can see that the only difference with respect to Figure 3.2 is the feature: one transition C^n - transition D^n sequence per database, which models the notion of exclusivity mentioned above. Apart from that, decentralized databases function in exactly the same fashion centralized ones do.

According to the description of centralized and decentralized database models made in this chapter, both configurations can be interpreted to introduce the same kind of modifications to the basic memoryless decisionmaking model. These modifications are the apparition of the same supplementary inputs, d^n_{SA} and d^n_{RS}, and the interpretation of the same existing variables to be supplementary outputs, in both database structures. This quite naturally leads to the same mathematical

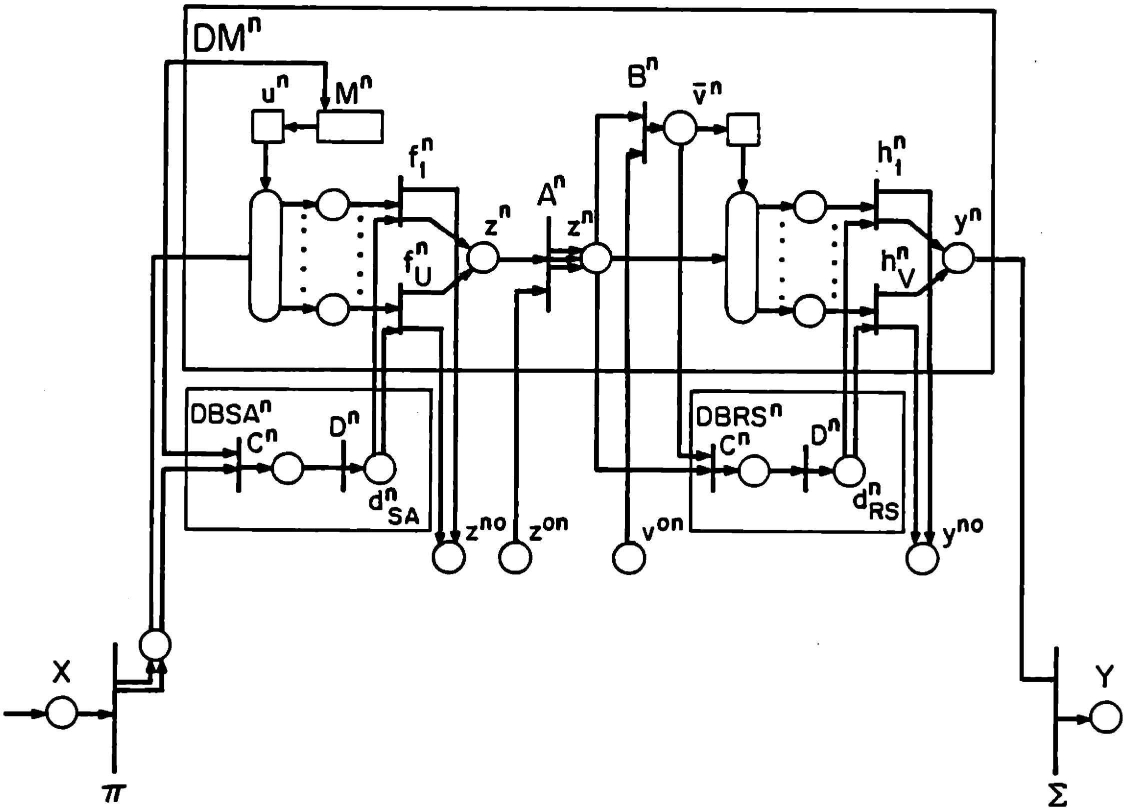


Figure 3.3 Petri Net Representation of DM^n Using Decentralized Databases

expressions for the various activity rate terms in both cases. However, the centralized database configuration leads to the following simplifications whenever the relationships between d^n_{SA} and d^n_{RS} and variables external to the system, like z^{on} and v^{on}, are concerned:

$$\bar{T}_z(x,d_{SA}:z^{on}) = \bar{T}_z(x:z^{on}) \qquad (3.19)$$

This is because the separation of the individual databases prevents z^{on} from depending on d^{n}_{SA} through a constraint between d^{n}_{SA} and d^{RO}_{SA}. Any relatedness between z^{on} and d^{n}_{SA} is comprised in, and accounted for by, the mutual constraint between x and z^{on}, which brings the simplification of Eq. (3.19). The same simplification similarly appears in another coordination term:

$$\bar{T}_{\bar{z}}(x,d_{SA},z^{on}:v^{on}) = \bar{T}_{\bar{z}}(x,z^{on}:v^{on}) \tag{3.20}$$

As a result of these simplifications, the term $\bar{T}(S^{I}:S^{A}:S^{B}:S^{II})$ is somewhat simpler for a decentralized database structure, and reads as follows:

$$\bar{T}(S^{I}:S^{A}:A^{B}:S^{II}) = \bar{H}(z) + \bar{H}(\bar{z}) + \bar{H}(\bar{z},\bar{v}) + \bar{T}_{\bar{z}}(x:z^{on}) + \bar{T}_{\bar{z}}(x,z^{on}:v^{on}) + \bar{T}_{\bar{z},\bar{v}}(x,z^{on},d_{SA},v^{on}:d_{RS}) \tag{3.21}$$

3.5 FIXED DATABASE AND THE MEMORYLESS MODEL

The results described in the previous sections were derived assuming d_{SA} and d_{RS} to be variable quantities. However, it might very well be the case that d_{SA} and d_{RS} are fixed, either because the databases are never updated or because the values taken by d_{SA} and d_{RS} remain valid during a very long time, compared to the mean input interarrival time. In this simple case, the database's direct contribution to the decisionmaker's activity rate is null, and the expressions developed above become similar to those derived in the basic memoryless decisionmaker case. They are derived by simply eliminating the variables d_{SA} and d_{RS} and the input variables to the databases from the equations, which shows that the reduction from the database-equipped model to the memoryless one is consistent:

Throughput Rate:

$$F_t = \bar{T}(x, z^{on}, v^{on} : z^{on}, y, y^{no}) \qquad (3.22)$$

Blockage Rate:

$$F_b = \bar{H}(x^n, z^{on}, v^{on}) - F_t \qquad (3.23)$$

Noise Rate:

$$F_n = \bar{H}(u) + \bar{H}_{\bar{z}}(v) \qquad (3.24)$$

Coordination Rate:

$$F_c = \sum_{i=1}^{U} (\bar{g}_c^{i} (p(x)) + \frac{\alpha_i}{\tau_{SA}} H(p_i) + \bar{H}(z, z^{no}) + \bar{g}_c^{A} + \bar{g}_c^{B}$$

$$+ \sum_{j=1}^{V} (\bar{g}_c^{U+j} (p(\bar{z})) + \frac{\alpha_{U+j}}{\tau_{RS}} H(p_j) + \bar{H}(y, y^{no}) \qquad (3.25)$$

These expressions are similar to those derived by Hall (1982). It is important, however, to note that an indirect contribution of the databases to the total activity rate is made through the numbers (α_i), (α_{U+j}). Indeed, the number of variables initialized upon the use of an algorithm accessing a database is larger than if this same algorithm does not access the database but has the data fixed in its structure (Hall, 1982).

3.6 COMPARISON OF PROPERTIES OF THE TWO DATABASE STRUCTURES

The three previous sections presented two different database structures: centralized and decentralized databases. From the point of view of Information Theory, it appeared that those two structures are almost similar. However, important differences are to be pointed out otherwise: first, the time associated with the query process, in transition D^n, is much shorter when the database is an individual one than when it is centralized. In effect, in the former case, no irrelevant information is to be scanned and then discarded, which happens in the latter case, and the system's answer to its stimuli is more timely. As noted earlier, timeliness is a vital element of the effectiveness measure of the system (Cothier, 1984) and can be incorporated in its performance evaluation.

An important advantage of a centralized database structure is that it allows for more convenient updating. It can be updated in one operation, providing all the decisionmakers with equally recent information, whereas decentralized databases require a much greater updating effort to obtain the same result. If the latter are not all updated at the same time, different decisionmakers will have simultaneously different pictures of their environment, which lowers the overall performance of the organization.

A final criterion for the comparison of the two database structures is of a tactical nature: one database is easier to protect than many but, if destroyed, paralyzes the whole organization, which cannot be the case when several databases are geographically spread out. This argument is of great importance for military organizations.

The notion of processing time introduced above brings out a distinct class of organizational problems, related to the existence of feasible protocols. The importance of these protocols as well as the interplay between the various comparative arguments presented above are show in the next chapter, where two generic organizational structures, parallel and hierarchical, are analyzed using particular models. In the following chapter, a numerical example of two such organizations is developed, and used to illustrate the properties of the model.

CHAPTER IV

PROTOCOLS AND THEIR APPLICATION TO ORGANIZATIONAL STRUCTURES

4.1 INTRODUCTION

Although distinct real-world organizational structures exist in great numbers, examining a few particular cases can provide considerable insight into the way organizations operate, and proves to be a satisfactory vehicle for testing theoretical results. In fact, two different organizational structures are used in this thesis, which present the advantage of modeling hierarchical and parallel decisionmaking in a simple but realistic way; this chapter will be concerned with the application of information storage principles and results to these particular configurations. The results evoked here are primarily related to taking into account a key variable in the decisionmaking process: time. Time is essential in the model developed from Chapter 2 to Chapter 5 because it provides a rigorous way of reframing the problem and putting it into a more realistic context, where the execution of a given task per se is not always more important than the date at which this task is performed or the time it took to be executed. The approach towards such a delicate notion will be constructed around a central element, organizational protocols. These protocols are defined in section 4.4 along with the establishment of some key properties through a proposition; the following sections are devoted to the interpretation and application of protocols in a way that brings out relevant characteristics of the various database-organization structures of interest here.

4.2 PARALLEL ORGANIZATIONAL STRUCTURE

In a parallel organizational structure, decisionmakers are linked by somewhat symmetrical relationships: they do not formally give orders to each other, and they can share information at all stages according to preestablished operating procedures. The parallel structure considered in this work is a three-person organization, called "organization P" from here on. It is represented in Figure 4.1 and it exhibits some simplifications with respect to the basic model presented in Chapter 3. Indeed, DM^1 and DM^3 use only one SA algorithm and two RS algorithms each, and DM^2 has the choice between two SA algorithms, whose output can be processed by only one RS algorithm. The command input v^{on} is absent from the model, due to the non-hierarchical structure; the decisionmakers do however share information about their situation assessments.

In Figure 4.1, organization P uses two centralized databases, DBSA and DBRS. A decentralized database structure can also be adopted, as shown in Figure 4.2 at the end of the section. Because of the computational similarity between the two database structures demonstrated in the previous chapter, only the centralized case will be dealt with insofar as activity computation is concerned. The decentralized information storage and access case will be examined in more detail in the sections concentrating on protocols and the consideration of time in measuring performance.

The various activity rate terms defined in Chapter 2 can be easily derived here by specifying the general equations of section 3.3:

<u>DM^1</u>

<u>Throughput Rate</u>: Eq. (3.1) becomes:

$$F_t^1 = \bar{T}(x^1, d_{SA}^1, z^{21}, d_{RS}^1 : z^{12}, \bar{z}^1, v^1, y^1) \tag{4.1}$$

<u>Blockage Rate</u>: Eq. (3.2) becomes:

$$F_b^1 = \bar{H}(x^1, d_{SA}^1, z^{21}, d_{RS}^1) - F_t^1 \tag{4.2}$$

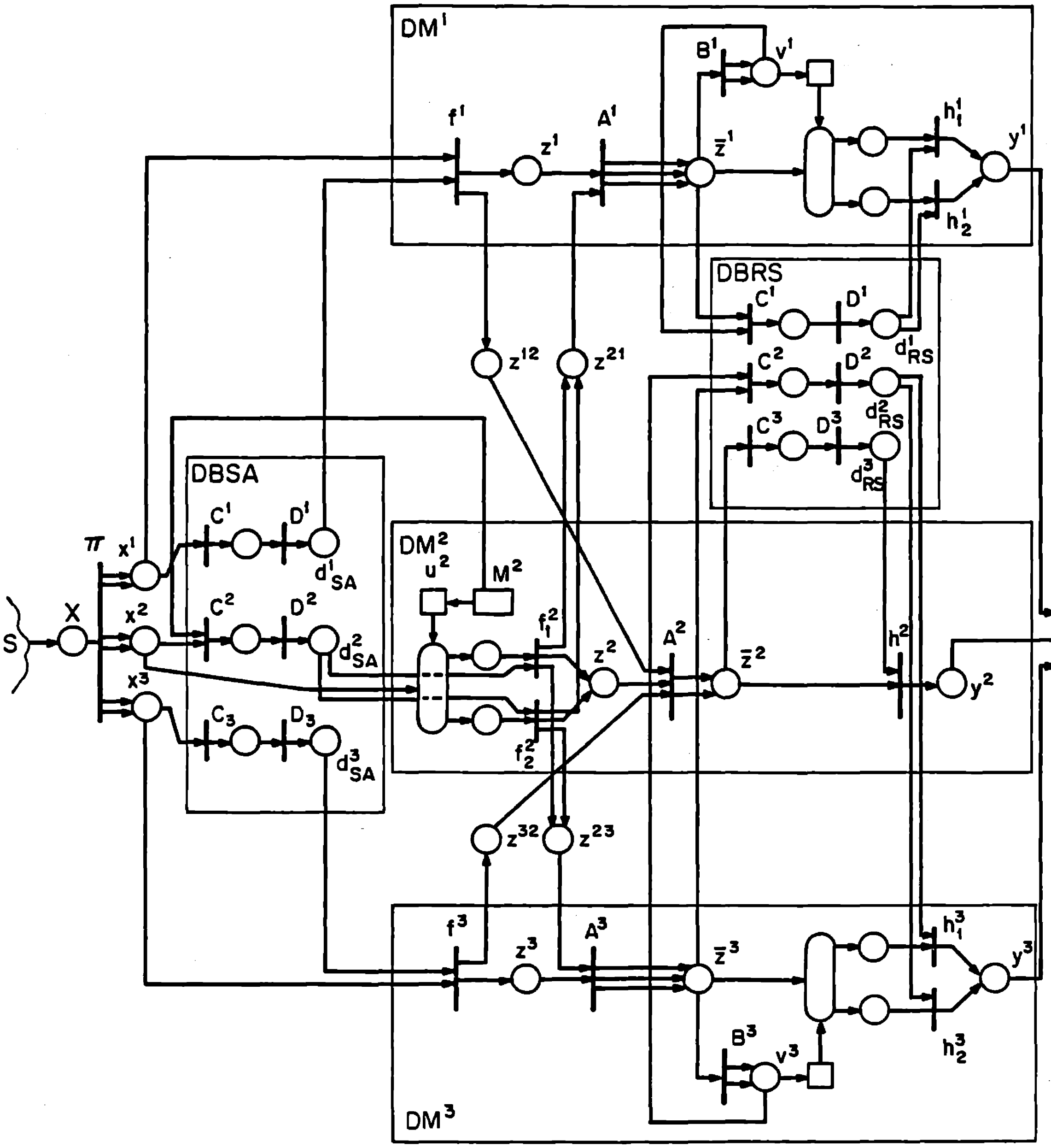


Figure 4.1 Petri Net Representation of Organization P
Using Centralized Databases

Noise Rate:

$$F_n^1 = \bar{H}_{\bar{z}^1}(v^1) \tag{4.3}$$

Coordination Rate: Eq. (3.14) becomes:

$$\begin{aligned} F_c^1 = {} & \bar{g}_c^1\ (p(x^1,d_{SA}^1)) + \bar{H}(z^1,z^{12}) \\ & + \bar{g}_c^A\ (p(z^1,z^{21})) + \bar{g}_c^B\ (p(\bar{z}^1)) \\ & + \sum_{j=1}^{2} (\bar{g}_c^{j+1}(p(\bar{z}^1,d_{RS}^1)) + \frac{\alpha_{j+1}}{\tau_{RS}} H(p_j)) + \bar{H}(y^1) \\ & + \bar{H}(z^1) + \bar{H}(\bar{z}^1) + \bar{T}_{z^1}(x^1,d_{SA}^1:z^{21}) + \bar{T}_{\bar{z}^1,v^1}(x^1,d_{SA}^1,z^{21}:d_{RS}^1) \end{aligned} \tag{4.4}$$

DM^2

Throughput Rate:

$$F_t^2 = \bar{T}(x^2,d_{SA}^2,z^{12},z^{32},d_{RS}^2:u^2,z^{21},z^{23},\bar{z}^2,y^2) \tag{4.5}$$

Blockage Rate:

$$F_b^2 = \bar{H}(x^2,d_{SA}^2,z^{12},z^{32},d_{RS}^2) - F_t^2 \tag{4.6}$$

Noise Rate:

$$F_n^2 = \bar{H}(u^2) \tag{4.7}$$

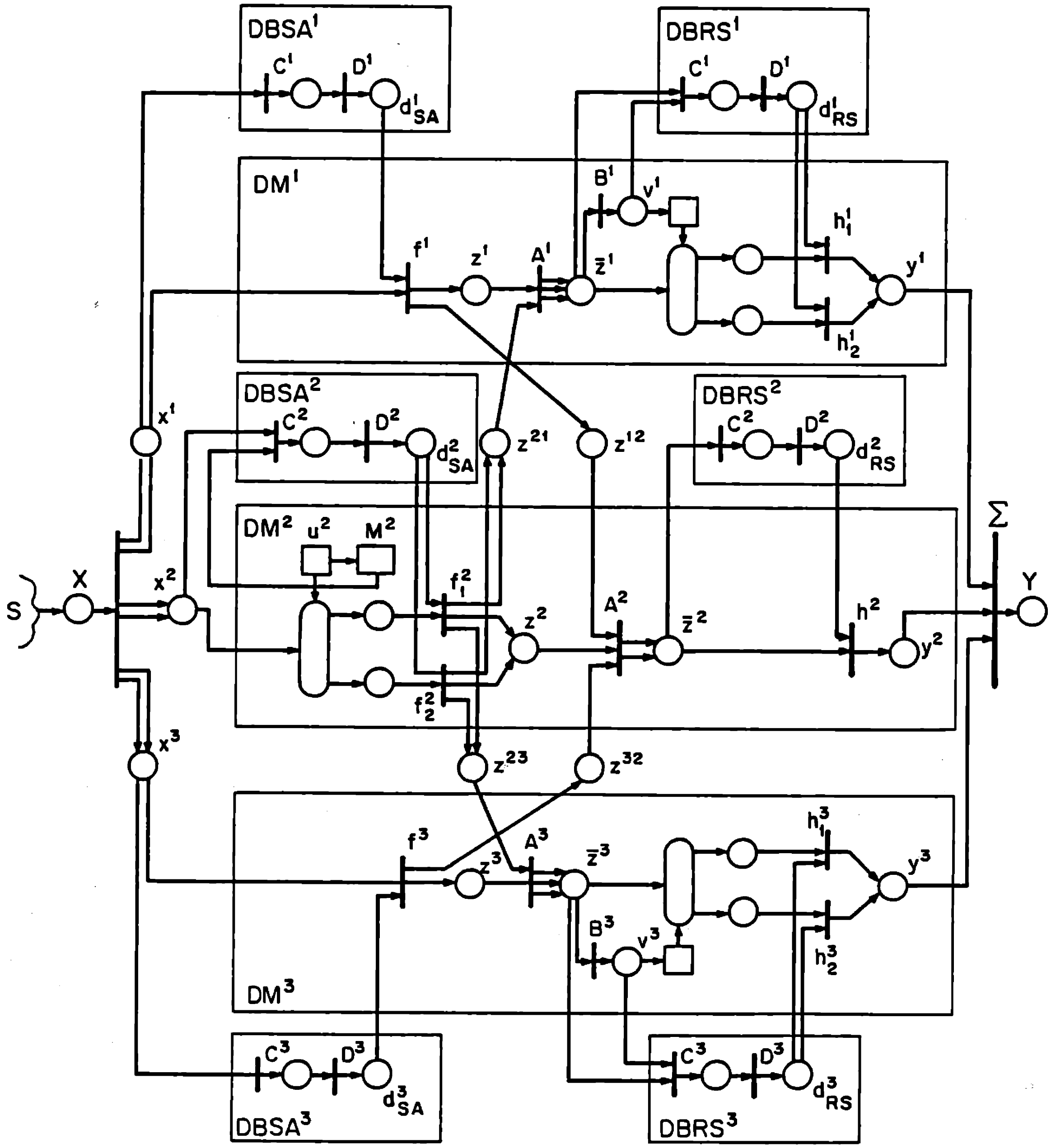


Figure 4.2 Petri Net Representation of Organization P Using Decentralized Databases

Coordination Rate:

$$F_c^2 = \sum_{i=1}^{2} (\bar{g}_c^i(p(x^2,d_{SA}^2)) + \frac{\alpha_i}{\tau_{SA}} H(p_i)) + \bar{H}(z^2,z^{21},z^{23})$$

$$+ \bar{g}_c^A (p(\bar{z}^2,z^{12},z^{32})) + \bar{g}_c^B(p(\bar{z}^2))$$

$$+ \bar{g}_c^3 (p(\bar{z}^2,d_{RS}^2)) + \bar{H}(y^2)$$

$$+ \bar{H}(z^2) + \bar{H}(\bar{z}^2) + \bar{T}_{z^2}(x^2,d_{SA}^2:z^{12},z^{32}) + \bar{T}_{\bar{z}^2}(x^2,d_{SA}^2,z^{12},z^{32}:d_{RS}^2)$$

(4.8)

DM^3:

The activity rate expressions for DM^3 are the same as for DM^1, provided 3 is substituted for 1 as a superscript in all the variables, due to the symmetry of these two decisionmakers' role with respect to DM^2's. This will be also true for organization H.

Note: It was shown in Chapter 3 that some terms are simpler in the event a decentralized database network is used. This result shows in the case at hand in the following way:

in F_c^1: $\bar{T}_{z^1}(x^1,d_{SA}^1:z^{21})$ becomes: $\bar{T}_{z^1}(x^1:z^{21})$

in F_c^2: $\bar{T}_{z^2}(x^2,d_{SA}^2:z^{12},z^{32})$ becomes: $\bar{T}_{z^2}(x^2:z^{12},z^{32})$

4.3 HIERARCHICAL ORGANIZATIONAL STRUCTURE

A hierarchical organizational structure allows decisionmakers to have an influence on each other's response selection. It was seen earlier that this influence is adequately represented by a command input, v^{on}. The hierarchical structure analyzed here is a three-person organization, known as organization H and represented when equipped with centralized databases in Figure 4.3.

Organization H consists of two decisionmakers who actually contribute to its output, DM^1 and DM^3, and one coordinating decisionmaker, DM^2, who analyzes DM^1's and DM^2's situation assessments in order to issue a command output towards them, that carries his instructions about the way the organization's response should be constructed. DM^2 is not in contact with the environment, therefore he does not need an SA stage, neither do DM^1 and DM^3 need an information fusion transition, A. The three decisionmakers in organization H have each two RS algorithms.

Here again, the various activity rate terms can be easily derived by applying Eq. (3.1), (3.2), (3.3) and (3.14) to each decisionmaker.

<u>DM^1</u>

<u>Throughput Rate</u>:

$$F_t^1 = \bar{T}(x^1, d_{SA}^1, y^{21}, d_{RS}^1 : z^{12}, z^1, \bar{v}^1, y^1) \tag{4.9}$$

<u>Blockage Rate</u>:

$$F_b^1 = \bar{H}(x^1, d_{SA}^1, y^{21}, d_{RS}^1) - F_t^1 \tag{4.10}$$

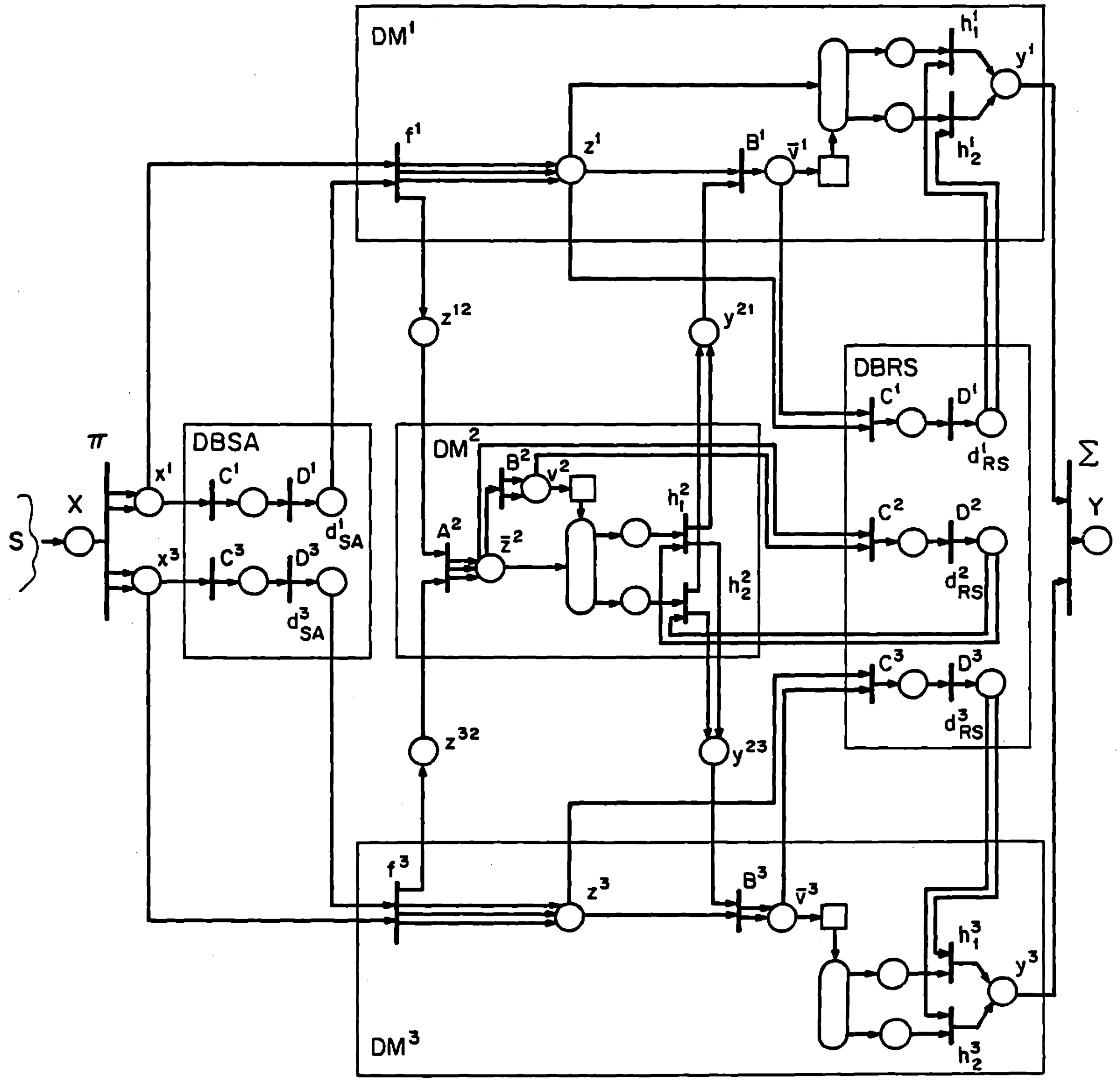


Figure 4.3 Petri Net Representation of Organization H Using Centralized Databases

Noise Rate:

$$F_n^1 = \bar{H}_{z^1}(v^1) \tag{4.11}$$

Coordination Rate:

$$\begin{aligned} F_c^1 = {} & \bar{g}_c^1(p(x^1, d_{SA}^1)) + \bar{H}(z^1, z^{12}) \\ & + \bar{g}_c^B\,(p(z^1, y^{21})) \\ & + \sum_{j=1}^{2} (\bar{g}_c^{j+1}(p(z^1, d_{RS}^1)) + \frac{\alpha_{j+1}}{\tau} H(p_j)) + \bar{H}(y^1) \\ & + \bar{H}(z^1) + \bar{H}(z^1, \bar{v}^1) + \bar{T}_{z^1, \bar{v}^1}(x^1, d_{SA}^1, y^{21} : d_{RS}^1) \end{aligned} \tag{4.12}$$

DM^2

Throughput Rate:

$$F_t^2 = \bar{T}(z^{12}, z^{32}, d_{RS}^2 : \bar{z}^2, v^2, y^{21}, y^{23}) \tag{4.13}$$

Blockage Rate:

$$F_b^2 = \bar{H}(z^{12}, z^{32}, d_{RS}^2) - F_t^2 \tag{4.14}$$

Noise Rate:

$$F_n^2 = \bar{H}_{\bar{z}^2}(v^2) \tag{4.15}$$

Coordination Rate:

$$F_c^2 = \bar{g}_c^A(p(z^{12},z^{32}) + \bar{g}_c^B(p(\bar{z}^2))$$

$$+ \sum_{j=1}^{2} (\bar{g}_c^j(p(\bar{z}^2,d_{RS}^2)) + \frac{\alpha_j}{\tau_{RS}} H(p_j)) + \bar{H}(y^{21},y^{23})$$

$$+ \bar{H}(\bar{z}^2) + \bar{H}(\bar{z}^2,v^2) + \bar{T}_{\bar{z}^2,v^2}(z^{12},z^{32}:d_{RS}^2) \tag{4.16}$$

Figure 4.4 is the Petri Net Representation of organization H in the case a decentralized database structure is used. The simplications brought to the centralized model by Eq. (3.19) and (3.20) are not relevant here because the terms to which they apply are null.

In the coming sections, new concepts originating in the consideration of time constraints in the functioning of the organization will be introduced. Apart from addressing a critical aspect of decisionmaking modeling, these concepts will constitute a new dimension along which centralized and decentralized databases can be compared.

4.4 DEFINITION OF PROTOCOLS AND DETERMINATION OF THEIR KEY VARIABLES

4.4.1 Introduction

One meaning of the word "protocol" according to the Wesbter's seventh New Collegiate Dictionary, is the following: "a code of diplomatic or military etiquette and precedence". This definition is more closely adapted to organizational operating procedures by stating that a protocol is the description of the chronological order in which elementary tasks have to be performed, within one decisionmaker as well as between two or

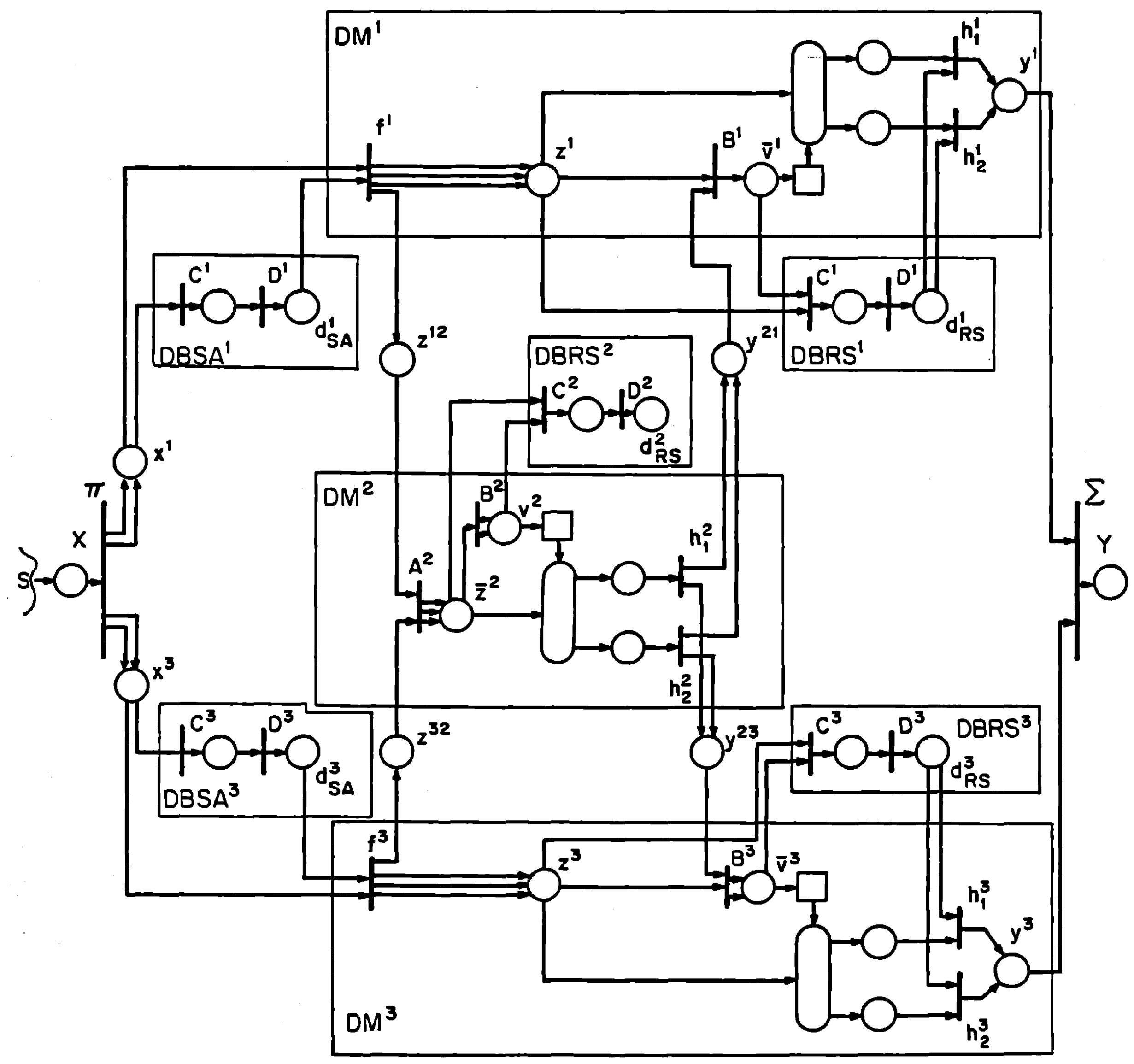


Figure 4.4 Petri Net Representation of Organization H Using Decentralized Databases

more of them. Procotols are a fundamental problem of organizations in general, and of updatable database-equipped ones in particular. Indeed, if the sequences of operations for each decisionmaker are not clearly defined, and if the updating tempo of the database does not take these sequences into account, chaos can result. In brief, the situation could arise where different decisionmakers would be accessing different databases at different times, with different levels of accuracy and relevance of the data, in order to process the same input.

One of the nicest features of a Petri Net representation is that it clearly illustrates the organization's protocol as defined above. This is due to two properties of Petri Nets: first, a transition cannot be fired as long as one of its input places is empty; second, a place remains empty until the transition of which it is an output place is fired. These basic properties make Petri Net theory an appropriate tool for introducing the concept of protocols in the information theoretic approach to organization evaluation.

Another key notion in the definition of a protocol is the amount of time involved at each step of the decisionmaking process. Therefore, in the remainder of the chapter, an acceptable protocol for a given organization will consist of its Petri Net representation supplemented with the allocation of a processing time to each transition. The processing time in fact represents the maximum allowable duration of a transition for the organization to function in an orderly fashion, following its operating protocol.

4.4.2 Protocol Acceptability

<u>Assumptions</u>: In devising an acceptable protocol for the kind of organizations dealt with here, the following assumptions are made:

(1) - the source emits the input X with a constant interarrival time (see section 3.3).

(2) - the various transitions have all a constant processing time.

(3) - communication between transitions is instantaneous.

(4) - any transition can process an incoming input as soon as it has finished processing the previous one, and no sooner.

(5) - no queueing is allowed at any stage of the process.

Assumptions (1) and (2) are a corollary of the broader assumptions that the whole system operates in a steady-state configuration. Assumption (3) states in fact that all the decisionmaking takes place within the transitions, and that no processing time is allocated to places. Assumption (4) is putting the "pipe-line effect" into words; this assures that the information flow through the system is continuous. Assumption (5) is a prerequisite to the application of Petri Net theory to the study of information theoretic decisionmaking models: in effect, when queueing takes place, two or more different tokens can coexist in the same place. Since transitions do not have any means of recognizing priorities in choosing one token as an input out of the same place, the queue cannot be managed, and the organization's protocol is transgressed.

<ins>Proposition</ins>: Under assumptions (1) and (5), two necessary conditions for an organization's protocol to be acceptable are:

- every transition in the system must have a processing time smaller than or equal to the mean input interarrival time.

- the total amount of time spent by a token in one place cannot exceed the mean input interarrival time. ■

The first necessary condition of the proposition will be demonstrated by showing that if any transition's processing time is greater than the system's input interarrival time, some of the proposition's assumptions are violated.

Consider a transition T. It receives an input every τ_i, it takes τ_p to process this input, and it produces an output every τ_o. It can be easily seen that the following always holds:

$$\tau_o = \max(\tau_i, \tau_p) \tag{4.17}$$

Now consider one transition T_λ whose processing time is greater than τ, the system's input interarrival time. If more than one transition qualify, consider the first one along the path followed by the information emitted by the source. All preceding transitions process their inputs in less than τ. Consider the very first of these preceding transitions: according to Eq. (4.17), its τ_o will be τ, because its input is directly emitted by the source and its τ_p is smaller than τ. Similarly, the τ_o of the next transition is also τ (one transition's output is another's input), and so on ... till the information gets to transition T_λ. Here, the following occurs: T_λ receives an input every τ, but takes τ_λ, τ_λ greater then τ, to process it. Since it cannot start processing an input before the previous one is completely dealt with (Assumption (4)), this means that at some point more than one input will be awaiting to be processed by T_λ, which contradicts Assumption (5).

The second necessary condition is a direct consequence of the first one. Indeed, in an acceptable protocol, it was just proven that any transition's processing time is smaller than or equal to τ. This, together with Eq. (4.17), means that every place receives a token from the transition of which it is an output place every τ. If any token spends more than τ in the same place, then this place will happen to hold at the same time two successive outputs of the same transition, waiting to be processed by the next one, which constitutes a violation of Assumption (5).

Before the more specific context of the two particular organizations analyzed in this chapter is approached, some remarks can be made about the proposition proposition demonstrated above:

1. The first result of the proposition is all the more interesting as it corresponds to the information theoretic interpretation of the bounded rationality principle, that states that the system must process its inputs at a rate at least equal to the rate at which they arrive. However, the proposition is more specific in that it states the same constraint for each transition, and not only for the system as a whole. One should nevertheless remain aware of the fact that this does not imply that the concept of bounded rationality applies as such to each transition alone. Again, it is an empirical result based on humans' reaction time and response to external stimuli, without any constraint on the way they use their time and energy during their internal decisionmaking process.

2. Both necessary conditions provide a symmetric analytical tool. Indeed, if the processing times of the transitions in the system are fixed, then the minimum admissible input interarrival time for the organization can be determined: it is equal to the greater of two quantities: the maximum processing time present on the Petri net diagram of that organization, and the maximum time any token spends in any place. Determining this minimum interarrival time is a very useful way of comparing the effectiveness of different organizational structures in a given context.

3. The second necessary condition applies in a nice way in cases of organizational interactions where one decisionmaker sends some information to another, and cannot proceed before receiving a message back. Thus, the proposition provides a way of determining the upper limit of the response time of this other

decisionmaker, everything else being fixed. This will be made clearer in the next section, where an acceptable protocol will be devised for organization H.

4. Organizations with complex protocols and intricate interrelationships between their members can be analyzed using the systematic approach develoepd by Tabak and Levis (1984). They used the Petri Net logic to construct matrices that link the time involved in each step of the decisionmaking process to the dependence of each transition on other transitions in different decisionmakers; these matrices are called the System Arrays of the organization.

5. As a last comment, one should realize that the use of the proposition stated in section 4.4.2 is not restricted to decisionmaking organizations. In fact, its arguments are relevant to any acyclical information processing structure where the notions of input and Petri Net apply and where Assumptions (1) to (5) are satisfied.

At this point, enough has been said about the general aspects of protocols to provide the tools for a more detailed study of the two organizational structures presented in the first sections of the chapter. In the next two sections, acceptable protocols will be developed for each of these organizations, in each database configuration; then, the use of these protocols to consider time as an evaluation criterion will be discussed. Lastly, the conclusions of the chapter will be presented in terms of the increased sophistication brought to the basic memoryless decisionmaking model.

4.5 THE CONSTRUCTION OF PROTOCOLS FOR SPECIFIC ORGNIZATIONAL STRUCTURES

In this section, the proposition will be used to develop protocols for both organizations, P and H, using a centralized database structure in section 4.5.1, and a decentralized one in 4.5.2. The main aspect of this exercise is the use of the symmetric argument of the proposition necessary conditions, and the existence of a global time constraint for the organization: the response time is the same for all decisionmakers who contribute to the global output, and it is the organization's response time.

Since transition Σ, which delivers the organization's output, will not fire until all individual responses are received, it follows that the goals of the organization are better served if all these responses are emitted towards Σ at the same time.

The basic elementary quantity for each organization is τ, the processing time of any SA or RS transition. It is assumed to be identical for all such transitions in both organizations, and it will be the unit used for all quantities computed here. Furthermore, τ is assumed to be greater than the processing time of other types of transitions, on the grounds that more decisionmaking takes place in SA and RS transitions than in the others. The database's response time is assumed to be τ as well for the centralized case, it can vary considerably, depending on the technical characteristics of the Decision Support System used by the organization.

4.5.1 Centralized Databases

Under the conditions outlined in the introductory paramgraph, an acceptable protocol for organization P can be derived. It is given in Figure 4.5. Its main characteristics are the minimum interarrival time (IT) it allows τ, and the organization's total response time (RT), $19\tau/3$.

Similarly, one acceptable protocol for organization H is that represented in Figure 4.6. The minimum interarrival time is much greater, $11\tau/3$. This is due to the relationship between transition f_1^1 and DM^2, where the information coming out of f_1^1 has to be processed by all DM^2's

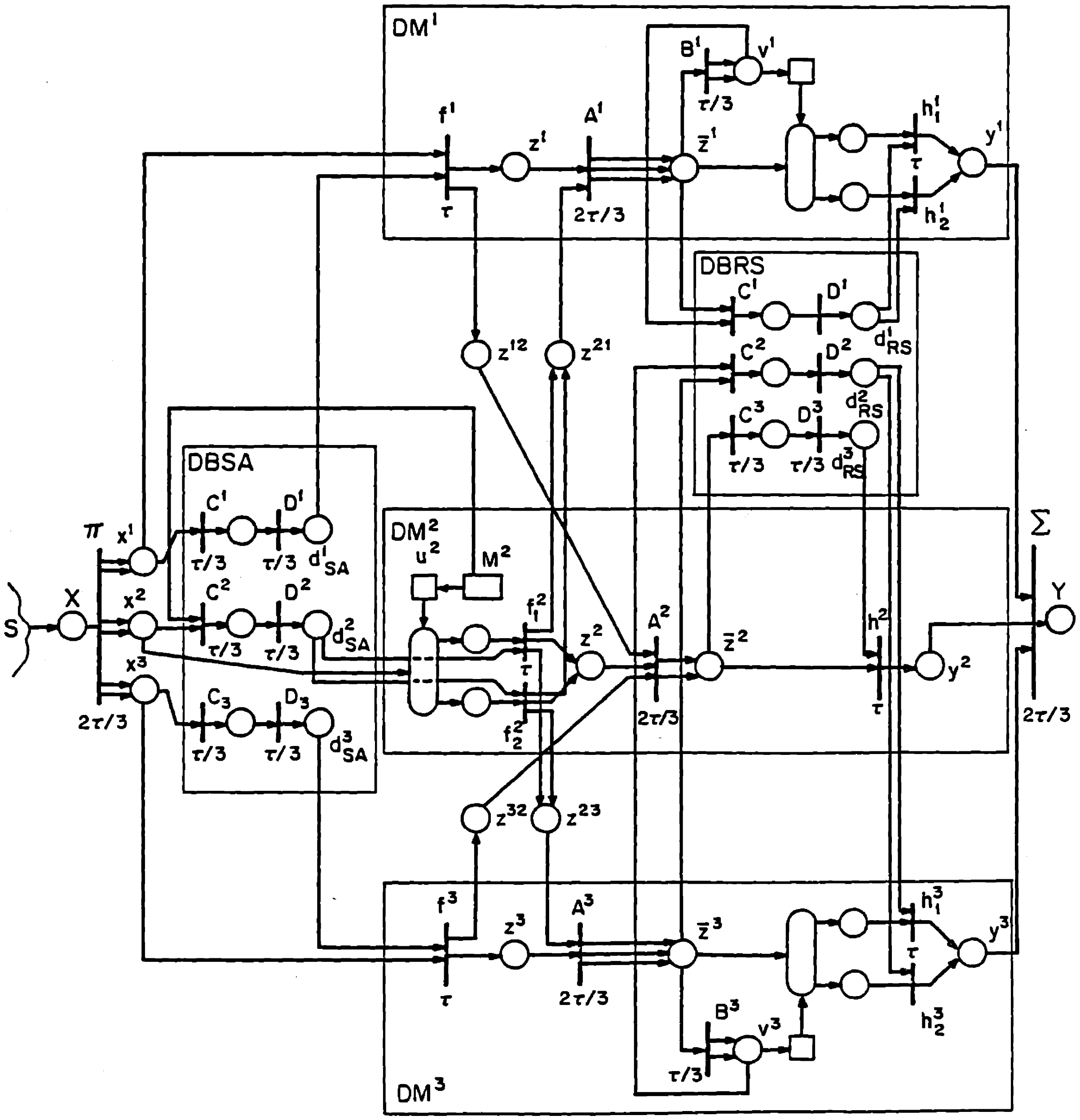


Figure 4.5 Protocol of Organization P Using Centralized Databases

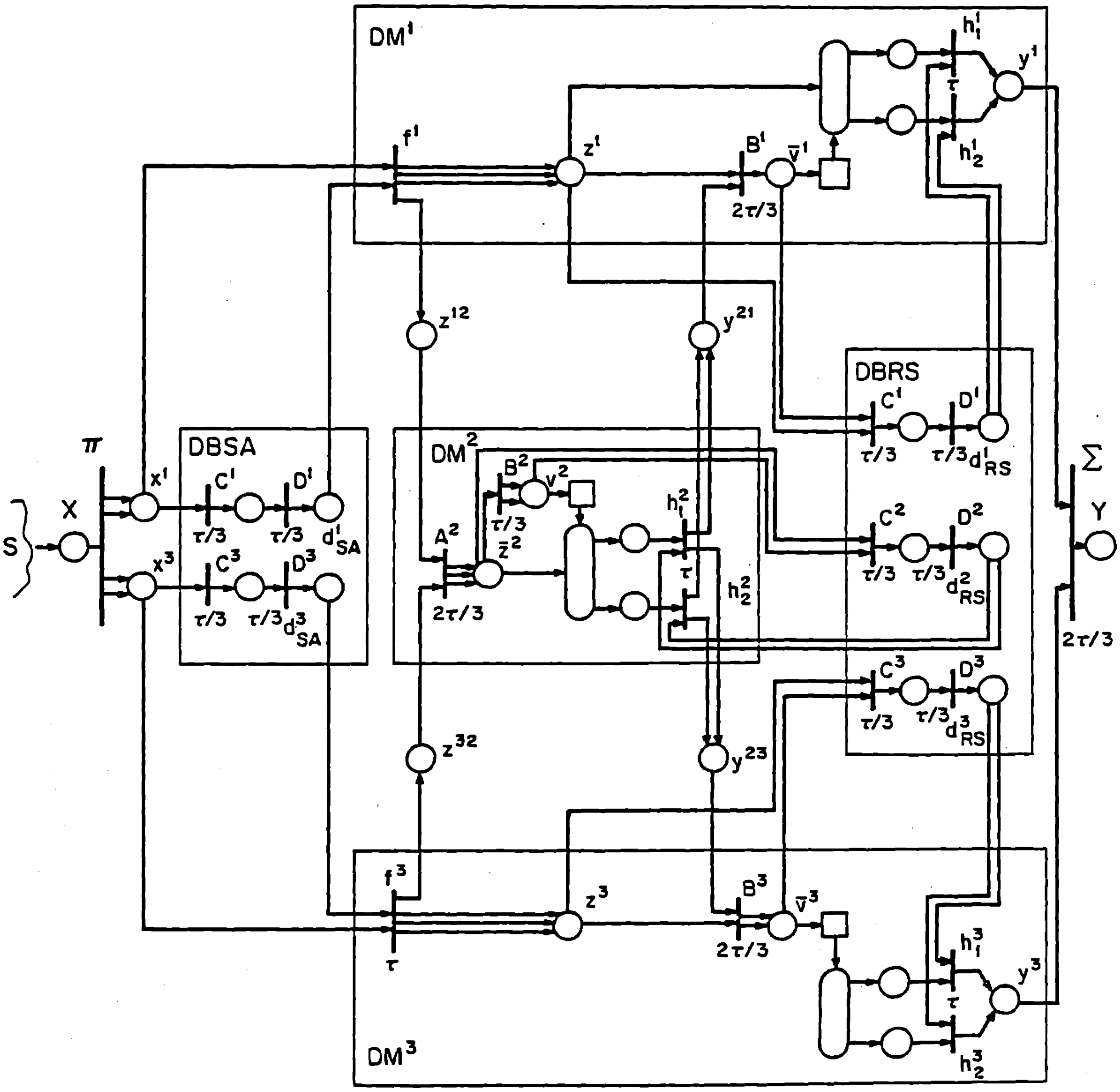


Figure 4.6 Protocol of Organization H Using Centralized Databases

transitions before transition B^1 can be fired and the last token leaves the place z^1. The application of the symmetric argument of the proposition's necessary conditions determined the mean interarrival time $11\tau/3$. The organization's response time is calculated quite simply in this case, by adding all processing times along the path followed by the original input, and is: 8τ. For more complex organizations, the System Array approach is preferable.

Once IT and RT are determined, a last issue, that of the updating instant (UD) of the two databases present in the system, must be addressed, if the protocol of the organization is to be consistently defined. The UD must be such that updating occurs when the database is not processing any query, and that the data provided to all decisionmakers have the same level of accuracy. It is assumed that updating occurs at every iteration performed by the system. In the cases at hand, UD is easily determined, due to the centralized structure, and the values it takes for each organization are:

Organization P: $\mathrm{UD(DBSA)} = \tau + 0$

$\mathrm{UD(DBRS)} = \tau + 3\tau$

Organization H: $\mathrm{UD(DBSA)} = 11\tau/3 + 0$

$\mathrm{UD(DBRS)} = 11\tau/3 + 3\tau$

In more complex cases, numerous iterations on IT may be required before a feasible set of values for the various UD can be reached. (This occurs when one decisionmaker needs data relevant to one input while another decisionmaker has not yet accessed the data relevant to the previous input. Furthermore, if the final IT is much greater than the one determined using the proposition, the adoption of the solution set of UD might entail considerable losses of time for the organization in terms of its ability to respond to external stimuli. A better solution consists then of breaking down the centralized database into individual units, for reasons discussed below.

4.5.2 Decentralized Databases

As far as protocols are concerned, it was pointed out in section 3.6 that the only salient differences between a centralized and a decentralized structure pertain to transition D's processing time and the establishment of satisfactory updating dates. In this section, transition D is assumed to require a total time of $\tau/3$, which is half what was needed in the centralized configuration. Again, this number depends greatly on the nature of the organization's DSS.

Acceptable protocols for both organizations are given in Figure 4.7 and Figure 4.8. The changes in the IT and RT are solely due to the shorter database response time. One gets for each organization:

Organization P: $IT = \tau$

$RT = 12\tau/3$

Organiation H: $IT = 10\tau/3$

$RT = 7\tau$

Although decentralized database structures have been determined to need a greater updating effort than centralized structures do to obtain the same results, they exhibit greater flexibility, which helps alleviate the

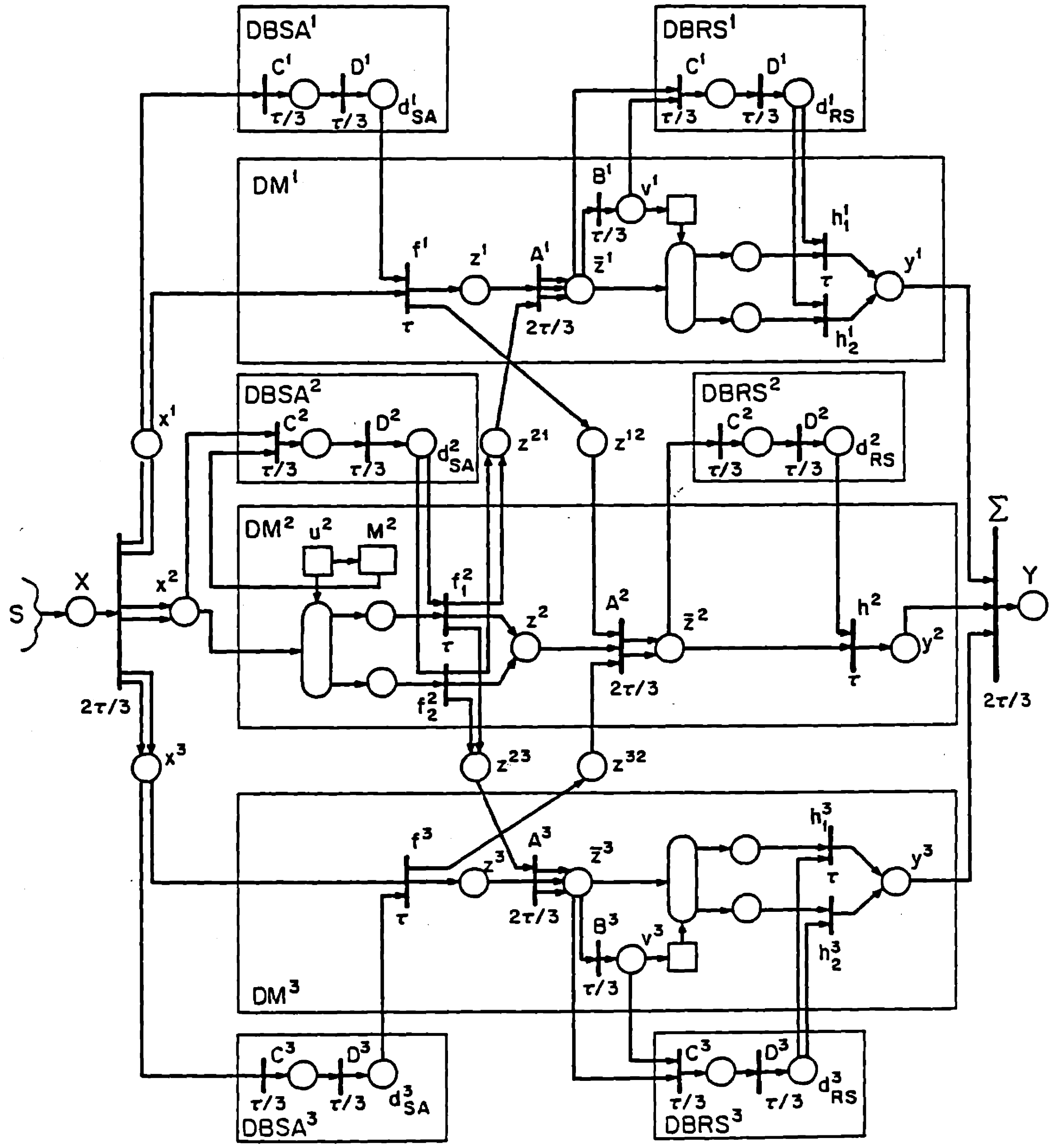


Figure 4.7 Protocol of Organization P Using Decentralized Databases

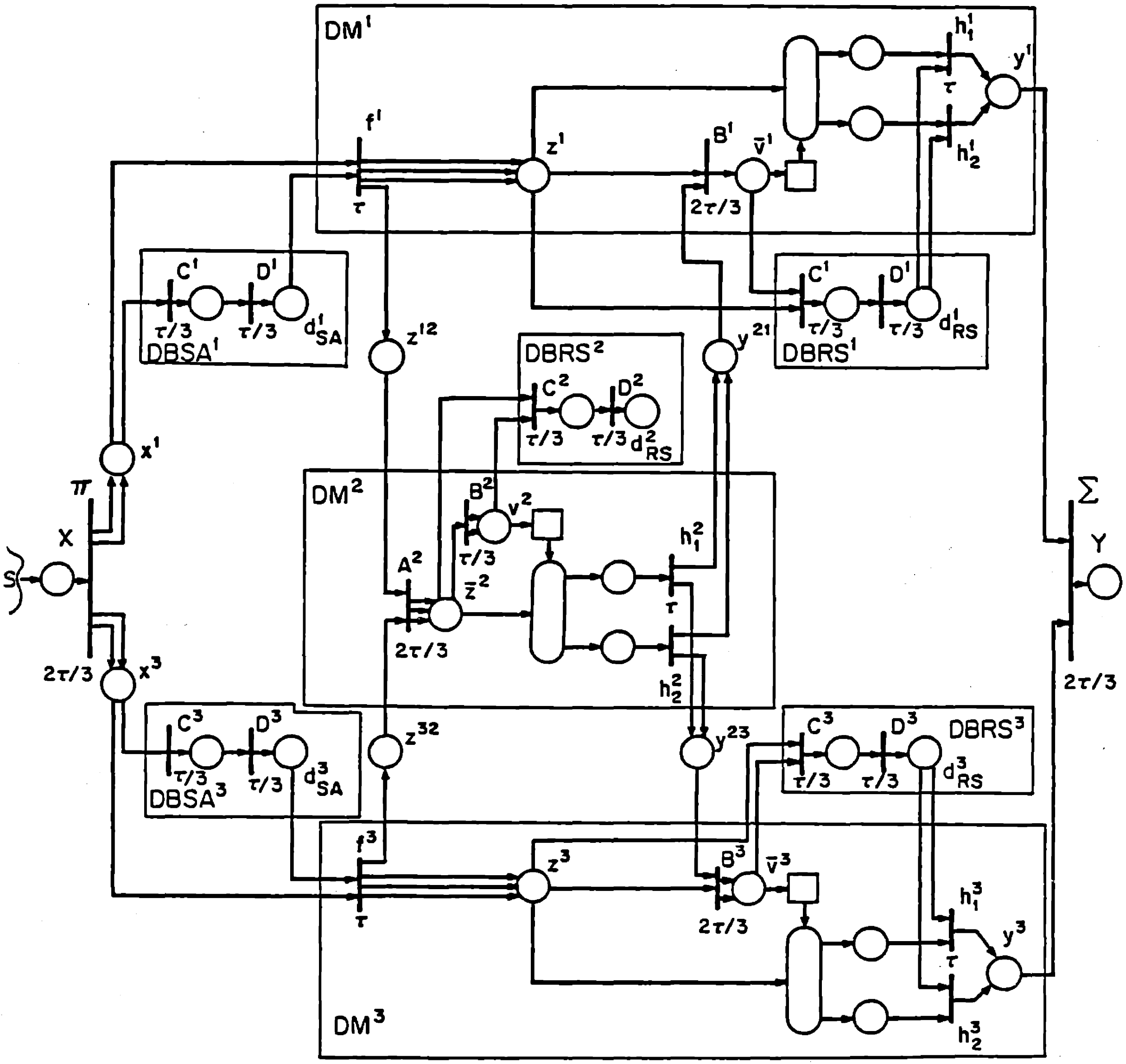


Figure 4.8 Protocol of Organization H Using Decentralized Databases

problems mentioned at the end of the last section. In effect, in the extreme case, each individual database can be updated when the need arises, independently of others. Such a procedure assures that at least one acceptable protocol does always exist and is easily obtainable. Its high cost, however, might induce an organization designer not to adopt it even when it is the only solution; the costs of such an approximation in terms of performance are analyzed in the next chapter.

For the particular organizations of interest here, updating can fortunately take place simultaneously for all variables databases of the same stage (SA or RS). Feasible UD sets are:

Organization P: $UD(DBSA^i,\ i=1,2,3) = \tau + 0$

$UD(DBRS^i,\ i=1,2,3) = \tau + 7\tau/3$

Organization H: $UD(DBSA^i,\ i=1,2,3) = 10\tau/3 + 0$

$UD(DBRS^i,\ i=1,2,3) = 10\tau/3 + 8\tau/3$

The protocols constructed in this section and the previous one are each only one of a quasi-infinity of possibilities. They have been derived under some particular conditions, however, to make different organizations and different database structures comparable along the same criteria. Such a comparison will be conducted in the next section.

4.6 THE CONSIDERATION OF TIME IN EVALUATING PERFORMANCE

As it was first pointed out in section 2.3, an organization's performance depends on two aspects of its response: its nature, and its timeliness. The first aspect of performance has been studied as a part of the basic decisionmaking model, and it will appear on the performance-workload loci presented in the numerical example analyzed in Chapter 5; it will be denoted type 1 performance. The present chapter has been concerned with the development of analytical tools for approaching the issue of timeliness and input interarrival time in organizations; these issues form the core of what will be called type 2 performance. (The type 1 - type 2 terminology will be used only when confusion may arise.) This section presents the results that can be gathered from the practical application of those tools to various database-organization structures in the previous sections. These results have two key characteristics: first, they are contingent upon using similar transition processing times for both organizational structures. Second, they depend in no direct fashion on the results derived during the analysis of type 1 performance - workload loci.

A first result that follows from the foregoing analysis is that the minimum allowable input interarrival time is much greater for a hierarchical organization than for a parallel one. This can be interpreted to follow from the more complex sequences of tasks that have to be performed in a hierarchical organization before a new input can be handled. The total response time is also greater for organization H than for organization P, and the difference is due to the same type of reasons invoked above.

The second result of this chapter is relevant to the database structure adopted by the organization. It was seen that, whatever the organization, a decentralized database structure leads to more performing time characteristics than a centralized one. In organization P, the decentralized structure leads to an 11% improvement in the response time, while in organization H its leads to improvements in both IT and RT of 9% and 13% respectively. These results are due to the basic premise that decentralized databases takes less time to perform the data query process than centralized ones do. (The numerical results of the above two paragraphs are summarized in Table 4.1.).

TABLE 4.1 TIME CHARACTERISTICS OF ORGANIZATION P AND H

	Centralized db	Decentralized db
IT(P)	τ	τ
IT(H)	$11\tau/3$	$10\tau/3$
RT(P)	$19\tau/3$	$17\tau/3$
RT(H)	8τ	7τ

IT = minimum allowable Interarrival Time; RT = Response Time

Type 2 performance cannot be assessed in an absolute way: the numbers themselves do not indicate which organization is intrinsically preferable. Such a choice depends primarily on the context in which the organization is evaluated. Also, the concept of satisficing performance does intervene in the design of an organization. As an example of that, consider a battery capable of shooting at the enemy over a certain range, covering a certain geographic area. It may be deemed satisficing for the designers of the system consisting of the battery and its servants to obtain an RT no smaller than a given fraction of the window of opportunity (for an extensive study of the notion of window of opportunity, see Cothier, 1984). It can also be decided that an IT no smaller than the rate at which enemy units can actually be sent towards the battery is enough to bring the battery's type 2 performance to an acceptable level.

Even when type 2 performance is evaluated after considering the environment and the goals of the organization at hand, type 1 performance still has to be a part of the final performance assessment of the system. An attempt to bring these two types of performance together is made in the next chapter.

4.7 THE DECISIONMAKING MODEL REVISITED

In Chapter 2, the basic organization model was presented; it was essentially a version of the single-decisionmaker model, extended to organizations. The main assumptions under which such a model could be used to represent reality to some extent were recapitulated; one of these was its being memoryless. Chapter 2 relaxed this constraint, arguing that decisionmakers were not memoryless, and that, furthermore, today's organizations used memory devices anyway. For those reasons, a first approach towards modeling the memory process was developed by equipping the organization with a database structure that in fact held information incoming from an external element, called the updating source. That actual functioning of such a database structure, both in a centralized and decentralized configuration, was commented on, and its impacts on the decisionmakers' workload derived by information theoretic means. At that point, concern about the performance of such an organization arose, and time-related connotations were added to the traditional evaluation of that performance. The present chapter's objective was to develop analytical tools for the evaluation of an organization along a time dimension. Procotols were defined, and their construction made explicit in four particular cases, after the proposition specifying their key charateristics: input interarrival time and response time, had been demonstrated. It was also shown that protocols, apart from providing an organization with a feasible and consistent sequencing of its elementary tasks, could be used to illustrate the impacts of information storage devices on the total response time and the minimum input interarrival time of that organization. These impacts are recognized to be critical; indeed, in organization P for instance, the introduction of centralized databases brings the response time to $19\tau/3$, as opposed to $13\tau/3$ in the memoryless case. (Assuming that τ does not change for the transitions.) Again this result should not be used to question the use of information storage devices, but rather to compare alternative structures.

In conclusion, it can be stated that the basic decisionmaking model has started providing for the storage and retrieval of external information, and that analytical tools for its evaluation along three of its primary attributes: decisionmakers' workload, type 1 performance level, and type 2 performance, are available. The next chapter turns to the illustration of the tradeoff existing between the first two attributes, and the suggestion of possible links between all three of them together.

CHAPTER V

AN ILLUSTRATIVE EXAMPLE

5.1 DESCRIPTION OF THE ORGANIZATIONS USED

The use of information storage and access devices by decisionmaking organizations has significant impacts on many aspects of the functioning of these organizations. These impacts have been qualitatively described in the previous chapters, as new concepts and tools were introduced. This chapter, however, is concerned with a more quantitative approach, relying on the construction of performance-workload loci for two particular organizations defined in the next section. The issues addressed in this chapter are the comparison of two different organizational structures, the use of input interarrival times in the qualitative evaluation of the organization, the problems raised by a lack of coordination between several individual databases, and a physical illustration of the concepts of timeliness introduced in the previous chapter.

5.1.1 Tactical Setting of the Organizations

The first organization is of the parallel type, as represented by organization P in Figure 4.1. It consists of three naval battle groups defending a maritime front. The first group, DM^1, holds one extremity of the front, DM^2 holds the center, and DM^3 the other end. The inputs received by the organization are signals emitted by unidentified platforms (submarines, surface ships, planes). The different decisionmakers' tasks are to attempt to identify the source of these signals (enemy or friends) in the SA stage, and to select the appropriate response (fire, request identification, or take all measures required to face a general attack) in the RS stage.

DM^1 and DM^3 communicate their first situation assessment, z^1 or z^3, to DM^2, and receive from him his own situation assessment, z^2. After the information fusion stage, A, each decisionmaker has determined whether both he and the adjacent decisionmaker are attacked, he or his neighbor is being attacked alone, or none of them is being attacked. The response is then selected based on $\bar{z}^n$, n = 1,2,3.

The databases contain and deliver information aimed at assisting the decisionmakers in their various choices and are used in a centralized structure. The SA database provides information, obtained from intelligence sources, that describes the codes the enemy could use when emitting the kind of signals received by organization P. This information will be compared to the actual input to determine the latter's identity. The RS database, DBRS, informs the decisionmakers about the level of alert present in their area at each iteration. This will have a great influence on the final response (indeed, in case of red alert, an enemy is shot at on sight, whereas a green alert allows for an identification request or a warning).

The second organization is built into a hierarchical structure. The context is the same as for organization P, but here only DM^1 and DM^3 actually receive any external signals or select an active response. DM^2 is a coordinator who, based upon the situation assessments received from DM^1 and DM^3, gives instructions about what RS algorithm should be selected by either of them. The organization's overall mission is the one defined for organization P.

The databases are again used in a centralized network for the first phase of the calculations. They provide the same information as they did in the previous case, except for d^2_{RS}. The data from $DBRS^2$, d^2_{RS}, is different, to conform to the different role played by DM^2: it reflects the degree of aggressiveness shown by the higher tactical command, or by the political authorities.

5.1.2 Basic Assumptions and the Methodology in Brief

Two primary features of the example in this chapter are its numerical simplicity and the importance of the insights it brings to information theoretic aspects of the organization. All the variables of the system are determined using a binary logic, based on the comparison of quantities, and not their actual computation. Detailed definition of the variables and of the algorithms in a numerical example of the decisionmaker has already been presented in the literature (see Boettcher, 1981) and is the subject of Appendix B for this example. The basic step in the computation of the pairs (J,G) in the performance-workload locus is determining the pure strategies present in the organization. (Levis and Boettcher, 1983.) In the cases at hand, each DM has two pure internal strategies, each obtained by the exclusive use of one algorithm (no decisionmaker here has, in any stage, more than two algorithms to choose among). In a three-person organization this leads to eight pure organization strategies ($2^3=8$). The points determined by each pure strategy and the corresponding type 1 performance level are first located on the graph; (these results are shown in several Tables in Appendix B). Then, the performance-workload (P-W) locus, of each DM is drawn, using the methodology developed by Boettcher (1981), considering all possible mixed strategies as linear combinations of the pure ones. The graphs thus obtained are the projections of the overall (P-W) locus of the organization on each of three planes: (G^1,J), (G^2,J), (G^3,J). Because the input is perfectly symmetric, as well as DM^1's and DM^3's roles, in each organization, only DM^1 and DM^2 are used to show the projections. The data sent by the databases is considered to be part of the global input to the organization emitted by a supersource, as defined by Stabile and Levis (1984).

Some other simplifying assumptions are the following:

- a centralized database structure is used by each organization

- all DMs receive the same data at a given stage (except DM^2 in organization H, as mentioned in the previous section)

- the data received from the databases does not depend on the input x^n
- the three individual inputs are assumed to be independent (the successive values they take are independent as well).

These assumptions have already been presented in the general description of the model made in the previous section. The first one, regarding the database structure used, will be relaxed in the last section of the chapter, to analyze a critical difference between centralized and decentralized databases.

5.2 PRESENTATION AND INTERPRETATION OF THE RESULTS

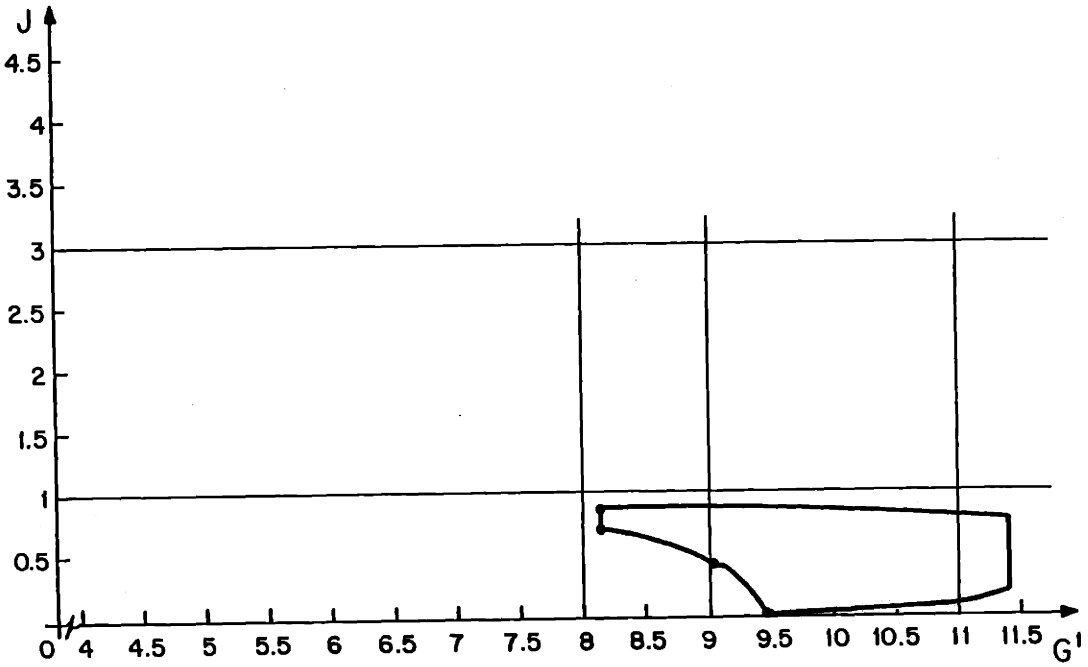
5.2.1 Comparison of (P-W) Loci for Organiations p and H

As far as timeliness was concerned, Chapter 4 showed that the organizations' characteristics were essentially structure dependent, whatever the mission assigned to the organization. Table 4.1 was derived using structural features of organizations P and H and the protocols developed in the same chapter. This section, however, points to the difficulty of establishing such general results with respect to the performance-workload loci of the two organizations. Indeed, in the example developed by Levis and Boettcher (1983), organization H appeared to possess strategies exhibiting lower workload and higher performance than any strategy in organization P does. This enabled organization H to be the preferred organizational structure for the more stringent constraints on workload and performance. In the example at hand, where the decisionmakers' tasks and the overall organization mission are somewhat different from the configuration quoted above, organization H's advantage is greatly reduced, as is demonstrated in this section.

The (P-W) loci for DM^1 and DM^2 in organizations P and H are represented in Figures 5.1 and 5.2 respectively. One can immediately see that organization P reaches higher performance levels than organization H does, but that H imposes a lower minimum workload on its members than P does. This is an indication that either organization P or H will be preferred in this example, depending on the workload and performance constraints. If, for instance, the higher admissible workload for any decisionmaker ($F_o\tau$, in Eq. (2.20)) is 8 and the lowest acceptable performance ($\bar{J}$ in Eq. (2.22)) is 1, Figures 5.1 and 5.2 show that neither organization P nor H is an acceptable solution. If $F_o\tau$ is identified to be 11 bits per symbol, then organization P is the configuration that provides for the best performance within this constraint, and is to be preferred. If $F_o\tau$ cannot possibly go beyong 8 bits per symbol and performance indices greater than 1 are acceptable, then organization H is the only solution to the problem. If $F_o\tau$ and $\bar{J}$ are such that both structures are feasible ($F_o\tau$ = 9 and $\bar{J}$ = 3 for instance), the preferred one is that for which the measure of mutual consistency is higher. (This measure, called Q, is the ratio of the volume occupied by the admissible strategies over that defined by all the organization's strategies, for a given set of constraints ($\bar{J},\tau$). For a detailed description of the concept, see Boettcher and Levis, 1983).

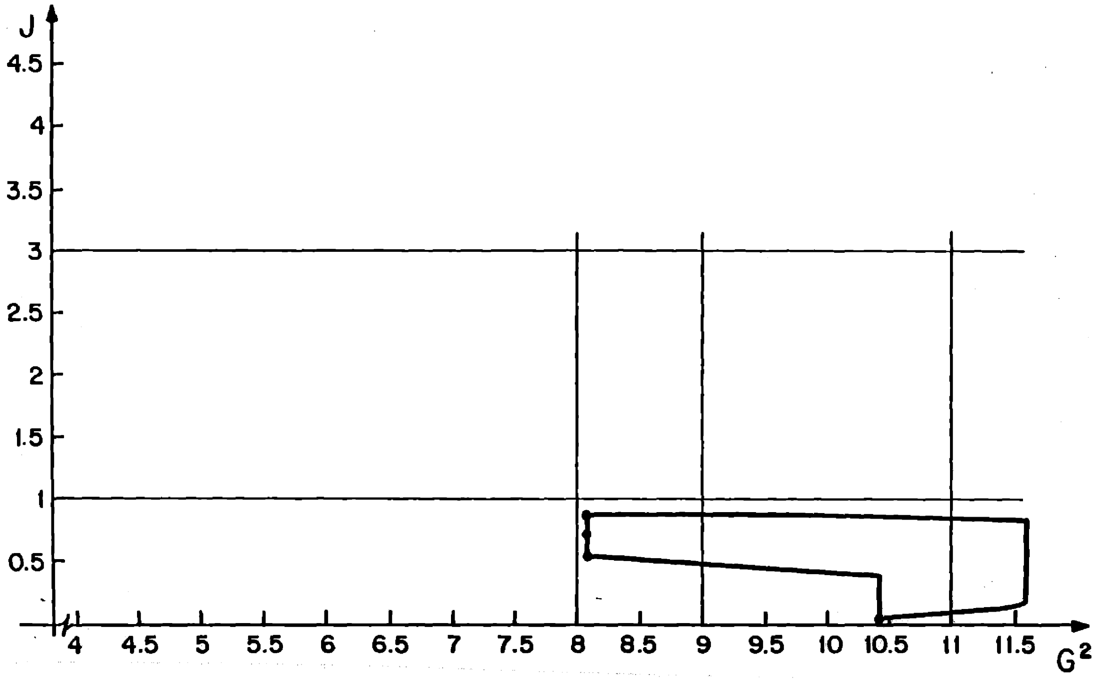
In summary, neither P nor H is a panacea in the example at hand; P is favored by relatively stringent performance requirements, whereas H is best when overload occurs for relatively low workloads, given the decisionmakers. This conclusion is different from the results obtained in the previous numerical study of organizations P and H mentioned above, where H had a clear superiority with respect to P in almost all cases.

A final comment here relates to the repartition of workload among the organization's members. In the example, the individual task definition is such that the workload is evenly distributed among the various decisionmakers. This is not always true, and it may very well be the case that G remains relatively low for all the organization's members but one, and reaches high values for that last member. This poses an important



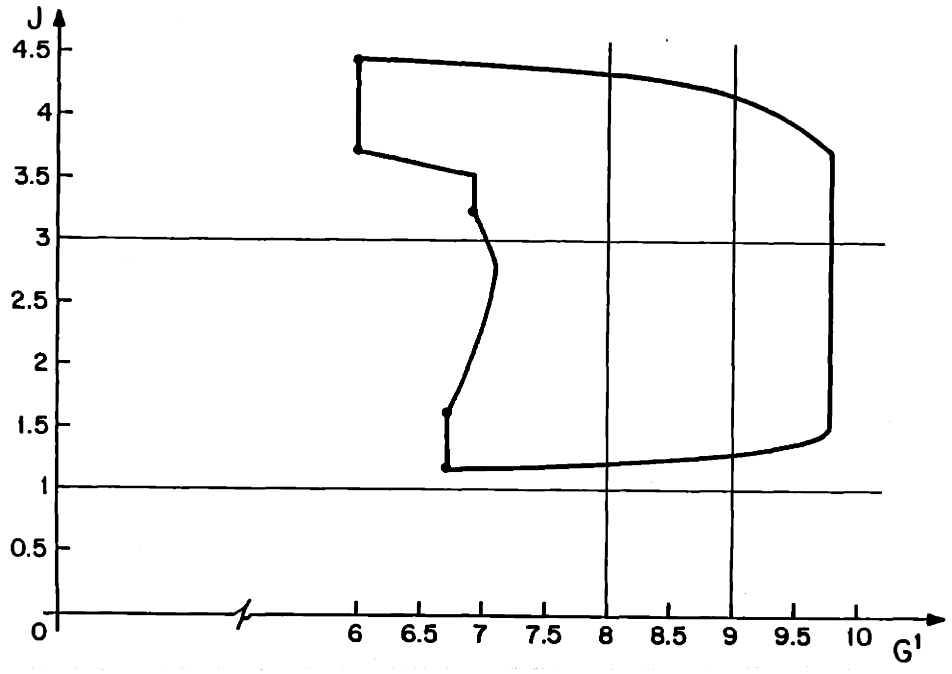
(a) DM^1

Figure 5.1 (P-W) Loci for Organization P with Centralized Databases



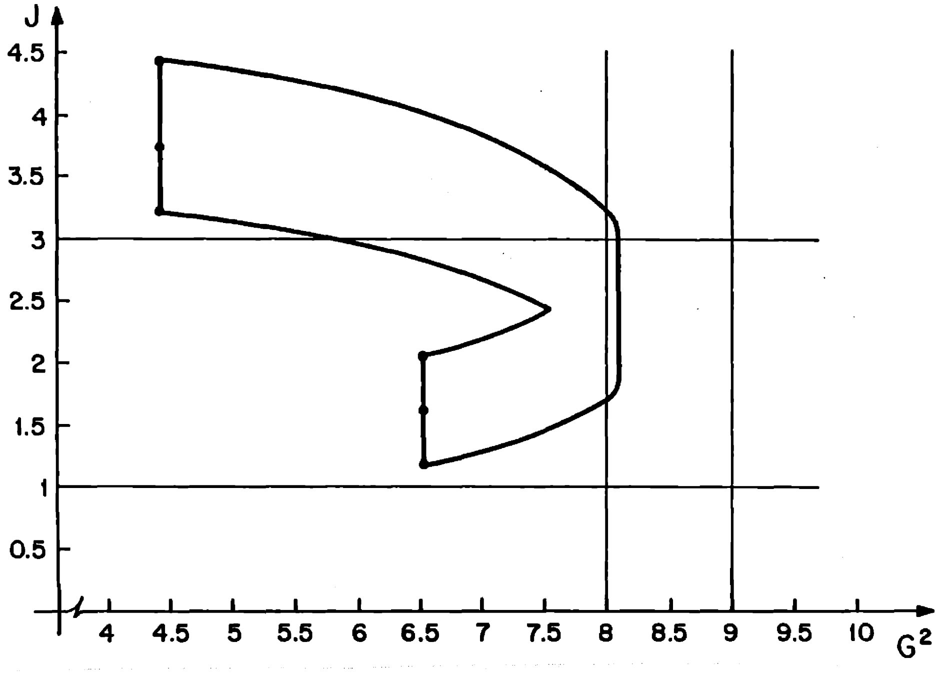
(b) DM^2

Figure 5.1 (P-W) Loci for Organization P with Centralized Databases



(a) DM^1

Figure 5.2 (P-W) Loci for Organization H with Centralized Databases



(b) DM^2

Figure 5.2 (P-W) Loci for Organization H with Centralized Datbases problem in organization design, because in such an event that decisionmaker constitutes a bottleneck for the organization, and impedes its proper functioning within reasonable workload and performance constraints.

5.2.2 The Importance of Input Interarrival Times in the Construction of (P-W) Loci

It was argued in the previous chapter that the introduction of time in the analysis of organizations was of critical importance because of the tool it provided for considering performance from a point of view more subject to the actual context in which the organization functions. This same argument holds for the analysis of workload, where activity rates are a closer indication of the pressure exercised on the decisionmakers than absolute activity, because of the time constraints present in real-world configurations. This section is concerned with the computation of activity rates for all decisionmakers and their use as a workload measure, in order to compare with and extend last section's results.

The entropy rate of a variable X was defined in Eq. (2.9). In the particular case at hand, where successive inputs are statistically independent and arrive at a rate τ, the entropy rate of the input is simply:

$$\bar{H}(x) = \frac{H(x)}{\tau} \tag{5.1}$$

Moreover, since the initialization of all variables in the system occurs at a rate equal to τ, as can be inferred from the proposition of section 4.4, the following equation applies here:

$$F^i = \frac{G^i}{\tau} \qquad i = 1,2,3 \tag{5.2}$$

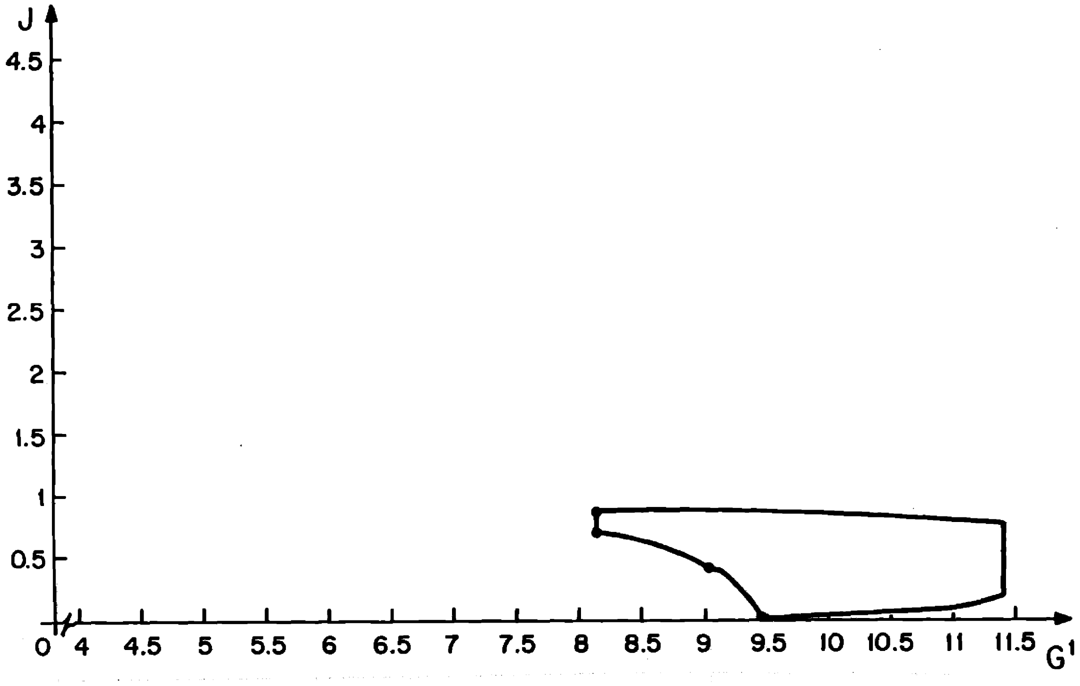
for either organization. The decisionmaker's activity rate will be called

$\bar{G}^i$ for practical purposes.

Equation (5.2) is used here by assuming $\tau = 1$ for organization P. Since centralized databases are used so far in this example, the results of section 4.5.1 show that the minimum input interarrival time for organization H is 11/3. The performance-workload loci for DM^1 and DM^2 are represented in Figure 5.3 (organization P) and Figure 5.4 (H), using $\bar{G}$ instead of G. These loci are easily derived from the diagrams in the previous section by reducing the locus along the workload axis by a factor 1 or 11/3 respectively. The performance levels do not change because they are calculated using only the nature of the response of the organization, independently of its time characteristics.

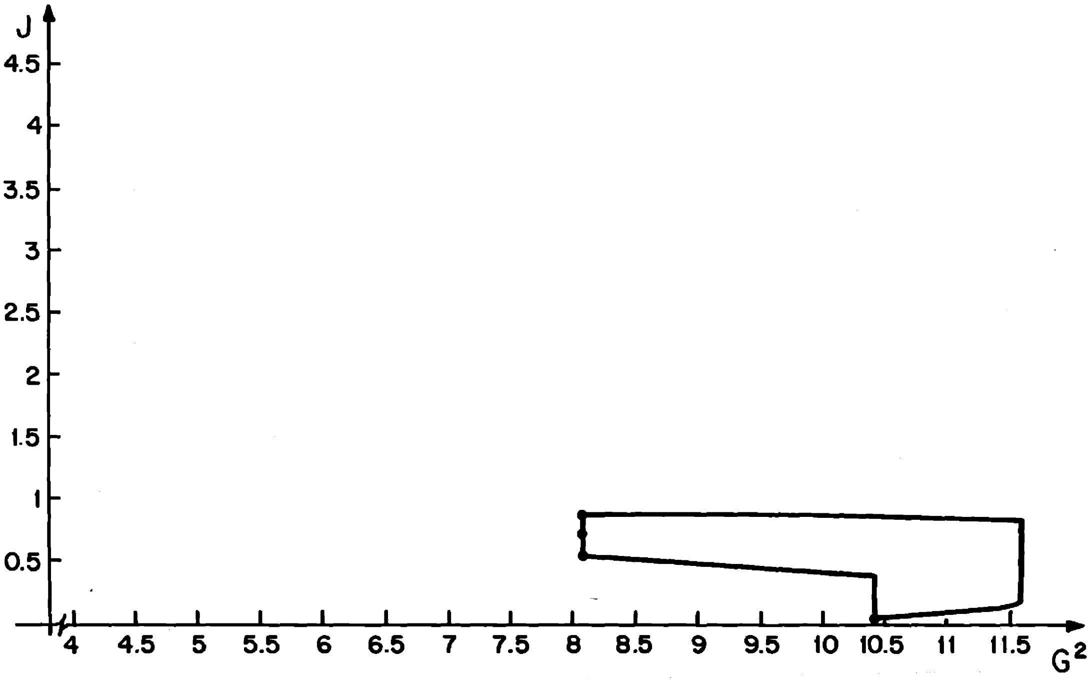
The use of activity rates in this instance has the effect of illustrating very accutely the tradeoff between timeliness or performance, and workload. Indeed, workload still varies between 8.1 and 11.6 and performance between 0 and 0.9 for organization P, whereas workload varies now between 1.2 and 2.7 and performance between 1.2 and 4.5 for organization H. In brief, measuring workload in terms of activity rates gives organization H a significant advantage as far as keeping decisionmakers' workload below a given maximum is concerned. However, another tradeoff appears here that involves the notion of timeliness. In effect, since in this section workload is a decreasing function of the mean interarrival time, low workload levels are obtained by allowing a high IT, which penalizes the organization in terms of its timeliness. This tradeoff is clearly perceived by looking at Table 4.1 and Figs. 5.2 and 5.4.

The previous section and this one so far have shown the general aspects of the tradeoffs between workload and both timeliness and performance at the organization level. The example, however, does not yield sufficient insight to the question of the possible correlations between performance and timeliness. Nevertheless, one can expect that an eventual correlation would depend on the structural characteristics of the organization (that determines its timeliness), on the task at hand (that



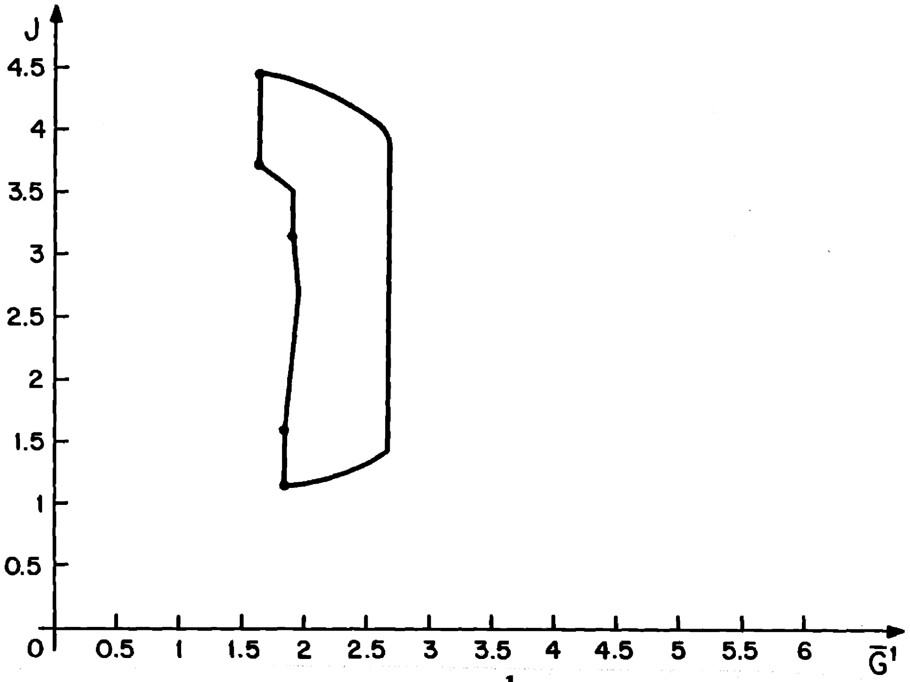
(a) DM^1

Figure 5.3 (P-W) Loci for P, Using Activity Rates



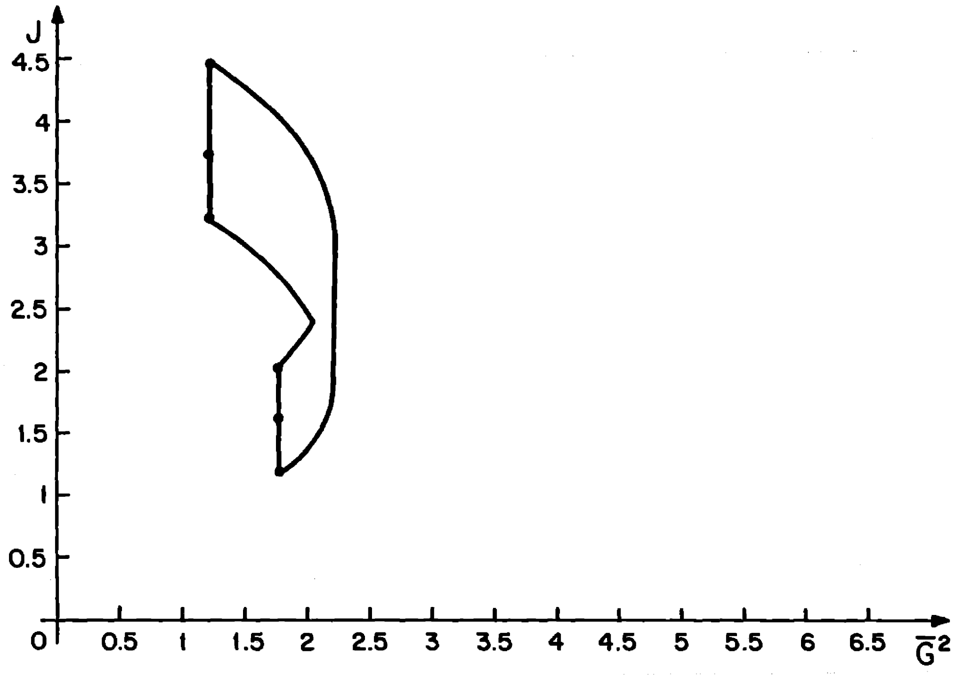
(b) DM^2

Figure 5.3 (P-W) Loci for P, Using Activity Rates



(a) DM^1

Figure 5.4 (P-W) Loci for H, Using Activity Rates



(b) DM^2

Figure 5.4 (P-W) Loci for H, Using Activity Rates influences greatly its performance, as seen in section 5.2.1), and on the criteria of satisficing performance and maximum allowable workload that delimit the regions of the (P-W) locus that are relevant to the study of a given organization.

Although evidence of the different tradeoffs pointed out in this section and the previous one is borne by the numerical study of two organizations equipped with a centralized database structure, these tradeoffs appear in the same way when a decentralized structure is used. In fact, they exist in the memoryless model of the organization as well, even though they are significantly enhanced by the introduction of databases, which increase both the overall workload level (as seen in Chapter 3), and the organization performance's sensitivity to time constraints (as seen in Chapter 4). In the next section, however, a result specific to the adoption of databases in a DSS structure will be presented.

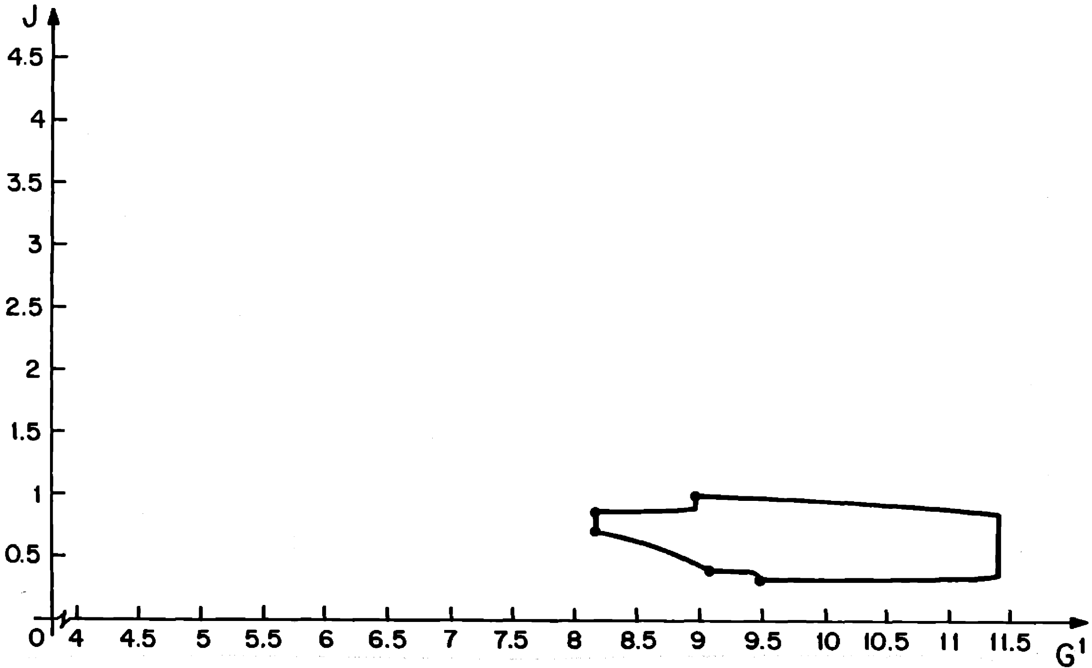
5.2.3 The Effect of Uncoordinated Databases on Performance

The major practical problem posed by the use of several individual databases has been mentioned to be the difficulty of updating all the information they contain at the same time, so that organization members can all dispose of data with the same level of relevance and accuracy. This problem is compounded in tactical configurations, where the databases can be largely disseminated and the data changes at a high rate. If proper updating is not assured, some decisionmakers could use consistenty wrong information as a support for making important decisions, which therefore could result in costly errors for the entire organization. Such a possibility is the subject of this section, where two different updating sequences are presented and their impacts on the type 1 performance of the organization shown on the (P-W) locus. (The same kind of problems can be encountered with centralized databases, if the information management program used by the DSS is not able to perfectly synchronize updating or if access to the database in real-time is not adequately managed).

Both examples are drawn from organization P; indeed, the problems raised by uncoordinated databases are similar whatever the organization, and the results derived in this section will be of general relevance to the study of databases. In the first scheme, DM^2's and DM^3's RS databases are assumed to be updated as stated in Chapter 4, at $\tau + 0$, in coordination with the input arrival. DM^1's DBRS, however, is updated at $\tau + \tau$, with a delay of τ over the input to which the data correspond. This means that DM^1 selects his response based on data he receives from $DBRS^1$, while this data is relevant to the previous selection he made ... In terms of type 1 performance, DM^1 will read wrong information every time the contents of the database should have been changed, but would not be until the next iteration. The probability of this happening is computed, using the characteristics of the data taken as an input to the system, and provides an increased risk for a mismatch between the actual and the desired responses to occur. New performance levels for each pure organization strategy are derived (Table B.3 in Appendix B), and a performance-workload locus can be drawn (see Fig. 5.5 (a)). The main results here are a move upwards of the original locus, and a reduction in the range of the performance levels, which vary now from 0.35 to 1.0 as opposed to a range of 0 to 0.9 in the perfectly coordinated (or the centralized) database case; this represents a drop of 29% in the average performance of the organization.

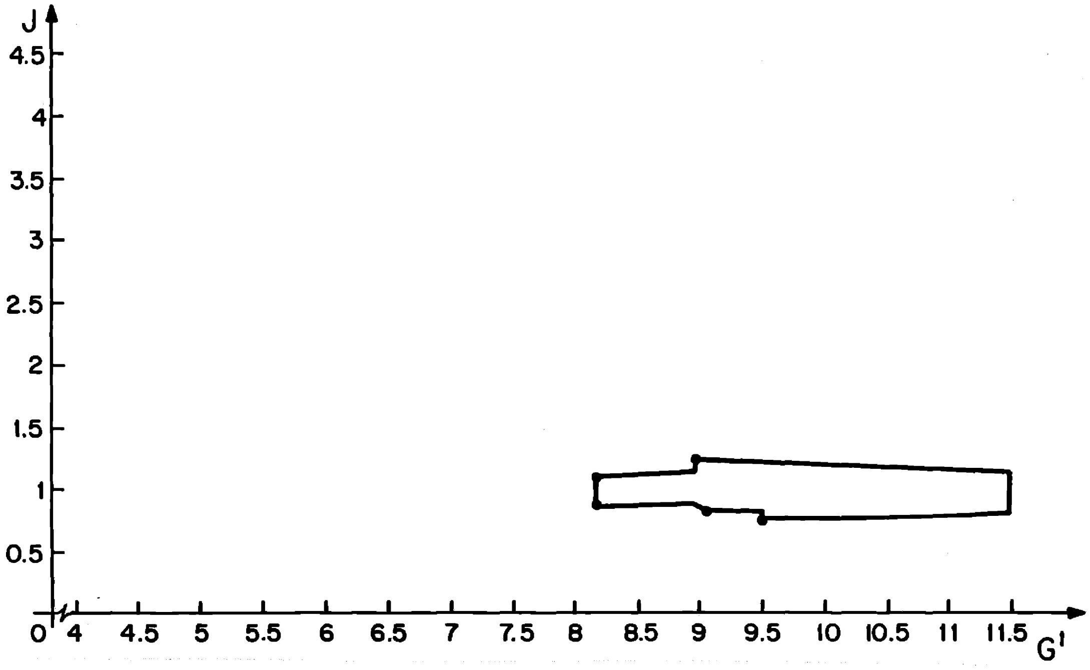
The second scheme exhibits a less coordinated updating sequence: $DBRS^2$ is updated at $\tau + 0$, $DBRS^1$ at $\tau + \tau$, and $DBRS^3$ at $\tau + 2\tau$. DM^1 and DM^3 both now have a greater propensity to make the wrong decision, and the resulting (P-W) locus is presented in Fig. 5.5 (b). The maximum performance is now 0.8, which is very close to what the minimum performance was in the coordinated case, and the lowest performance level is 1.2. The range of possible performance levels has further shrunk, and the drop in average performance with respect to the original case is 68%.

This section provides a quantitative measure for the evaluation of DSS structures. The measure yields an answer to the question: what configuration to adopt, when timeliness considerations do not point to a



(a) One Uncoordinated Database

Figure 5.5 (P–W) Loci for DM^1, in P with Uncoordinated Databases



(b) Two Uncoordinated Databases

Figure 5.5 (P–W) Loci for DM^1 in P with Uncoordinated Databases clear superiority of any structure. In effect, even when databases are updated in a totally coordinated fashion, the organization designer has to consider the consequences of a failure in the updating process, that would result in the kind of errors described above. Of course, these problems are all the more costly as the updating process is originally designed to be uncoordinated, for reasons of economic or technical feasibility. It can be predicted , however, that lack of coordination leads to less costly errors in commercial or business organizations, as in a case of slowly changing inventories for instance, than in military organizations where the global tactical context can be totally reversed between one iteration and the next.

5.2.4 The Evaluation of Organizations P and H Against Time Characteristics

Timeliness was proven to be a function of the structural characteristics of the organization, and not to be influenced by the mission or the context as far as the comparison of organization P and H is concerned. Indeed, Table 4.1 shows that the only quantity that depends on the specific tasks performed by the decisionmakers is τ, the transition processing time (Table 4.1 is reproduced below as 5.1); but since interarrival time and response times are scaled on τ, a change in τ does not question the relative advantages of P and H over each other. This section purports to put the general results of section 4.6 into perspective, and show their meaning in the example at hand.

At this point, any confusion about the definition of τ should be avoided: in this section τ is the transition processing time, and not the actual input interarrival time, as was the case in Section 5.2.1. It is fixed, and considered as a characteristic of the system; it determines the minimum allowable interarrival time, IT, which is another structural feature of the organization. $(IT)^{1}$ represents the maximum rate at which the organization can porocess its inputs. If the actual input interarrival time is smaller than IT, then the decisionmakers will have to ignore part of or all the signals, even though they might not be overloaded. The concept used in this section is different from what determined the workload constraint in Section 5.2.1. In that instance, τ was the actual input interarrival time, and varying it did have an influence on the decisionmakers' workload, because of Eq. (2.20). Taking IT into account in examining the functioning of the organization adds significant insights to the evaluation process, as discussed below.

TABLE 5.1 TIME CHARACTERISTICS OF ORAGNIZATIONS P AND H

	Centralized db	Decentralized db
IT(P)	τ	τ
IT(H)	$11\tau/3$	$10\tau/3$
RT(P)	$19\tau/3$	$17\tau/3$
RT(H)	8τ	7τ

IT = minimum allowable Interarrival Time; RT = Response Time

In what follows, τ will be assumed to be equal to 1 second. This is to say that IT(P) will be 1 second as well, as IT(H) 11/3 seconds (or 10/3 sec., depending on the database configuration). In any event, IT(H) is much greater than IT(P) and may handicap organization H if it is attacked by units arriving at a high rate. As an example, consider a wave of enemy planes attacking the battlegroup: If no defense can be initiated without processing every input through the SA stage, then the anti-aircraft batteries cannot shoot at a rate higher than one missile every 3.6 sec. If a new plane arrives once every second, then P is an adequate structure, which H is far from being. A supplementary handicap of organization H appears when response times are taken into account: the battlegroup will need 8 seconds to fire on a plane after it is detected; this might be too long if the plane is very close to the battlegroup when its presence is suspected.

When the platforms that the organization has to deal with are slower units, like submarines or surface ships, organization H's timeliness disadvantage is less critical, because of the longer time available for constructing an adequate response and because of the smaller target arrival rate. In fact, the latter can be so small as to make any difference between IT(P) and IT(H) seem irrelevant. Since the organizations designed in this example have to deal with both slow and fast platforms, one has to consider the relative probability of being attacked by planes or ships, and weigh it by the expected costs in each alternative during the evaluation of the two organizational structures.

CHAPTER VI

CONCLUSIONS AND DIRECTIONS FOR FUTURE RESEARCH

6.1 CONCLUSIONS

One limitation of most of the decisionmaking models developed so far within the Information Theory framework was their being memoryless. This thesis was concerned with relaxing that basic constraint, and allowing decisionmakers to take advantage of a primary human function: memory. For this purpose, the use of database networks was introduced into the organization, storing and delivering data received from an external updating source. Information storage and access devices were presented in two alternative structures, centralized and decentralized, and the activity expressions as defined in the Partition Law of Information Rates were derived in each case.

A major non-information theoretic aspect of databases was identified to be their high sensitivity to time constraints and the necessity for an organization to set precise rules and instants for their access and their updating. Protocols were defined and guidelines for their construction were proposed, to provide an analytical framework within which time characteristics can be examined. The two main time-related attributes of an organization, its mean input interarrival time (IT) and its databases' updating instants (UD), were determined for four distinct database-organization structures, made of either a hierarchical or a parallel organizational structure along with either a centralized or decentralized database network. These examples were the basis for the introduction of a new performance evaluation criterion, relying on the organization's time characteristics mentioned above, and called type 2 performance.

The last chapter's objective was to address the quantitative aspects of any comparison between parallel and hierarchical decisionmaking structures, based on studying two specific organizations. It was shown that no comparison could be made in an absolute fashion, and that the relative advantages and drawbacks of organizations P and H were mission-dependent. It was also shown that the use of the total activity rate as a workload measure shed more light on the tradeoff between workload and both timeliness and performance, respectively. A third result was that the possibility of loss in updating coordination caused a considerable deterioration of the average organization performance. Finally, both organizations' timeliness was examined as it applies to the specific context of the example, illustrating the importance of minimum allowable interarrival times and response times in terms of the risk to which they can expose military organizations.

6.2 DIRECTIONS FOR FUTURE RESEARCH

Although the basic organization model is now equipped with a means of incorporating the use of databases into the decisionmaking process, considerable research still has to be conducted regarding the extension of the admissible sources of data, to inclued internally generated information. This will encounter major theoretical obstacles within the Information Theory framework, which does not presently provide any tool for accounting for the order in which inputs are received. This order intervenes in a most critical fashion in the generation of data that reflects the successive iterations of the system, and carries lessons that can be learned. This kind of data is stored in what is called "recursive databases". A comprehensive approach to the modeling of recursive databases cannot be undertaken without relaxing the steady-state assumptions under which the present model is valid. It is, however, possible to obtain an unambiguous assessment of the limitations of the present framework for this endeavor. This could be achieved by analyzing the sensitivity of information theoretic results to the assumed dependence of variables' entropies on each other's chronological order (especially for successive data inputs), as a first step towards a more flexible interpretation of the steady-state research framework.

On a different level in the study of organizations, the aparition of a third evaluation criterion, namely timeliness, raises two issues: the first one is related to the design of a quantitative method of measuring it, as is the case for performance; the second issue is that of the relationships between timeliness and the two already existing criteria: workload and performance.

In section 5.2.4, a first attempt towards considering timeliness as a measurable attribute was made. It was argued that time characteristics were to be analyzed with respect to the different inputs received by the system, and that an organization's absolute timeliness advantage could be either enhanced or made irrelevant by its specific context. More research has to be conducted in order to establish desired interarrival and response times, as opposed to the actual quantities, and a probability distribution related to the nature of the input, with a cost assigned to each error. Thus, timeliness could be expressed on a scale that is dependent on the mission solely; that would be a realistic measure for comparing alternative organizational structures in a given context.

The second issue raised by the consideration of timeliness is the investigation of possible links between it and the traditional performance attribute. These links would be inherent to the structure of the organization, where critical parameters could be identified and varied in an instructive way. (It is important to differentiate between the effects of input characteristics and those of the organization's structual parameters). If the two types of performance can be related to each other unambiguously, then a further refinement of organization evaluation will be the creation of a common scale on which both could be measured. Even better, the three criteria present in this thesis could be some of the main attributes in the analysis and evaluation of an organization using the System Efefctiveness methodology. Indeed, this framework allows for the global appraisal of a system according to the different preferences of its designers, its users and the high-level decisionmakers - as opposed to the organization's members, called "users" here - and to its specific mission and context. (For a detailed description of the methodology and its roots in Utility Theory , see Karam, 1985).

An adaptation of System Effectiveness to the specific field of military organizations will provide a quantitative means of considering all the criteria of interest here, among many others, in designing an organization best suited to the mission at hand, subject to the satisfaction of the authorities, and its adaptation to its users, the human decisionmakers.

REFERENCES

Ashby, W. R., "Measuring the Internal Information Exchange in a System," Cybernetica, Vol. 8, 1965, pp. 5-22.

Bailey, R. W., Human Performance Engineering, Prentice-Hall, Inc. , Englewood Cliffs, NJ, 1982, pp. 1-190.

Boettcher, K. L., "An Information-Theoretic Model of the Decisionmaker," MS Thesis, LIDS-TH-1096, Laboratory for Information and Decision Systems, MIT, Cambridge, MA, July 1981.

Boettcher, K. L. and A. H. Levis, "Modeling the Interacting Decisionmaker with Bounded Rationality," IEEE Trans. on Systems, Man, and Cybernetics, Vol. SMC-12, No. 3, May/June 1982.

Boettcher, K. L., and A. H. Levis, "Modeling and Analysis of Teams of Interacting Decisionmakers with Bounded Rationality," Automatica, Vol. 19, No. 6, 1983, pp. 703-709.

Chyen, G. H-L, "Information Theoretic Models of Preprocessors and Decision Aids," MS Thesis, LIDS-TH-1377, Laboratory for Information and Decision Systems, MIT, Cambridge, MA, June 1984.

Conant, R. C., "Laws of Information Which Govern Systems," IEEE Trans. on Systems, Man, and Cybernetics, Vol. SMC-6, No. 4, April 1976.

Cothier, P. H., "Assessment of Timeliness in Command and Control," MS Thesis, LIDS-TH-1391, Laboratory for Information and Decision Systems, MIT, Cambridge, MA , August 1984.

De Waele, M., "Managerial Style and the Design of Decision Aids," Omega, Vol. 6, No. 1, 1978, pp. 5-13.

Gallager, R. G., Information Theory and Reliable Communication, John Wiley and Sons, Inc., New York, 1968.

Hall, S. A., "Information Theoretic Models of Storage and Memory," MS Thesis, LIDS-TH-1232, Laboratory for Information and Decision Systems, MIT, Cambridge, MA, June 1982.

Huber, G. P., "The Nature of Organizational Decisionmaking and the Design of Decision Support Systems," MIS Quarterly, June 1981.

Huber, G. P., "Cognitive Style as a Basis for MIS and DSS Designs: Much ADO About Nothing," Management Science, Vol. 29, No. 5, May 1983.

Karam, J. G., "Effectiveness Analysis of Evolving Systems," MS Thesis, Technology and Policy, MIT, Cambridge, MA, January 1985.

Keen, P. G. W., "Value Analysis: Justifying Decision Support Systems," MIS Quarterly, March 1981.

Keen, P. G. W., and M. Scott-Morton, Decision Support Systems: An Organizational Perpesctive, Addison-Wesley, Reading, MA, 1958.

Levis, A. H., "Information Processing and Decisionmaking Organizations: A Mathematical Description," Proc. 6th MIT/ONR Workshop on C^3 Systems, LIDS-R-1354, MIT, Cambridge, MA, July 25-29, 1983.

Levis, A. H. and K. L. Boettcher, "Decisionmaking Organizations with Acyclical Information Structures," IEEE Trans. on Systems, Man and Cybernetics, Vol. SMC-13, No. 3, May/June 1983.

March, J. G., "Bounded Rationality, Ambiguity, and the Engineering of Choice," Bell Journal of Economics, Vol. 9, 1978, pp. 587-608.

McGill, W.J., "Multivariate Information Structures," Psychrometrika, Vol. 19, No. 2, 1954, pp. 97-116.

Peterson, J. L., Petri Net Theory and the Modeling of Systems, Prentice-Hall, Inc., Englewood Cliffs, NJ, 1981.

Tabak, D. and A. H. Levis, "Petri Net Representation of Decision Models,"Proc. 7th MIT/ONR Workshop on C^3 Systems, San Diego, CA, June 11-14, 1984.

Sen, P. and R. F. Drenick, "Information Processing in Man-Machine Systems," Proc. Fourth MIT/ONR Workshop on C^3 Systems, LIDS-R-1159, Laboratory for Information and Decision Systems, Cambridge, MA, October 1981.

Shannon, C. E. and W. Weaver, The Mathematical Theory of Communication, University of Illinois Press, 1949.

Sheridan, T. B., and W. R. Ferrell, Man-Machine Systems, MIT Press, 1974.

Simon, H. A., Administrative Behavior, Mc Millan, New York, 1947.

Simon, H. A., "Rational Decision Making in Business Organizations," American Economic Review, September 1979.

Simon, H. A. and J. G. March, Organizations, John Wiley and Sons, Inc., New York, 1958.

Sprague, Jr., R. H., "A Framework for the Development of DSS," MIS Quarterly, December 1980.

Sprague, Jr., R. H. and E. D. Carlson, Building Effective Decision Support Systems, Prentice-Hall, Englewood Cliffs, NJ, 1982.

Stabile, D. A. and A. H. Levis, "The Design of Information Structures: Basic Allocation Strategies for Organizations," Large-Scale Systems, Vol. 6, 1984, pp. 123-132.

APPENDIX A

DERIVATION OF THE ACTIVITY RATE TERMS FOR THE DATABASE EQUIPPED ORGANIZATION

Since it is proven in Chapter 3 that the activity rate expressions for decentralized databases are similar to those for the centralized structure, this Appendix will be concerned with developing the latter terms only. The model of the decisionmaker using a centralized database is reproduced here to help understand the relationships between the variables of the system (Fig. A.1). Also, the variables and quantities in what follows are all related to DM^n, except when specified otherwise; the variable y will refer to both y^n and y^{no}:$y = \{y^n, y^{no}\}$.

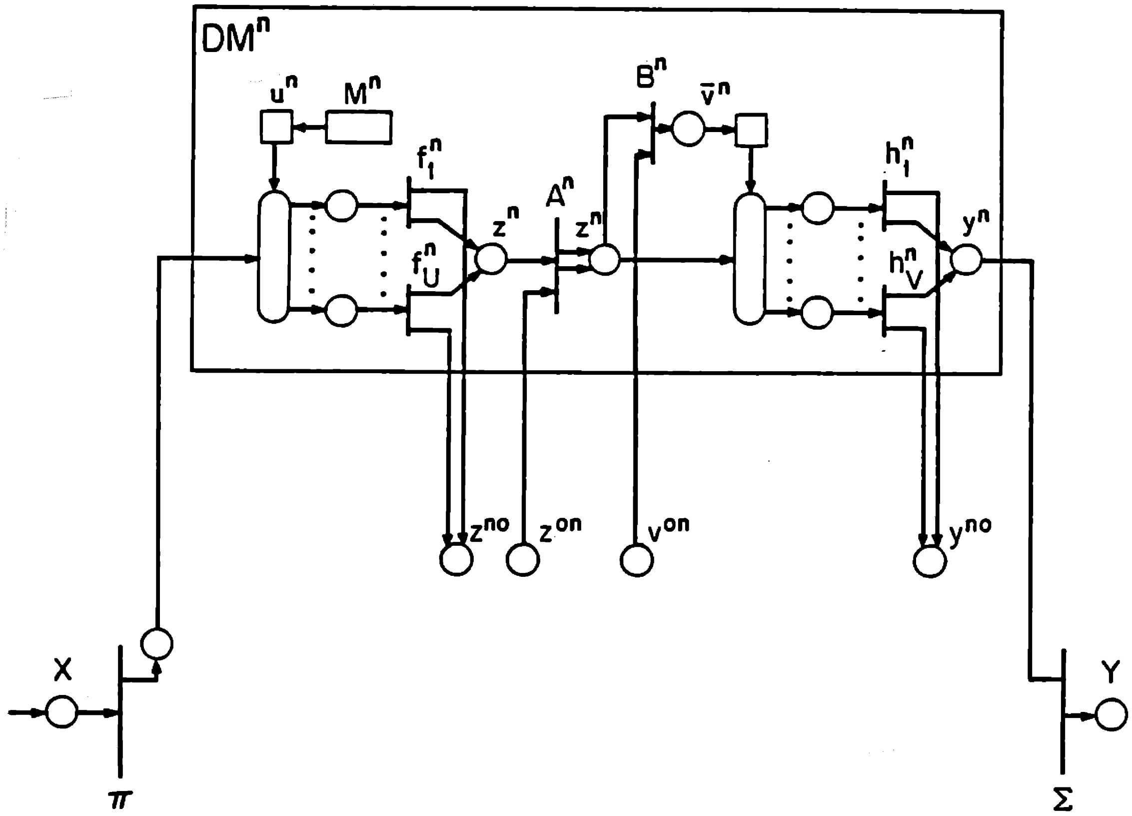


Figure A.1 Petri Net Representation of DM^n Using Centralized Databases

Centralized Database

Throughput Rate: F_t measures the total relatedness of the decisionmaker's output to his input. In this case, it is computed in the following way:

$$F_t = \bar{T}(x,d_{SA},z^{on},v^{on},d_{RS}:u,z^{no},\bar{z},\bar{v},y) \tag{A.1}$$

Blockage Rate: The blockage rate is obtained by applying the auxiliary equation to the PLIR, Eq. (2.17), written here as follows:

$$F_t + F_b + F_r = \bar{H}(x,d_{SA},z^{on},v^{on},d_{RS}) \tag{A.2}$$

The rejection rate has been assumed to be zero for all algorithms, in Chapter 2. That gives the blockage rate:

$$F_b = \bar{H}(x,d_{SA},z^{on},v^{on},d_{RS}) - F_t \tag{A.3}$$

Noise Rate: By definition, the noise rate present in DM^n is given by:

$$F_n = \bar{H}_{x,d_{SA},z^{on},v^{on},d_{RS}}(u,W^1,\ldots,W^U,z,W^A,W^B,\bar{v},W^{U+1},\ldots W^{U+V},y) \tag{A.4}$$

$$F_n = \bar{H}_{x,d_{SA},z^{on},v^{on},d_{RS}}(u) + \bar{H}_{x,d_{SA},z^{on},v^{on},d_{RS},u}(W^1,\ldots W^U,z)$$

$$+ \bar{H}_{x,d_{SA},z^{on},v^{on},d_{RS},u,W^1,\ldots,W^U,z}(W^A,W^B,\bar{v})$$

$$+ \bar{H}_{x,d_{SA},z^{on},v^{on},d_{RS},u,W^1,\ldots,W^U,z,W^A,W^B,\bar{v}}(W^{U+1},\ldots,W^{U+V},y) \tag{A.5}$$

Equation (A.5) is obtained by using the identity in Eq. (2.7). The second and last terms of Eq. (A.5) are zero because the conditioning variables determine in fact the variables between parentheses. Furthermore, the decision strategy p(u) has been assumed to be independent of all other variables. Equation (A.5) can then be rewritten as:

$$F_n = \bar{H}(u) + \bar{H}_{x,d_{SA},z^{on},v^{on},d_{RS},u,W^1,\ldots,W^U,z}(W^A)$$

$$+ \bar{H}_{x,d_{SA},z^{on},v^{on},d_{RS},u,W^1,\ldots,W^U,z,W^A}(W^B,\bar{v}) \tag{A.6}$$

W^A is fully determined by the knowledge of z and z^{on}. In addition, W^B is comprised of: v, and $\{W^B - v\} = \tilde{W}^B$ (see Boettcher, 1981), and the conditioning in the last term of Eq. (A.6) is equivalent to conditioning only on $\bar{z}$ and v^{on}. That gives for Eq. (A.6):

$$F_n = \bar{H}(u) + \bar{H}_{v^{on},\bar{z}}(v,\tilde{W}^B,\bar{v}) \tag{A.7}$$

$$F_n = \bar{H}(u) + \bar{H}_{\bar{z},v^{on}}(v) + \bar{H}_{\bar{z},v^{on},v}(\tilde{W}^B,\bar{v}) \qquad (A.8)$$

The last term in Eq. (A.8) is zero since knowledge of v^{on},$\bar{z}$ and v determines all the variables in B, and $\bar{v}$ as well. The second term of Eq. (A.8) becomes $\bar{H}_{\bar{z}}(v)$ because the decision strategy in the RS stage has been assumed to depend only on the final situation assessment, $\bar{z}$. The noise rate term is then:

$$F_n = \bar{H}(u) + \bar{H}_{\bar{z}}(v) \qquad (A.9)$$

<u>Coordination Rate</u>: For computational purposes, the total system is divided into the following four subsystems:

$$S^I = \{u,W^1,\ldots,W^U,z\} \qquad (A.10)$$

$$S^A = \{W^A\} \qquad (A.11)$$

$$S^B = \{W^B\} \qquad (A.12)$$

$$S^{II} = \{\bar{v},W^{U+1},\ldots,W^{U+V},y\} \qquad (A.13)$$

In this case, the following hold (see Eq. (2.6)):

$$S = S^I \cup S^A \cup S^B \cup S^{II} \qquad (A.14)$$

and

$$F_c = F_c^I + F_c^A + F_c^B + F_c^{II} + \bar{T}(S^I:S^A:S^B:S^{II}) \tag{A.15}$$

The first coordination term, F_c^I, is:

$$F_c^I = \bar{T}(u:w_1^1:w_2^1:\ldots:w_{\alpha_1}^1:w_1^2:\ldots:w_{\alpha_2}^2:\ldots:w_{\alpha_U}^U:z) \tag{A.16}$$

$$F_c^I = \sum_{i=1}^{U}\sum_{j=1}^{\alpha_i} \bar{H}(w_j^i) + \bar{H}(u) + \bar{H}(z,z^{no}) - \bar{H}(W^1,W^2,\ldots,W^U,u,z) \tag{A.17}$$

The joint uncertainty term can be rewritten, using Eq. (2.7):

$$\begin{aligned}\bar{H}(W^1,W^2,\ldots,W^U,u,z) &= \bar{H}(u) + \bar{H}_u(W^1) + \bar{H}_{u,W^1}(W^2)\\ &\quad +\ldots+ \bar{H}_{u,W^1,\ldots,W^{U-1}}(W^U) + \bar{H}_{u,W^1,\ldots W^U}(z)\end{aligned} \tag{A.18}$$

The last term of Eq. (A.18) is zero since knowledge of all variables of the SA stage determines z. Consider now any term $\bar{H}_{u,W^1,\ldots,W^j}(W^{j+1})$, $1 < j \leq U-1$. The algorithm variable sets have been assumed to be disjoint: $W^i \cap W^j = \phi$, $i \neq j$. Furthermore, data supplied by DBSA comes from a source external to the system, which prevents the values taken by W^j on iteration (t) from depending on values taken by W^i, $i \neq j$ on iteration (t-1). This shows that any conditioning of $\bar{H}(W^j)$ on u and W^i, $i \neq j$, is equivalent to conditioning only on u, because u is the variable that determines whether an algorithm is activated or not. This gives for Eq. (A.18):

$$\bar{H}(W^1,W^2,\ldots,W^U,u,z,) = \bar{H}(u) + \sum_{i=1}^{U} \bar{H}_u(W^i) \tag{A.19}$$

The coordination rate F_c^I becomes:

$$F_c^I = \sum_{i=1}^{U} \sum_{j=1}^{\alpha_i} \bar{H}(w_j^i) - \sum_{i=1}^{U} \bar{H}_u(W^i) + \bar{H}(z,z^{no}) \tag{A.20}$$

Add and subtract:

$$\sum_{i=1}^{U} \sum_{j=1}^{\alpha_i} \bar{H}_u(w_j^i)$$

to get:

$$\begin{aligned} F_c^I = & \sum_{i=1}^{U} [\sum_{j=1}^{\alpha_i} \bar{H}_u(w_j^i) - \bar{H}_u(W^i)] \\ & + \sum_{i=1}^{U} [\sum_{j=1}^{\alpha_i} (\bar{H}(w_j^i) - \bar{H}_u(w_j^i))] + \bar{H}(z,z^{no}) \end{aligned} \tag{A.21}$$

$$F_c^I = \sum_{i=1}^{U} [\sum_{j=1}^{\alpha_i} \bar{H}_u(w_j^i) - \bar{H}_u(W^i)] + \sum_{i=1}^{U} \bar{T}(w_j^i:u) + \bar{H}(z,z^{no}) \tag{A.22}$$

Define $\bar{g}_c^i$ to be the internal coordination rate of algorithm i:

$$\bar{g}_c^i = \sum_{j=1}^{\alpha_i} \bar{H}_u(w_j^i) - \bar{H}_u(W^i) \tag{A.23}$$

The second term of Eq. (A.22) has been proven by Hall (1982, Appendix A) to be:

$$\sum_{i=1}^{U} \sum_{j=1}^{\alpha_i} \bar{T}(u:w_j^i) = \sum_{i=1}^{\alpha_i} \alpha_i H(p(u=i)) \tag{A.24}$$

with:

$$H(p(u=i)) = H(p_i) = p_i \log_2 p_i + (1-p_i) \log_2 (1-p_i) \tag{A.25}$$

and α_i being the number of variables of algorithm i to be reinitialized at each interation. The result of Eq. (A.24) is true if inputs to the SA stage arrive every second. If inputs arrive every τ_{SA} seconds in general, it is straightforward to show that:

$$\sum_{i=1}^{U} \sum_{j=1}^{\alpha_i} \bar{T}(u:w_j^i) = \sum_{i=1}^{U} \frac{\alpha_i}{\tau_{SA}} H(p_i)) \tag{A.26}$$

Using the results derived in Eq. (A.23) and (A.26), F_c^I becomes:

$$F_c^I = \sum_{i=1}^{U} (\bar{g}_c^i(x,d_{SA}) + \frac{\alpha_i}{\tau_{SA}} H(p_i)) + \bar{H}(z,z^{no}) \tag{A.27}$$

The second term of Eq. (A.15), F_c^A, can be written:

$$F_c^A = \bar{T}(w_1^A, w_2^A, \ldots, w_{\alpha_A}^A) \tag{A.28}$$

$$F_c^A = \sum_{i=1}^{\alpha_A} \bar{H}(w_i^A) - \bar{H}(W^A) \tag{A.29}$$

$$F_c^A = \bar{g}_c^A\,(p(z, z^{on})) \tag{A.30}$$

Note here that the internal coordination rate of algorithm A is not conditioned on any switch. Similarly, one gets:

$$F_c^B = \bar{g}_c^B\,(p(\bar{z}, v^{on})) \tag{A.31}$$

Finally, F_c^{II} can be computed using the technique developed in the derivation of F_c^I, and can be written as:

$$F_c^{II} = \sum_{j=1}^{V} (\bar{g}_c^{U+j}(p(\bar{z}, d_{RS})) + \frac{\alpha_{U+j}}{\tau_{RS}} H(p_j)) + \bar{H}(y) \tag{A.32}$$

which gives:

$$F_c^I + F_c^A + F_c^B + F_c^{II} = \sum_{i=1}^{U} \left(\bar{g}_c^i(p(x,d_{SA})) + \frac{\alpha_i}{\tau_{SA}} H(p_i)\right) + \bar{H}(z,z^{no})$$

$$+ \bar{g}_c^A \, (p(z,z^{on})) + \bar{g}_c^B \, (p(\bar{z},v^{on}))$$

$$+ \sum_{j=1}^{V} \left(\bar{g}_c^{U+j}(p(\bar{z},d_{RS})) + \frac{\alpha_{U+j}}{\tau_{RS}} H(p_j)\right) + \bar{H}(y) \qquad (A.33)$$

The last term of Eq. (A.15) can be developed in the following way (see Boettcher, 1981):

$$\bar{T}(S^I:S^A:S^B:S^{II}) = \bar{T}(S^I:S^A) + \bar{T}(S^I,S^A:S^B) + \bar{T}(S^I,S^A,S^B:S^{II}) \qquad (A.34)$$

The first term in Eq. (A.34) is, by definition:

$$\bar{T}(S^I:S^A) = \bar{H}(S^A) - \bar{H}_{S^I}(S^A) \qquad (A.35)$$

It is equivalent to using only the variables that determine all the others, in both subsystems:

$$\bar{T}(S^I:S^A) = \bar{H}(z,z^{on}) - \bar{H}_{u,x,d_{SA}}(z,z^{on}) \qquad (A.36)$$

$$\bar{T}(S^I:S^A) = \bar{H}(z) + \bar{H}_z(z^{on}) - \bar{H}_{u,x,d_{SA}}(z) - \bar{H}_{u,x,d_{SA},z}(z^{on}) \qquad (A.37)$$

The third term in Eq. (A.37) is null, because z is fully determined by the conditioning variables. Equation (A.37) becomes:

$$\bar{T}(S^I:S^A) = \bar{H}(z) + \bar{T}_z(u,x,d_{SA}:z^{on}) \quad (A.38)$$

But the set-up of the organization is such that u and z^{on} are independent. Equation (A.38) becomes then:

$$\bar{T}(S^I:S^A) = \bar{H}(z) + \bar{T}_z(x,d_{SA}:z^{on}) \quad (A.39)$$

The second term in Eq. (A.34) is:

$$\bar{T}(S^I,S^A:S^B) = \bar{H}(S^B) - \bar{H}_{S^I,S^A}(S^B) \quad (A.40)$$

It is equivalent to:

$$\bar{T}(S^I,S^A:S^B) = \bar{H}(\bar{z},v^{on},v) - \bar{H}_{u,x,d_{SA},z^{on}}(\bar{z},v^{on},v) \quad (A.41)$$

$$\bar{T}(S^I,S^A:S^B) = \bar{H}(\bar{z}) + \bar{H}_{\bar{z}}(v^{on}) + \bar{H}_{\bar{z},v^{on}}(v) - \bar{H}_{u,x,z^{on},d_{SA}}(\bar{z})$$

$$- \bar{H}_{u,x,d_{SA},z^{on},\bar{z}}(v^{on}) - \bar{H}_{u,x,d_{SA},z^{on},\bar{z},v^{on}}(v) \quad (A.42)$$

The third and last terms are equal, because v depends only on $\bar{z}$. The fourth term is null, because $\bar{z}$ is fully determined by the conditioning variables. Equation (A.42) becomes:

$$\bar{T}(S^I,S^A:S^B) = \bar{H}(\bar{z}) + \bar{T}_{\bar{z}}(u,x,z^{on},d_{SA}:v^{on}) \tag{A.43}$$

$$\bar{T}(S^I,S^A:S^B) = \bar{H}(\bar{z}) + \bar{T}_{\bar{z}}(x,z^{on},d_{SA}:v^{on}) \tag{A.44}$$

The third term in Eq. (A.34) can be written as:

$$\bar{T}(S^I,S^A,S^B:S^{II}) = \bar{H}(S^{II}) - \bar{H}_{S^I,S^A,S^B}(S^{II}) \tag{A.45}$$

$$\bar{T}(S^I,S^A,S^B:S^{II}) = \bar{H}(\bar{z},\bar{v},d_{RS}) - \bar{H}_{u,x,z^{on},d_{SA},v^{on},v}(\bar{z},\bar{v},d_{RS}) \tag{A.46}$$

$$\begin{aligned}\bar{T}(S^I,S^A,S^B:S^{II}) = {} & \bar{H}(\bar{z},\bar{v}) + \bar{H}_{\bar{z},\bar{v}}(d_{RS}) - \bar{H}_{u,x,z^{on},d_{SA},v^{on},v}(\bar{z}) \\ & - \bar{H}_{u,x,z^{on},d_{SA},v^{on},v,\bar{z},\bar{v}}(d_{RS})\end{aligned} \tag{A.47}$$

The third and the fourth terms of Eq. (A.47) are zero, because the variables between parentheses are determined by the conditioning variables. Equation (A.47) reduces to:

$$\bar{T}(S^I,S^A,S^B:S^{II}) = \bar{H}(\bar{z},\bar{v}) + \bar{T}_{\bar{z},\bar{v}}(u,x,z^{on},d_{SA},v^{on}:d_{RS}) \tag{A.48}$$

Once $\bar{z}$ and $\bar{v}$ are known, d_{RS} is independent of u. Equation (A.48) becomes then:

$$\bar{T}(S^I,S^A,S^B:S^{II}) = \bar{H}(\bar{z},\bar{v}) + \bar{T}_{\bar{z},\bar{v}}(x,z^{on},d_{SA},v^{on}:d_{RS}) \tag{A.49}$$

The final result for the coordination rate between subsystems is:

$$\bar{T}(S^I:S^A:S^B:S^{II}) = \bar{H}(z) + \bar{H}(\bar{z}) + \bar{H}(\bar{z},\bar{v}) + \bar{T}_z(x,d_{SA}:z^{on})$$

$$+ \bar{T}_{\bar{z}}(x,d_{SA},z^{on}:v^{on}) + \bar{T}_{\bar{z},\bar{v}}(x,z^{on},d_{SA},v^{on}:d_{RS}) \tag{A.50}$$

The final form for the coordination rate of the system is obtained by writing:

$$F_c = F_c^I + F_c^A + F_c^B + F_c^{II} + \bar{T}(S^I:S^A:S^B:S^{II})$$

and substituting Eq. (A.33 and A.50) into Eq. (A.15).

APPENDIX B

VARIABLE DEFINITION AND NUMERICAL RESULTS FOR THE EXAMPLE

B.1 VARIABLE DEFINITION

Organization P:

DM^1:

x^1 : can be 0 or 1 ; $p(1) = 0.9$ $p(0) = 0.1$

d_{SA}^1 : can be 0 or 1 ; $p(1) = 0.3$ $p(0) = 0.7$

$z^1 = z^{12}$: can be 0 or 1 ; $z^1 = 0(1)$ means that the input is sent by an enemy (friend). The value of z^1 is determined by the following list of possible combinations for (x^1, d_{SA}^1, z^1):

(0,0,0) ; (0,1,1) ; (1,0,1) ; (1,1,0)

z^{21}: can be 0 or 1, with the same meaning for each as above. It is independent of z^1.

$\bar{z}^1$: can be 0, 1 or 2. 2 means that the front is under a general attack. The possible values of $(z^1, z^{21}, \bar{z}^1)$ are:

(0,0,2) ; (0,1,0) ; (1,0,1) ; (1,1,1)

v^1: can be 1 or 2, depending on the decision strategy

d_{SA}^1: can be 0 or 1 ; $p(1) = 0.8$ $p(0) = 0.2$ (if $v^1=1$). 0 means that there is a red alert. 1 is for no alert. If $v^1=2$, $d_{RS}^1 = \square$ (no data accessed)

y^1: can be 0, 1 or 2. 0 means: fire. 1 means: request identification. 2 means: order all the planes off.
The value of y^1 is determined by $(\bar{z}, v^1, d^1_{RS}$, according to the following possibilities for $(\bar{z}, v^1, d^1_{RS}, y^1)$:

(0,1,0,0) ; (0,1,1,1) ; (0,2,□,1) ; (1,1,0,1) ; (1,1,1,1) ;
(1,2,□,1) ; (2,1,0,2) ; (2,1,1,1) ; (2,2,□,2)

<u>DM^2</u>: The values taken by the various variables have the same meaning as for DM^1, unless when specified otherwise.

x^2: can be 0 or 1 ; p(1) = 0.8, p(0) = 0.2

u^2: can be 1 or 2, depending on the decision strategy

d^2_{SA}: if $u^2 = 1$, can be 0 or 1, with p(1) = 0.3, p(0) = 0.7. If $u^2=2$, then $d^2_{SA} = \square$.

$z^2=z^{21}=z^{23}$: can be 0 or 1, depending on u^2, x^2, d^2_{SA}. The possibilities for $(x^2, u^2, d^2_{SA}, z^2)$ are:

(0,1,0,0) ; (0,1,1,1) ; (0,2,□,0) ; (1,1,0,1) ;
(1,1,1,0) ; (1,2,□,1)

z^{12}: can be 0 or 1, see DM^1

z^{32}: can be 0 or 1, see DM^3

$\bar{z}^2$: can be 0, 1, or 2. The possible combinations for $(z^2, z^{12}, z^{32}, \bar{z}^2)$ are:

(0,0,0,2) ; (0,0,1,2) ; (0,1,0,2) ; (0,1,1,0) ;
(1,0,0,2) ; (1,0,1,1) ; (1,1,0,1) ; (1,1,1,1)

d^2_{RS}: can be 0 or 1 ; with $p(1) = 0.8$, $p(0) = 0.2$

y^2: can be 0, 1, or 2; it is determined by $(\bar{z}^2, d^2_{RS})$. The possibilities for $(\bar{z}^2, d^2_{RS}, y^2)$ are:

$(0,0,0)$; $(0,1,1)$; $(1,0,1)$; $(1,1,1)$; $(2,0,2)$; $(2,1,1)$

<u>DM^3</u>: All variables are identical to those of DM^1.

<u>Pure Internal Strategies</u>: DM^1 is assumed to have 2 pure decision strategies, each consisting of the exclusive use of one RS algorithm, irrespective of $\bar{z}^1$. Dm^2 has 2 pure decision strategies, each consisting of the exclusive use of one SA algorithm. DM^3 has 2 pure strategies (see DM^1). In all, organization P has 8 (2*2*2) pure organization strategies.

<u>Organization H</u>: The values taken by the various variables have the same meaning they had in organization P, with four exceptions, however:

$d^2_{RS} = 0$ (1) means that the higher authorities have instructed the field commander to practice an aggressive (cautions) policy.

$y^{1(3)} = 2$ corresponds to giving the order to firing on the target in addition to ordering all the planes off.

$\bar{z}^2 = 0$ means that the front is being attacked; $\bar{z}^2 = 1$ is simply the negation of that.

$y^2 = 0$ means that DM^2 is giving DM^1 and DM^3 the permission to use whatever RS algorithm they deem adequate; y^2=1(2) is a firm order to use algorithm 1 (2).

$\underline{\mathrm{DM}^1}$:

x^1: can be 0 or 1, with $p(1) = 0.9$, $p(0) = 0.1$

d^1_{SA}: can be 0 or 1, with $p(1) = 0.3$, $p(0) = 0.7$

$z^1 = z^{12}$: can be 0 or 1. The possible combinations for (x^1, d^1_{SA}, z^1) are:

$(0,0,0)$; $(0,1,1)$; $(1,0,1)$; $(1,1,0)$

v^1: can be 1 or 2, depending on the internal decision strategy

y^{21}: can be 0, 1 or 2 (see DM^2)

$\bar{v}^1$: can be 1 or 2, depending on (v^1, y^{21}). The possible values for $(v^1, y^{21}, \bar{v}^1)$ are:

$(1,0,1)$; $(1,1,1)$; $(1,2,2)$; $(2,0,2)$; $(2,1,1)$; $(2,2,2)$

d^1_{RS}: can be 0 or 1, with $p(1) = 0.8$, $p(0) = 0.2$ if $\bar{v}^1 = 1$.
If $\bar{v}^1 = 2$, then $d^1_{RS} = \square$

y^1: can be 0, 1 or 2, depending on $(z^1, \bar{v}^1, d^1_{RS})$; the values for $(z^1, \bar{v}^1, d^1_{RS}, y^1)$ are:

$(0,1,0,0)$; $(0,1,1,1)$; $(0,2,\square,2)$; $(1,1,0,1)$;
$(1,1,1,1)$; $(1,2,\square,1)$

$\underline{\mathrm{DM}^2}$:

z^{12}: can be 0 or 1 (see DM^1)

z^{32}: can be 0 or 1 (see DM^3)

$\bar{z}^2$: can be 0 or 1, depending on (z^{12},z^{32}). The possibilities for $(z^{12},z^{32},\bar{z}^2)$ are:

(1,1,1) ; (1,0,1) ; (0,1,1) ; (0,0,2)

v^2 : can be 1 or 2, depending on the decision strategy

d^2_{RS}: can be 0 or 1, with p(1) = 0.7 and p(0) = 0.3 if $v^2=1$. If $v^2=2$, then $d^2_{RS} = \square$

$y^2=y^{21}=y^{23}$: can be 0, 1 or 2, depending on $(\bar{z}^2,v^2,d^2_{RS})$. The possible combinations for $(\bar{z}^2,v^2,d^2_{RS},y^2)$ are:

(0,1,0,2) ; (0,1,1,1) ; (0,2,0,2)
(1,1,0,1) ; (1,1,1,□) ; (1,2,□,0)

DM^3: The variables in DM^3 are determined in the same fashion as in DM^1.

Pure Internal Strategies: Organization H has 8 pure organization strategies. This result is obtained by using the same rationale used for organization P above.

B.2 METHODOLOGY AND ASSUMPTIONS

In Section 2.1.3, the total information theoretic activity of a system was defined to be the sum of the entropies of its internal variables (Eq. (2.13)). This definition will be used here to compute the activity and activity rate of each DM in each organization.

$$G^i = \sum_{j=1}^{N} H(w_j) \qquad i = 1,2,3 \tag{B.1}$$

The algorithms are defined in a simple way, due to the binary logic used. Their variables consist of their inputs and their outputs, if they belong to a stage which uses more than one algorithm (RS^1 and SA^2 in P for instance). If they belong to a single algorithm stage, like A^1 in organization P, their variables are their inputs only (the output does not need to be an internal variable in this case). An example of that is algorithms f_1^1 and (h_1^1, h_2^1) in P:

$$f_1^1: W^1 = \{x^1, d_{SA}^1\}$$

$$h_1^1: W^2 = \{\bar{z}^1, d_{RS}^1, y^1\}$$

$$h_2^1: W^3 = \{\bar{z}^1, d_{RS}^1, y^1\}$$

All the algorithms in both organizations are defined in the same way, which gives the following expressions for G^1 and G^2 in organizations H and P:

Organization P:

$$G^1 = H(x^1) + H(d_{SA}^1) + 2H(z^1) + H(z^2) + 3H(\bar{z}^1) + H(v^1)$$

$$+ H(d_{RS}^1) + 2H(y^1) \tag{B.2}$$

$$G^2 = H(x^2) + H(d^2_{SA}) + 3H(z^2) + H(z^1) + H(z^3) + 2H(\bar{z}^2)$$

$$+ H(d^2_{RS}) + H(y^2) + H(u^2) \qquad (B.3)$$

<u>Organization H</u>:

$$G^1 = H(x^1) + H(d^1_{SA}) + 3H(z^1) + H(v^1) + H(y^2) + H(\bar{v}^1)$$

$$+ H(d^1_{RS}) + 2H(y^1) \qquad (B.4)$$

$$G^2 = H(z^1) + H(z^3) + 3H(\bar{z}^2) + H(v^2) + H(d^2_{RS}) + 2H(y^2) \qquad (B.5)$$

<u>Assumptions</u>: The above expressions for the decisionmakers' activities are computed under the following simplifying assumptions:

- the individual inputs to the DMs are independent as well as their successive values

- the data received by the different DMs at each stage are identical, except for d^2_{RS} as seen above

- in fact, all variables that are not determined by one another according to the variable definition section are independent.

Under the above assumptions, the entropy of every variable in the system can be computed. The probability of its taking respective values is calculated using the list of possible combinations for the variable concerned and the ones that determine it, presented in the previous section. For example, consider the computation of $H(z^2)$ in organization P:

$$p(z^2=0) = p(x^2=0) * p(u^2=1) * p(d^2_{SA}=0) + p(x^2=0) * p(u^2=2)$$

$$+ p(x^2=1) * p(u^2=1) * p(d^2_{SA}=1) \qquad (B.6)$$

$$p(z^2=1) = p(x^2=0) * p(u^2=1) * [p(d^2_{SA}=1) + p(d^2_{SA}=0)]$$

$$+ p(x^2=1) * p(u^2=2) \qquad (B.7)$$

$$H(z^2) = - [p(z^2=0)\log_2 p(z^2=0) + p(z^2=1)\log_2 p(z^2=1)] \qquad (B.8)$$

If the variables' entropies are computed in the order they appear in the logic of the system, all the probabilities needed at some point will have been calculated in an earlier step, and the variables will be fully determined.

Performance Evaluation: The performance level for each pure strategy is derived by weighing the probabilities of having a mismatch between the desired and actual responses by the cost to the organization of this particular error occuring, and summing all the probabilities over all the possible inputs. An input is constituted of the traditional input, x, and the data inputs, d_{SA} and d_{RS}.

$$X = \begin{pmatrix} x^{1}_{2} & d^{1}_{SA} & d^{1}_{RS} \\ x^{2} & d^{2}_{SA} & d^{2}_{RS} \\ x^{3} & d^{3}_{SA} & d^{3}_{RS} \end{pmatrix} \tag{B.9}$$

The possible errors, for both organizations, are given by the presentation of the mapping between the actual and desired responses. The costs are shown next to each response, for each organization:

Organization P:

0 → 0 : 0
0 → 1 : 2
0 → 2 : 3
1 → 0 : 6
1 → 1 : 0
1 → 2 : 2
2 → 0 : 3
2 → 1 : 4
2 → 2 : 0

Organization H:

0 → 0 : 0
0 → 1 : 2
0 → 2 : 3
1 → 0 : 6
1 → 1 : 0
1 → 2 : 8
2 → 0 : 3
2 → 1 : 6
2 → 2 : 0

The pure organization strategies for organization P and organization H are given by the triplets (v^1,u^2,v^3) and (v^1,v^2,v^3) respectively. The numbers of the pure strategies for both organizations correspond to the same values of the triplets:

s_1 : (1,1,1)
s_2 : (1,1,2)
s_3 : (1,2,1)
s_4 : (1,2,2)
s_5 : (2,1,1)
s_6 : (2,1,2)
s_7 : (2,2,1)
s_8 : (2,2,2)

The desired responses for organization P are obtained by using the pure strategy s_1, which thus has a performance indice of 0. The desired responses for organization H are obtained by first restructuring the input matrix in the following way:

$$X(H) = \begin{pmatrix} x^1 & d^1_{SA} & d^1_{RS} & d^2_{RS} \\ x^3 & d^3_{SA} & d^3_{RS} & d^2_{RS} \end{pmatrix} \qquad (B.10)$$

and then by deriving these desired responses by using only the submatrix:

$$\begin{matrix} x^1 & d^1_{SA} & d^1_{RS} \\ x^3 & d^3_{SA} & d^3_{RS} \end{matrix}$$

according to the rationale developed for organization P. In summary, the desired response $y^1(H)$ corresponds to the first two elements of the vector $y^1(P)$ when the submatrix shown above corresponds to the first two lines in the input matrix of organization P.

B.3 NUMERICAL RESULTS:

Tables B.1 and B.2 show the activities and performance for each pure strategy in organizations P and H.

TABLE B.1 ACTIVITIES AND PERFORMANCE LEVEL FOR PURE STRATEGIES IN ORGANIZATION P

	G^1	G^2	J
s_1 (1,1,1)	9.48	10.42	0
s_2 (1,1,2)	9.48	10.42	0.41
s_3 (1,2,1)	8.97	8.09	0.54
s_4 (1,2,2)	8.97	8.09	0.70
s_5 (2,1,1)	9.03	10.42	0.41
s_6 (2,1,2)	9.03	10.42	0.82
s_7 (2,2,1)	8.15	8.09	0.70
s_8 (2,2,2)	8.15	8.09	0.87

TABLE B.2 ACTIVITIES AND PERFORMANCE LEVEL FOR PURE STRATEGIES IN ORGANIZATION H

	G^1	G^2	J
s_1 (1,1,1)	6.76	6.53	1.17
s_2 (1,1,2)	6.76	6.53	1.61
s_3 (1,2,1)	6.94	4.43	3.20
s_4 (1,2,2)	6.94	4.43	3.70
s_5 (2,1,1)	8.69	6.53	1.61
s_6 (2,1,2)	8.69	6.53	2.05
s_7 (2,2,1)	5.99	4.43	3.70
s_8 (2,2,2)	5.99	4.43	4.46

Table B.3 presents the differences in performance levels for organization P due to the lack of coordination in database updating. Schemes 1 and 2 correspond to the updating sequences considered in Section 5.2.3.

TBALE B.3 EFFECT OF LACK OF UPDATING COORDINATION ON PERFORMANCE LEVELS IN ORGANIZATION P

	Perfect Coordination	Scheme 1	Scheme 2
s_1	0	0.35	0.83
s_2	0.41	0.76	0.76
s_3	0.54	0.84	1.23
s_4	0.70	1.00	1.00
s_5	0.41	0.41	0.89
s_6	0.82	0.82	0.82
s_7	0.70	0.70	1.10
s_8	0.87	0.87	0.87