RAWIŃSKI TOMASZ
Shipresearch Institute, Technical University of Gdańsk

A THEORY OF DOCUMENT PROCESSING MACHINES - PRAXIOLOGICAL AND CYBERNETICAL APPROACH TO INFORMATION SYSTEMS.

The practice of information systems application and development is definitely in need of scund theoretical foundations - the task to develop them was formulated and undertaken by IFIP. The document processing machines /dcpm/ theory is an attempt to accomplish the task. The paper presents the theory in a very condensed way and discusses: task formulated by IFIP, enviroment and scope of dcpm usage, activities of dcpm users, architecture and organization of full-utility dcpms. The theory has originated of the practice.

1. NEED FOR INFORMATION SYSTEMS THEORY.

In recent years the data processing community realized that information systems application and development is definitely in need of scund theoretical foundations. The foundations are to cover principles of architecture, organization and development of the systems Projects were undertaken to establish the foundations In 1975 IFIP initiated [1] a new Technical Committee TC 8 with the name "Information Systems". In 1976-77 two working groups were set up:

WG 8.1: Design and Evaluation of Information Systems.
WG 8.2: The Interaction of Information Systems and the Organization.

The aims of the working groups, as presented in Preface to [1], include:

"1. to identify concepts and to develop theories relevant to the design of information systems;
3. to develop methods for the specification of information needs within an enterprise, with emphasis on interface aspects;"

The third aim is more precisely expressed in [2] where it is

2. AN ENVIROMENT OF ACTIVITY AND PURPOSE OF DCP MACHINES.

The existence of every organization, being a separate entity, is the effect of permenent exchange of documents between units of the organization. The exchanged documents are created within organization units and are of two kinds:

1. source documents - these depict events occuring within the organization and its enviroment,

2. issued documents - these are created within organization units and derived from documents which the unit stores and maintains.

There are two kinds of documents stored within units: permenent documents and temporary documents.
Permenent documents are integrated images of the segments of reality being of importance for the organization. These images are kept true using source documents. Temporary documents are sets of temporary and fragmentary images necessary to update permanent documents.
Documents handling activities of organization units are performed in response to tasks assigned to the units - these tasks are called the "document processing tasks", in short "dcp tasks". A dcp task usually is of the following form:

- keep true such-and-such permanent documents using such-and-such source dokuments,

- produce such-and-such issued documents deriving them from such-and-such stored (permanent and temporary) documents.

To perform their dcp tasks the organization units employ several tools, to be called "document processing tools", in short "dcp tools". Units employing dcp tools are "dcp tools users". A specific kind of dcp tools is called "computer-based information system" or "information system" - the systems are tools of high efficiency and complexity. Such efficient and complex tools are in common language called "machines", why to ignore it? Therefore a "computer-based information system" is a machine to process documents - the only consistent name for such a machine is a "document processing machine".
The term reveals the real nature of the object named and promotes the development of the theory. On the contrary, the term "information system" leads to obscure problems of sustems and information.

stated that for systems description is necessary " a language which is application oriented, to be understood by the application engineers, but must also be formal, with well understood semantics, to offer more precision than do natural languages. The description language shold be at an abstract level, for the way in which the system is implemented is not relevant at this stage. Classical informatics mainly used algorithms written in procedural languages for the specification of actions; however, such an approach would not be appropriate here,as we require a definition of the effect, not the method of achieving it."
To create the language relevant theoretical foundations are necessary, that is information system theory formulating ideas and concepts enabling system description which:

1. presents the system in terms of effects of actions, i.e. in terms of system users and applications,

2. does not contain any concepts dependent on system implementation, that is any hardware-software concepts,

3. is exast and fit to be formalized and univocally translated into a system description in hardware-software terms.

The purpose of this paper is to present briefly such a theory named "theory of document processing machines". In a more comprehensive way the theory is presented in [3]. The theory, as the author believes, is a step forward towards reaching the above mentioned aims of TC 8 working groups.

An object identified with the phrase "the document processing machine", in short "dcp machine" or "dcpm", is to be an abstract model of information system. The model is an abstract object possessing all essential features of information system; independent of implementation. A justification for employing the phrase "dcp machine" as a name for an information system model is presented in next chapter. The justification is based on two definitions from "Longmans English Larousse" Longmens 1969 where it is stated that:

machine is "an apparatus, made of organized, interacting parts, which takes in some form of energy, modifies it, and delivers it in a more suitable form for a desired function; a thing or system resembling such an apparatus in acting with regularity as a result of the interaction of its component parts."

document is "an official paper, a certificate; anything written that gives inform,ation or supplies evidence."

A dcp machine is a full-utility dcp machine for some dcp task if the machine:

- is capable to execute all dop activities occuring when the dcp task is performed and all a-d activities not requiring comparing documents with reality,

- ensures that the dcp task is performed with the lowest effort of organization units - users of the machine.

The lowest effort of users is ensured by the dcp machine which:

1. provides for each user a direct and independent access to dcpm

2. provides that each user can safely employ the dcpm without taking notice of activities of other users. It means that the dcpm prohibits any damages which could arise from independent and uncoordinated activities of many users.

3. is controlled with a language which:
   - is task-oriented, i.e. includes the task relevant concepts and terms only and does not include any terms and concepts relevant to implementation,
   - is as simple as possible and undertakes to prevent misunderstandaing.

4. protects users against their own errors by refusing to execute unconsistent and damaging orders.

5. is sufficiently reliable, i.e. ensures that sufficient percentage of users orders is completed before acceptable time limit elapses.

Architecture and organization of full-utility dcp machines are presented in next chapters.

In the past practice low-utility machines were mainly used - it results from the fact that the dcpm creators and developers are very little aware of the full-utility dcp machines facilities and principles, as well as how to implement them.

It is particularly easy to implement a full-utility dcpm basing on an ICL 1900/GEORGE 3 installation. And a dcp machine with features very close to those of full-utility dcpm was completed under author's direction, in 1982 at Gdańsk Regional Computing Centre. The dcp machine emerged as a result of several years of author's experience with development and operating so-called "control software". A control software is the software to control an application system run at GEORGE 3 installation.

The control software forms a main implementation substance to build a full-utility dcpm.

3. APPLICATION OF DCP MACHINES BY ORGANIZATION UNITS.

The problem is of praxiological nature and the conceptual framework employed is very similar to that J. Konieczny defined in [4].

One dcp task is usually performed by many units of the organization. To perform the task each unit performs a sequence of activities. The sequence is composed of activities of two kinds:

- analytic-decisive activities - a-d activities,
- document processing activities - dcp activities.

A dcp activity consists in:

- altering permanent documents according to source documents and/or
- creating issued documents according to stored (permanent and temporary) documents.

An a-d activity consists in analysing relevant documents content and deciding which activity is to be performed next.
After each dcp activity, an a-d activity is performed, next an activity decided upon by the latter - it can be an a-d or dcp activity. To faciliate their work the organization units employ dcp tools when performing the dcp tasks.

The dcp machine is a specific tool which can provide the greatest ease in performing a dcp activity owing to following features:

- a dcp machine can execute the entire dcp activity without any need for participation of an organization unit, as well as such a-d activity that does not require comparing any documents with reality,

- a dcp machine can be a "language directed machine", it means a user get the machine executing the required activity simply by transmitting to it an order in a "machine control language", without need of any energy-consuming interaction. In general, the machine is directed by dialogue in the machine control language.

Not each dcp machine provides all possible features to make performing their dcp tasks by organization units easy - to compare the different machines a concept of the "dcp machine utility for dcp task" is necessary. The utility is measured by the level up to which the dcp machine facititates performing the document handling activities of organization unit.

4. ARCHITECTURE OF FULL-UTILITY DCP MACHINES.

A process of a dcp machine utilization is a sequence of user activities necessary to perform a dcp task. The activities are performed with application of the dcp machine.

A dcp machine architecture is a set of dcp machine features which influence a progress of the machine utilization process, or the features taken into account by a user when employing a dcp machine to perform the dcp task.

A dcp machine executes dcp activities, which were described in previous chapters. To execute the dcp activities a dcp machine must have following abilities:

- ability to receive orders and source documents from user,
- ability to store and maintain permanent documents,
- ability to deliver issued documents to user.

To possess the abilities a dcp machine has to be equiped with following architectural elements:

- transreceiving or input/output boards, in short i/o boards,
- storing rooms.

Through an i/o board a user delivers orders and source documents and obtains issued documents. At storing rooms permanent and temporary documents are stored. In a full-utility machine the architectural elements are used in a specific way:

- every user has a separate, independent i/o board for his exclusive use and a separate, own room to store his documents,

- orders to control a dcp machine are delivered by special document named "The processing instruction",

- to employ a dcpm a user put a batch on an i/o board and it starts handling the batch by the dcpm. The batch is composed of "The processing instruction" and a set of source documents packets. "The processing instruction" comprises orders necessary to process enclosed document packets. With one batch one or number of dcp activities can be executed. Communicating with the dcpm using the batch ensures that the user effort is minimized and the machine is more proof against user errors.

The dcp machine equiped with the architectural elements is not necessarily the full-utility dcpm - it also has to work in specific way. The way is related to an internal organization of the machine - the problems will be discussed now.

5. INTERNAL ORGANIZATION OF THE FULL-UTILITY DCP MACHINES.

The organization of dcp machines will be defined according to cybernetical method. W. R. Ashby defines in [6] "cybernetics is a "theory of machines" which does not deal with an object, but with a behaviour of it." and "Cybernetics deals with all forms of the behaviour as far as they are regular, defined and repetitive. Materiality is insignificant...". Therefore the organization will be defined in purely behavioural, functional terms, in full abstaction from implementation concepts, from material features of the dcp machines. To define the functional organization functional blocks of the dcp machine and their interactions will be defined.

A functional block is an abstract, in material object which is a part of dcp machine and is necessary, together with other functional blocks, to get the machine working in suitable way. The functional organization of a dcp machine is as follows:

- every i/o board is connected to a couple of functional blocks: control block and processing block,
- all storing rooms are connected to one room access block.

Actions and interactions of the blocks are following:

- a control block receives a batch from i/o board, then analyses and validates it. If the batch is valid, the control block determines what activity is to be executed as the first to process the batch, prepares necessary conditions for it and sends suitable order to a processing block. After the order execution is completed, the processing block sends a message to the control block.This block receives the message and determines next activity. It can be a dcp or a-d activity. The a-d activity is executed by the control block itself, to execute the dcp activity the control block sends order to the processing block. Then working the control block checks state of documents at his own and other's storing rooms - the rooms are accessed through the room access block;

- a processing block receives consecutive orders from the control block and executes ordered dcp activities. Documents stored at storing rooms are read and written through the access block. When execution of order is completed the processing block sends a message to the control block and halts waiting for next order

Only a part of obtained results was discussed, the results especially practical, are, in fact, much more comprehensive - to present them here is impossible for the lack of place.
The theory presented here emerged as a generalization of the experience gained at GEORGE 3 installation. A more detailed presentation of the theory, as well as problems of dcp machines creation at GEORGE 3/ICL 1900 installation can be found in [3] and [5].

REFERENCES

[1] Schneider H. J. /ed/: Formal models and practical tools for information systems design, N. Holland, Amsterdam, 1979

[2] Goldsack S., Haase V.,Halling V.: A step towards application oriented specifications, in Meyer E. /ed/: Real-Time Data Handling and Process Control, N. Holland, Amsterdam, pp. 525-533, 1980

[3] Braniecki A., Rawiński T.: The document processing machine as a information processing device in shiprepair yard /in Polish/, Shipresearch Institute Report, Gdańsk, 1982

[4] Konieczny J.: Praxiological Models - Praxiological Modelling of Systems of Action, in Gasparski W., Pszczołowski T. /ed/: Praxiological Studies. Polish Contributions to the Science of Efficient Action, PWN, Warszawa and D. Reidel PC, Dordrecht, pp. 179-194, 1983

[5] Rawiński T.: The subsystems controlling data processing systems operations at GEORGE 3 installation /in Polish/, "Problemy Informatyki" Series, Warszawa, to be published.

[6] Ashby W.R.: An introduction to cybernetics /in Polish/, PWN, Warszawa, 1963