Business Process Engineering

Author: Włodzimierz Dąbrowski, PhD.

Lesson 6 DFD diagrams

Learning outcomes

In this lesson, you will achieve the following learning outcomes:

Knowledge (knowledge and understanding) - student:

- *Recognizes the elements of DFD diagram notation;*
- *recognizes types of diagrams;*
- *Characterizes the correct features of DFD diagrams.*

Skills - *Student:*

- *Knows how to evaluate the correctness of DFD diagrams;*
- *Knows how to construct a DFD diagram for a selected business problem.*

Justification

The competencies gained in this lesson are needed to understand the DFD diagrams used to model business processes.

Block 1 DFD diagram elements

New knowledge/skills

Data Flow Diagrams (DFDs) are used to model economic processes taking place in an organization, and are also used in economic and strategic planning. The DFD technique was first used in the second half of the twentieth century and is still very much used today to model systems. With the help of DFD diagrams, we can describe a system in terms of the processes that occur in it and the data that flows between these processes.

- **Components of the DFD diagram**

DFD diagrams include:

(1) Processes
(2) Flows
(3) Magazines
(4) Terminators

- **Process**

In the most general terms, we can say that processes are the transformation of something. From a formal point of view, processes generate results on the basis of received input data. On the diagram, a process is symbolized (identified) by its name, which should correspond to the activity that is performed during the process (e.g. ordering goods, making payments). On DFD diagrams there is no specific order in which processes are executed. Processes can be executed sequentially, in parallel, synchronously or asynchronously. A process in DFD diagrams is symbolized as a circle with the name of that process written in the center (Yourdon-DeMarco notation) or a rectangle with rounded sides (Gane-Sarson notation):

Figure 1. Designation of the process according to the notation (a) Yourdon-DeMarco; (b) Gen-Sarson; (c) SSADM

- **Flow**

The main task of flows is to model the exchange of data between individual processes, as well as warehouses, external systems and terminators. We distinguish the following possible data flows:

- transfer - occurs when data is exchanged between two processes,
- read/write/modify - occurs when data is exchanged between a process and a data repository.

Flows can contain individual data, or groups of data in the form of uniform structures (e.g., an order, an invoice, an employment contract). If the data flow performs the function of transferring data then it should contain a name specifying the data that is transferred, while. An exception to this rule is when the data is transferred to a data warehouse, then the name is not required, because the name of the warehouse identifies the data that is transferred (e.g. data transfer to a filing cabinet with sales invoices6 will identify just the

transfer of an invoice to this data warehouse). Sometimes there is a situation when only part of the data is sent to the data repository then it is necessary to specify exactly what data is being sent.

Note that the same data flow can occur between different processes on the diagram (e.g., a sales invoice may first be sent between the customer and the goods warehouse, where the goods are sold, and then the same invoice goes to the accounting department). The arrowhead on the flow determines the direction of data movement between the various objects on the diagram. When the arrowhead is on either side of the flow we are dealing with a dialogue between objects.

In DFD diagrams, data flow must not be placed directly between the data store and the terminator. There must always be a process between the two objects that will properly transfer the data.

It is possible to introduce divergence and convergence of processed data. Divergence means that the data is copied and transferred to different objects. In the diagram, we denote this by data branching. Data convergence, on the other hand, means that data flowing out of several processes are transferred to one object and are further processed there.

Figure 2: Flow designation

- **Magazine**

According to the convention adopted in DFD diagrams, warehouses are otherwise known as data repositories, which are used to store data. In offices, for example, these are archives where business documents are stored. As the name of a data warehouse, a plural noun is most often adopted to denote the elements stored in this object. Note that it is possible to combine partial data flowing into the warehouse. Four basic operations can be performed on the data warehouse:

- Searching - that is, reading specific information from the warehouse. This information can then be used in the process;
- Add - create new data in the warehouse;
- deletion - deletion of redundant or unnecessary data;
- updating - bringing the data to a new state that we desire

.

Operations such as adding, deleting and updating are associated with changing the contents of the data store. In the diagram, these operations are denoted by a data flow line ending with an arrow on one side. The search operation, on the other hand, belongs to the data warehouse dialogue operation, which involves first sending search criteria and then returning results. This operation is denoted by a data flow line ending with an arrow on both sides.

The storage symbol is the area defined by two parallel horizontal segments (Yourdon-DeMarco notation) or a rectangle without a right side (Gane-Sarson notation) in the center of which is the name of the data storage:

Figure 3. : Storage designation according to the notation (a) Yourdon-DeMarco; (b) Gen-Sarson; (c) SSADM.

- **Terminator**

The last element of the DFD diagram is the terminator. It is used to depict external objects in the modeled system. Terminators are used to communicate between the system and the environment. These can be objects such as a customer, supplier, office, as well as other systems. It should be remembered that it is not important how the object functions, which is contained in the terminator. Terminators can be not only objects external to the modeled organization, terminators can also be objects internal to the company, such as individual departments with which processes communicate. Data flows must not be placed directly between the terminator and the data repository, the two should always be connected by a process. It is possible to repeat the terminator name on the diagram, but only if it represents the same object, and the purpose of this repetition is to increase the readability of the diagram.

Terminators are symbolized by a square (Yourdon-DeMarco notation) or rectangle (Gane-Sarson notation), with the terminator name in the center:

Figure 4. : Terminator designation according to (a) Yourdon-DeMarco notation; (b) Gen-Sarson; (c) SSADM.

Current control

Question 1 Which of the following is a component of DFD diagrams (a
) flow
(b) warehouse
(c) terminator

Three buttons **[a][b][c]**. Any of them can be selected. [Confirm] button.

After marking the wrong answer, the information:

Review the components of DFD diagrams.

After marking the correct one:

Yes, this is the correct answer.

Block 2 Types of diagrams

New knowledge/skills

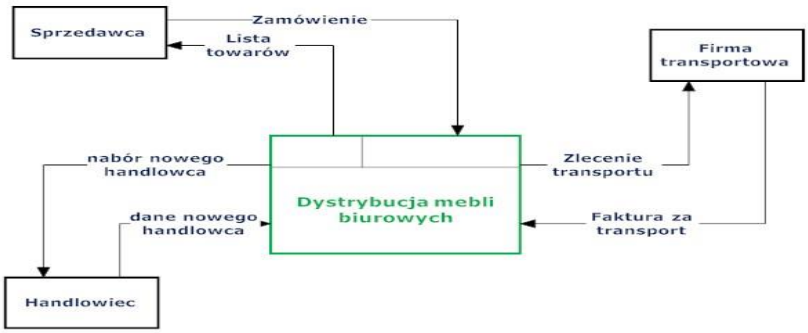
The DFD method defines two special types of diagrams.

Context diagram

The context diagram maps the boundaries and scope of the system. Its main task is to show how the modeled system is related to the environment. In this diagram, the system is depicted as one process with, while all objects that interact with the system are marked.

An example of a context diagram is shown in Figure 5.

Figure 5: Example of a context diagram

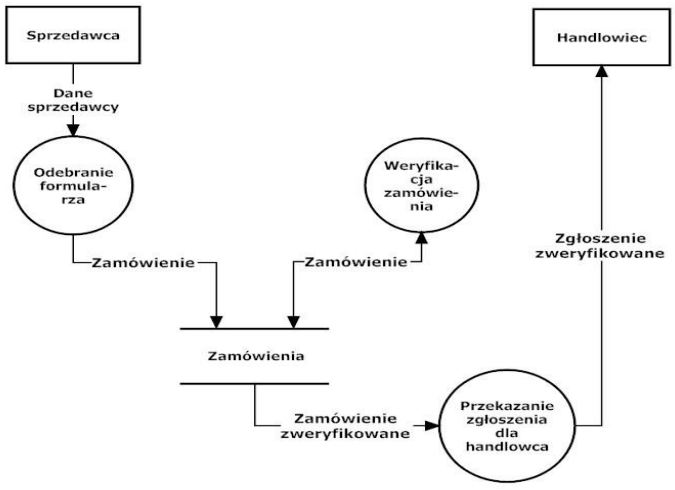


Zero Diagram.

A null diagram is otherwise known as a system diagram. It is a general diagram that shows the system from the most general point of view possible. By looking at the modeled system in this way, it allows us to isolate diagrams that are at lower levels. Through such a construction we get the possibility to define elementary processes. Note that there are diagrams of subsequent levels, which detail the general processes from the zero diagram.

An example of a lower level diagram is shown in the following figure.

Figure 6: Example of a lower level diagram



When modeling a business system using a DFD diagram, it is important to pay attention to the correctness of such a diagram. We take into account two types of correctness:

Syntactic correctness

Syntactic correctness means that you should check whether there is at least one process in the diagram, whether the elements have correct names, whether the repeated elements definitely represent the same business objects, whether each process has at least one input and output flow, whether there are no flows directly between the terminator and the data store.

Semantic correctness

The main tasks that need to be done to check semantic validity include whether the level of detail is adequate for the system being modeled, whether the input and output data guarantee the success of processing operations by the process, and whether the specific processes have been well separated into sub-processes.

Current control

Question 1A system diagram is otherwise known as a diagram:
(a) contextual
(b) zero

Two buttons [a]**[b]**. Either of them can be selected. [Confirm] button.

After marking the wrong answer, the information:

Review the description of the two diagrams: contextual and null.

After marking the correct one:

Yes, this is the correct answer.

Block 3 Improving DFD diagrams

New knowledge/skills

To improve the creation of models using the DFD method, several improvements have been made. First of all, the hierarchical structure of the modeled processes should be mentioned. Well, each process can be defined more precisely by breaking it down into several lower-level processes, until we reach elementary processes that can no longer be divided further. A process that is described by a lower level flow diagram is marked with three dots at the top. Thanks to hierarchization, we get a clearer structure of the process network. A good practice is to draw no more than seven processes on one diagram, because, as research has shown, the human brain can grasp such a number of diagrams with its thought. Along with hierarchization comes the concept of complexity of the business process being modeled. The following degrees of complexity of the business system are assumed:

low - up to 3 levels,

Medium - from 3 to 5 levels,

complex - more than 5 levels.

It is a good practice to number the processes at the top level using numbers, and then number the processes at lower levels using the number of the parent process along with adding a dot and the process number on the current diagram. For example, when we have a process numbered 2 on the parent diagram and the subdiagram shows three processes that describe the parent process numbered two, then give the following numbers to these processes: 2.1, 2.2, 2.3.

Current control

Question 1. We can categorize a business system with five levels of complexity as follows
(a) low complexity
(b) medium
(c) complex

Three buttons [a]**[b]**[c]. Any of them can be selected. [Confirm] button.

After marking the wrong answer, the information:

Familiarize yourself with the complexity classification of business systems.

After marking the correct one:

Yes, this is the correct answer.

Block 4 Typical errors

New knowledge/skills

When creating data flow diagrams, the basic principles of their construction are often forgotten. Typical mistakes involve: using the wrong component names (Figure 5, 6, 7, 8), incorrectly combining components (Figure 9, 10, 11), using ghost processes and well processes (Figure 12, 13).

The basic misconstructions of the DFD diagram include:

Using the wrong names of Terminator

Figure 7: Incorrect terminator name

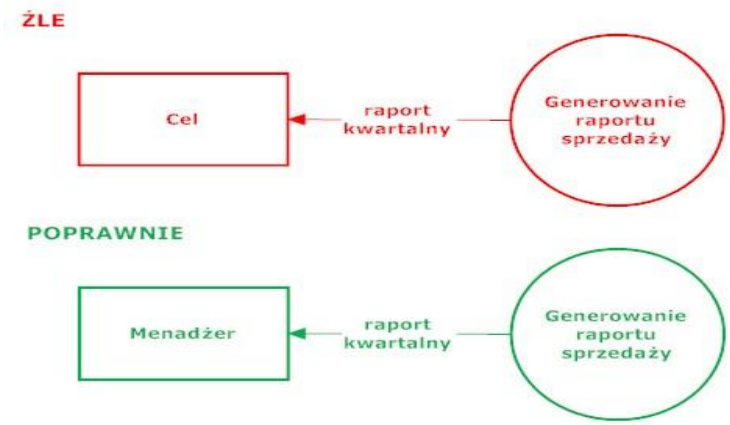


Figure 8: Incorrectly named data flow

Data flow

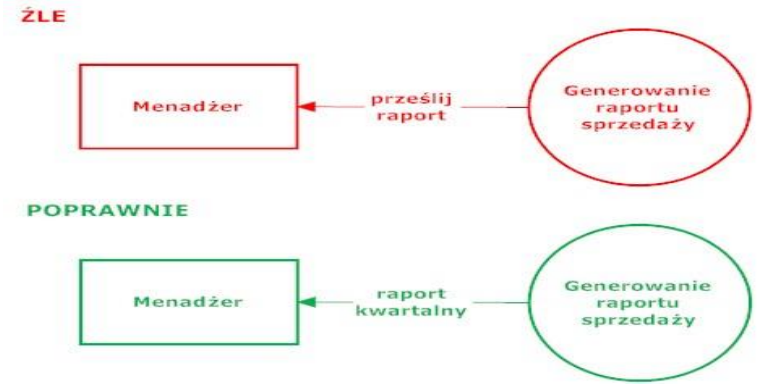


Figure 9: Incorrect process name

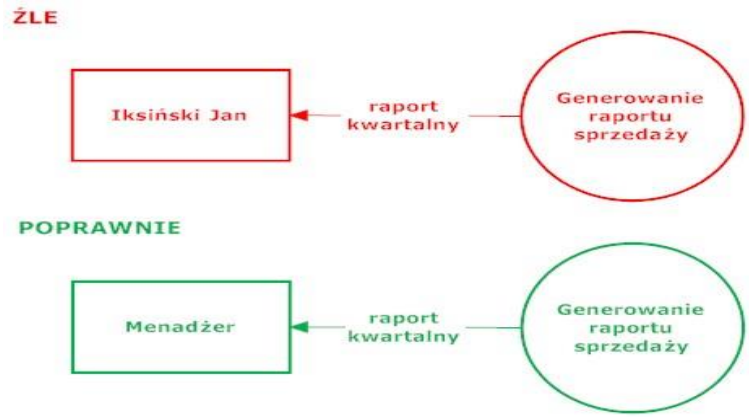


Figure 10: Incorrect name of data warehouse

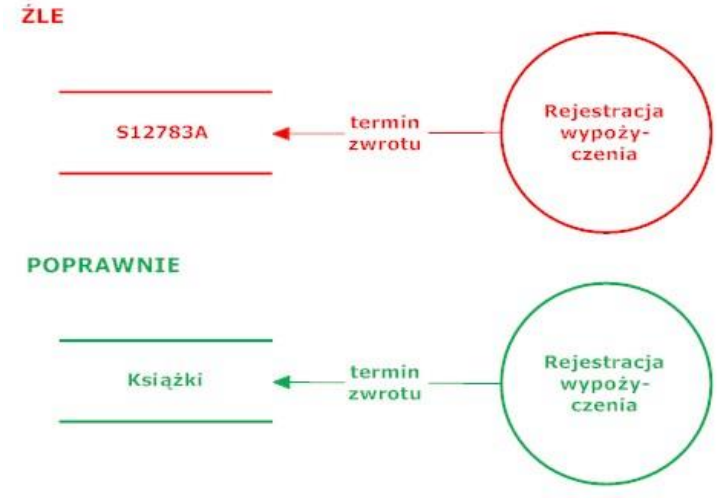


Figure 11: Incorrect connection of two terminators

Improper connection of components of Two Terminators

Figure 12: Incorrect connection between terminator and data store

Terminator with data storage

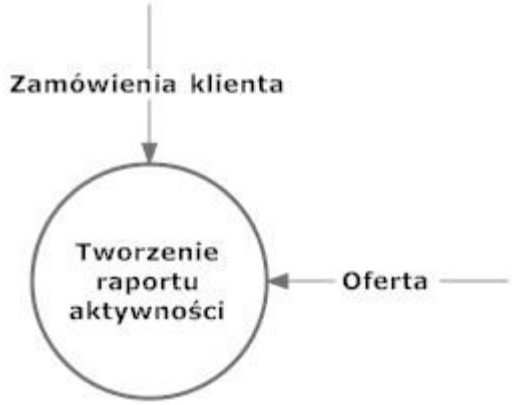
Figure 13: Incorrect connection of two warehouses

Two data stores

(3) Erroneous processes

The "well" type

Figure 14: Error well process



The "ghost" type

Figure 15: The erroneous "ghost" process

Current control

Question 1 The error in Figure 7 is:
(a) the name of the terminator is too short
(b) the name of the terminator is a noun
(c) the terminator's name does not reflect its role in the process

Three buttons [a][b]**[c]**. Any of them can be selected. [Confirm] button.

After marking the wrong answer, the information:

Consider what information you can read about the process from the diagram shown.

After marking the correct one:

Yes, this is the correct answer.

Block 5 Rules of correctness

New knowledge/skills

The rules of correctness of the DFD diagram are as follows:

DFD diagrams are organized hierarchically: context diagram, null (system) diagram, detailed diagrams;

The DFD diagram must not be larger than A4 format and should contain three to seven processes. Excessive complexity of the DFD should be avoided;

The data warehouse (data store) must be used by at least two processes;

The category names specified in the diagram hierarchy are unique;

All categories from the subordinate level must be shown on the superordinate level;

The names of flows to and from data stores (data repository) can be unnamed if full information is stored or retrieved;

The arrow to the data store (data repository) means that a specific change (entry, update, deletion) is made, while the arrow from the data store (data repository) means that data is read;

flow between the data store (data repository) and the terminator is not allowed;

The DFD diagram includes both manual and automated processes;

Each process must have at least one input flow and one output flow;

verify that the DFD is internally inconsistent with other related DFDs.

The SSADM methodology is used as a required standard in government administrations of some Western countries. There are some differences from the Yourdon-DeMarco and Gane-Sarsone notations and they are as follows:

External terminators can appear in the diagram of any level;

data stores, repeated in this diagram, are marked with a vertical line;

elementary (non-decomposable) processes, the specification of which is the specification of this process and not the DFD diagram, are marked with a special sign in the lower right corner of the process symbol (diagonal dash + asterisk);

multiple process is allowed, meaning multiple occurrences of the same process or simultaneous processes;

Among flows, there are two types: data flow and resource flow;

data flow between external objects is allowed (which is forbidden in Yourdon notation).

-
-

Current control

Question 1 Which of the following is
a true sentence?
a) the DFD diagram should be arranged hierarchically;
(b) the data store in the DFD diagram should be used by at least two processes;
(c) the names of the categories on the diagram should be unique.

Three buttons **[a][b][c]**. Any of them can be selected. [Confirm] button.

After marking the wrong answer, the information:

Review the rules for correctness of DFD diagrams again

After marking the correct one:

Yes, this is the correct answer.

Block 6 Data dictionary

New knowledge/skills

A very important modeling tool is the data dictionary (Data Directory). A data dictionary is an ordered list of all elements having to do with the system, along with their detailed description. Since in most systems the elements or data packages are very complex, it is therefore necessary to describe them by simpler elements. In turn, these simpler elements are defined by the units and values they can take.

Data elements are defined by the data dictionary by describing the following elements:

-flows and meaning in DFD diagrams;

-value properties that portions of information can take in data compositions and flows;

-building of complex packets (consisting of several elements) transmitted along flows'

-building data packages found in data warehouses;

-relationships between data compositions from the DFD model shown in the data model (in ERD diagrams).

2.5.1. Data Dictionary Notation.

Since the description of the elements of the DD model - data dictionary is not systematized, so there are many notations for describing the components of the system. The notation given below is one of the simplest (most intuitive) and most widely used methods for describing system components. It uses several simple symbols:

=- Data element definition ("consists of", "is defined as"),

+- "i",

**- Comment,

{ }- Iteration - used to indicate the repetition of a component of a data element ("zero or

more speeches"),

[ ] - Selection - indicates that the data element contains one of several alternatives

possibilities ([a|b]),

I- divider between alternative choices in the construction of [ ],

( ) - Optional occurrence of an element, specifies possibility but not necessity

occurrences of an element as a component of a composite data element,

@- Identifier for the data composition.

Current control

Question 1 The data dictionary should include:
(a) a list of all items having to do with the system;
(b) a detailed description of the elements in (a);
(c) a list of all terms used in the model.

Three buttons **[a][b]**[c]. Any of them can be selected. [Confirm] button.

After marking the wrong answer, the information:

Review the data dictionary definition again.

After marking the correct one:

Yes, this is the correct answer.

Final inspection

Question 1 Explain the differences between process and flow used in modeling with DFD diagrams.

A field for a text response of up to 2,000 characters. The [Send for evaluation] button. After its approval, the student's statement goes to the teacher. The teacher can provide a text feedback message for this assignment

Question 2 Explain in what situation a null type diagram should be used in modeling with DFD diagrams

A field for a text response of up to 2,000 characters. The [Send for evaluation] button. After its approval, the student's statement goes to the teacher. The teacher can provide a text feedback message for this assignment

Question 3 Explain what the error shown in Figure 9 is.

A field for a text response of up to 2,000 characters. The [Send for evaluation] button. After its approval, the student's statement goes to the teacher. The teacher can provide a text feedback message for this assignment

Question 4 Explain what the error shown in Figure 11 is. Suggest the correct solution.

Box for text and graphic response up to 2000 characters. The button [Send for evaluation]. After its approval, the student's statement goes to the teacher. The teacher can provide a text feedback message for this assignment

Question 5 Propose a definition of any dictionary term and write it using the notation discussed in Block 6.

A field for a text response of up to 2,000 characters. The [Send for evaluation] button. After its approval, the student's statement goes to the teacher. The teacher can provide a text feedback message for this assignment