# 1

# Failure Modes: Building Reliability Networks



## 1.1 Failure Modes

According to a commonly accepted definition (IEC, 1991), *reliability* is 'the ability of an entity to perform a required function under given conditions for a given time interval'. A system or component is said to have a failure *if the service it delivers to the user deviates from the specified one*, for example, if the system stops production. System failures or component failures usually require immediate corrective action (e.g. intervention for repair or replacement), in order to return the system or component into operating condition. Each failure is associated with losses due to the cost of intervention, the cost of repair and the cost of lost production.

*Failure mode* is the way a system or a component fails to function as intended. It is the effect by which failure is observed. The physical processes leading to a particular failure mode will be referred to as *failure mechanism*. It is important to understand that the same failure mode (e.g. fracture of a component) can be associated with different failure mechanisms. Thus, the fracture of a component could be the result of a *brittle fracture* mechanism, *ductile fracture* mechanism or *fatigue failure* mechanism involving nucleation and slow propagation of a fatigue crack. In each particular case, the failure mechanism behind the failure mode 'fracture' is different.

Apart from fracture, other examples of failure modes are 'short circuit', 'open circuit', 'overheating of an electrical or mechanical component', excessive noise and vibration, leakage from a seal, excessive deformation, excessive wear, misalignment which causes a loss of precision, contamination, etc.

Design for reliability is about preventing failure modes from occurring during the specified lifetime of the product. Suppose that the space of all design parameters is denoted by $\Omega$ (see Figure 1.1) and the component is characterised by $n$ distinct failure modes. Let $A_1$, $A_2$, …, $A_n$ denote the domains of values for the design variables which prevent the first failure mode, the second failure mode and the $n$th failure mode, respectively.

The intersection $A_1 \cap A_2 \cap \cdots \cap A_n$ of these domains will prevent all failure modes from occurring. An important objective of the design for reliability is to specify the design variables so that they all belong to the intersection domain. This prevents from occurring any of the identified failure modes.

In order to reduce the risk of failure of a product or a process, it is important to recognise their failure modes as early as possible in order to enable execution of design modifications and specific actions

*Reliability and Risk Models: Setting Reliability Requirements*, Second Edition. Michael Todinov.

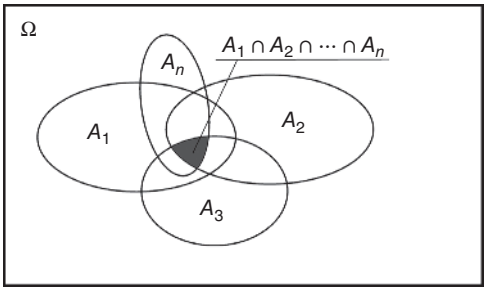


**Figure 1.1** Specifying the controllable design variables to be from the intersection domain will prevent all $n$ failure modes

reducing the risk of failure. The benefits from identifying and eliminating failure modes are improved reliability of the product/process, improved safety, reduced warranty claims and other potential losses from failures. It is vital that identifying the failure modes and the required design modifications for their elimination is made during the *early stages* of the design. Design modifications during the early stages of the design are much less costly compared to design modifications executed during the late stages of the design.

Systematic procedures for identifying possible failure modes in a system and evaluating their impact have already been developed. The best known method is the failure mode and effects analysis abbreviated as FMEA, developed in 1963 by NASA (National Aeronautics and Space Administration) for the Apollo project. The method has subsequently been applied in aerospace and aeronautical engineering, nuclear industry, electronics industry, automotive industry and software development. Many literary resources concerning this method are related to the American Military Standard (MIL-STD-1629A, 1977). The fundamental idea behind FMEA is to discover as many as possible potential failure modes, evaluate their impact, identify failure causes and outline controls and actions limiting the risks associated with the identified failure modes. The extension of FMEA which includes *criticality analysis* is known as failure mode and effects criticality analysis (FMECA):

- The inductive approach is an important basic technique for identifying possible failure modes at a system level. It consists of considering sequentially the failure modes of all parts and components building the system and tracking their effect on the system's performance.
- The deductive approach is another important basic technique which helps to identify new failure modes. It consists of considering an already identified failure mode at a system level and investigating what else could cause this failure mode or contribute to it.

Other techniques for identifying potential failure are:

- A systematic analysis of common failure modes by using check lists. An example of a simple check list which helps to identify a number of potential failure modes in mechanical equipment is the following:
  Are components sensitive to variations of load?
  Are components resistant against variations of temperature?
  Are components resistant against vibrations?
  Are components resistant to corrosion?
  Are systems/assemblies robust against variation in their design parameters?
  Are parts sensitive to precise alignment?
  Are parts prone to misassembly?
  Are parts resistant to contamination?
  Are components resistant against stress relaxation?

- Using past failures in similar cases. For many industries, a big weight is given to databases of the type 'lessons learned' which help to avoid failure modes causing problems in the past. Lessons learned from past failures have been useful to prevent failure modes in the oil and gas industry, the aerospace industry and nuclear industry.
- Playing devil's advocate. Probing what could possibly go wrong. Asking lots of 'what if' questions.
- Root cause analysis. Reveals processes and conditions leading to failures. Physics of failure analysis is a very important method for revealing the genesis of failure modes. The root cause analysis often uncovers a number of unsuspected failure modes.
- Assumption analysis. Consists of challenging and testing common assumptions about the followed design procedures, manufacturing, usage of the product, working conditions and environment.
- Analysis of the constraints of the systems. The analysis of the technical constraints of the system, the work conditions and the environment often helps to discover new failure modes.
- Asking not only questions about what could possibly go wrong but also questions how to make the system malfunction. This is a very useful technique for discovering rare and unexpected failure modes.
- Using creativity methods and tools for identifying failure modes in new products and processes (e.g. brainstorming, TRIZ, lateral thinking, etc.)

Before discovering failure modes is attempted, it is vital to understand the basic processes in the system and how the system works. In this respect, building a functional block diagram and specifying the required functions of the system are very important.

The functional diagram shows how the components or process steps are interrelated.

For example, the required system function from the generic lubrication system in Figure 1.2 is *to supply constantly clean oil at a specified pressure, temperature, debit, composition and viscosity to contacting moving parts*. This function is required in order to (i) reduce wear, (ii) remove heat from friction zones and cool the contact surfaces, (iii) clean the contact surfaces from abrasion particles and dirt and (iv) protect from corrosion the lubricated parts. Not fulfilling any of the required components of the system function constitutes a system failure.

The system function is guaranteed by using components with specific functions. The sump is used for the storage of oil. The oil filter and the strainer are used to maintain the oil cleanliness. Maintaining the correct oil pressure is achieved through the pressure relieve valve, and maintaining the correct oil temperature is achieved through the oil cooler. The oil pump is used for maintaining the oil debit, and the oil galleries are used for feeding the oil to the contacting moving parts.

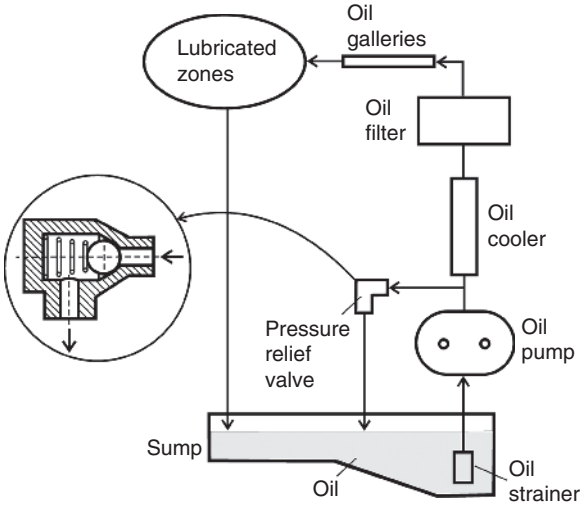


**Figure 1.2** Functional block diagram of a lubrication system

The inductive approach for discovering failure modes at a system level starts from the failure modes of the separate components and tracks their impact on the system's performance. Thus, a clogged oil filter leads to a drop of the oil pressure across the oil filter and results in low pressure of the supplied lubricating oil. A low pressure of the supplied lubricating oil constitutes a system failure because supplying oil at the correct pressure is a required system's function.

A mechanical damage of the oil filter prevents the retention of suspended particles in the oil and leads to a loss of the required system function 'supply of clean oil to the lubricated surfaces'.

If the pressure relief valve is stuck in open position, the oil pressure cannot build up and the pressure of the supplied oil will be low, which constitutes a system failure. If the pressure relief valve is stuck in closed position, the oil pressure will steadily build up, and this will lead to excessive pressure of the supplied oil which also constitutes a system failure. With no pressure relief mechanism, the high oil pressure could destroy the oil filter and even blow out the oil plugs.

A cooler lined up with deposited plaques or clogged with debris is characterised by a reduced heat transfer coefficient and leads to decreased cooling capability and a 'high temperature of the supplied oil' which constitutes a system failure. Failure of the cooling circuit will have a similar effect. Clogging the cooler with debris will simultaneously lead to an increased temperature and low pressure of the supplied oil due to the decreased cooling capability and the pressure drop across the cooler.

Excessive wear of the oil pump leads to low oil pressure, while a broken oil pump leads to no oil pressure. Failure of the sump leads to no oil pressure; a blocked oil strainer will lead to a low pressure of the supplied oil.

Blockage of the oil galleries, badly designed oil galleries or manufacturing defects lead to loss of the required system function 'delivering oil at a specified debit to contacting moving parts'.

Oil contamination due to inappropriate storage, oil degradation caused by oxidation or depletion of additives and the selection of inappropriate oil lead to a loss of the required system function 'supplying clean oil with specified composition and viscosity'.

The deductive approach for discovering failure modes at a system level starts with asking questions what else could possibly cause a particular failure mode at a system level or contribute to it and helps to discover contributing failure modes at a component level.

Asking, for example, the question what can possibly contribute to a too low oil pressure helps to discover the important failure mode 'too large clearances between lubricated contact surfaces due to wear out'. It also helps to discover the failure mode 'leaks from seals and gaskets' and 'inappropriate oil with high viscosity being used'.

Asking the question what could possibly contribute to a too high oil pressure leads to the cause 'incorrect design of the oil galleries'. Asking the question what could possibly contribute to a too high oil temperature leads to the cause 'a small amount of circulating oil in the system' which helps to reveal the failure modes 'too low oil level' and 'too small size of the sump'. Undersized sumps lead to a high oil temperature which constitutes a failure mode at the system level.

A common limitation of any known methodology for identifying failure modes is that there is no guarantee that all failure modes have been identified. A severe limitation of some traditional methodologies (e.g. FMEA) is that they treat failure modes of components independently and cannot discover complex failure modes at system level which appear only if a combination of several failure modes at a component level is present.

Another severe limitation of some traditional approaches is that they (e.g. FMEA) cannot discover failure modes dependent on the timing or clustering of conditions and causes. If a number of production units demand independently specified quantity of particular resource (e.g. water steam) for a specified time, the failure mode 'insufficient resource supply' depends exclusively on the clustering of random demands during the time interval and the capacity of the generator centrally supplying the resource.

**Exercise**

Discover the failure modes of the clevis joint in the figure. The clevis is subjected to a constant axial tensile loading force $P$ (Figure 1.3).

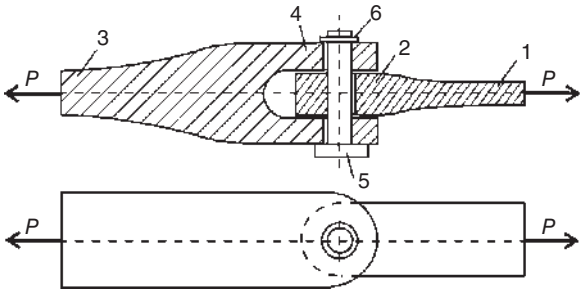


**Figure 1.3** A clevis joint

***Solution***
*Shear failure modes*:

- Shear failure of the pin 5
- Shear failure of the eye 2
- Shear failure of the clevis 4

*Compressive failure modes*:

- Compressive failure of the pin 5 due to excessive bearing pressure of the eye 2
- Compressive failure of the pin 5 due to excessive bearing pressure of the clevis 4
- Compressive failure of the clevis 4 due to excessive bearing pressure of the pin 5
- Compressive failure of the eye 2 due to excessive bearing pressure of the pin 5

*Tensile failure modes*:

- Tensile failure of the blade in zone 1, away from the eye 2
- Tensile failure in zone 3, away from the clevis 4
- Tensile failure of the blade in the area of the eye 2
- Tensile failure in the area of the clevis 4

*Other failure modes*:

- Bending of the pin 5
- Failure of the clip 6

Thirteen failure modes have been listed for this simple assembly. The analysis in Samuel and Weir (1999), for example, reported only eight failure modes. Preventing all 13 failure modes means specifying the controllable design variables to be from the intersection of the domains which prevent each listed failure mode (Figure 1.1)

## 1.2 Series and Parallel Arrangement of the Components in a Reliability Network

The operation logic of engineering systems can be modelled by reliability networks, which in turn can be modelled conveniently by graphs. The nodes are notional (perfectly reliable), whereas the edges correspond to the components and are unreliable.

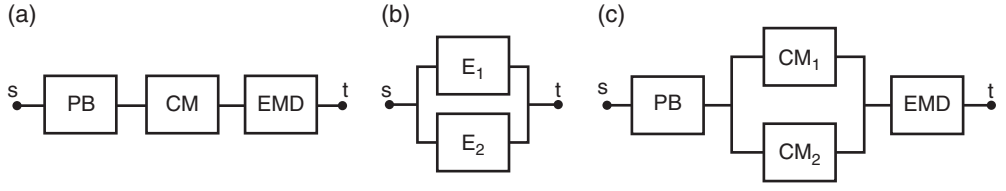


**Figure 1.4** (a) Reliability network of a common system composed of a power block (PB), a control module (CM) and an electromechanical device (EMD). (b) Reliability network of a system composed of two power generators $E_1$ and $E_2$; the system is working if at least one of the power generators is working. (c) Reliability network of a simple production system composed of power block (PB), two control modules ($CM_1$ and $CM_2$) and an electromechanical device (EMD)

The common system in Figure 1.4a consists of a power block (PB), control module (CM) and an electromechanical device (EMD).

Because the system fails whenever any of the components fails, the components are said to be logically arranged in series. The next system in Figure 1.4b is composed of two power generators $E_1$ and $E_2$ working simultaneously. Because the system is in working state if at least one of the generators is working, the generators are said to be logically arranged in parallel.

The simple system in Figure 1.4c fails if the power block (PB) fails or if the electromechanical device (EMD) fails or if both control modules $CM_1$ and $CM_2$ fail.

However, failure of control module $CM_1$ only does not cause a system failure. The redundant control module $CM_2$ will still maintain control over the electromechanical device and the system will be operational.

The system is operational if and only if in its reliability network a path through working components exists from the start node s to the terminal node t; (Figure 1.4).

Reliability networks with a single start node (s) and a single end node (t) can also be interpreted as single-source–single-sink flow networks with edges with integer capacity. The system is in operation if and only if, on demand, a unit flow can be sent from the source s to the sink t (Figure 1.4). In this sense, reliability networks with a single start node and a single end node can be analysed by the algorithms developed for determining the reliability of the throughput flow of flow networks (Todinov, 2013a).

## 1.3 Building Reliability Networks: Difference between a Physical and Logical Arrangement

Commonly, the reliability networks do not match the functional block diagram of the modelled system. This is why an emphasis will be made on building reliability networks.

The fact that the components in a particular system are logically arranged in series does not necessarily mean that they are logically arranged in series. Although the physical arrangement of the seals in Figure 1.5a is in series, their logical arrangement with respect to the failure mode 'leakage in the environment' is in parallel (Figure 1.5b). Indeed, leakage in the environment is present only if both seals fail.

Conversely, components may be physically arranged in parallel, with a logical arrangement in series. This is illustrated by the seals in Figure 1.6. Although the physical arrangement of the seals is in parallel, their logical arrangement with respect to the failure mode *leakage in the environment* is in series. Leakage in the environment is present if at least one seal stops working (sealing).

Reliability networks are built by using the top-down approach. The system is divided into several large blocks, logically arranged in a particular manner. Next, each block is further detailed into several

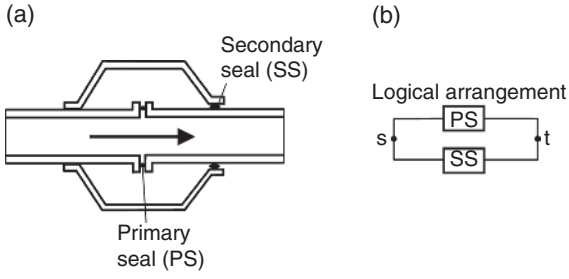


**Figure 1.5** Seals that are (a) physically arranged in series but (b) logically arranged in parallel

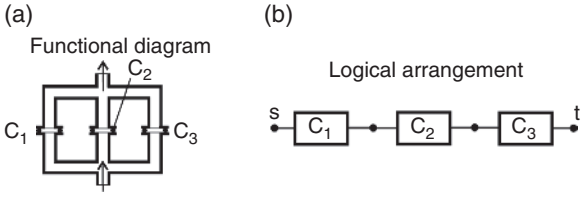


**Figure 1.6** The seals are (a) physically arranged in parallel but (b) logically in series

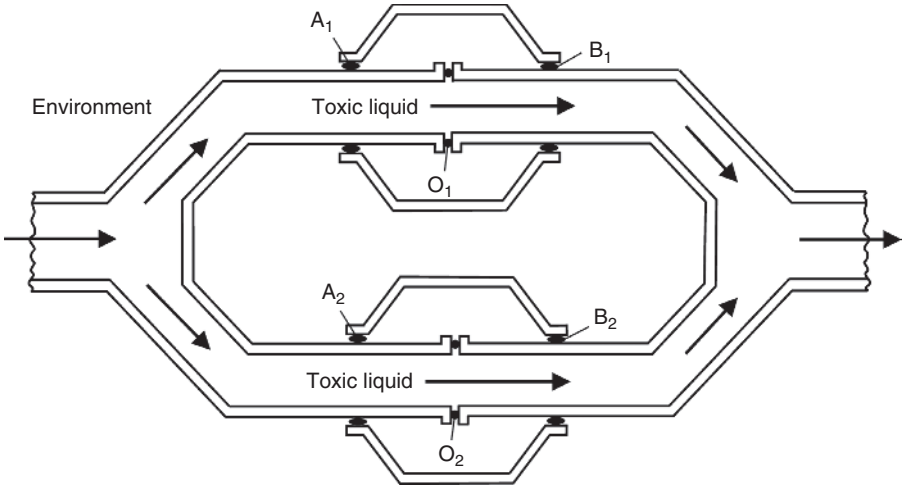


**Figure 1.7** A functional diagram of a system of seals isolating toxic liquid from the environment

smaller blocks. These blocks are in turn detailed and so on, until the desired level of indenture is achieved for all blocks.

This approach will be illustrated by the system in Figure 1.7, which represents toxic liquid travelling along two parallel pipe sections. The O-ring seals '$O_1$', and '$O_2$' are sealing the flanges; the pairs of seals ($A_1$, $B_1$) and ($A_2$, $B_2$) are sealing the sleeves.

The first step in building the reliability network of the system in Figure 1.7 is to note that despite that physically, the two groups of seals ($O_1$, $A_1$, $B_1$) and ($O_2$, $A_2$, $B_2$) are arranged in parallel, they are arranged logically in series with respect to the function 'preventing a leak to the environment' because both of the two groups of seals must prevent the toxic liquid from escaping in the environment

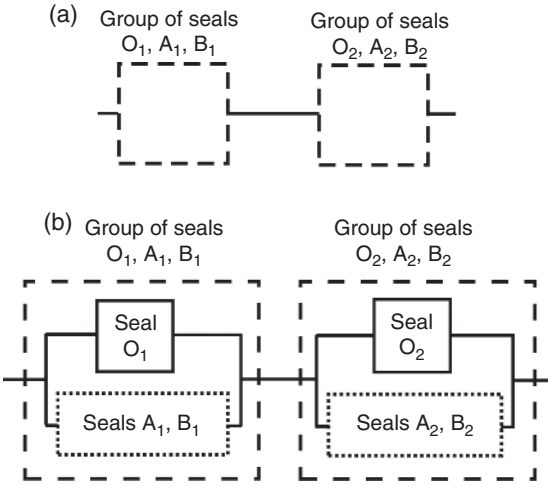


**Figure 1.8** (a) First stage and (b) second stage of detailing the reliability network of the system in Figure 1.7

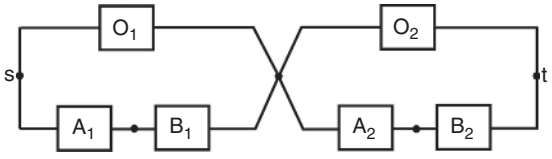


**Figure 1.9** A reliability network for the system of seals in Figure 1.7

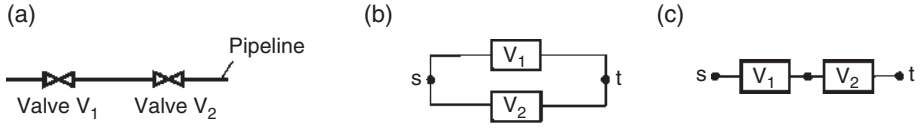


**Figure 1.10** Physical and logical arrangement of (a) two valves on a pipeline with respect to the functions. (b) Stopping the production fluid and (c) 'enabling the flow through the pipeline'

(Figure 1.8a). Failure to isolate the toxic liquid is considered at the highest indenture level – the level of the two groups of seals.

Within each of the two groups of seals, the O-ring seal is logically arranged in parallel with the pair of seals (A, B) on the sleeves (Figure 1.8b). Indeed, it is sufficient that the O-ring seal '$O_1$' works or the pair of seals ($A_1$, $B_1$) works to guarantee that the first group of seals ($O_1$, $A_1$, $B_1$) will prevent a release of toxic liquid in the environment.

Finally, within the pair of seals ($A_1$, $B_1$), both seals '$A_1$' and '$B_1$' must work in order to guarantee that the pair of seals ($A_1$, $B_1$) works. The seals $A_1$ and $B_1$ are therefore logically arranged in series. This reasoning can be extended for the second group of seals, and the reliability network of the system of seals is as shown in Figure 1.9.

The next example features two valves on a pipeline, physically arranged in series (Figure 1.10). Both valves are initially open. With respect to stopping the production fluid in the pipeline, on demand, the valves are arranged in parallel (Figure 1.10b). Now suppose that both valves are initially closed. With respect to enabling the flow through the pipeline, on demand, the valves are logically arranged in series (Figure 1.10c).

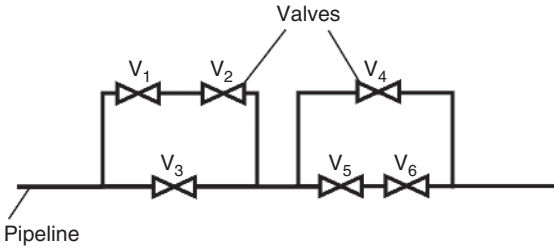


**Figure 1.11** A functional diagram of a system of valves

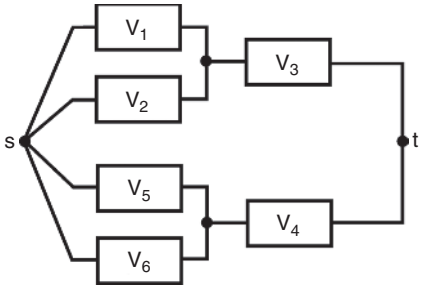


**Figure 1.12** The reliability network of the system in Figure 1.9

Indeed, to stop the flow through the pipeline, at least one of the valves must work on demand; therefore, the valves are logically arranged in parallel with respect to the function 'stopping the production fluid'. On the other hand, if both valves are initially closed, to enable the flow through the pipeline, both valves must open on demand; hence, in this case, the logical arrangement of the valves is in series (Figure 1.10c).

**Example**

Figure 1.11 features the functional diagram of a system of pipes with six valves, working independently from one another, all of which are initially open. Each valve is characterised by a certain probability that if a command for closure is sent, the valve will close and stop the fluid passing through its section. Construct the reliability network of this system with respect to the function 'stopping the flow through the pipeline'.

***Solution***

The reliability network related to the function stopping the flow in the pipeline is given in Figure 1.11. The blocks of valves ($V_1$, $V_2$, $V_3$) and the block of valves ($V_4$, $V_5$, $V_6$) are logically arranged in parallel because the flow through the pipeline is stopped if either block stops the flow. The block of valves ($V_1$, $V_2$, $V_3$) stops the flow if both groups of valves ($V_3$) and ($V_1$, $V_2$) stop the flow in their corresponding sections. Therefore, the groups ($V_1$, $V_2$) and $V_3$ are logically arranged in series. The group of valves ($V_1$, $V_2$) stops the flow if either valve $V_1$ or $V_2$ stops the flow in the common section. Therefore, the valves $V_1$ and $V_2$ are logically arranged in parallel.

Similar reasoning applies to the block of valves $V_4$, $V_5$ and $V_6$. The reliability network of the system in Figure 1.11 is given in Figure 1.12.

The operational logic of the system has been modelled by a set of perfectly reliable nodes (the filled circles in Figure 1.12) and unreliable edges connected to the nodes.

Interestingly, for the function stopping the fluid in the pipeline, valves or blocks of valves arranged in series in the functional diagram are arranged in parallel in the reliability network. Accordingly, valves or blocks arranged in parallel in the functional diagram are arranged in series in the reliability network.

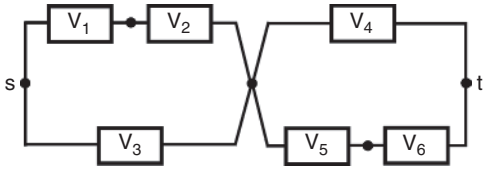


**Figure 1.13** The reliability network of the system in Figure 1.9, with respect to the function 'letting flow through the pipeline'

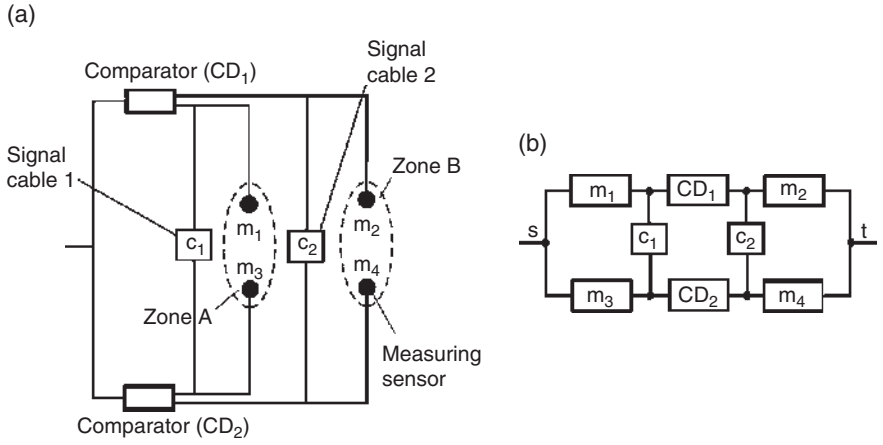


**Figure 1.14** (a) A safety-critical system based on comparing measured quantities in two zones and (b) its reliability network

There are also cases where the physical arrangement coincides with the logical arrangement. Consider again the system of valves in Figure 1.11, with all valves initially closed. With respect to the function 'letting flow (any amount of flow) through the pipeline' (the valves are initially closed), the reliability network in Figure 1.13 mirrors the functional diagram in Figure 1.11.

## 1.4 Complex Reliability Networks Which Cannot Be Presented as a Combination of Series and Parallel Arrangements

Many engineering systems have reliability networks that cannot be described in terms of combinations of series–parallel arrangements. The safety-critical system in Figure 1.14a is such a system. The system compares signals from sensors reading the value of a parameter (pressure, concentration, temperature, water level, etc.) in two different zones. If the difference in the parameter levels characterising the two zones exceeds a particular critical value, a signal is issued by a special device (comparator).

Such generic comparators have a number of applications. If, for example, the measurements indicate a critical concentration gradient between the two zones, the signal may operate a device which eliminates the gradient. In the case of a critical differential pressure, for example, the signal may be needed to open a valve which will equalise the pressure. In the case of a critical temperature gradient measured by thermocouples in two zones of the same component, the signal may be needed to interrupt heating/cooling in order to limit the magnitude of the thermal stresses induced by the thermal gradient. In the case of a critical potential difference measured in two zones of a circuit, the signal may activate a switch protecting the circuit.

The complex safety-critical system in Figure 1.14a compares the temperature (pressure) in two different zones (A and B) measured by the sensors ($m_1$, $m_2$, $m_3$ and $m_4$). If the temperature (pressure) difference is greater than a critical value, the difference is detected by one of the comparators (control devices) $CD_1$ or $CD_2$, and a signal is sent which activates an alarm. The two comparators and the two pairs of sensors have been included to increase the robustness of the safety-critical system. For the same purpose, the signal cables $c_1$ and $c_2$ have been included, whose purpose is to increase the connectivity between the sensors and the comparators. If, for example, sensors $m_1$, $m_2$ and comparator $CD_2$ have failed, the system will still be operational. Because of the existence of signal cables, the measured parameter levels by the remaining operational sensors $m_3$ and $m_4$ will be fed to comparator $CD_1$ through the signal cables $c_1$ and $c_2$ (Figure 1.14a). If excessive difference in the parameter levels characterising the two zones exists, the comparator $CD_1$ will activate the alarm. If sensors $m_1$ and $m_4$ fail, comparator $CD_1$ fails and signal cable $c_1$ fails, the system is still operational because the excessive difference in the measured levels will be detected by sensors $m_3$ and $m_2$ and through the working signal cable $c_2$ will be fed to comparator $CD_2$.

The system will be operational whenever an s–t path through working components exists in the reliability network in Figure 1.14b. The reliability network in Figure 1.14b cannot be reduced to combinations of series, parallel or series–parallel arrangements. Telecommunication systems and electronic control systems may have very complex reliability networks which cannot be represented with series–parallel arrangements.

## 1.5 Drawbacks of the Traditional Representation of the Reliability Block Diagrams

### *1.5.1 Reliability Networks Which Require More Than a Single Terminal Node*

Traditionally, reliability networks have been presented as networks with a single start node s and a single terminal node t (Andrews and Moss, 2002; Billinton and Allan, 1992; Blischke and Murthy, 2000; Ebeling, 1997; Hoyland and Rausand, 1994; Ramakumar, 1993). This traditional representation, however, is insufficient to model the failure logic of many engineering systems. There are systems whose logic of failure description requires more than a single terminal node. Consider, for example, the safety-critical system in Figure 1.15 that consists of a power supply (PS), power cable (PC), block of four switches ($S_1$, $S_2$, $S_3$ and $S_4$) and four electric motors ($M_1$, $M_2$, $M_3$ and $M_4$).

In the safety-critical system, all electric motors must be operational on demand. Typical examples are electric motors driving fans or pumps cooling critical devices, pumps dispensing water in case of fire, life support systems, automatic shutdown systems, control systems, etc. The reliability on demand of the system in Figure 1.15a can be improved significantly by making the inexpensive low-reliability components redundant (the power cable and the switches) (Figure 1.15b). For the system in Figure 1.15b, the electric motor $M_1$, for example, will still operate if the power cable PC or the switch $S_1$ fails because power supply will be maintained through the alternative power cable PC′ and the switch $S_1'$. The same applies for the rest of the electric motors. The power supply to an electric motor will fail only if both power supply channels fail. The reliability network of the system in Figure 1.15b is given in Figure 1.16. It has one start node s and four terminal nodes $t_1$, $t_2$, $t_3$ and $t_4$. The system is in working state if a path through working components exists between the start node s and each of the terminal nodes $t_1$, $t_2$, $t_3$ and $t_4$.

The reliability network in Figure 1.16 is also an example of a system which cannot be presented as a series–parallel system. It is a system with complex topology.

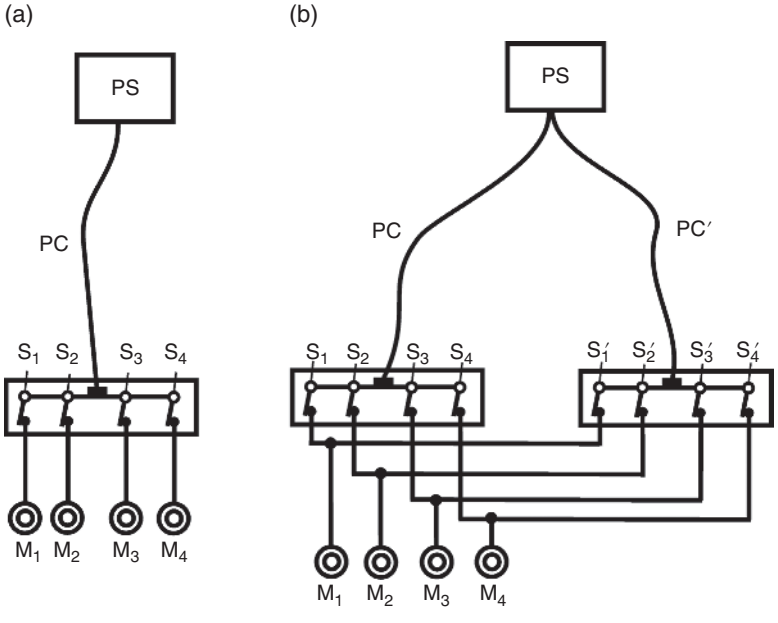


**Figure 1.15** A functional diagram of a power supply to four electric motors (a) without redundancy and (b) with redundancy

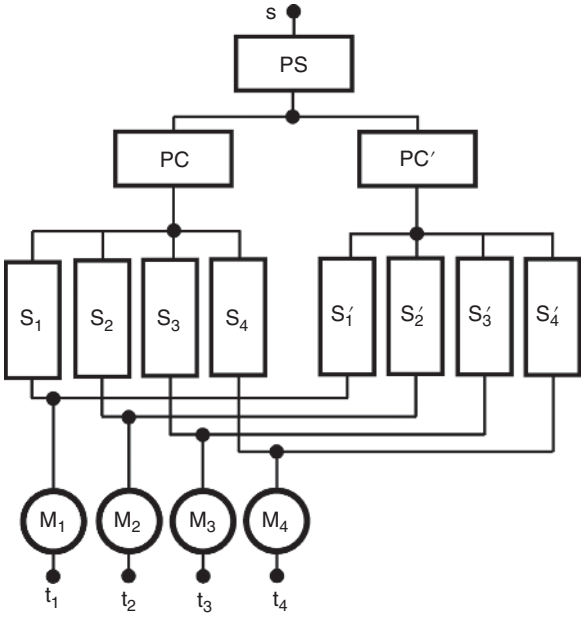


**Figure 1.16** A reliability network of the safety-critical system from Figure 1.15b

### 1.5.2 *Reliability Networks Which Require the Use of Undirected Edges Only, Directed Edges Only or a Mixture of Undirected and Directed Edges*

Commonly, in traditional reliability networks, only undirected edges are used (Andrews and Moss, 2002; Billinton and Allan, 1992; Blischke and Murthy, 2000; Ebeling, 1997; Hoyland and Rausand, 1994; Ramakumar, 1993). This traditional representation is often insufficient to model correctly the logic of system's operation and failure. Often, introducing directed edges is necessary to emphasise that the edge can be traversed in one direction but not in the opposite direction. Consider, for example, the electronic control system in Figure 1.17a, which consists of a control module CM, electronic control switches $K_1$–$K_4$ and four controlled devices $S_1$–$S_4$.

Assume for the sake of simplicity that the connecting cables are perfectly reliable. As a result, the reliability of the system in Figure 1.17 is determined by the reliability of the control module, the electronic control switches and the controlled devices. Suppose that a signal sent by the control module must reach all four controlled devices $S_1$–$S_4$. The reliability of the system is defined as 'the probability that a control signal from the control module CM will reach every single controlled device and all controlled devices will be in working state'.

Similar to the power supply system from Figure 1.15, the reliability of the control system in Figure 1.17a can be improved significantly by making some of the components redundant (e.g. the control module and the electronic control switches) and by providing dual control channels to each controlled device. As a result, from the system in Figure 1.17a, the system in Figure 1.17b is obtained. For the system in Figure 1.17b, for example, the controlled device $S_1$ will still receive the controlling signal if the control module $CM_1$ or the switch $K_1$ fails. The control signal will be received through the alternative control module $CM_2$ and the switch $K_5$. The same applies to the rest of the controlled devices. The control signal will not be received only if both control channels fail.

Despite the seeming similarity between the reliability network in Figure 1.18 of the control system and the reliability network in Figure 1.16 of the power supply system, there are essential differences. The power supply system in Figure 1.15b, for example, will be fully operational after the failure of power cable PC′ and switches $S_2$, $S_3$ and $S_4$ (see Figure 1.19a). In contrast, after the failure of control module $CM_2$ and switches $K_2$, $K_3$ and $K_4$, only device $S_1$ will receive the control signal. This is because unlike the current in the power supply system, the control signal transmitted to device $S_1$ cannot reach the other controlled devices by travelling backwards, through the electronic control switch $K_5$. This backward path has been forbidden *by introducing directed edges in the reliability network.*

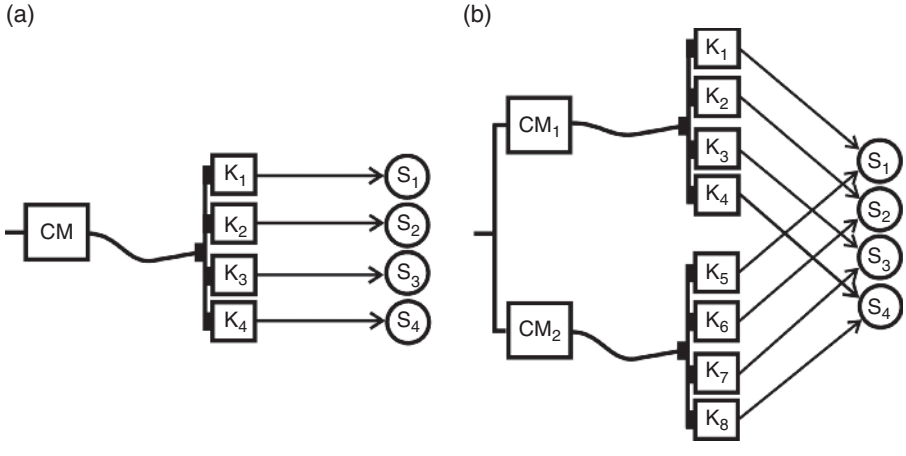


**Figure 1.17** An example of a control system including control modules, switches and controlled devices: (a) a single-control system and (b) a dual-control system

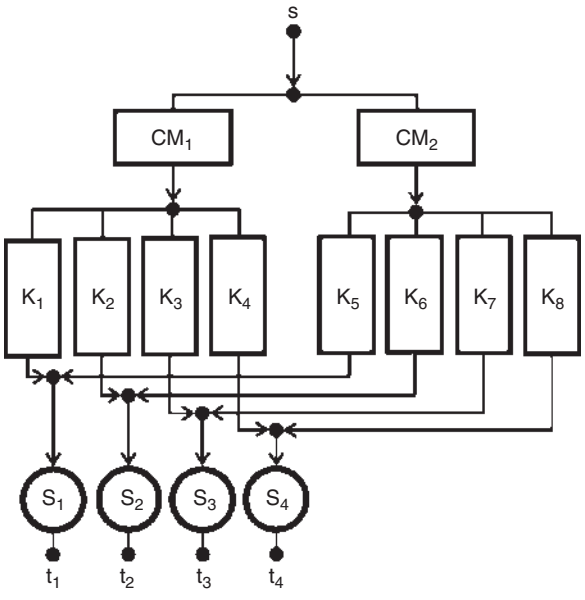


**Figure 1.18** A reliability network of the control system from Figure 1.17b

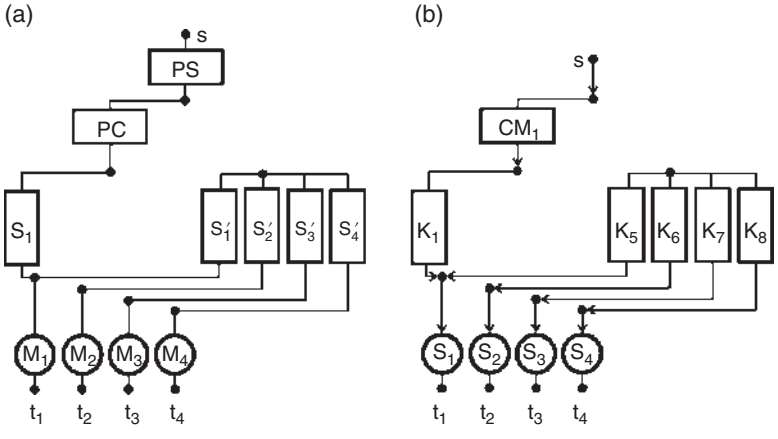


**Figure 1.19** An illustration of the difference between the reliability networks in (a) Figures 1.16 and (b) 1.18

A unique sequence of edges between the start node s of the reliability network and any of the terminal nodes will be referred to as *a path*. Edges which point into the direction of traversing the path will be referred to as *forward edges*, edges without direction will be referred to as *undirected edges*, while edges pointing in the opposite direction of the path traversal will be referred to as *backward edges*. A valid path in a reliability network connecting the start node with any of the terminal nodes can have forward edges or undirected edges or both, but it cannot have backward edges.

Thus, in Figure 1.20a, edge (i, j) is a forward edge, while edge (j, k) is a backward edge, and no transition can be made from node j to node k. The sequence of edges between the start node s and the terminal node t of Figure 1.20a is not a valid connecting path. The sequence of edges in Figure 1.20b, however, is a valid s–t path because it consists of forward and undirected edges only.

The next example features a system where both directed and undirected edges are necessary for describing correctly the logic of system operation. The safety-critical system in Figure 1.21 features two power

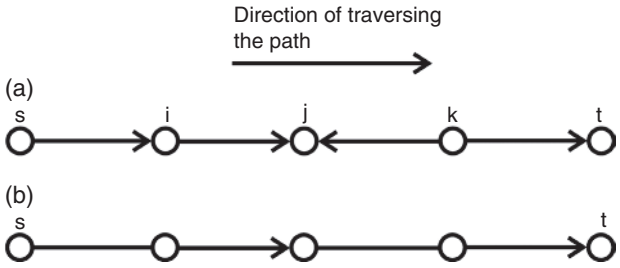


**Figure 1.20** The sequence of edges in (a) does not constitute a valid connecting path because of the backward edge (j, k). The sequence of edges in (b) constitutes a valid connecting path

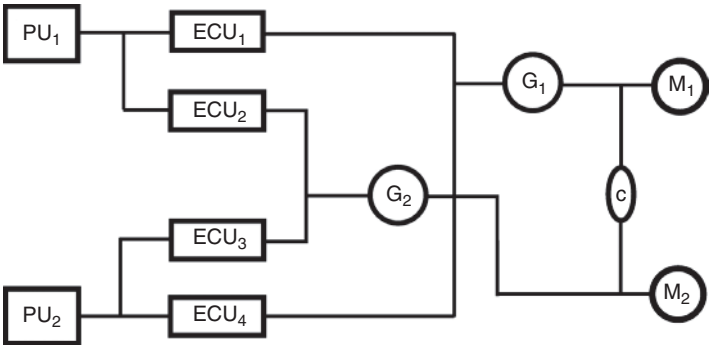


**Figure 1.21** Two power generators $G_1$ and $G_2$ powering two electric motors $M_1$ and $M_2$. The power generators are controlled by four electronic control units $ECU_1$–$ECU_4$, powered by the units $PU_1$ and $PU_2$

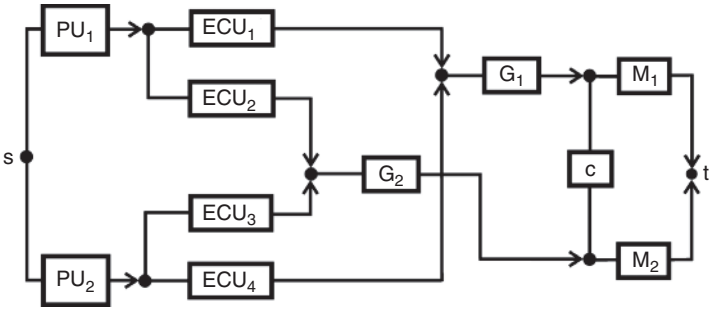


**Figure 1.22** Reliability network of the system in Figure 1.21. Both directed and undirected edges are necessary to correctly represent the logic of system's operation

generators $G_1$ and $G_2$ delivering current to two electric motors $M_1$ and $M_2$. The system is in operation when at least a single electric motor is in operation. The identical, independently working power generators $G_1$ and $G_2$ are controlled by four identical electronic control units $ECU_1$, $ECU_2$, $ECU_3$ and $ECU_4$ powered by two power units $PU_1$ and $PU_2$ (Figure 1.21). The redundant electronic control units guarantee that the control over the generators will be maintained even if some of the control units have failed.

To further reduce the risk of system failure, a bridge (power cable) c has also been included. The bridging power cable 'c' guarantees the system's operation in the case where both the electric motor $M_1$ and the power generator $G_2$ are in failed state at the end of a specified time interval or in the case where both the electric motor $M_2$ and the power generator $M_1$ are in failed state.

The reliability network of the system from Figure 1.21 is given in Figure 1.22.

As can be verified, both directed and undirected edges are necessary to represent correctly the logic of system's operation. The electronic control units, for example, cannot be represented by undirected edges. Otherwise, this would mean that a control signal will exist for generator $G_1$ if the power unit $PU_1$ and $ECU_4$ are in failed state and the power unit $PU_2$ is in working state and the electronic control units $ECU_1$, $ECU_2$ and $ECU_3$ are in working state. This is not possible because the control signal cannot travel from $ECU_3$ to $G_1$ through $ECU_2$ and $ECU_1$. The directed edges are necessary *to forbid* such redirection. On the other hand, the bridge 'c' in Figure 1.22 cannot be represented by a directed edge, because the current must travel in both directions of the bridge, from $G_1$ to $M_2$ and from $G_2$ to $M_1$. The edge representing the bridge 'c' must be undirected edge.

### *1.5.3 Reliability Networks Which Require Different Edges Referring to the Same Component*

In the traditional reliability block diagrams, different edges always correspond to different components. The next example, however, reveals that sometimes, the description of the logic of operation and failure, even for simple mechanical systems, cannot avoid using different edges referring to the same component.

The mechanical system in Figure 1.23 consists of a plate connected through the pin joints $a_2$, $b_2$, $c_2$ and $d_2$ and the struts A, B, C and D to the supports $a_1$, $b_1$, $c_1$ and $d_1$. For a strut to support the plate, it is necessary that the strut and its pin joints to be all in working condition. Therefore, the strut and its pin joints are logically arranged in series. For the sake of simplicity, the strut and both of its pin joints are aggregated and treated as a single component called 'strut assembly'.

The structure in Figure 1.23 is stable if all four strut assemblies are in working state, if any three of the strut assemblies are in working state or if strut assemblies A and B are in working state. In the rest of the cases, the structure collapses. For example, if only strut assemblies C and D are in working state,

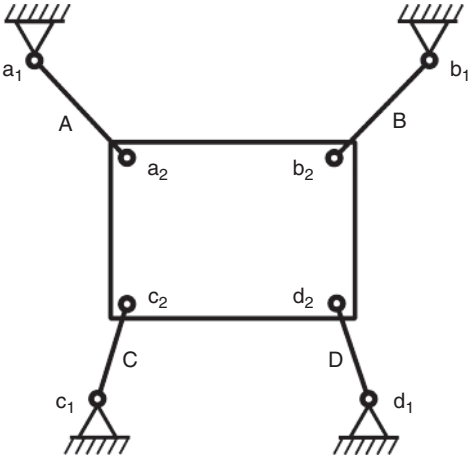


**Figure 1.23** A simple mechanical structure

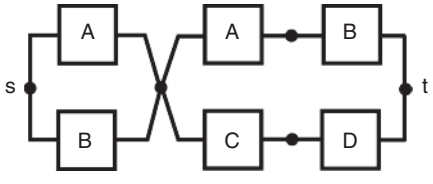


**Figure 1.24** The reliability network of the structure in Figure 1.23

the structure collapses. The structure also collapses if only strut assemblies C and B are in working state or if only strut assemblies D and A are operational.

The reliability block diagram of the mechanical structure is shown in Figure 1.24. As can be seen, even for this simple mechanical system, to represent correctly the logic of reliable operation, it is necessary that different edges refer to the same components A and B in the reliability network.

It must be pointed out that in the reliability network from Figure 1.24, the edges marked by A and B *cannot be treated as statistically independent components* because whenever an edge labelled A is in a failed/working state, the other edge also labelled A is in a failed/working state. The same applies to the edges labelled B. Consequently, the reliability of this system cannot be determined through the well-known analytical relationships working for systems with parallel–series arrangement. The reliability of such systems however can be determined easily by using the Monte Carlo simulation technique described in Chapter 10.

### *1.5.4 Reliability Networks Which Require Negative-State Components*

Traditional reliability block diagrams do not deal with negative-state components – components which provide connection between their nodes in the reliability network only if they are in a failed state. An example of a reliability network which requires a negative-state component can be given with the system for transportation of toxic gas in Figure 1.25 through parallel pipes with flanges. The system includes a pump (P) control module (CM) toxic gas sensors ($TS_1$ and $TS_2$) and seals ($O_1$, $O_2$). To protect personnel in the case of toxic gas release from the seals $O_1$ and $O_2$ of the flanges, an enclosure sleeve ES has been added, sealed by the seals $K_1$, $K_2$ and $K_3$. If a toxic gas escapes in the enclosure sleeve ES from the flange seals $O_1$ or $O_2$, it is expected that sensor $TS_1$ or sensor $TS_2$ will detect the toxic gas release and through the power cut off control module CM will cut the power to the pump (P) and the supply of toxic gas will stop. Stopping the toxic gas supply by cutting the power to the pump prevents the formation of dangerous concentration of toxic gas in the environment. Only one working sensor is needed for the activation of the control module. If the active protection system based on sensors fails to operate, the only remaining barrier to the formation of a dangerous concentration of toxic gas and the environment are the seals $K_1$, $K_2$ and $K_3$.

It is assumed that the enclosure sleeve ES is a perfectly reliable component.

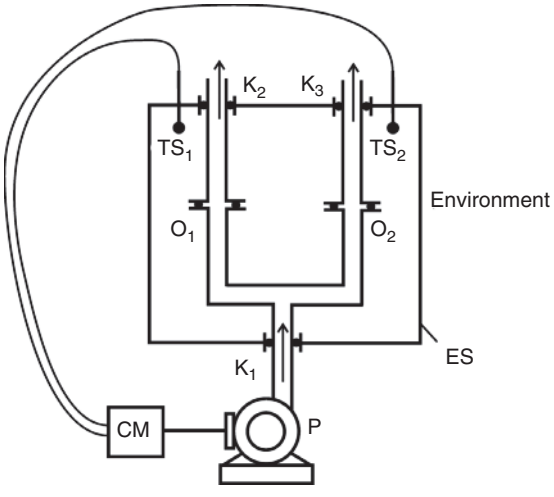


**Figure 1.25** A system supplying toxic fluid

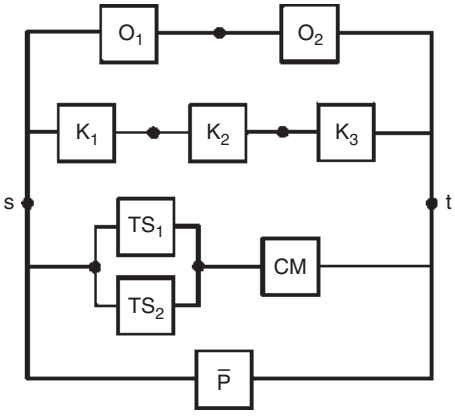


**Figure 1.26** Reliability network of the system from Figure 1.25

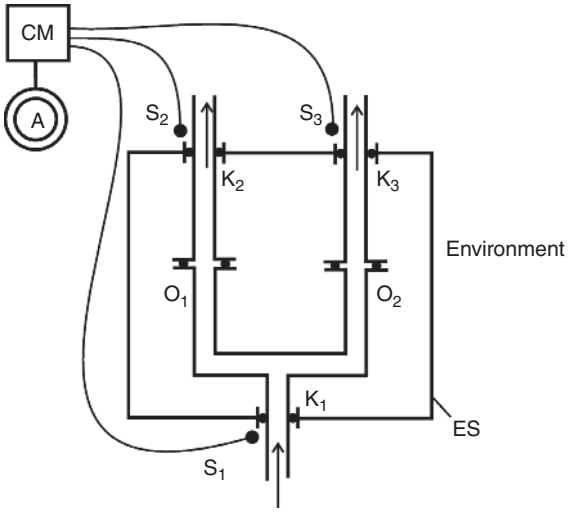


**Figure 1.27** A system supplying toxic gas with three sensors and an alarm

In order to isolate the toxic fluid from the environment, either both seals $O_1$ and $O_2$ work (seal) or the toxic fluid release is sensed and the power to the pump is cut off or all three seals $K_1$, $K_2$ and $K_3$ work. The power to the pump is cut off if at least one of the sensors $TS_1$ or $TS_2$ detects the toxic fluid release and the control module works. The state of the pump does not affect the reliability network of the switching off branch, and this is why the pump is not present there. The state of the pump however does affect the reliability network with respect to the function "prevention of a toxic fluid release in the environment". If the pump is in a failed state, the environment is automatically protected because toxic fluid is no longer supplied. The pump is therefore logically arranged in parallel, as a negative-state component which provides connection between the start node s and the terminal node t only when the pump is not working (Figure 1.26).

Consider now a modification of the system in Figure 1.25. In the case of a leak of toxic gas from the flanges and from the seals $K_1$, $K_2$ or $K_3$, the role of the sensors $S_1$, $S_2$ and $S_3$ is to detect the toxic gas release and to trigger the control module CM into activating the alarm A. The sensors can detect a toxic gas release only locally, in the immediate vicinity of the seal they are attached to (Figure 1.27).

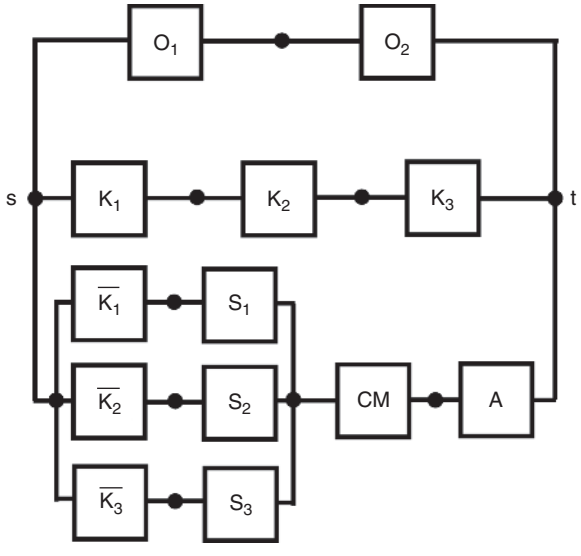


**Figure 1.28** Reliability network of the system from Figure 1.27

In order to isolate the toxic gas from the environment, either both seals $O_1$ and $O_2$ work (seal) or all three seals $K_1$, $K_2$ and $K_3$ work. Therefore, the block of O-seals and the block of K-seals are logically arranged in parallel. In the case of failure of any of the K-seals, the alarm can be activated if the control module CM, the alarm and the corresponding sensor are in working state. The correct logical arrangement of the components is given in Figure 1.28. The components $\overline{K}_1$, $\overline{K}_2$ and $\overline{K}_3$ are negative-state components. They provide connection between their corresponding nodes only when component $K_1$, $K_2$ or $K_3$ is in failed state, respectively. When component $K_1$, $K_2$ or $K_3$ is in working state, the negative-state component provides no connection between its nodes. Because of the statistical dependence of a component and its negative-state counterpart, Monte Carlo simulation methods are needed for analyzing reliability networks where components and their negative-state counterparts are both present.