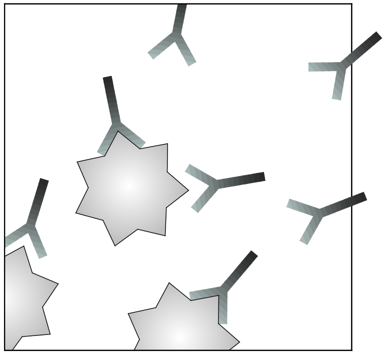
Immune Systems



Companion slides for the book *Bio-Inspired Artificial Intelligence: Theories, Methods, and Technologies* by Dario Floreano and Claudio Mattiussi, MIT Press

Introduction

Biological immune systems

- Living organism must protect themselves from the attempt of other organisms to exploit their resources
- Some would-be exploiter (*pathogen*) is much smaller than its target (the *host*)
  - The organs that the host uses to interact with the environment are poorly suited to the detection and elimination of potential pathogens
  - The pathogen can reproduce much faster than the typical host and can rapidly evolve new strategies of attack
- Physical barriers, alteration of physiological conditions, and avoidance of dangerous environments are only a partial solution
- The host needs a set of countermeasures which operate at the same scale and which can keep the evolutionary pace of the pathogens.
- This collection of countermeasures constitutes the *immune system* of the host.

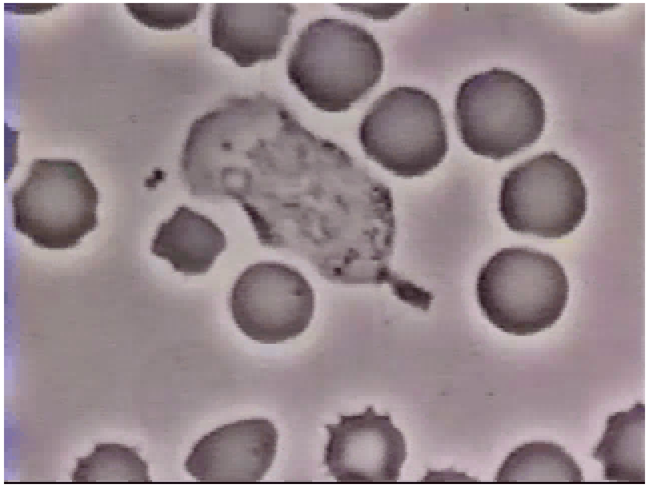
Functions of the immune systems

The immune system must:

- *Detect* the pathogens once they have entered the host body
- *Eliminate* the pathogens with *minimal cost* in terms of resources employed and damage done to the host
- Initiate the repair of the damages done by the pathogen

Additionally, the immune system can be asked to

- Detect and repair the malfunctioning and failures of individual host cells (e.g., damaged, mutated, and cancerous cells)



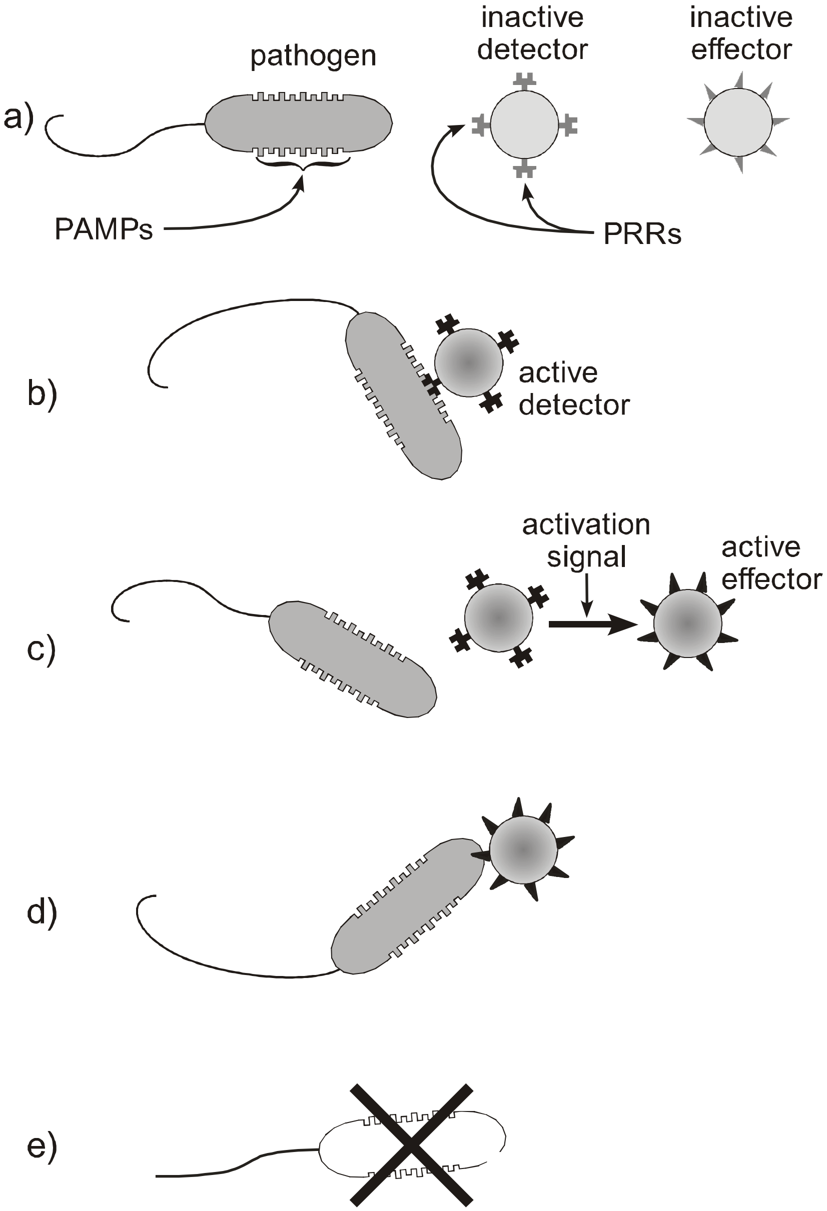
Artificial immune systems

- Human-built systems (e.g., networked computers) must also be protected from the attempt of exploitation of their resources (computational power, data, identity…)
- Some non-authorized operations are executed at a low level in the hierarchy of software levels of the computer system
  - Their effect is not immediately apparent at the scale of the computer user or network administrator interface
  - The strategies of attack can change rapidly
- Isolation of the computing system is seldom an option
- The current solution is the use of antivirus and intrusion detection programs designed and updated by specialized software firms
- A better solution would be a protection system capable of autonomously detecting and opposing the attempts to intrusion and exploitation, that is, an *artificial immune system* (AIS).
- An AIS might also detect and correct (sub)system malfunctioning

How biological immune systems work

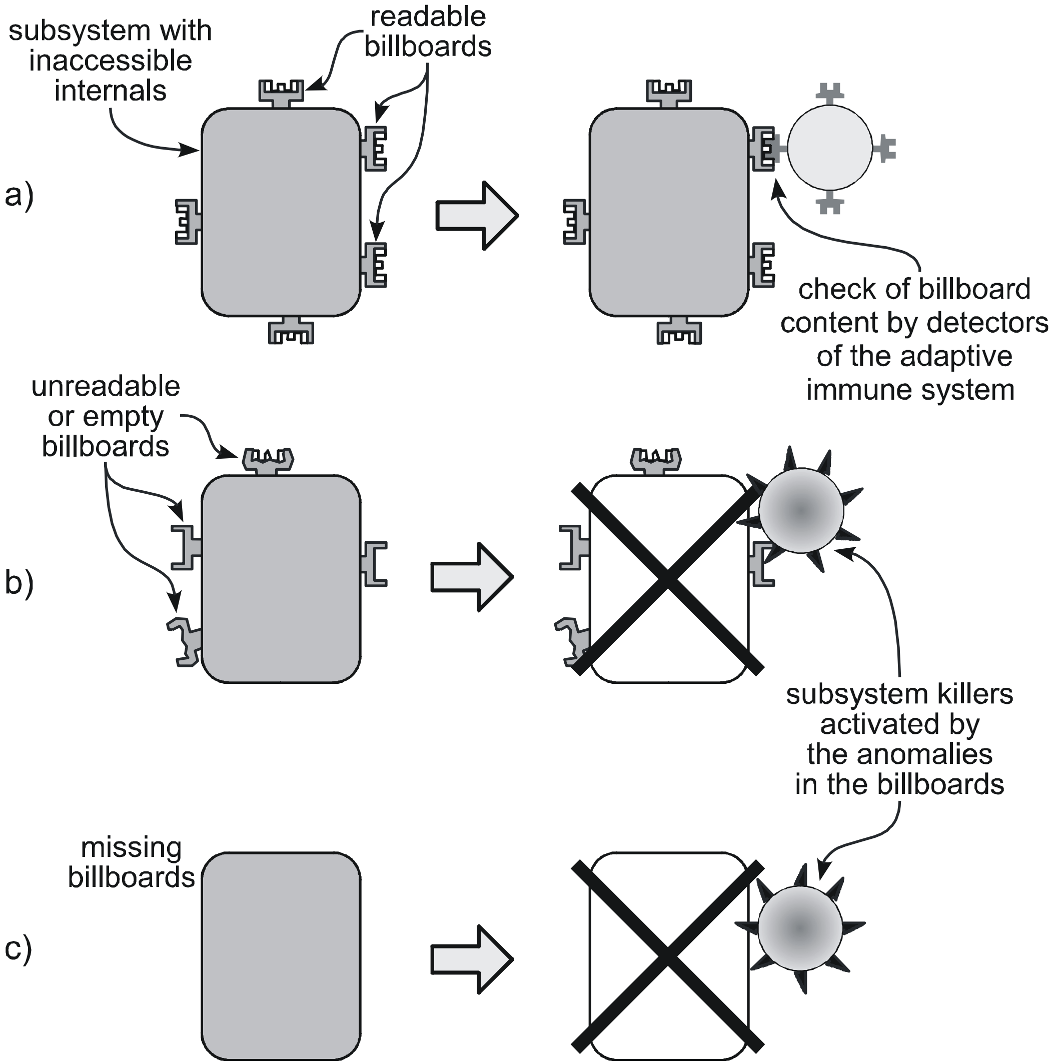
Innate immunity

- It is the set of immune countermeasures that do not change during the lifetime of the host
- It is based on a collection of *immune detectors* and *effectors* (immune “elements”) distributed in the host body
  - The immune detectors carry *pattern recognition receptors* (PRR) that can recognize molecular structures called *antigens*
  - Host antigens are called *autoantigens*
  - Antigens that permit the detection of pathogens are called *pathogen associated molecular patterns* (PAMP)



Monitoring of subsystems

- Some pathogens do not circulate in the host body but "hide" within the host cells, which are not accessible to the immune detectors
- Inaccessible subsystems must report on their internal activity using specialized interfaces ("billboards")
- There is some variability in populations in the way the internal activity of subsystems is reported


subsystem with inaccessible internals
readable billboards
a)
check of billboard content by detectors of the adaptive immune system
unreadable or empty billboards
b)
subsystem killers activated by the anomalies in the billboards
missing billboards
c)

The limits of innate immunity

It cannot change during the lifetime of the host, thus

- It must target PAMPs that the pathogen cannot hide in order to escape detection (e.g., flagella and essential constituents of cell wall of bacteria)

However, pathogens can evolve and

- Change the patterns that are accessible for inspection by immune detectors
- Change the pathogen structures that are exploited by the immune effectors to gain access to the pathogen and destroy it

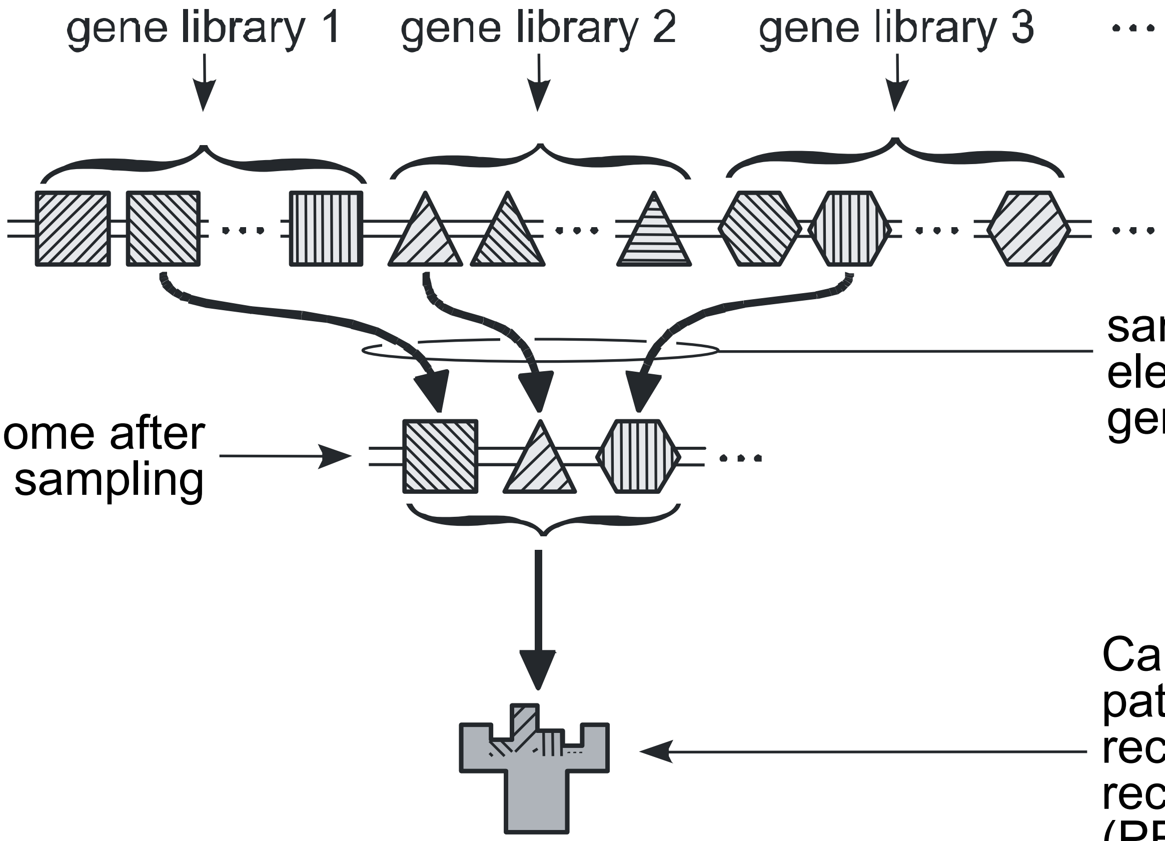
Note that

- An innate immune system capable of recognizing and attacking all the structures carrying patterns that are not found in the healthy host (self/nonself discrimination) has many drawbacks (excessive number of PRRs; intolerance for harmless substances; limitations to changes in the host during, evolution, development, and aging; tolerance of fetus during pregnancy…)

Adaptive immunity

- The *adaptive immune system* (vertebrates) is also based of a collection of immune detectors and effectors
- Both detectors and effectors can change during the lifetime of the host
- This requires the definition of a strategy for the generation of detectors and effectors that
  1. Are effective against pathogens, but
  2. Do not interfere with the normal activity of the host tissues
- The adaptive immune system uses a multi-stage process
  1. Generation of *inactive* elements by random recombination of gene libraries
  2. *Tolerization*, i.e., elimination of *autoreactive* elements by *negative selection* and of non-reactive elements by limiting their lifespan
  3. *Positive selection* of the best non-autoreactive elements
  4. *Activation* of immune elements according to a notion of *context*
  5. Maintenance of a pool of *memory* elements

Generation via gene libraries



genome before the random sampling

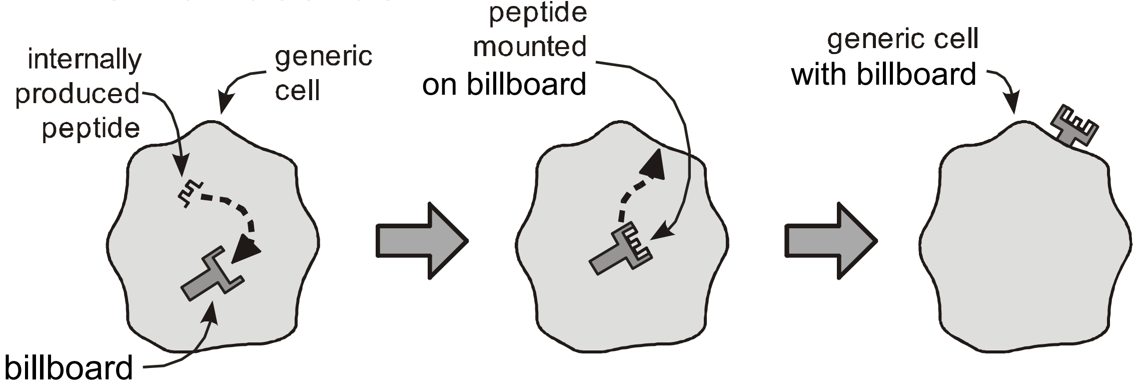
sampling of elements from gene library

genome after the sampling

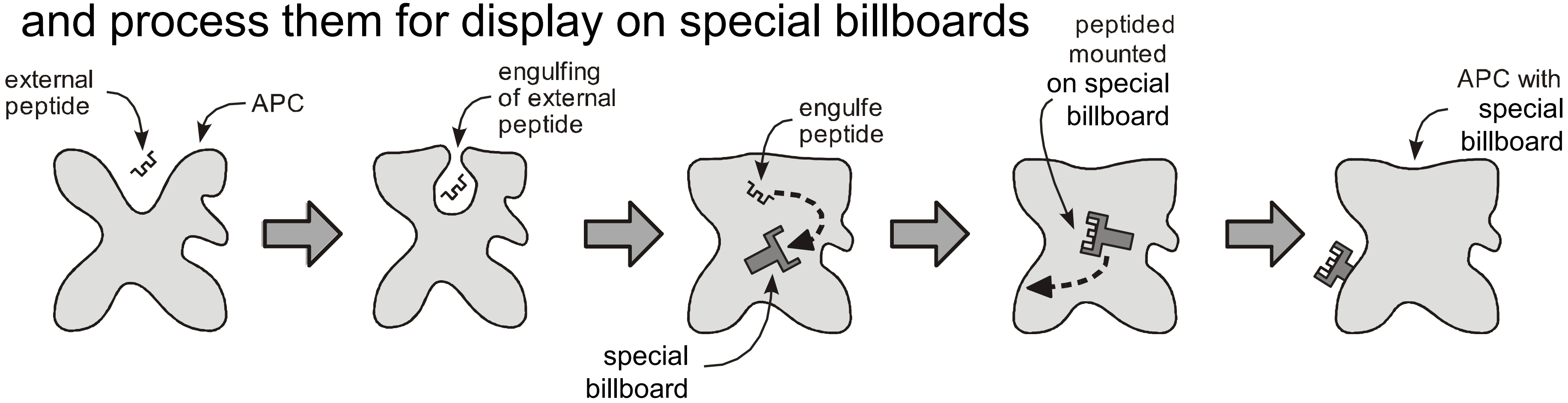
Candidate pattern recognition receptor (PRR)

Antigen presentation

All cells process internally produced molecules (proteins) and display fragments of them on their billboards

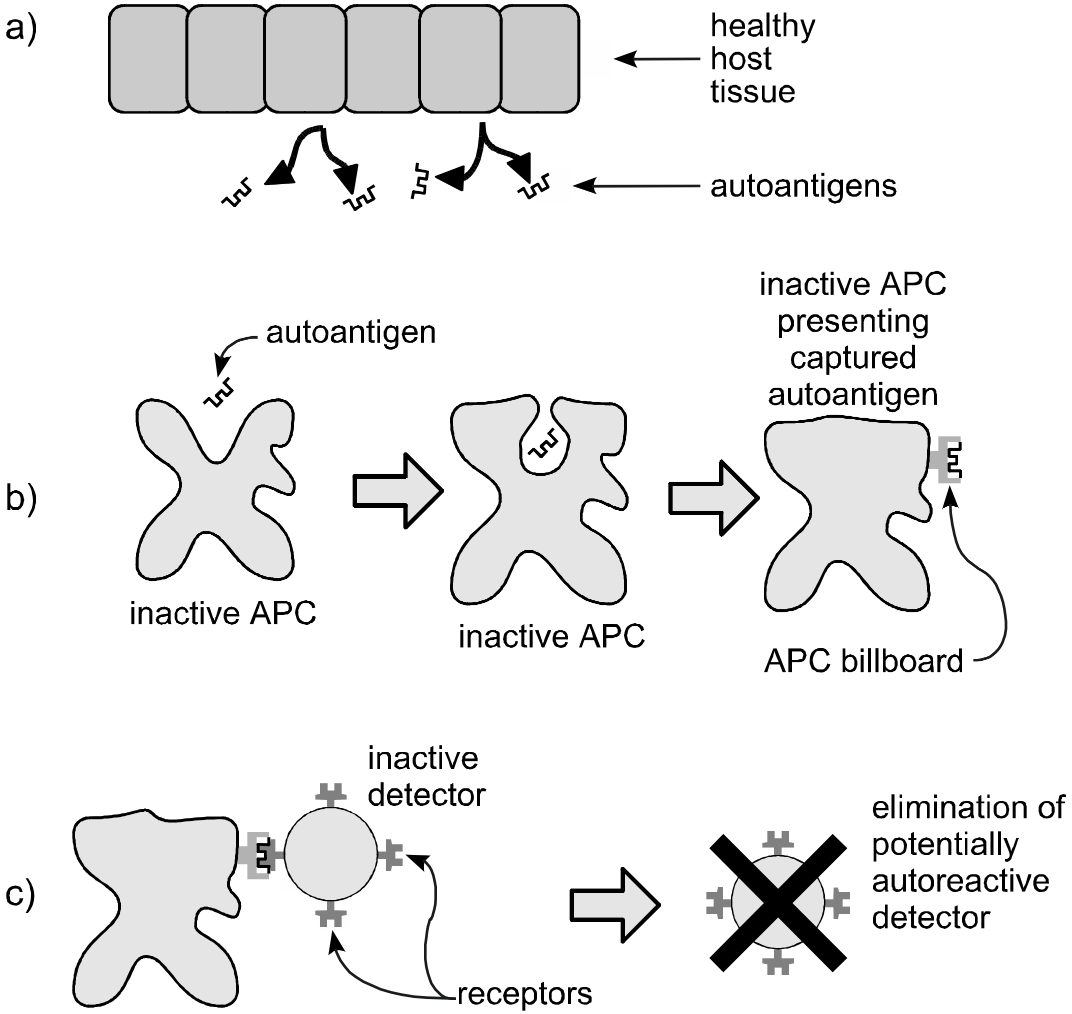

internally produced peptide
generic cell
billboard
peptide mounted on billboard
generic cell with billboard

Specialized antigen presenting cells (APCs) capture external molecules and process them for display on special billboards


external peptide
APC
engulfing of external peptide
engulfe peptide
special billboard
peptided mounted on special billboard
APC with special billboard

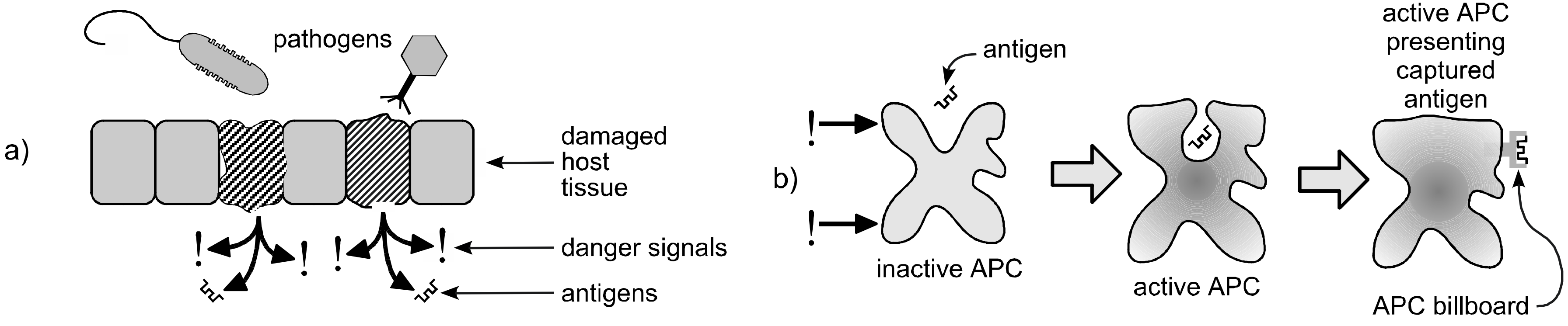
Tolerization

- It is based on the activity of *antigen presentation* of generic cells and of specialized *antigen presenting cells* (APCs)
- A *central tolerization* is performed in specialized host regions after the generation of adaptive immune elements
- *Peripheral tolerization* is performed while adaptive immune elements circulate in the host body

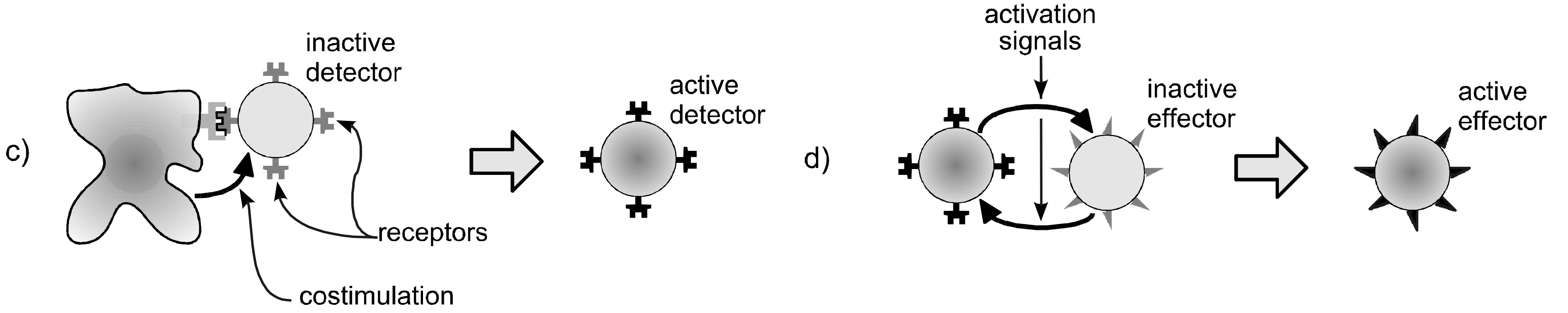


Activation of adaptive immune elements

The damage produced by pathogens results in *danger signals* which activate the Antigen Presenting Cells (APCs)

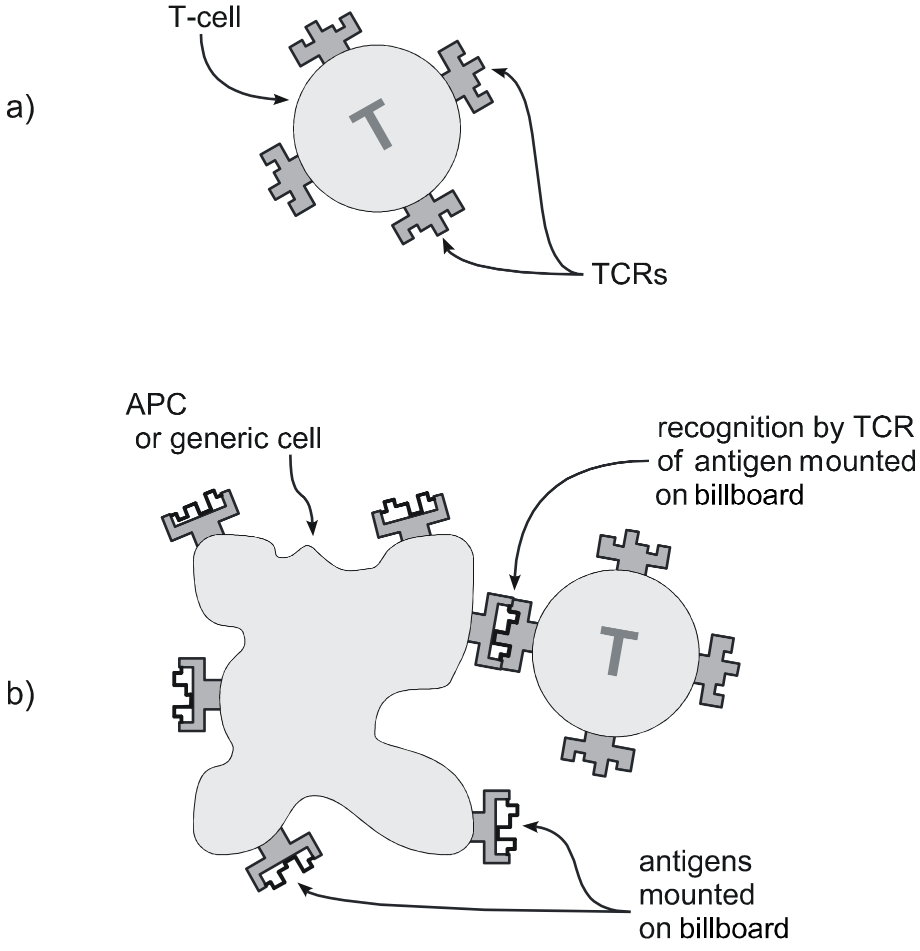


Active APCs activate by *costimulation* the immune elements which recognize the antigens presented by the active APC

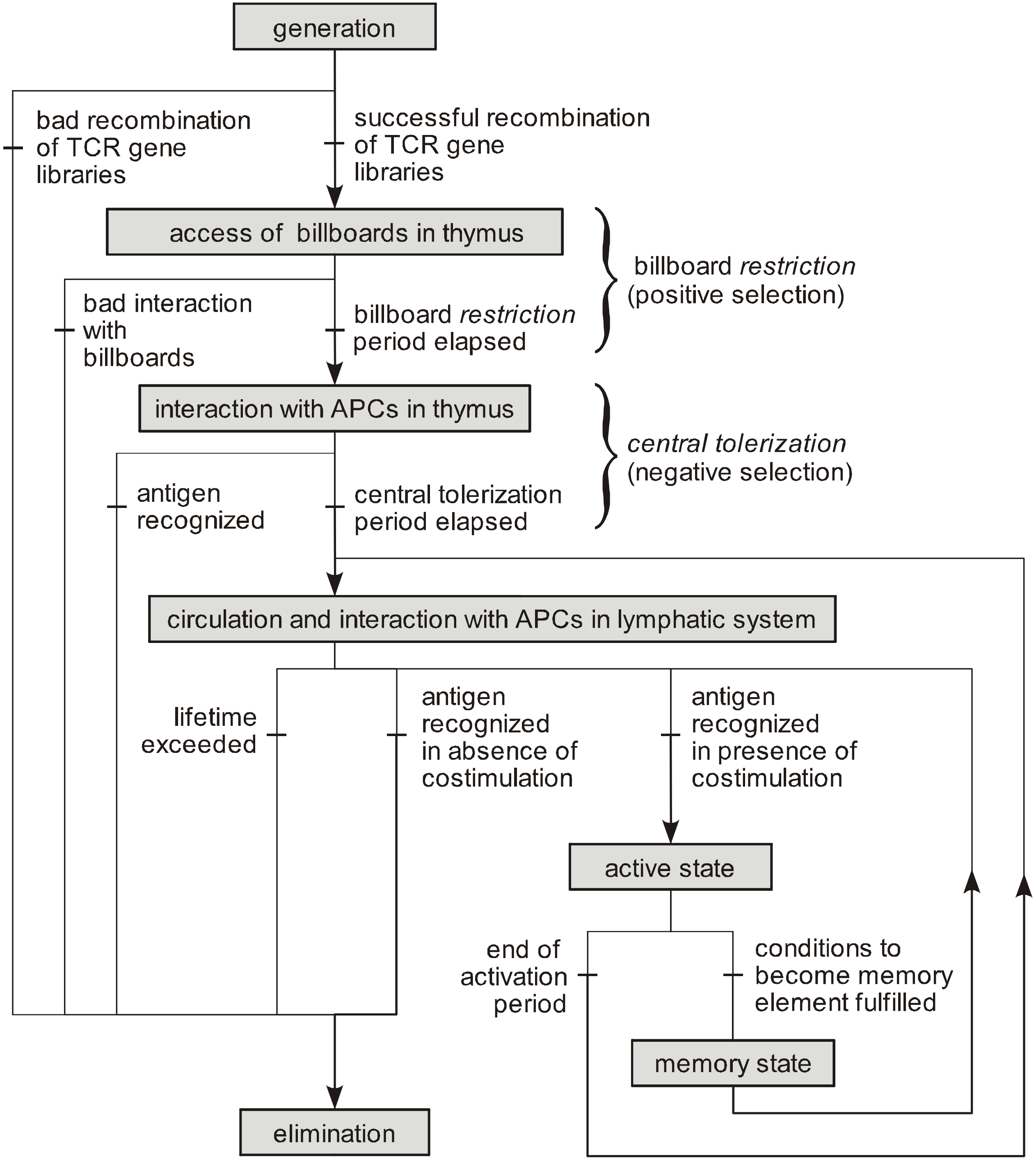


T-cells: “digital” recognition

- We can distinguish a “digital” modality of recognition of antigens from an “analog” one
- *T-cells* work in the “digital” modality: they inspect the billboards with their T-cell receptors (TCRs). When activated they are in charge of:
  - Killing cells that display antigens they recognize
  - Activating the elements that work in the “analog” modality
  - Become *memory T-cells*



Life cycle of T-cells



generation

bad recombination of TCR gene libraries

successful recombination of TCR gene libraries

access of billboards in thymus

billboard *restriction* (positive selection)

bad interaction with billboards

billboard *restriction* period elapsed

interaction with APCs in thymus

central tolerization (negative selection)

antigen recognized

central tolerization period elapsed

circulation and interaction with APCs in lymphatic system

lifetime exceeded

antigen recognized in absence of costimulation

antigen recognized in presence of costimulation

active state

end of activation period

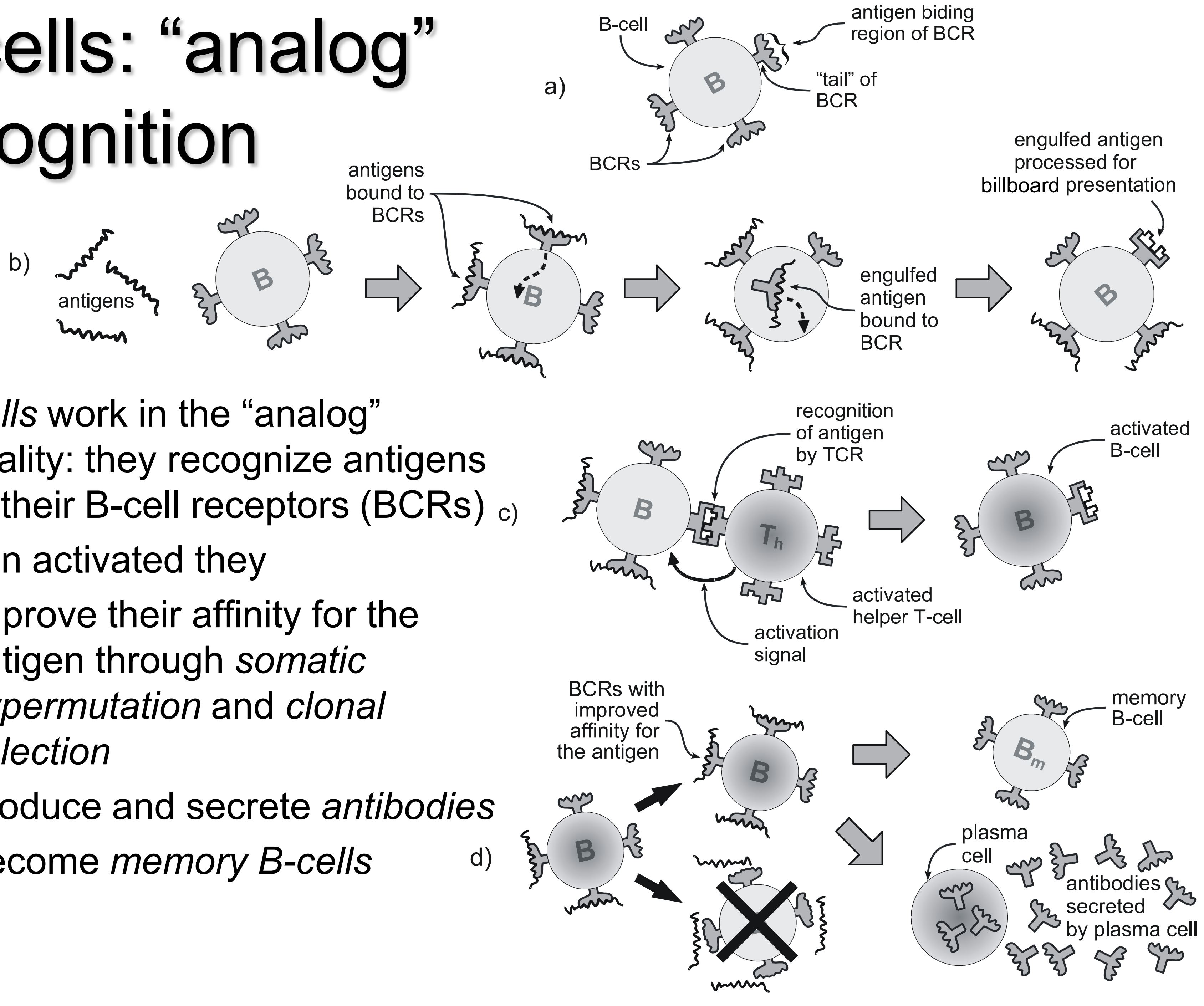
conditions to become memory element fulfilled

memory state

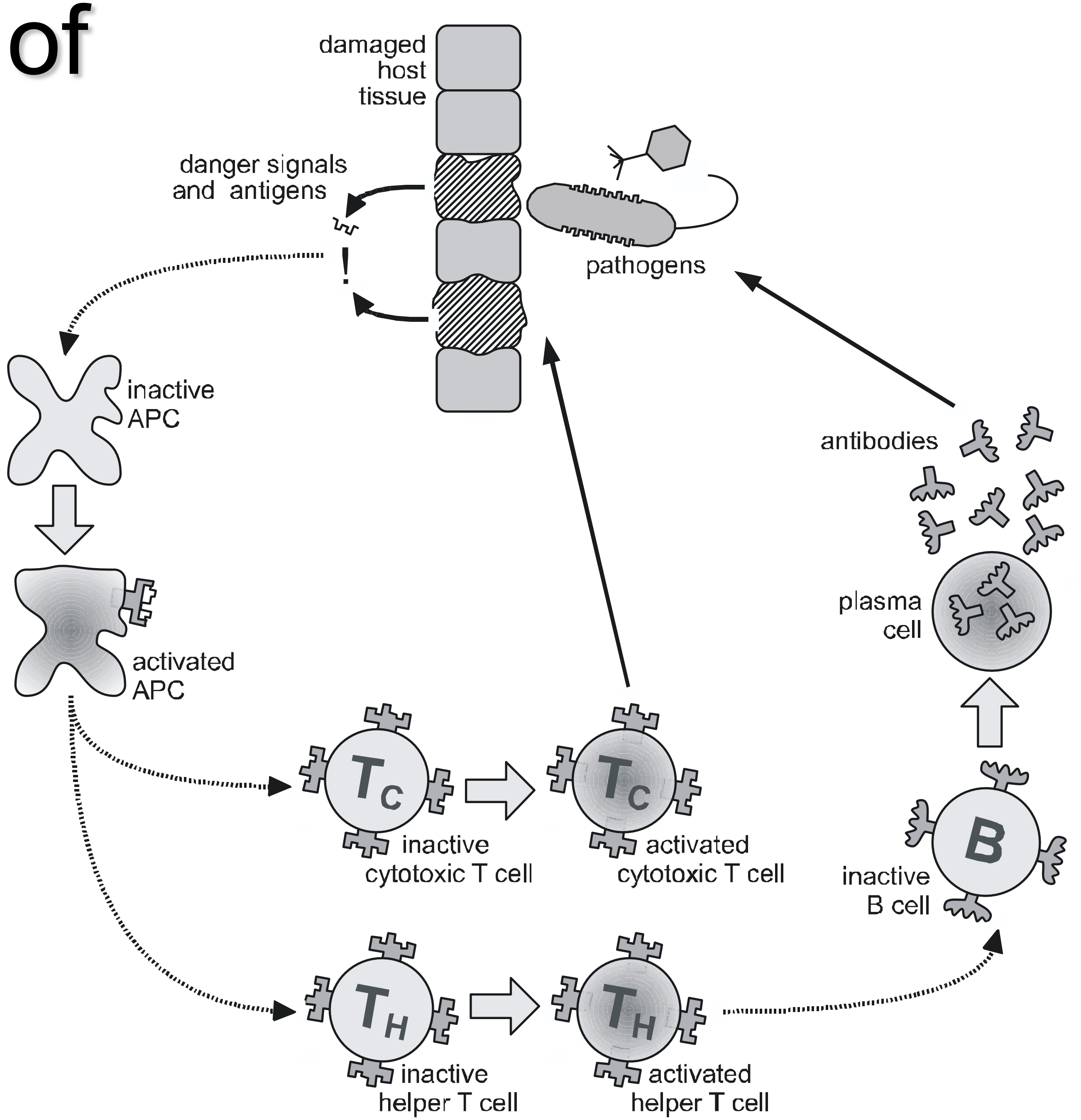
elimination

B-cells: “analog” recognition

- *B-cells* work in the “analog” modality: they recognize antigens with their B-cell receptors (BCRs)
- When activated they
  - Improve their affinity for the antigen through *somatic hypermutation* and *clonal selection*
  - Produce and secrete *antibodies*
  - Become *memory B-cells*

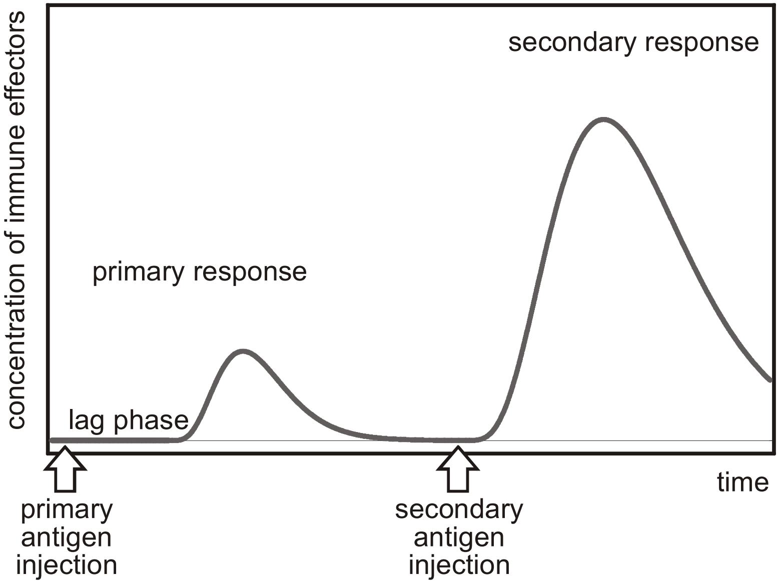

a)
B-cell
antigen biding region of BCR
B
“tail” of BCR
BCRs
b)
antigens
B
antigens bound to BCRs
B
engulfed antigen bound to BCR
engulfed antigen processed for billboard presentation
B
c)
B
recognition of antigen by TCR
T_h
activated B-cell
B
activated helper T-cell
activation signal
d)
BCRs with improved affinity for the antigen
B
B
memory B-cell
B_m
plasma cell
antibodies secreted by plasma cell

Overview of adaptive immunity



The limits of adaptive immunity

- There is a delay in generating an effective response when challenged by a new pathogen (however, further challenges by the same or similar pathogen are met rapidly)
- Possibility of escape by rapid *antigenic variation* (e.g., HIV)
- Possibility of *autoimmune diseases* due to attack of host cells
  - that carry antigens found in a danger zone
  - That carry antigens similar to those of the pathogens
- Possibility of tolerization with respect to pathogens and mutated cells (e.g., cancerous cells) which do not produce early danger signals



Concepts for Artificial Immune Systems

Lessons from biological immune systems

1. **Cost:** An adaptive immune system can be expected to be very expensive in term of resources (explanation of placebo effect?)
2. **Damage and regeneration:** The operation of an immune system can inflict damage to the host; the host must be able to generate new subsystems to replace the ones destroyed
3. **Design for immunity:** The host is explicitly designed to cooperate with the immune system (e.g., by generating danger signals)
4. **Distributedness, decentralization, self-protection and robustness:** The immune system is a self-organizing distributed system composed by autonomous agents. The control of the immune activity is decentralized. The elements of the immune system can control each other. The immune system is self-protecting and robust to the malfunctioning of individual agents.
5. **Parallel operation and scalability:** scaling to different sizes and complexities of the host requires merely the adaptation of the number of immune elements, not their "reprogramming"

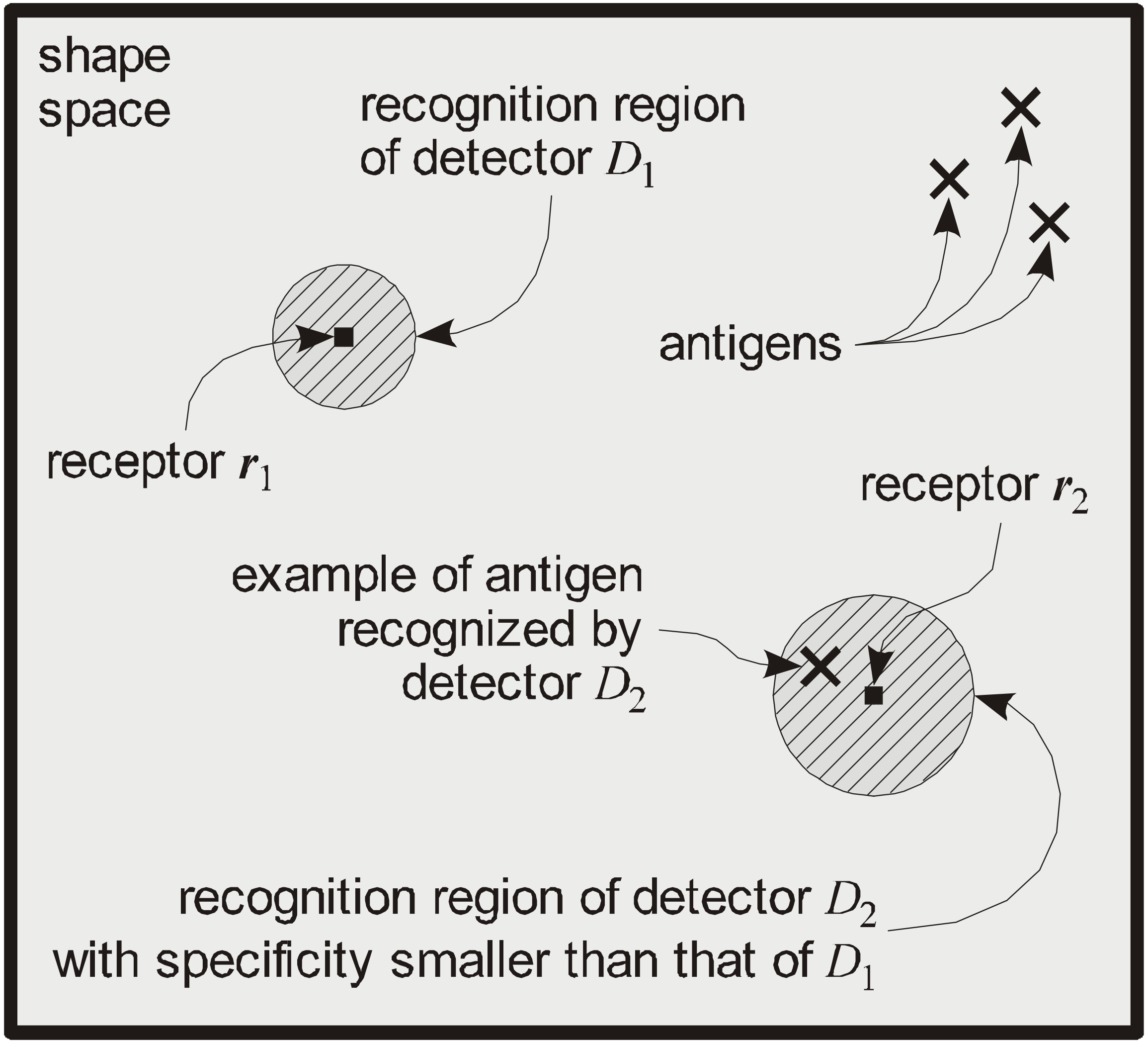
Lessons from biological immune systems

6. **Mechanisms of adaptation and tolerance:** mechanism of central tolerance based on a process of positive and negative selection; mechanism of peripheral tolerance based on danger signals
7. **Risk of autoimmunity**
8. **Dynamic allocation of resources and self-limitation:** the resources available to the immune system are dynamically allocated in terms of type of elements and distribution in the host body; the limited lifetime of most immune elements tapers the response when no longer needed
9. **Mechanisms of generation of diversity:** recombination of elements of genetically encoded libraries of building blocks rather than random generation of receptors from scratch
10. **Strategies of detection:** Use of detectors with different modalities of recognition and different specificity; presentation of multiple “views” of the pathogen
11. **Learning and memory**

Shape space

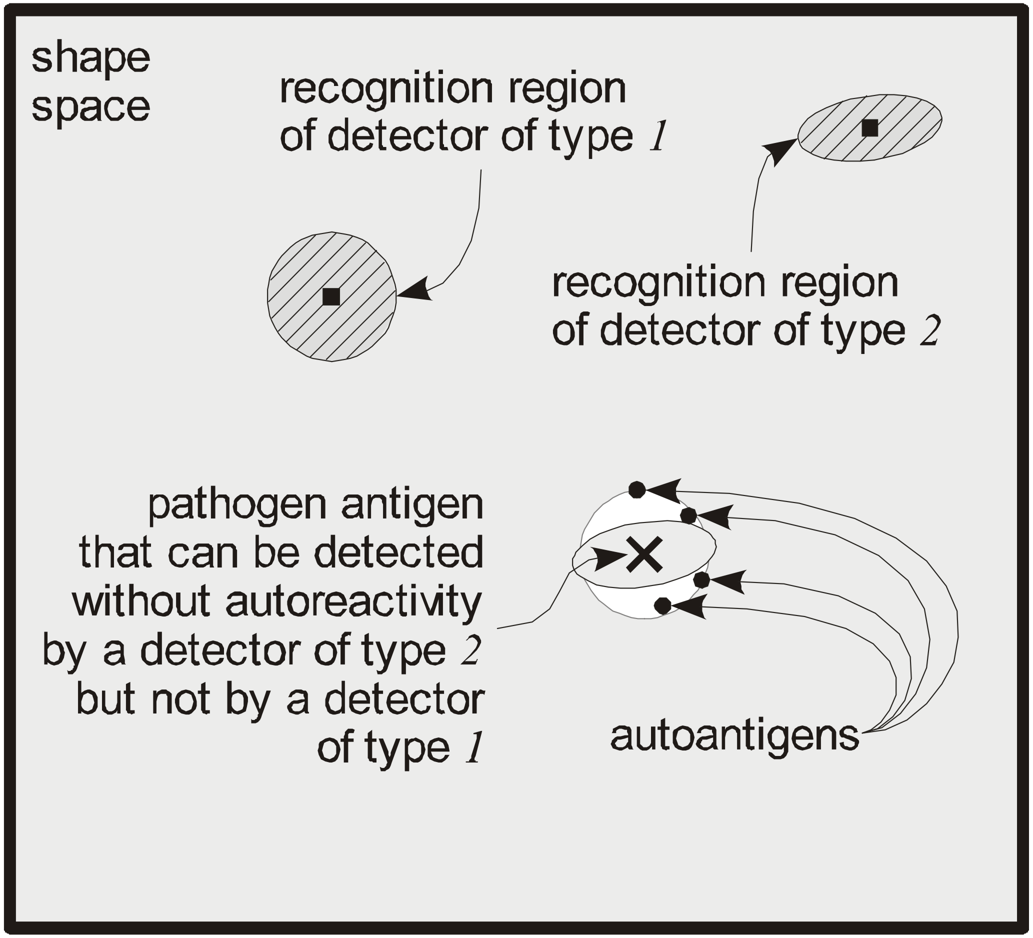
- It is an abstraction that gives a geometrical interpretation to the process of recognition of an antigen from the part of an immune detectors.
- The goal is to obtain a simplified model of the action of the immune system to be used in the analysis and design of artificial immune systems.
- The properties of the antigen and of the receptor are represented with a list of *l* parameters called the *generalized shape*
- We specify a measure of the *affinity* between the receptor **r** and the antigen **a** by defining a distance *d*(**a**,**r**) in the shape space (Euclidean distance, Hamming distance…)
- We say that a detector D equipped with receptors of type **r** recognizes an antigen **a** if *d*(**a**, **r**) is below a certain threshold *θ.*
  - The value of the threshold determines the *specificity* of the detector
  - The region of shape space thus defined is the *recognition region* of D
- The union of the recognition regions of all the detectors of an immune system is called its *immune repertoire*

Shape space


shape space
recognition region of detector D_1
antigens
receptor r_1
receptor r_2
example of antigen recognized by detector D_2
recognition region of detector D_2 with specificity smaller than that of D_1

Holes in immune repertoire

- Ideally, the immune repertoire should cover all the regions of space that do not correspond to autoantigens.
- If this is not the case, we say that the immune repertoire has *holes* that can be potentially exploited by a pathogen to escape detection.
- A technique reduce the possibility of there being holes consists in implementing several distinct distance functions


shape space
recognition region of detector of type *1*
recognition region of detector of type *2*
pathogen antigen that can be detected without autoreactivity by a detector of type *2* but not by a detector of type *1*
autoantigens

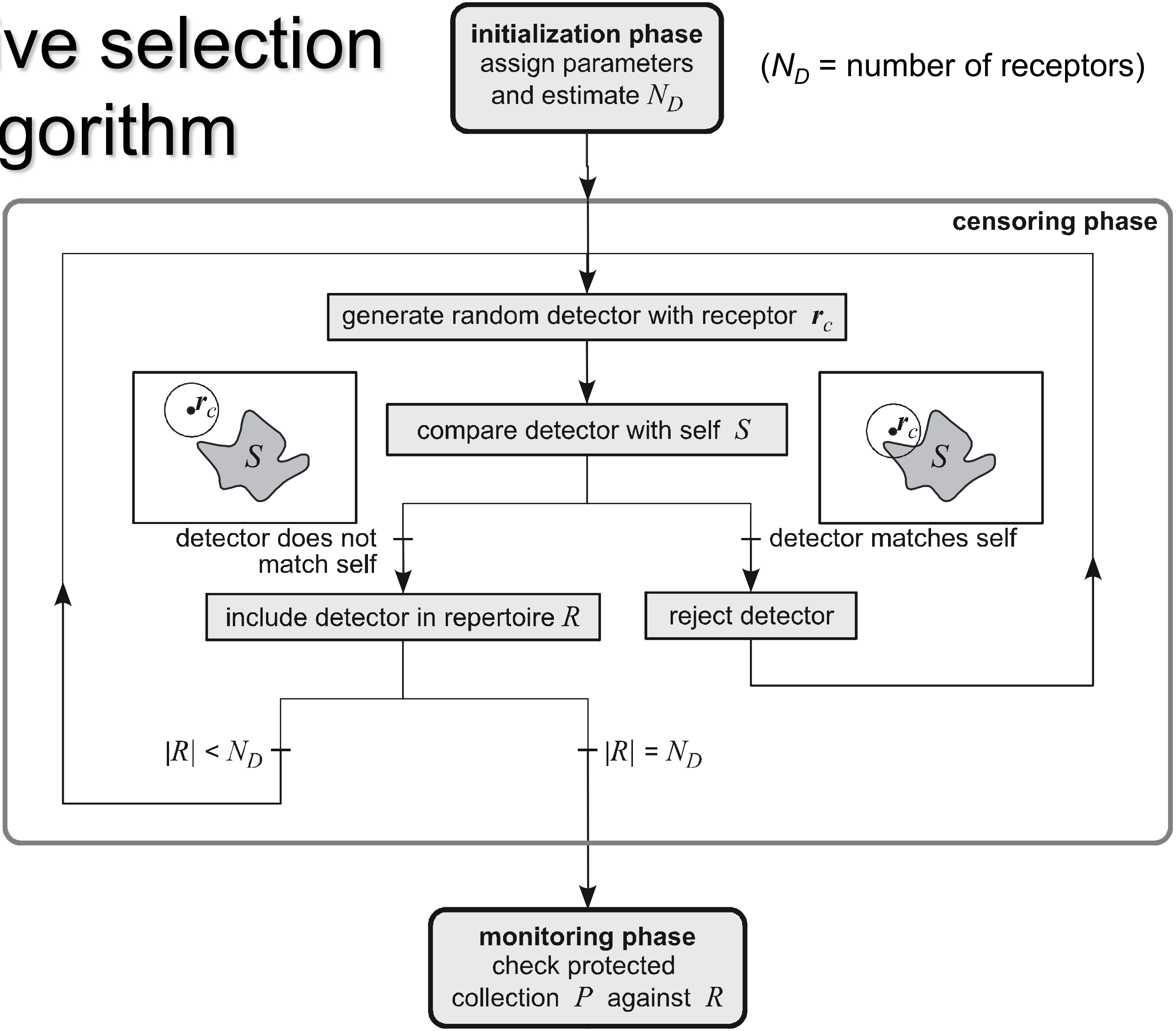
Companion slides for the book *Bio-Inspired Artificial Intelligence: Theories, Methods, and Technologies* by Dario Floreano and Claudio Mattiussi, MIT Press

Algorithms

Negative selection algorithm

- One of the strategies used by the vertebrate immune system for the generation of an immune repertoire is the random generation of receptors and a process of *negative selection* that removes the receptors that match autoantigens.
- The *negative selection algorithm* for AIS assumes that there is a collection *P* of fixed-length strings of symbols (e.g., data and program files) which must be protected from unauthorized change with respect to a reference collection *S* called the *self.*
- The goal of the algorithm is to generate a set of detectors that can signal the appearance in *P* of any string that does not belong to *S*, that is, the appearance in *P* of any nonself string.
- The algorithm starts by assigning a similarity function m(·, ·) for pairs of strings, a detection threshold θ, a mechanism of generation of candidate receptor strings, and the maximum acceptable probability *p* of detection failure. It then estimates the required number of receptors N_D in the repertoire *R.*

Negative selection algorithm



(N_D = number of receptors)

initialization phase
assign parameters and estimate N_D

censoring phase

generate random detector with receptor $\boldsymbol{r}_c$

compare detector with self S

detector does not match self

detector matches self

include detector in repertoire R

reject detector

$|R| < N_D$

$|R| = N_D$

monitoring phase
check protected collection P against R

Clonal selection algorithm

- In the vertebrate immune system, the B-cells that survive the process of negative selection undergo a further process of *affinity maturation* based on *clonal selection* that improves their recognition performance.
- A *clonal selection algorithm* based on the characteristics of this biological process has been proposed for pattern function optimization problems.

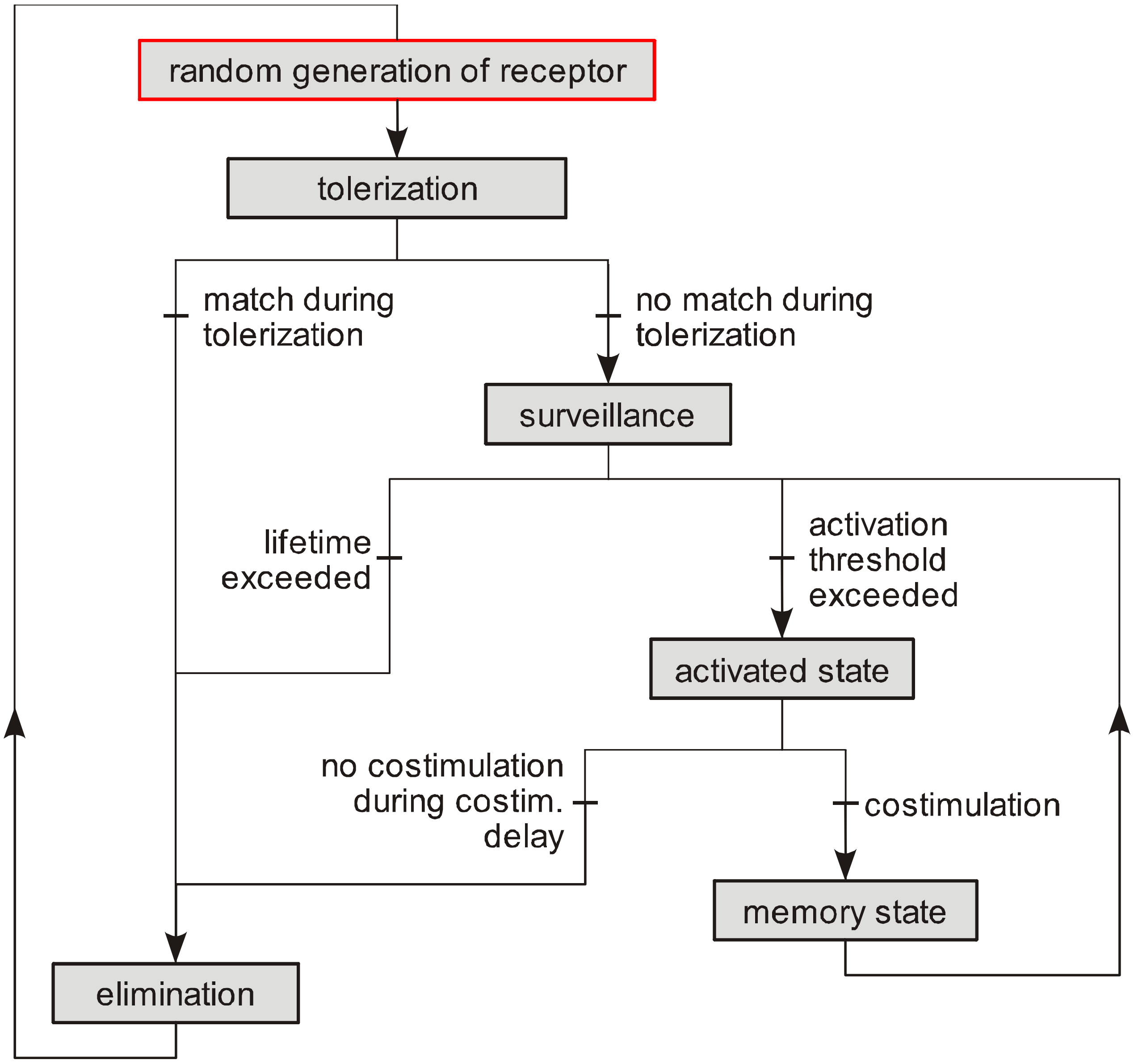
```
Randomly initialize a population R of tentative solutions (receptors)
Repeat
    For each receptor in R
        Determine the function value (affinity)
        Select receptors with highest affinity
        Clone the selected elements and mutate them with rate of
        mutation inversely proportional their affinity, obtaining R’
        generate new mutants R’’
    Select the highest affinity receptors in R, R’, and R’’ to form the new R
Until a stopping criterion is met
```

Examples of Artificial Immune Systems

ARTificial Immune System (ARTIS)

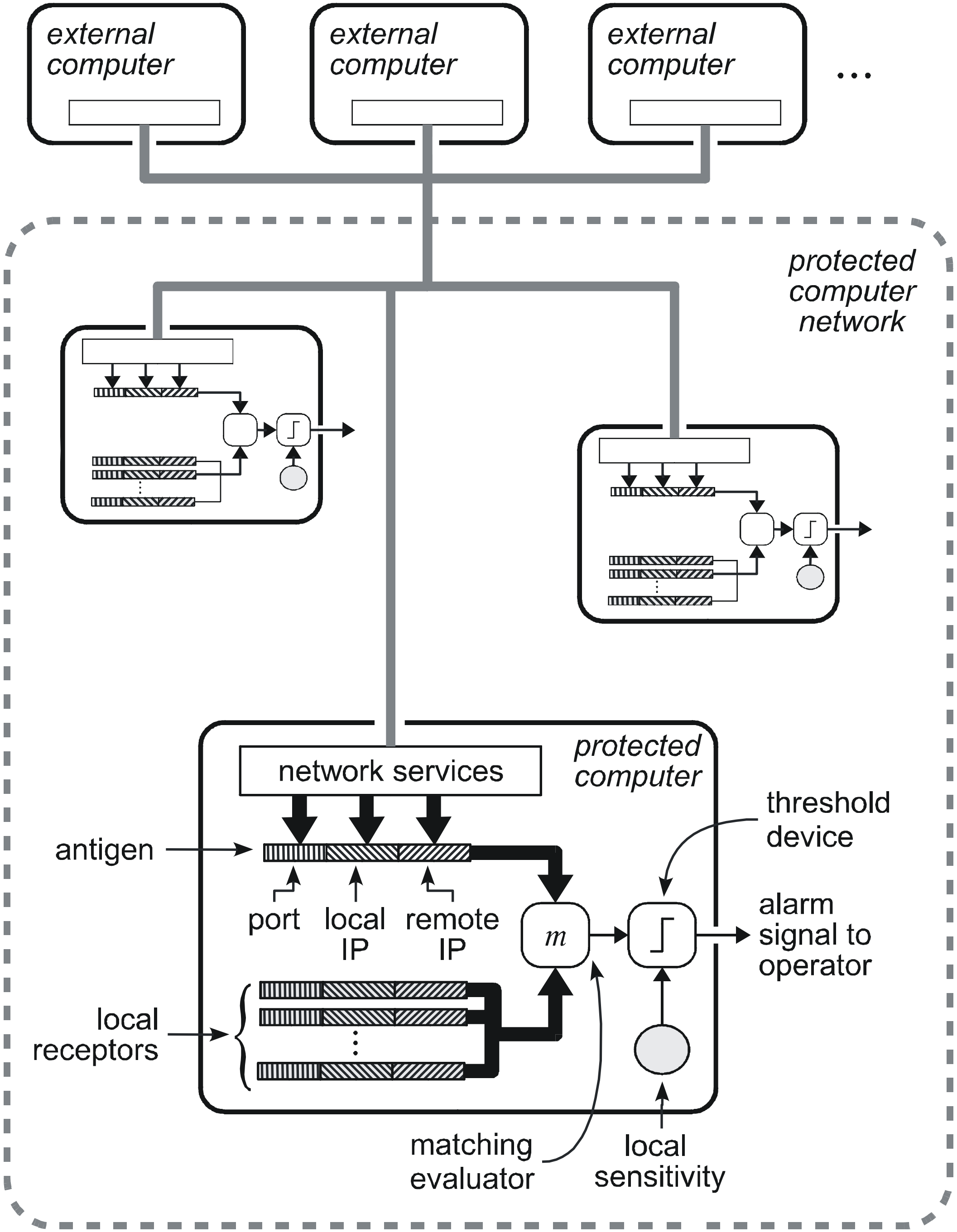
- ARTIS is an AIS framework that models many of the processes and properties of the vertebrate immune system.
- The goal of ARTIS is to specify the elements of a general adaptive distributed system without reference to any specific application.
- In ARTIS the system to be monitored is a distributed environment represented by nodes that can exchange information.
  - At each node a collection of fixed-length strings is defined, which are the target of the security monitoring. For example, the nodes can be computers in a network and the collection of strings can represent the network traffic.
- The goal of the monitoring is the detection of the appearance of anomalous strings in the collection.
- To implement *tolerization*, ARTIS uses a distributed version of the negative selection algorithm.
- Upon activation of a detector a human operator can provide a *costimulation signal* that confirms the dangerous nature of the event.

Life cycle of ARTIS receptors


random generation of receptor
tolerization
match during tolerization
no match during tolerization
surveillance
lifetime exceeded
activation threshold exceeded
activated state
no costimulation during costim. delay
costimulation
memory state
elimination

Example

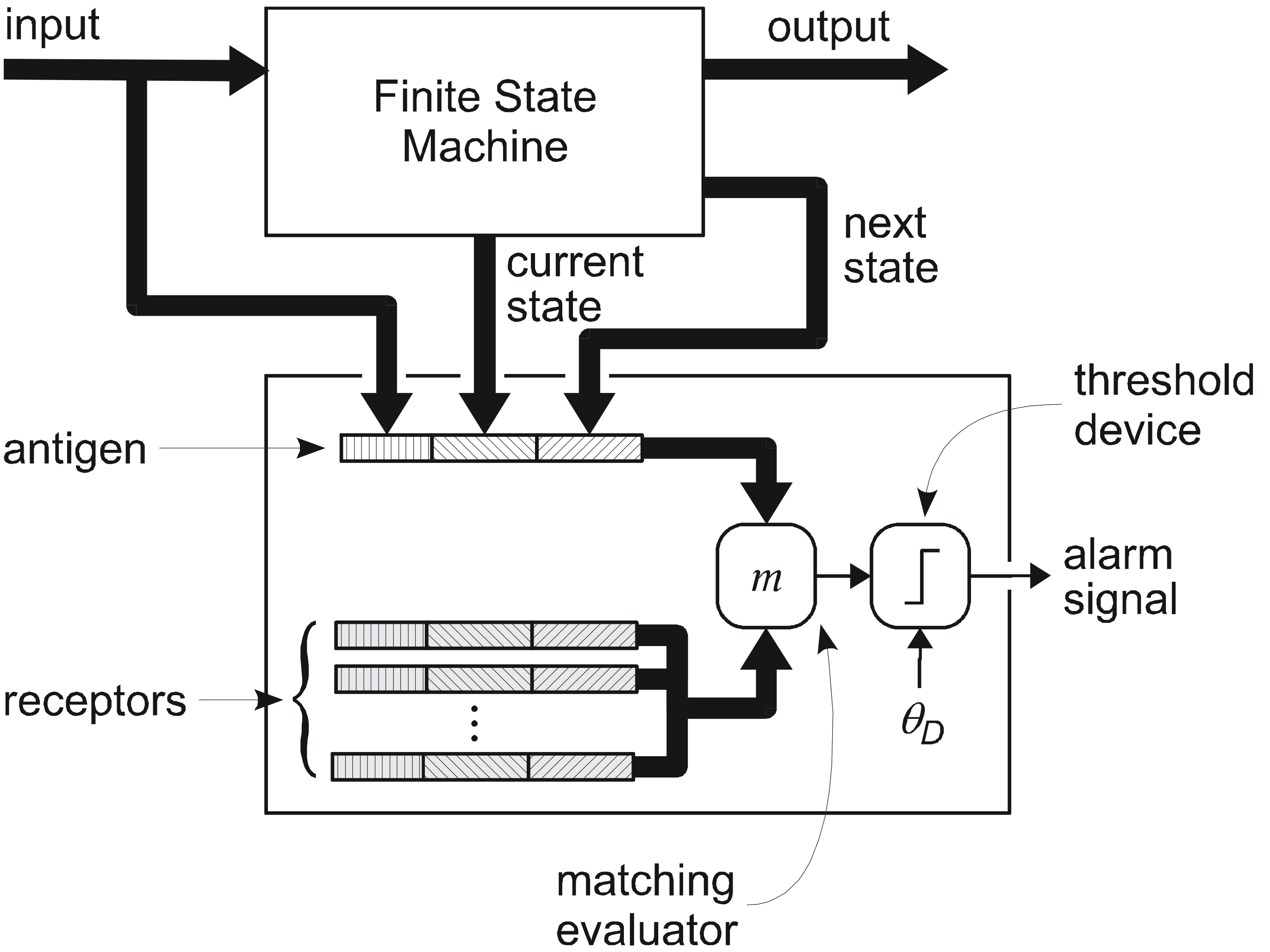
- LISYS (Lightweight Intrusion detection SYStem) is a network intrusion detection system based on ARTIS.
- The strings that are monitored summarize the information about the connections that concern the nodes. Each string contains the identity (IP addresses) of the connected nodes and the specification of the kind of service requested.
- The system was tested with data collected from real computer networks which contained known intrusions and was able to detect all the intrusion attempts, apart from very short ones



Immunotronics

- Immunotronics is an application of AIS concepts to the detection and recovery of faults in digital electronic systems.
- The classical approaches to fault detection and recovery in digital systems are redundancy and the addition of protection systems that check and possibly corrects the validity of the system state
- The immunotronics applies the immune system concept of self/nonself discrimination to automate the generation of the verification criteria used by the protection system.
- The immunotronics approach applies to finite state machines (FSM), a class of systems where the operation is modeled in terms of states and transitions between them. The self can be defined as the collection of strings that represent the legal transitions between the states of the machine
  - The self can be generated by observing the operation of the system in its fault-free condition.

Immunotronics



Closing remarks

- The majority of examples of AIS described in the literature focus on the implementation of one or at most a few of the concepts observed in biological immune systems
  - the full potential of the immune system concept emerges when all the elements work together.
- One of the principal facts that hinder the attempts at the realization of a full-blown AIS is the scarcity of systems that are designed from the beginning to operate in collaboration with an AIS (e.g., by generating danger signals). Typically the current approach is instead to try retrofitting existing systems with immune protection
  - The activation of the protection system following the generation of a danger signal from the part of the protected system implies that some damage has possibly already been done to the system. Current engineering practices, prefer a scenario where the protective action precedes the damage.
  - The current technology does not permit the regeneration of damaged subsystems.