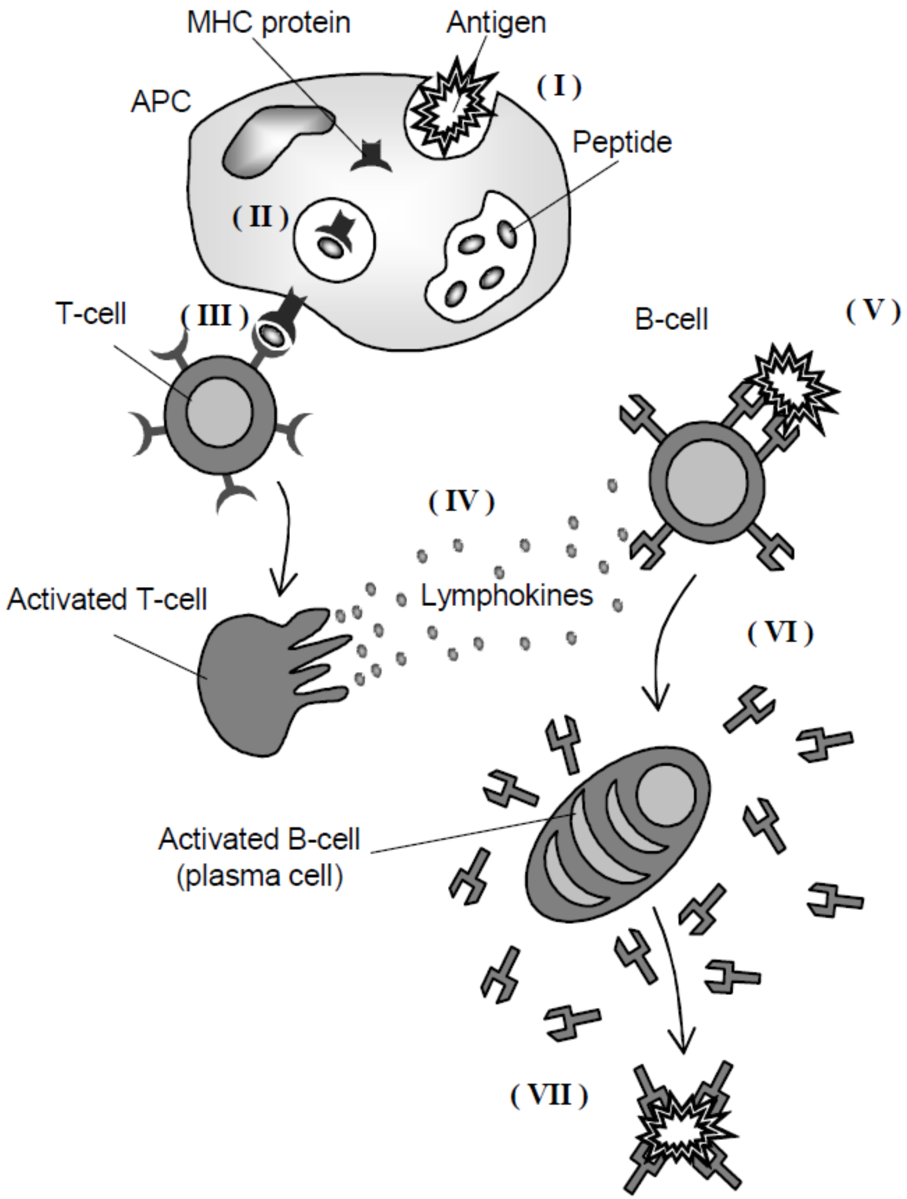
VALIS, an immune-inspired classification algorithm

Peter Karpov
PeterKarpov@inversed.ru

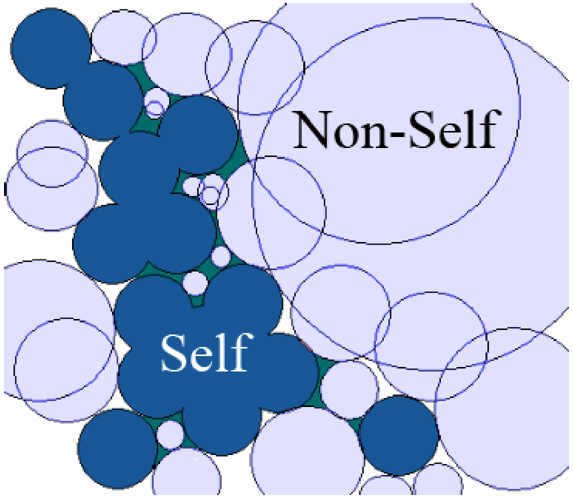
Natural immune system



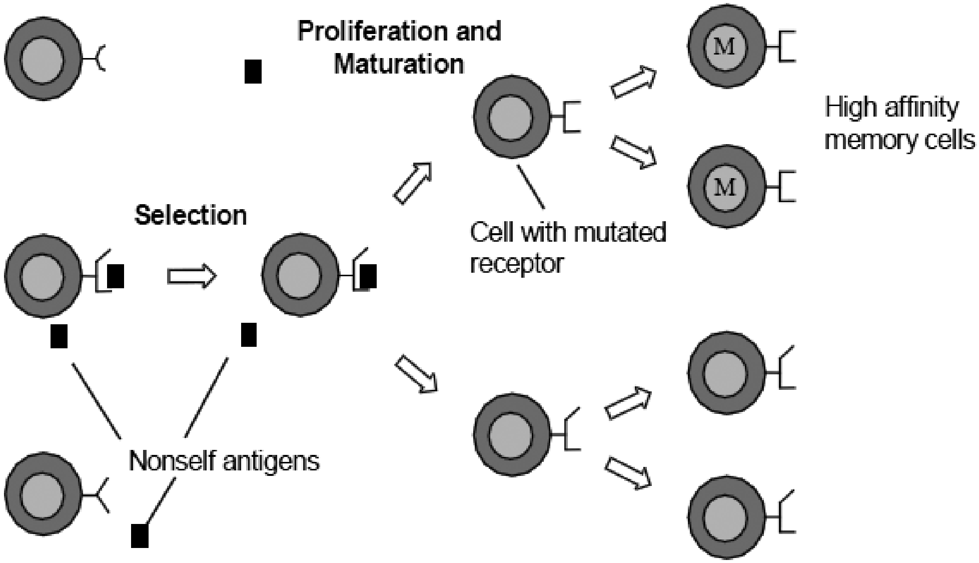
- Antigen presentation
- T-cell activation
- B-cell activation
- Antigen matching
- Antibody production
- Pathogen destruction

Immune-inspired algorithms

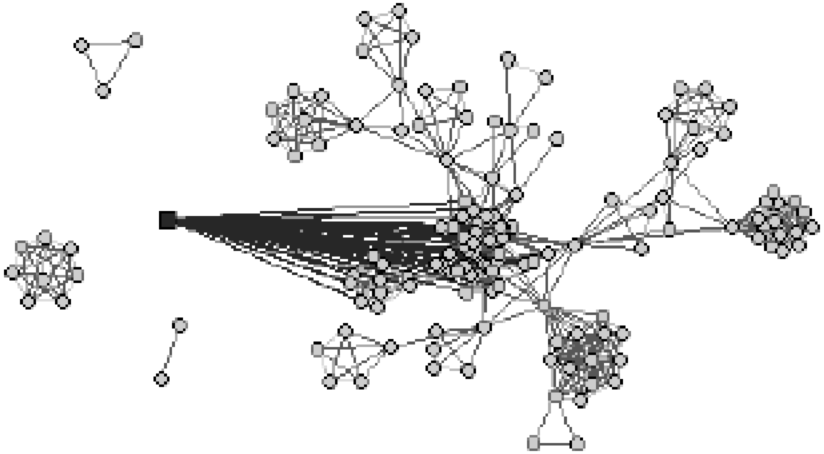
Negative selection



Clonal selection


Proliferation and Maturation
High affinity memory cells
M
M
Selection
Cell with mutated receptor
Nonself antigens

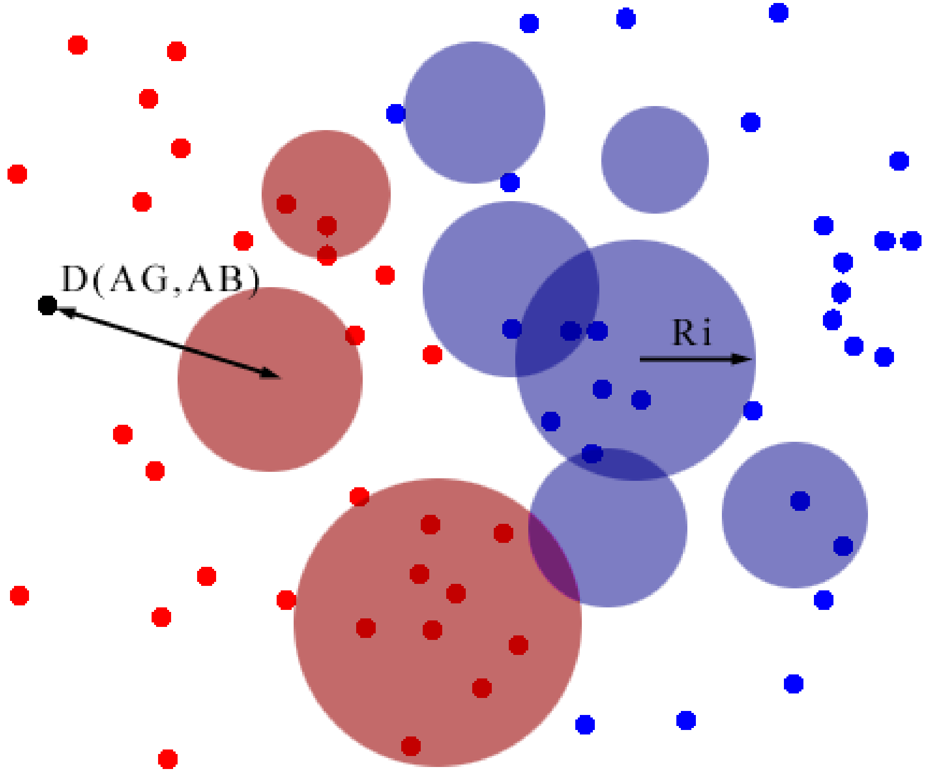
Immune networks



Dendritic cells,
danger theory,
… … …

Immune system — classification analogy

Antigen	←--→	Data sample AG
Antibody	←--→	Data structure AB
Immune response	←--→	Classification result



Binding:

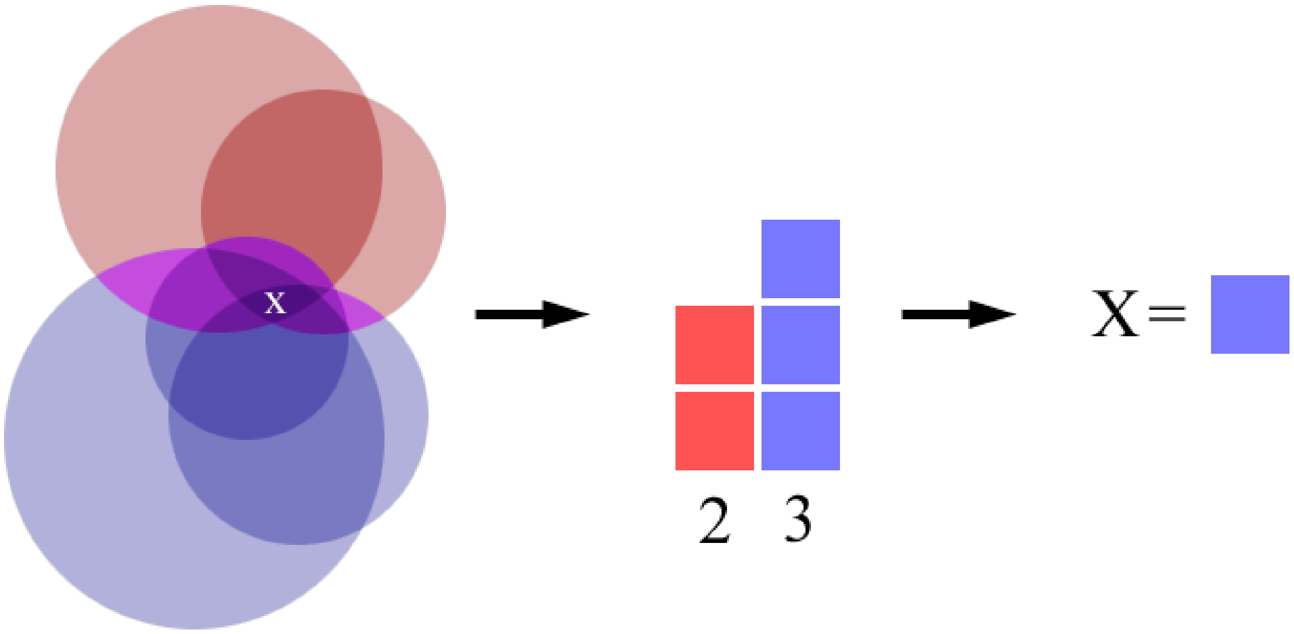
$D(AB_i, AG) < R_i$,

D — distance function,

R_i — binding threshold

Classification by voting

VALIS — Vote Allocating Immune System



Bound antibodies vote for their class. Majority of votes determines the classification result.

Training algorithm

1. Initialization

Create *PopSize* random antibodies

2. Dataset presentation

Present dataset to the system, for each antibody calculate:

- Number of bound and correctly classified antigens
- Probability of correct classification *PCorrect*
- Overlapping coefficient *KShare*
- Fitness $f = PCorrect / KShare$

3. Class assignment

Set antibody classes according to the majority of bound antigens

4. Reproduction

Create $NChildren = LearnRate \times PopSize$ children by crossover and mutation

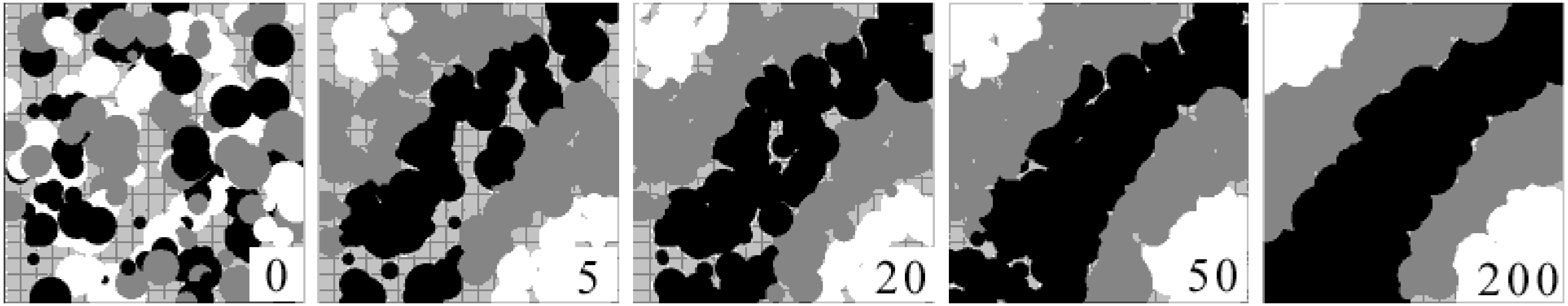
5. Replacement

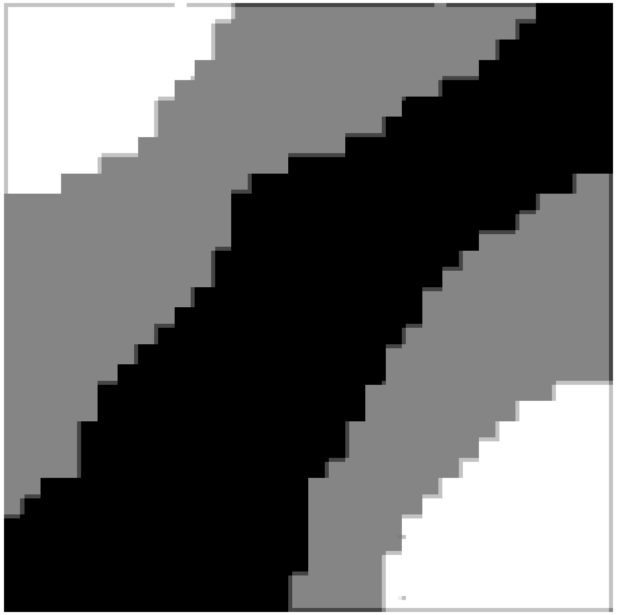
Replace *NChildren* lowest fit antibodies with children

6. End of one generation

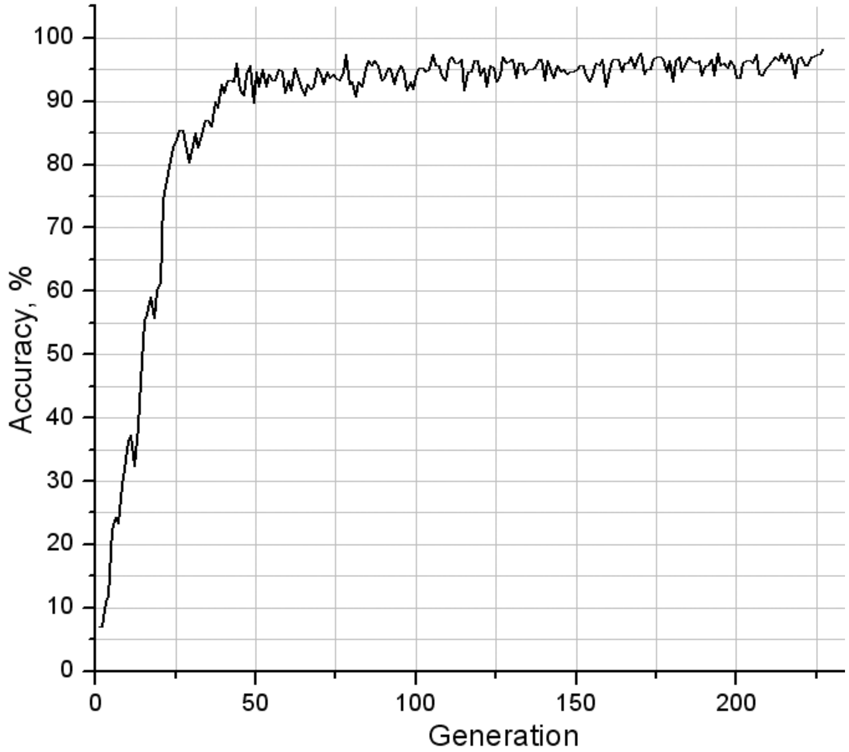
Go to step 2 until a termination criterion is satisfied

Test problem





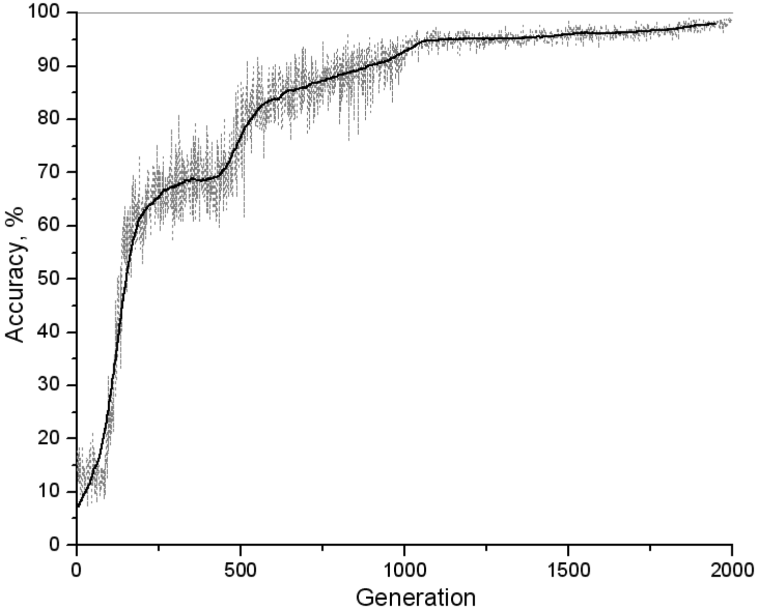
True class distribution



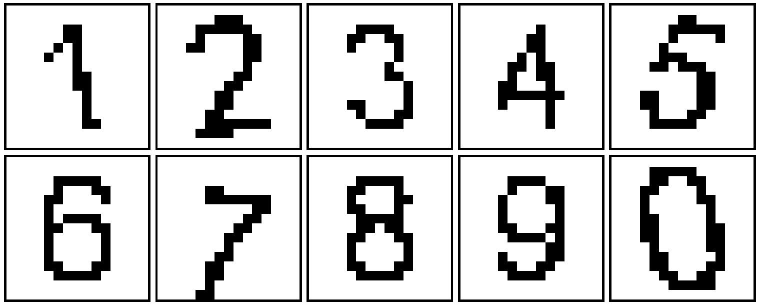
Training dynamics

Character recognition

Antigens	Numerals from 0 to 9, randomly scaled and rotated
Antibodies	Binary arrays
Classification accuracy	99%

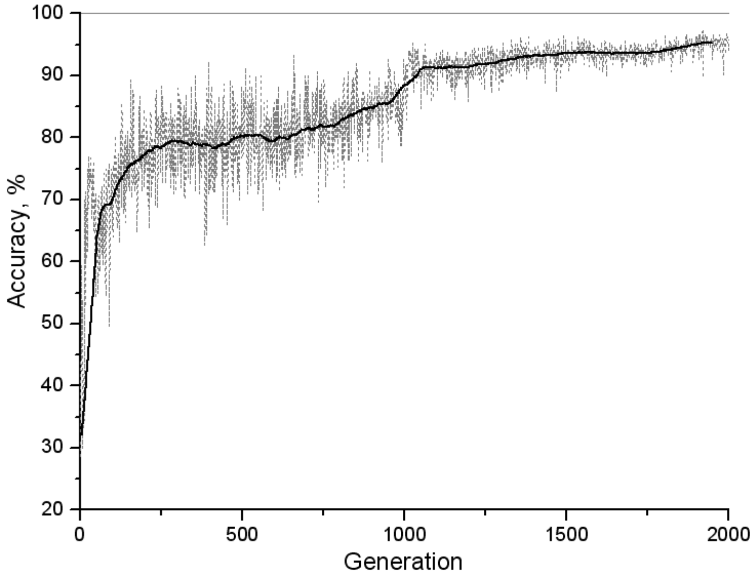


Generalization ability: recognition invariance to scale and rotation



Text classification

Antigens	Pieces of source code in one of 6 programming languages
Antibodies	Variable length strings
Classification accuracy	96%



Typical antibody elements:

Assembler	@, mov, lf
Basic	$
C	fp, {#13, }, <, \n
Fortran	#10c
Lisp	;;,),))#13
Pascal	en, in#13, :=

Conclusion

Features of the proposed system:

- Effective training
- High generalization ability
- Decentralized memory
- Collective classification
- Robust to antibody deletion

Future work:

- Testing on various problems
- Variable learning rate
- Soft binding
- Antibody population postprocessing

Project's web page is available at:
inversed.ru/AIS.htm