

Cognitive techniques for Control of Dynamic Object Behavior in Group

**Koupreenkov Ivan
Prof. Stankevich L.A.**

06.04.2008

1



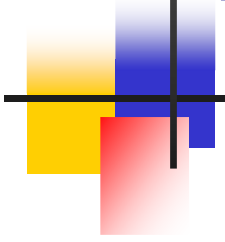
CONTENTS

1. Introduction
2. Cognition-based and multi-agent approaches
3. Architectures of the agent for teamwork
4. On-line Agent Team Work Training Using Immunological Network Model



Introduction

- At present a large interest challenges an use of **cognitive and multi-agent approaches** to increase behavioral capabilities of intelligent systems. Although the cognitive multi-agent systems **require large computational resources**, they can provide of human like behavioral capabilities since the systems able to learn from and adapt to their environment.
- In this work we want to show how the cognitive agents and learnable components for implementing individual and collective behaviors of their agents, can be built.



2. Cognitive and multi-agent approaches

Concepts of the cognitive systems

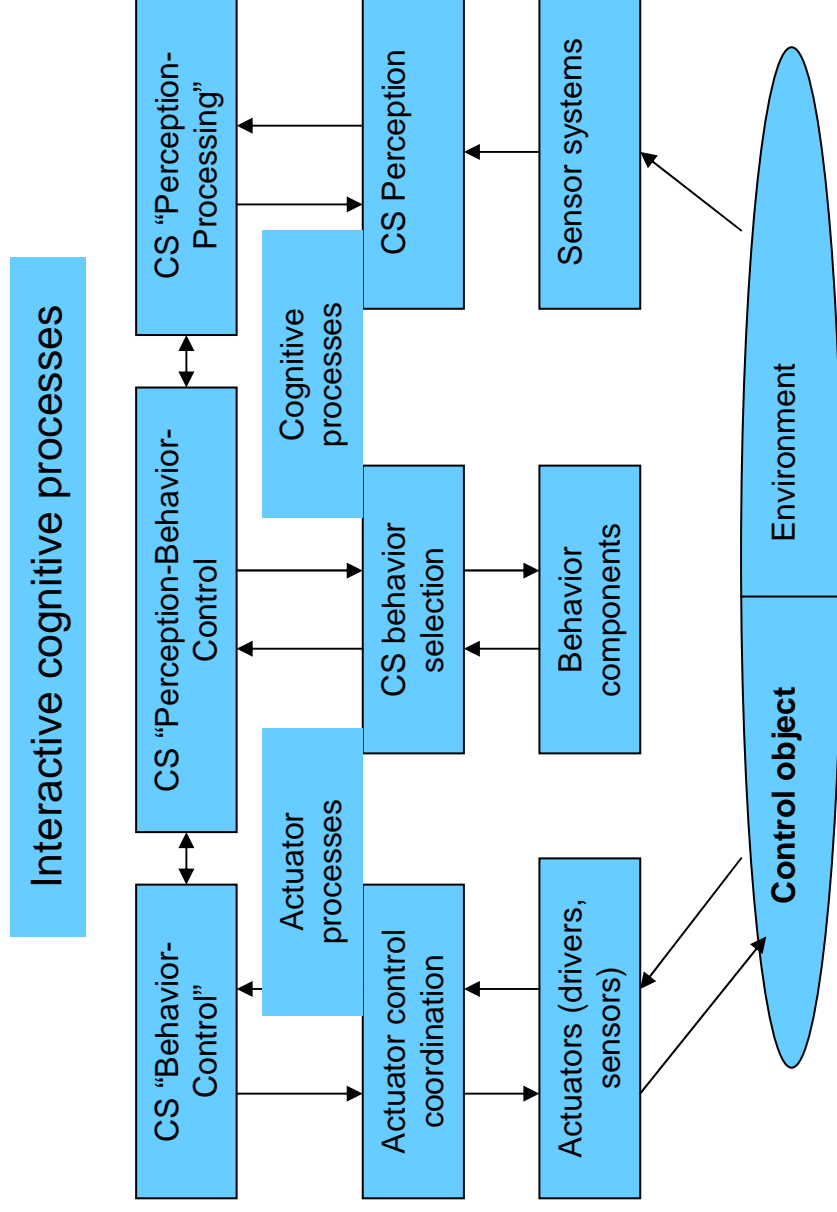
- **Intelligence:** main principles of the cognitive systems are usage of formal knowledge for description of problem domain and appropriate procedures of reasoning
- **Self-organization:** automatic constructing systems for solving tasks of control for multiple objects with complex behavior and operative correction of the knowledge
- **Nervous-system organization:** usage principles of human nervous systems for structural, functional, and behavioral organization of the cognitive systems
- **Cognition:** in the cognitive systems are modeled technically useful cognitive functions and processes coupled with knowledge acquisition and mental processing information.



Cognitive Function & Process

- Cognitive agent behavior is defined by set of **cognitive functions (CF) and processes (CP)**. Each of CF is formed from set of examples of correct behavior and maps of transformations of input to output variables, i.e. $X \rightarrow y$ (vector-scalar) or $X \rightarrow Y$ (vector-vector). It is not mathematical function but it is some approximate mapping transformation of the variables defining correct behavior. CF can be as simple or complex ones. The simple CF map primitive behaviors but complex CF are compositions of simple CF.
- CP consist of temporal sequence of CF and map development of behavior under time in changing conditions.

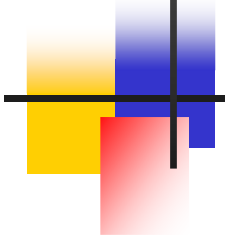
Cognitive structures (CS) and processes



Multi-agent approach

■ **Agent** can be separated into five distinct dimensions: communication, competence, self, planner, and environment. Here the metaphor uses that an individual agent is an information processing brain, where functionality emerges by combining different models that cooperate.

- **Multi-agent system:** an intelligent system is designed as based on collective of agents that can interact from other to another to perform a work with common goals. Such systems are named multi-agent systems.
- **Cognitive multi-agent system** combines both the cognitive and multi-agent approaches and are power mean for human like behavior realization in technical applications.



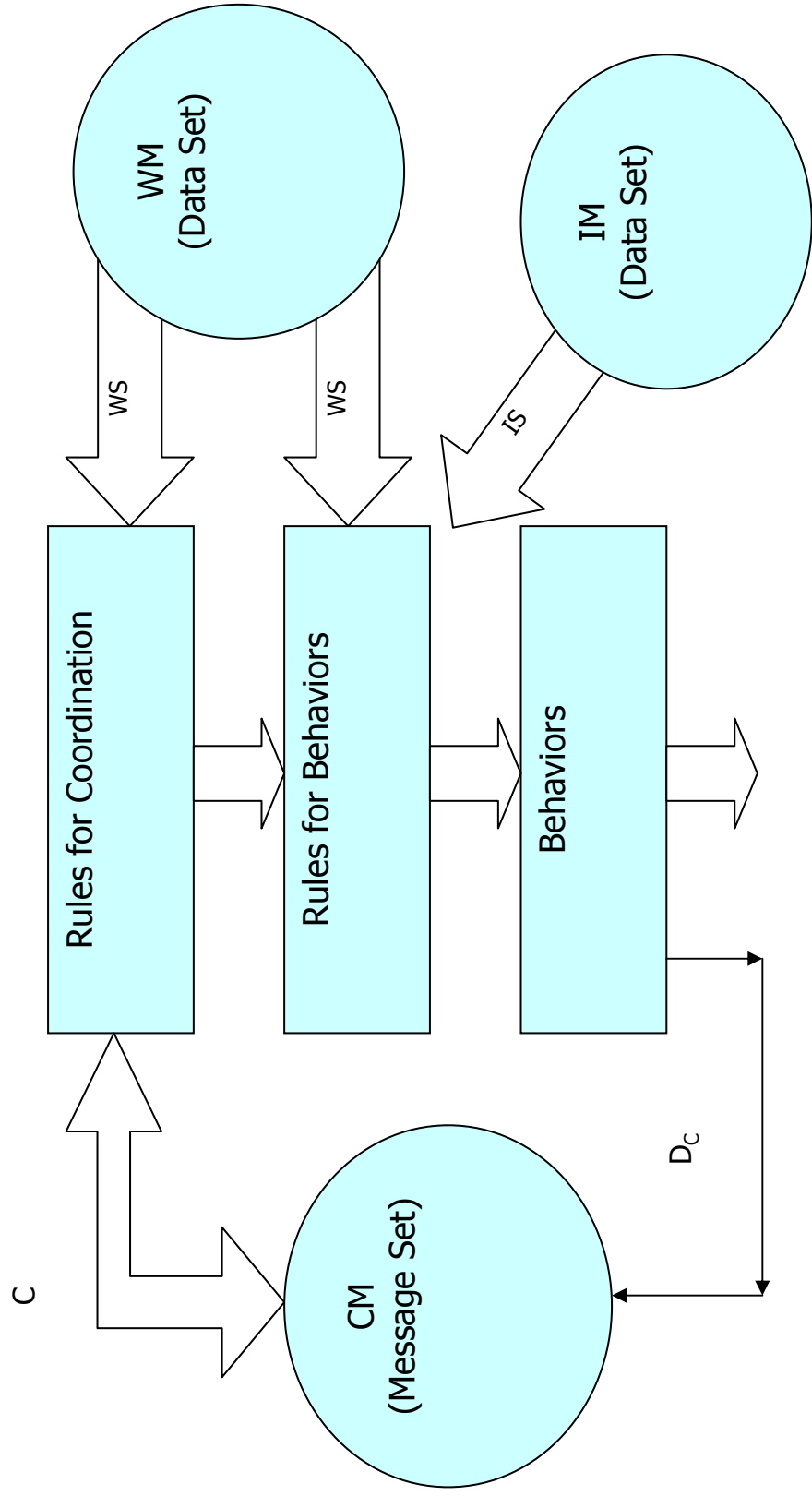
3. Architectures of agents for teamwork



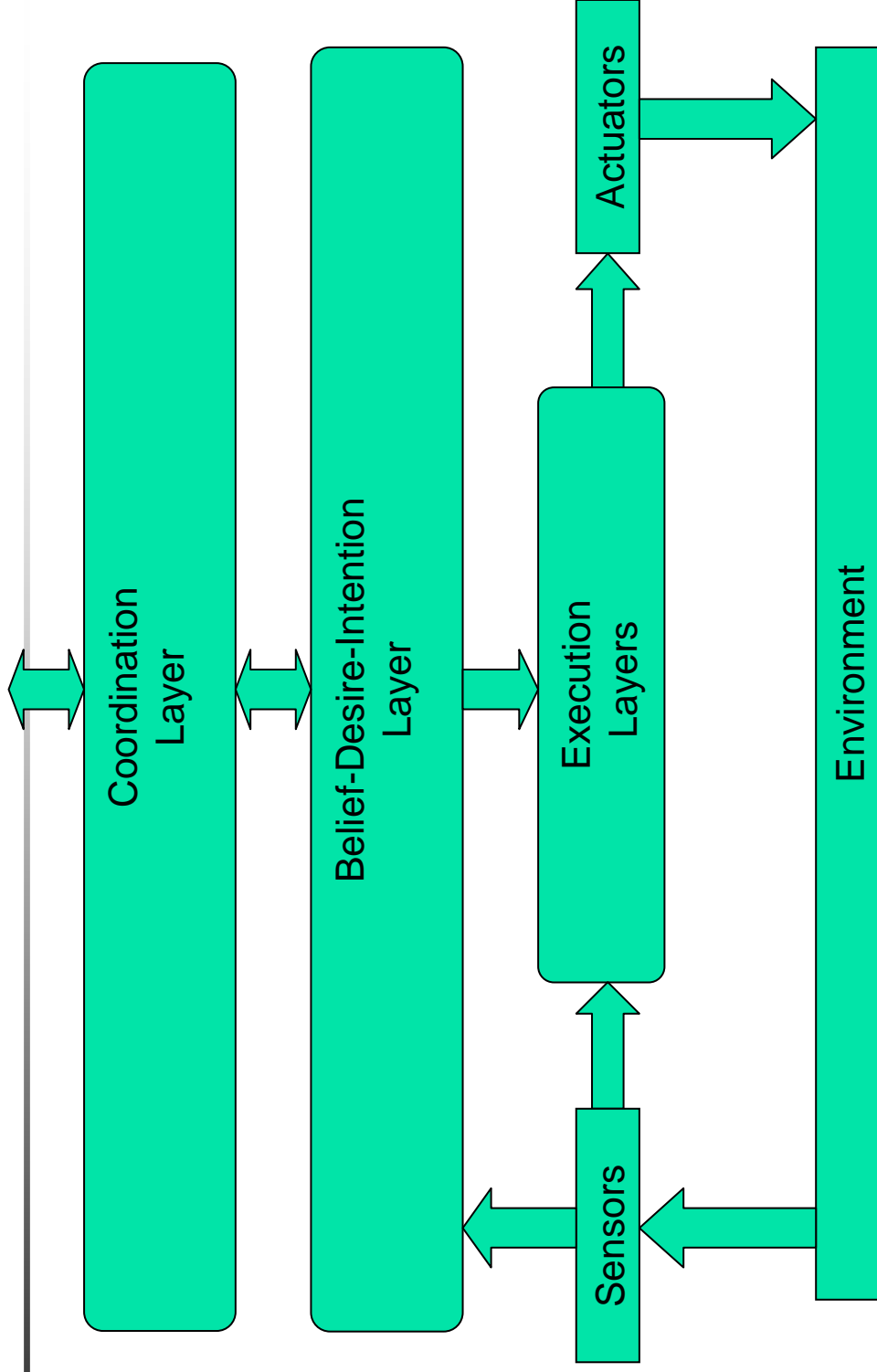
Architectures

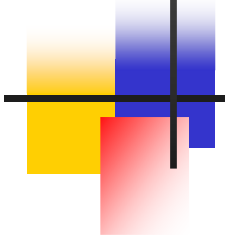
- Reactive
- Deliberative
- Cognitive

Reactive



Deliberative architecture



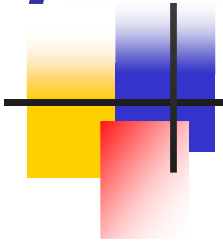


4. On-line Agent Teamwork Training Using Immunological Network Model

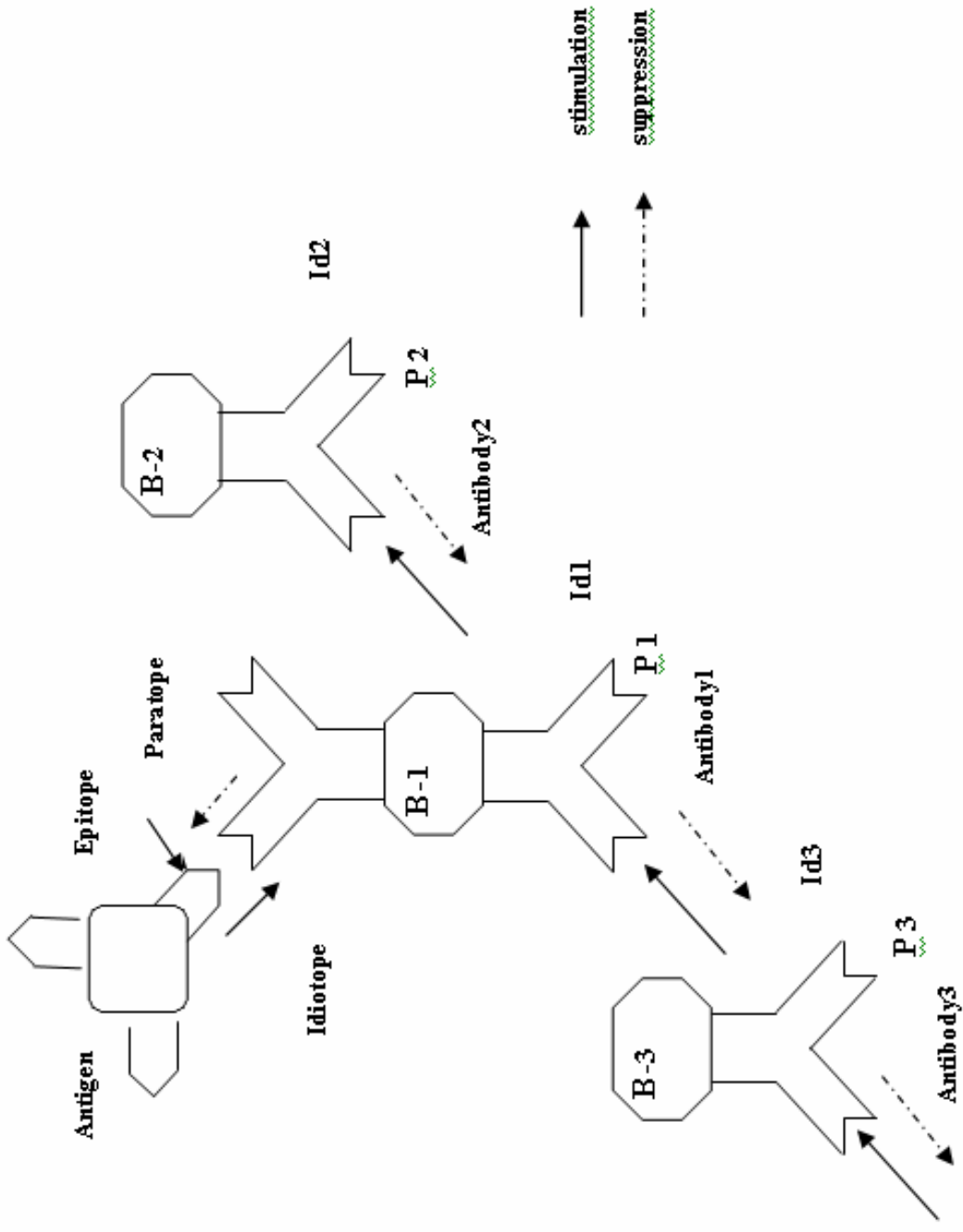
On-line training agents for teamwork



- On-line training agents for teamwork can be based on use of adaptive modules capable to fast learning, for example, implemented on neural networks, fuzzy logic, and their hybrids. However such modules can be used to train for simple such as behaviors.
- There are several models of artificial immune systems models based on idiotopic network hypothesis, clonal-selection theory, and spatial immune network model. These models can be used to implement adaptive modules capable to train complex behavior of objects in the changing environment.
- The adaptive modules based on immune systems along with reinforcement learning are proposed to use for training of the agents with complex individual and scenario behavior. The modules are based on immunological (idiotopic) network model.



Jerne's idiotopic network model





Immunological network dynamics (Watanabe et al., 1999)

Immunological network dynamics concerning with changing concentration of antibodies. The concentration of i -th antibody can be calculated using following equations:

$$\frac{dA_i(t)}{dt} = \left\{ \alpha \sum_{j=1}^N m_{ji} a_j(t) - \alpha \sum_{k=1}^N m_{ik} a_k(t) + \beta m_i - k_i \right\} a_i(t)$$
$$a_i(t+1) = \frac{1}{1 + \exp(0.5 - A_i(t))}$$

Here, N is number of antibodies, m_{ij} and m_i denote affinities of antibodies i and j and between antibody i and detected antigen, respectively. The first and second terms of right hand side of the first equation denote the stimulation and suppression from other antibodies, respectively. The third term represents stimulation from antigen, but fourth term is dissipation factor (for example essential death). The second equation is a squashing function used to ensure the stability of the concentration.

Adjustment mechanism for training

- Adjustment mechanism can be realized by the use special procedure of calculation of degrees of stimuli, which are described in each idiotope. The mechanism starts from the situation where idiotopes of the prepared antibodies are undefined, and then obtains idiotopes using reinforcement learning (RL).
- Reinforcement signals of positive rewards and negative rewards (penalties) are used in order to calculate parameters and m_{ji} of each antibody.
- For explanation of RL procedure, let's assume that antigens 1 and 2 invade in immune network and each antigen simultaneously stimulates antibody 1 and 2. Consequently, the concentration of each antibody increases. However, since the priority between antibodies is unknown (because idiotopes are initially undefined), in this case either of them can be selected randomly. Now, let's assume that the network randomly selects antibody 1 and then receives a positive reinforcement signal as reward. To make the network tend to select antibody 1 under the same or similar antigens (situation), we record the number of antibody 1 in idiotope of antibody 2 and increase a degree of stimulation .



Adjustment mechanism for training

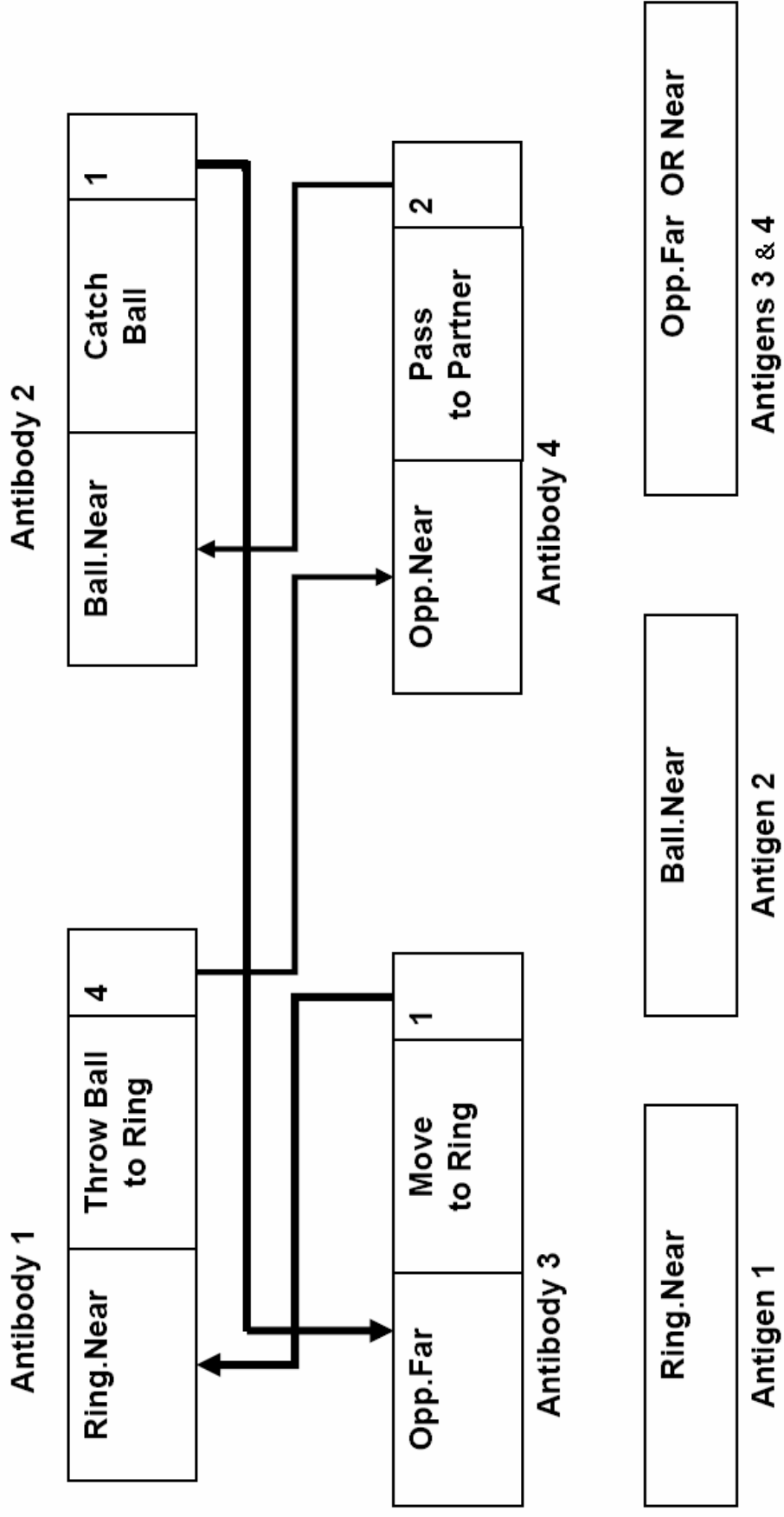
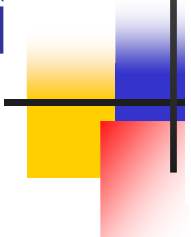
To modify the degrees of stimulation and suppression can be used such equations:

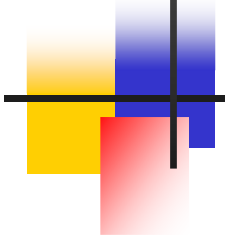
$$m_{12} = \mu_{Af}^{St} = \max(\mu_{R1}^p, \mu_{R2}^r)$$

$$m_{21} = \mu_{Af}^{Sup} = \max(\mu_{R2}^p, \mu_{R1}^r)$$

Here μ_{Af}^{St} and μ_{Af}^{Sup} are degrees of membership for terms of Stimulation (*St*) and Suppression (*Sup*) accordingly for fuzzy variable Affinity (*Af*); μ_{r1}^p , μ_{r2}^p , μ_{r1}^r , and μ_{r2}^r are degrees of membership for fuzzy terms of Penalty (*p*) and Reward (*r*) for reinforcement signals *R1*, *R2* (antibody 1 and antibody 2). Last four degrees of membership can be calculated using given membership functions for the terms *p* and *r* of the variables *R1*, *R2* through number of times of obtaining penalties and rewards when antibody 1 or antibody 2 are selected.

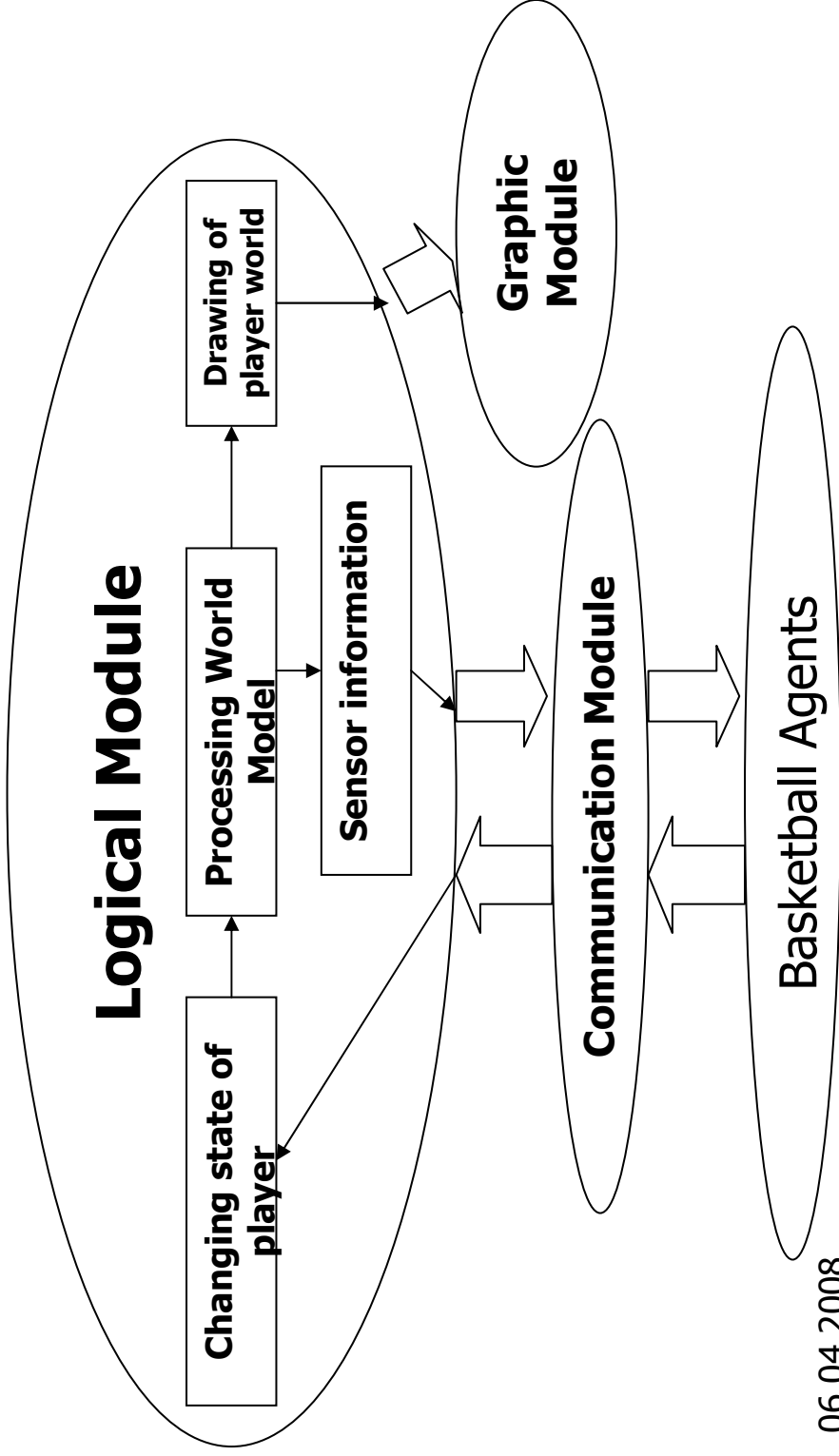
Example of the simple situation





RoboFIBA Server

Structure and Interaction of Modules





Basketball Agents Actions

SHOOT (power *Pow*, direction *DirXY* and *DirZ*). The player shoots the ball with the power *Pow*, in direction of horizontal plane *DirXY* and in direction of vertical plane *DirZ*.

PASS (power *Pow*, direction *DirXY* and *DirZ*). The player passes the ball with power *Pow*, in direction of horizontal planes *DirXY* and in direction of vertical plane *DirZ*. The ball, moving with the power *Pow* in direction *DirXY* and *DirZ*, is switched in state «FREE». **RUN(power *Pow*).** The player runs with power *Pow* in current direction.

TURN_DIRECTION (direction to *Dir XY*). The player changes its body direction

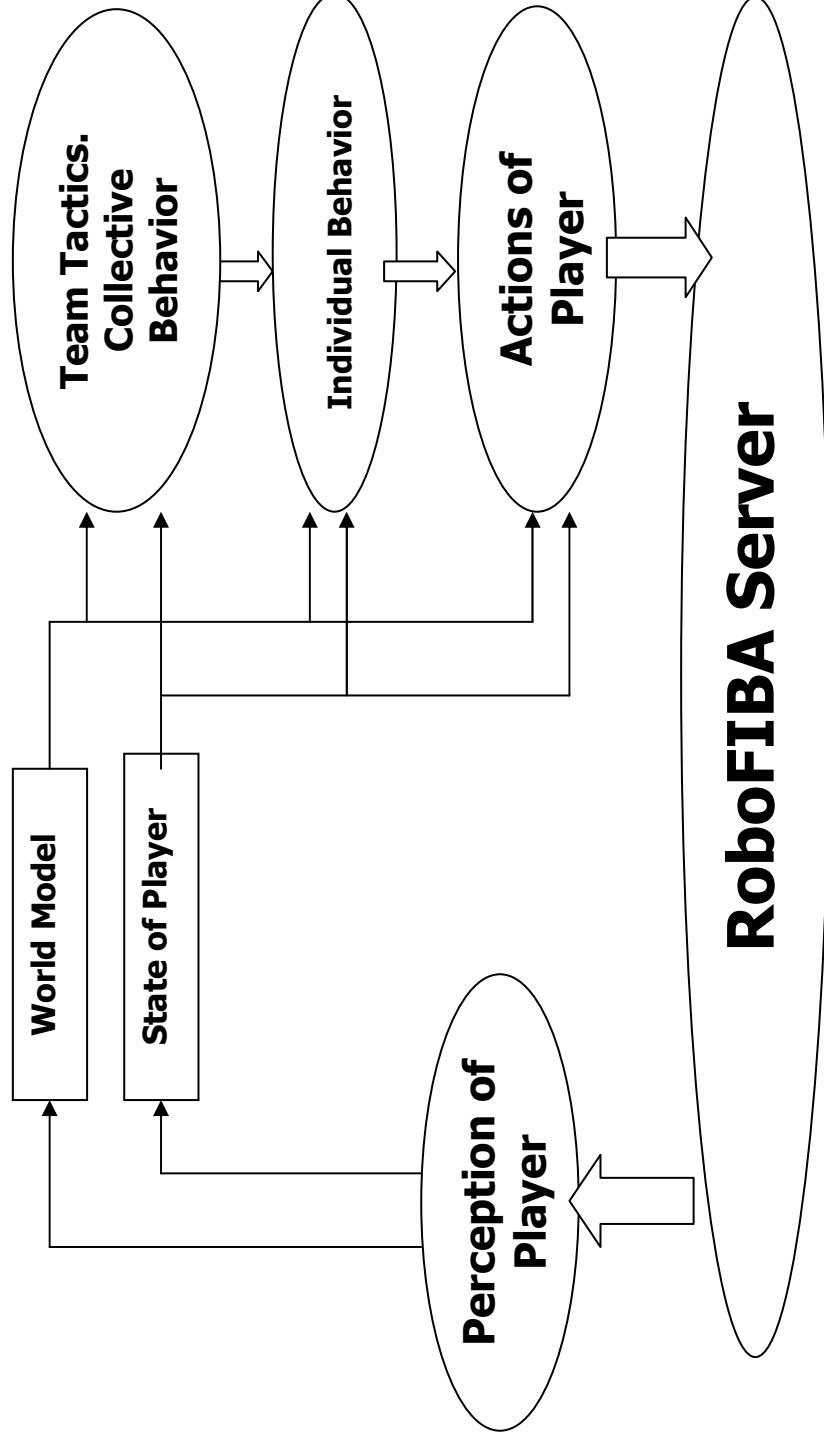
CATCH. The player captures the ball. If distance between the ball and the player is less than *CatchableDist*, the ball belongs to the player. If more than one player is within distance *CatchableDist* to the ball, the ball will go to the nearest player. Catch action is executed only when the ball is free.

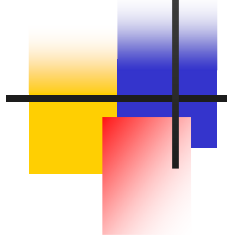
BLOCK SHOOT. The player tries to block throw, pass, or ball free flight. After ball block success, the ball changes flight directional.

JUMP(power *Pow*). Player jumps vertically up with initial power *Pow*.

RoboFIBA Agent

Agent's Structure





Agent Behavior Rules

Example of rule composed by hand

$$((BALL = FREE) \wedge (t_0 \leq t_1) \wedge (t_0 \leq t_1)) \Rightarrow (Gotoball),$$

else $(t_1 \leq t_2) \Rightarrow (Gotoattack),$
else $(Gotodefence);$

t_0 - distance between player and ball,

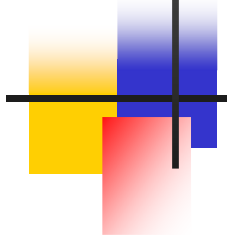
t_1 - distance between ball and partner closest to ball,

t_2 - distance between ball and opponent closest to ball.



Set of Antigens

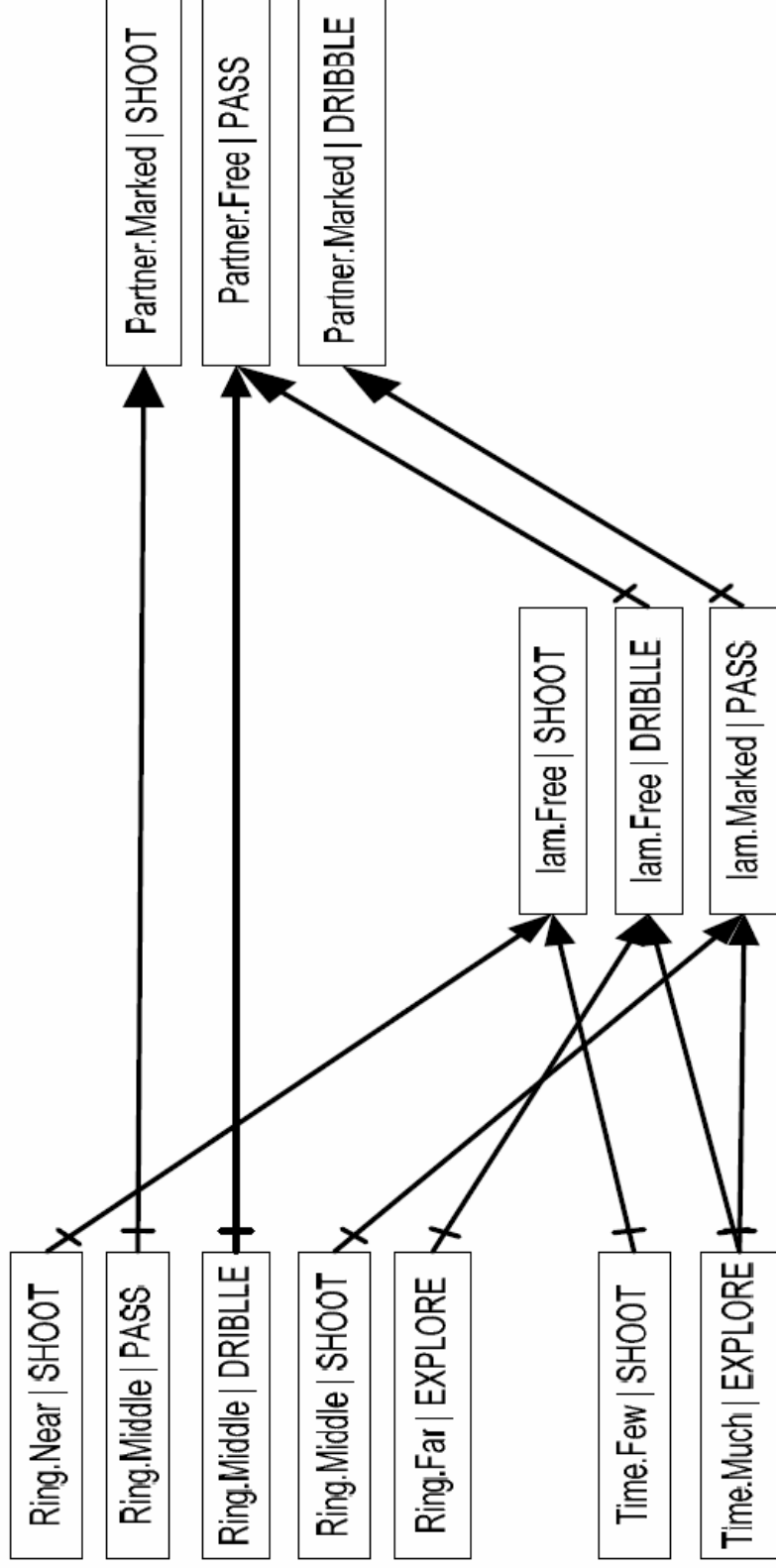
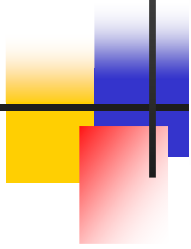
- I have the ball
- Partner has the ball
- Opponent has the ball
- The ball is free
- The ball is near
- Partner is closer to the ring than opponent
- Partner is closer to our ring
- Opponent with the ball is near
- I have the ball
- Opponent is on track to ring
- I'm jumping



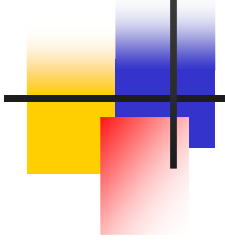
Set of Antibodies

- To throw to the ring of opponent
- To move to the opponent ring in parallel to field
- To stand on the spot
- To intercept the path of the opponent
- Hass to partner ahead
- Pass to the nearest partner
- To jump
- To step left/right (randomly)
- To catch the ball

Example of the complex situation



Conclusion



- Cognitive agent have advantage, that they can change its previous behavior.
- Special modules capable to be trained using on-line learning procedure based on immunological network model along with reinforcement learning are more appropriate for agent teamwork in dynamically changing environments than rule based modules.