Game-theoretic model on cognitive maps and its tolerance to errors in input data

Kulivets Sergei,
Trapeznikov Institute of Control Sciences RAS.

Linear cognitive map

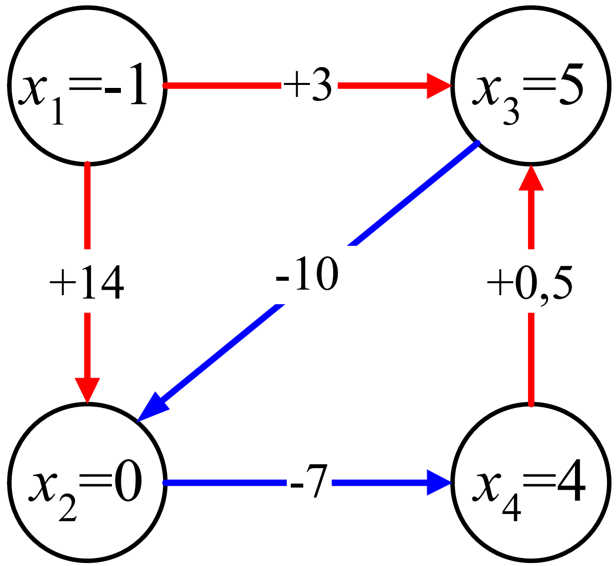
- ***Linear cognitive map*** is a cognitive map with the rule of changes of concept values according the rule (R. Axelrod, F. Roberts, 1976):

$$x_i(t+1) = x_i(t) + p_i(t)$$

$$p_i(t) = \begin{cases} p_i(0), & t = 0 \\ \sum_{j \in M} w_{ji} \cdot p_j(t-1), & t = 1,2,3,\ldots \end{cases}$$

where

- W – adjacency matrix of the digraph
- M – set of all concepts
- ***Controlled concept*** is the concept which can be controlled by decision maker at the moment time 0 (with manipulating values by external pulse $p(0)$)
- ***Target concept*** is the concept which has "desired value" for the moment T for decision maker



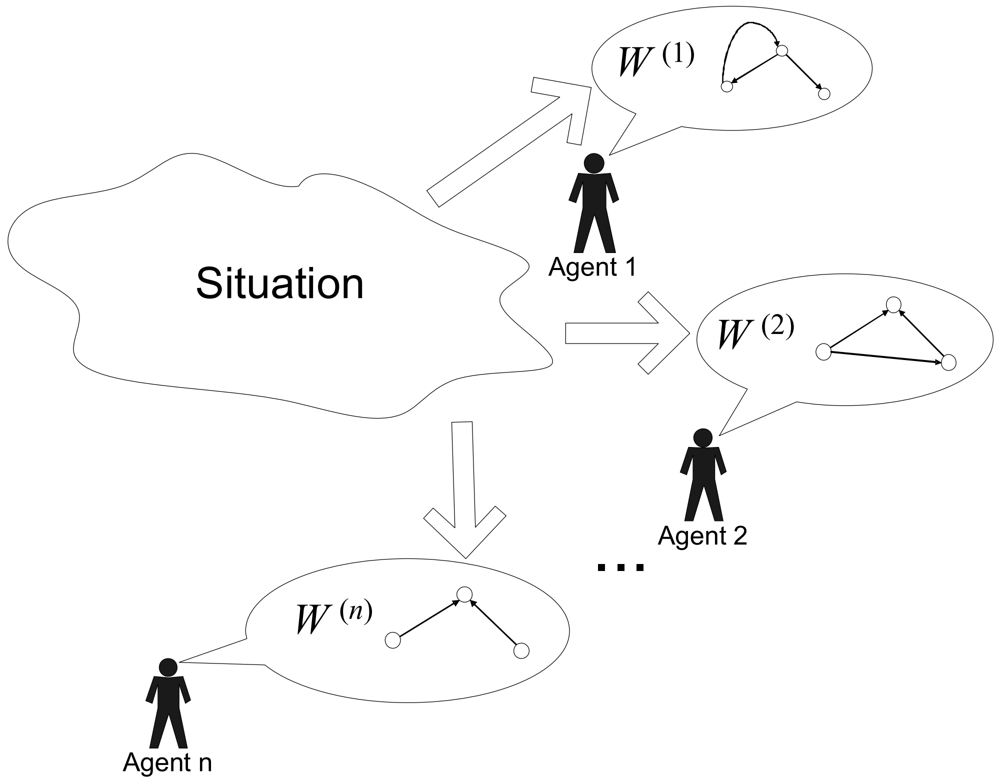
$x(0)=(x_1(0), x_2(0), \ldots, x_m(0))$ – the initial concept vector

$p(0)=(p_1(0), p_2(0), \ldots, p_m(0))$ – the external pulse to each node at the zero time point.

$p_i(0) \in [p_i^{min}, p_i^{max}]$

A game of agents {1, 2, …, *n*} = *N* with different beliefs

$$\Gamma_C = \left\{N, \{S_i\}_{i\in N}, \{f_i\}_{i\in N}, \left\{W^{(i)}\right\}_{i\in N}\right\}$$



- Set of strategies:

$$S_i = \mathop{\times}_{j\in M_i} [p_j^{\min}, p_j^{\max}]$$

- f_i – utility function of agent *i*, $i \in N$.
- $W^{(i)}$ – linear cognitive map of agent *i*.

Utility function f_i

The control target *of* *i*-th agent is the maximization the function f_i. If the *i*-th agent want to increase (alternatively, decrease) in the value of the concept x_j, then it is desirable for him to maximize the expression

$$(x_j(T) - x_j(0)) \qquad (\text{similarly } -(x_j(T) - x_j(0))).$$

If the agent *i* can define desirable values for several concepts, then the weighted sum should be maximized according to the above stated expressions for such concepts. Each coefficient is interpreted as an “importance percentage” of restrictions on the corresponding concept.

$$f_i\big(x_1(T), x_2(T), ..., x_m(T)\big) = \sum_{j \in M} \gamma_{ij} \cdot (x_j(T) - x_j(0))$$

$|\gamma_{ij}|$ -- is the “importance percentage” of the *j*-th concept value for the *i*-th agent

Game solution

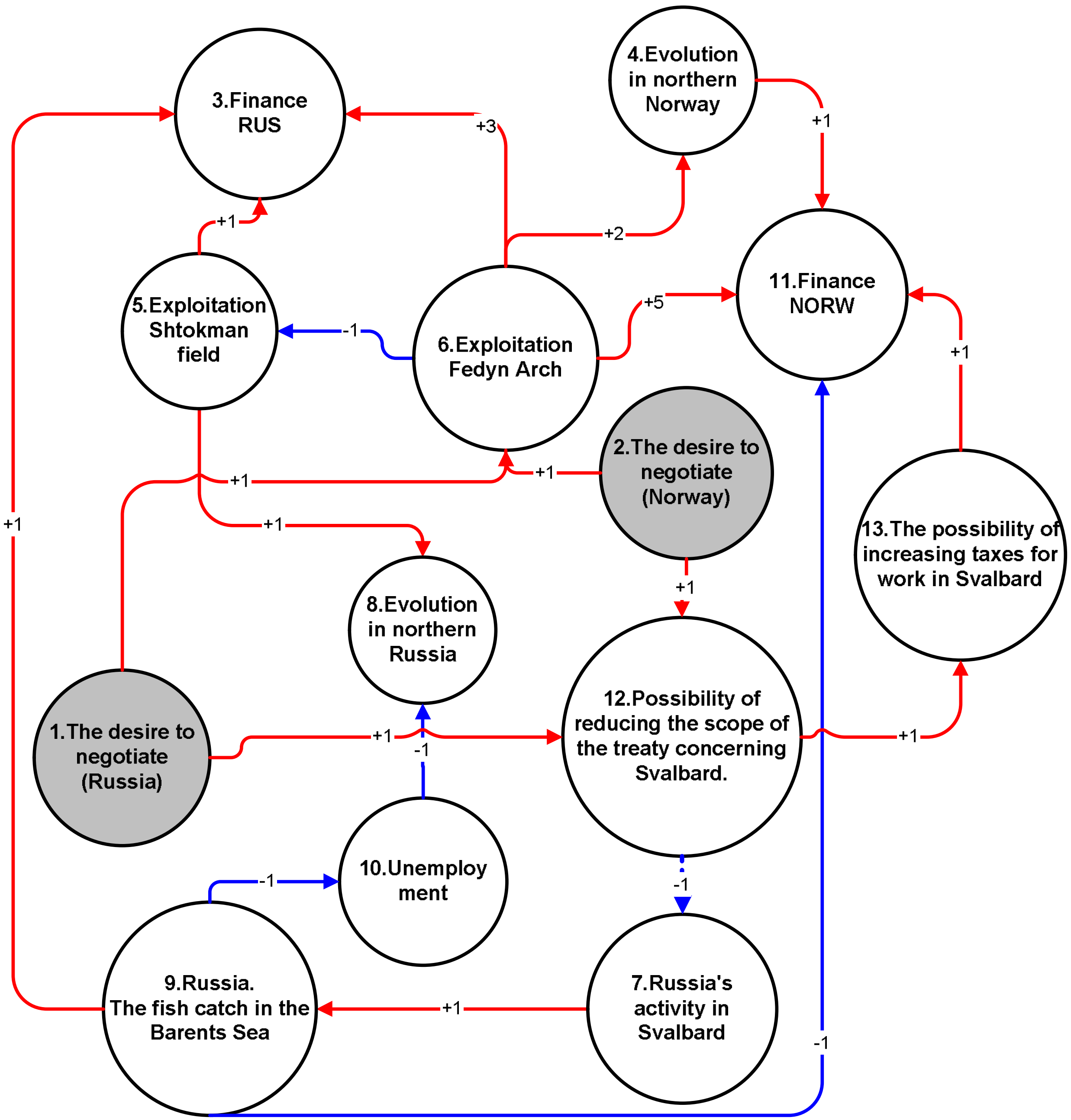
✓ There are an equilibrium in dominant strategies in *strategic-form game with different agent beliefs* with linear utility function.

$$x(T) = x(0)+p(0)\cdot(E+W+\ldots+W^{T-1}) = x(0)+p(0)\cdot{}_TQ,$$

$$f_i = \sum_{j\in M} \gamma_{ij}(x_j(T) - x_j(0)) = \sum_{k\in M}\left(\sum_{j\in M}\gamma_{ij}\,{}_Tq_{kj}\right)p_k = \sum_{k\in M} {}_T\alpha_{ik}\,p_k$$

$$p_k = sign({}_T\alpha_{ik}), k \in M_i$$

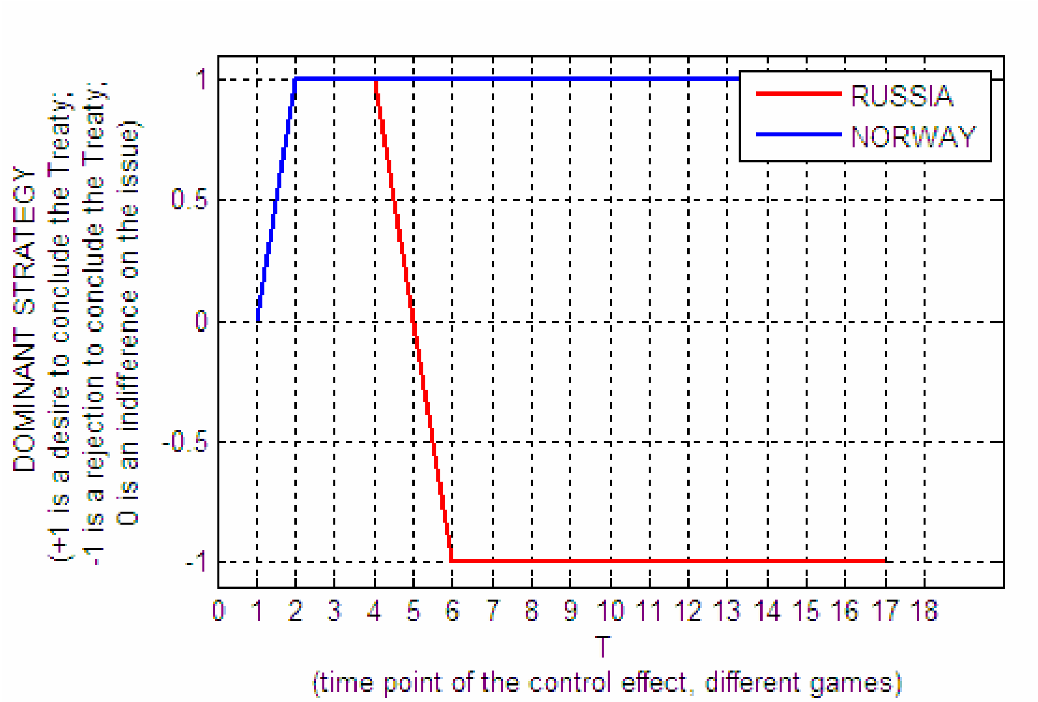
Example



Cognitive map that reflects the causal links between concepts in the problem of the disputed territory in the Barents Sea.

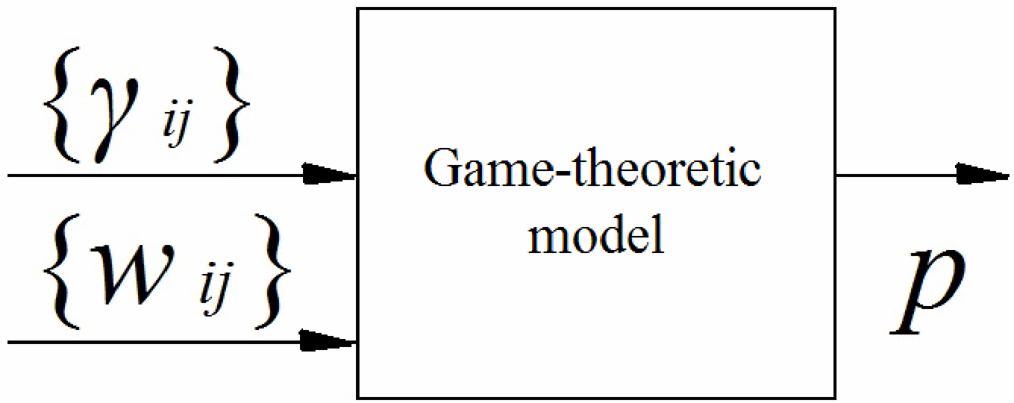
The *control concept* for **Russia** is **#1** to **Norway** is **#2**
The *target concepts* for **Russia** is **#3 and #8** to **Norway** is **#4 and #11**

Example



- the axis “X” is the different values of the target time T,
- the vertical axis “Y” corres-ponding to the equilibrium strategies of agents:
  - •Russia (red) and
  - •Norway (blue)

Model tolerance to errors in input data



$$\hat{p} = p\left(\{\gamma_{ij}\}, \{w_{ij}\}\right)$$

$$\hat{p} = p\left(\{\gamma_{ij} \pm \varepsilon\}, \{w_{ij}\}\right) \qquad \hat{p} = p\left(\{\gamma_{ij}\}, \{w_{ij} \pm \varepsilon\}\right)$$

Estimate of tolerance to errors in a target coefficient γ_{ij}

Let consider the situation where the expert make an error in one of the weights coefficients γ_{is} in the utility function (UF). If expert did not make an error, the value of coefficient would be

$$ {}_T\alpha_{ik} = \gamma_{i1}\, {}_Tq_{k1} + \gamma_{i2}\, {}_Tq_{k2} + \ldots + \gamma_{is}\, {}_Tq_{ks} + \ldots + \gamma_{i,m}\, {}_Tq_{k,m} $$

Because of the error the value of changed to

$$ {}_T\alpha_{ik}^{\ \varepsilon} = \gamma_{i1}\, {}_Tq_{k1} + \gamma_{i2}\, {}_Tq_{k2} + \ldots + (\gamma_{is} \pm \varepsilon_{ks}^{i})\, {}_Tq_{ks} + \ldots + \gamma_{i,m}\, {}_Tq_{k,m} $$

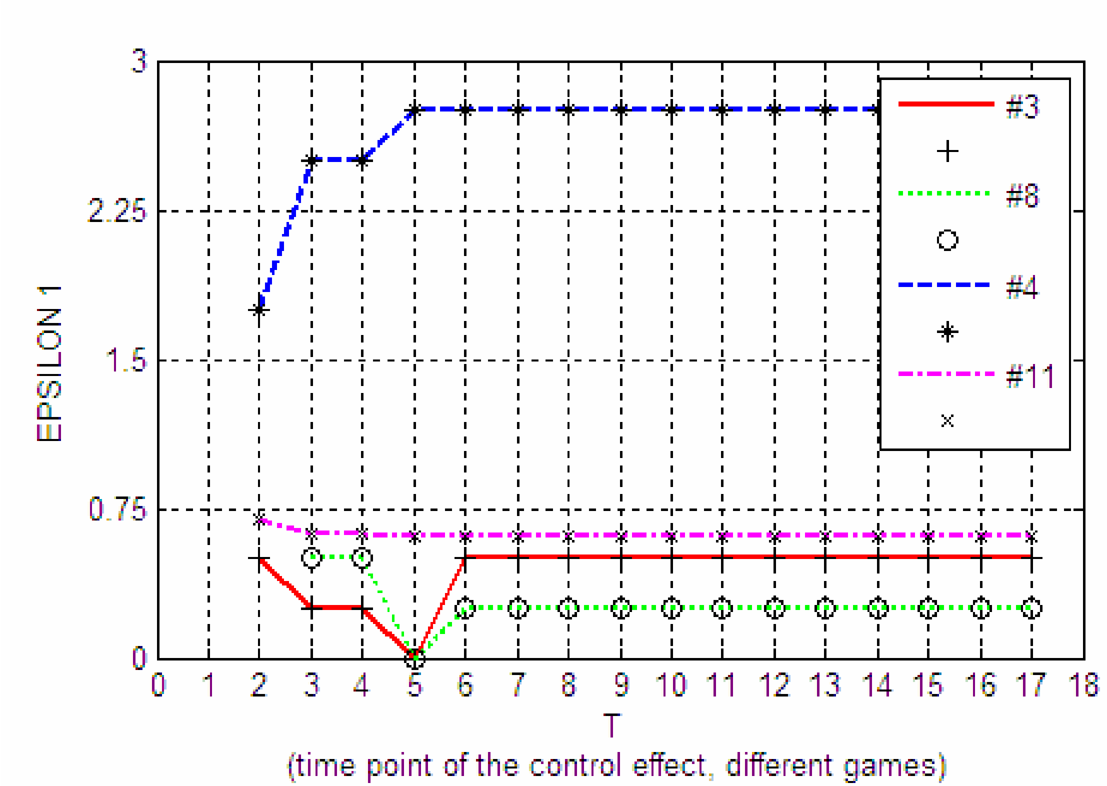
From this condition we obtain an estimate of error in one coefficient in UF

$$ {}_T\alpha_{ik}\, {}_T\alpha_{ik}^{\ \varepsilon} = {}_T\alpha_{ik}\left({}_T\alpha_{ik} \pm \varepsilon_{ks}^{i} \cdot {}_Tq_{ks}\right) > 0 \Rightarrow \varepsilon_{ks}^{i} < \left|\frac{{}_T\alpha_{ik}}{{}_Tq_{ks}}\right| $$

If deviation in <u>one</u> coefficient γ_{ij} of utility function f_i does not exceed the value of (1), the optimal strategy of agent *i* remains unchanged

$$ \varepsilon_{s}^{i} = \min_{k \in M_i} \left|\frac{{}_T\alpha_{ik}}{{}_Tq_{ks}}\right| \qquad (1) $$

Estimate of tolerance to errors in one target coefficient



10

Estimate (lower bound) of tolerance to error in all target coefficients

Let consider the situation where the expert make an error in all of the weights coefficients in the utility function (UF). If expert did not make an error, the value of coefficient would be

$$_T\alpha_{ik} = \gamma_{i1}\, {}_Tq_{k1} + \gamma_{i2}\, {}_Tq_{k2} + \ldots + \gamma_{is}\, {}_Tq_{ks} + \ldots + \gamma_{i,m}\, {}_Tq_{k,m}$$

Because of the error the value of changed to

$$_T\alpha_{ik}^{\varepsilon} = (\gamma_{i1} \pm \varepsilon_k^i)\, {}_Tq_{k1} + (\gamma_{i2} \pm \varepsilon_k^i)\, {}_Tq_{k2} + \ldots + (\gamma_{is} \pm \varepsilon_k^i)\, {}_Tq_{ks} + \ldots + (\gamma_{i,m} \pm \varepsilon_k^i)\, {}_Tq_{k,m}$$

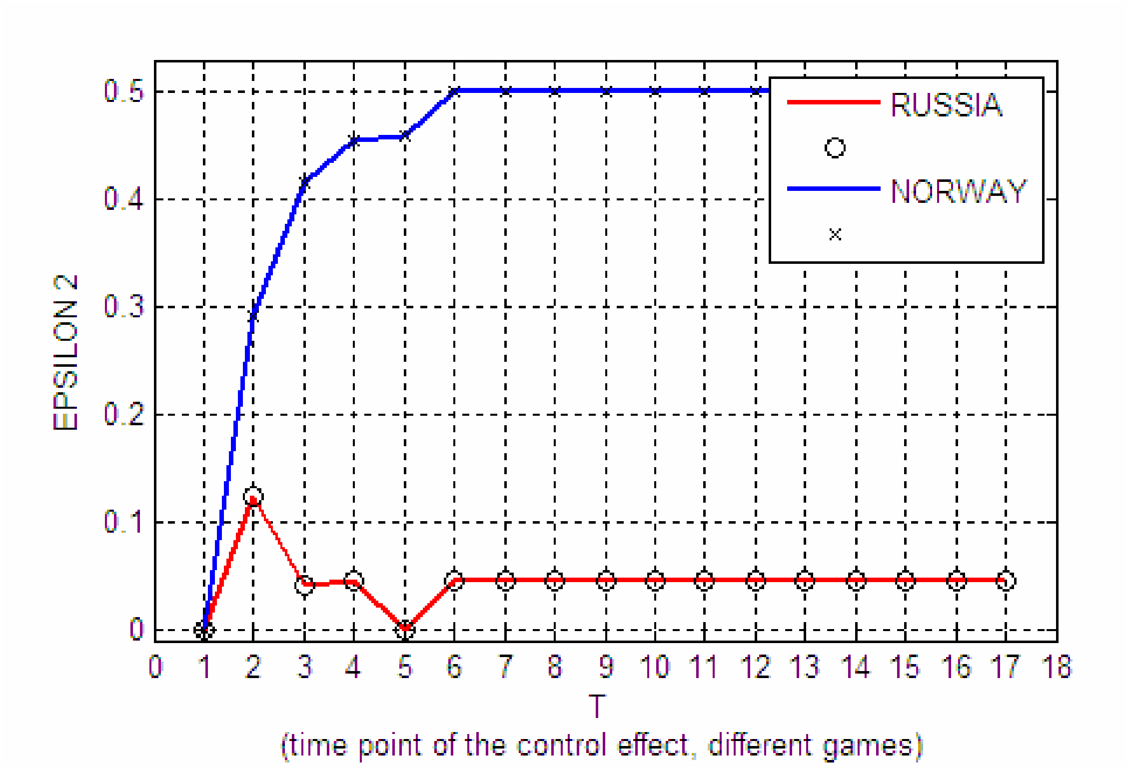
From this condition we obtain an estimate of error (lower bound) in all coefficients in UF

$$_T\alpha_{ik}\, {}_T\alpha_{ik}^{\varepsilon} = {}_T\alpha_{ik}\left({}_T\alpha_{ik} \pm \varepsilon_k^i \sum_{s=1}^{M} {}_Tq_{ks}\right) > 0 \Rightarrow \varepsilon_k^i < \left| \frac{{}_T\alpha_{ik}}{\sum_{s=1}^{M} {}_Tq_{ks}} \right|$$

If deviation in <u>all</u> coefficients of utility function f_i does not exceed the value of (2), the optimal strategy of agent *i* remains unchanged

$$\varepsilon^i = \min_{k \in M_i} \left| \frac{{}_T\alpha_{ik}}{\sum_{s=1}^{m} {}_Tq_{ks}} \right| \qquad (2)$$

Estimate of tolerance to errors in all target coefficients



Estimate (lower bound) of tolerance to error in a weight of arc in digraph of cognitive map

Let consider the situation where the expert make an error on δ in a weight of arc w_{rs} in cognitive map. In this case, the adjacency matrix of a digraph to be the next:

$$W_\delta = \begin{pmatrix} w_{11} & \dots & w_{1s} & \dots & w_{1m} \\ \dots & \dots & \dots & \dots & \dots \\ w_{r1} & \dots & w_{rs} \pm \delta & \dots & w_{rm} \\ \dots & \dots & \dots & \dots & \dots \\ w_{m1} & \dots & w_{ms} & \dots & w_{mm} \end{pmatrix}$$

In this situation the error δ is reason for errors in elements of *matrix of an influence reachability by the time T* ${}_TQ$

$$_TQ_\varepsilon = (E + W_\delta + W_\delta{}^2 + \dots + W_\delta{}^{T-1})$$

The elements of the matrix (${}_TQ_\varepsilon$ - ${}_TQ$) correspond to a change caused by an error δ in adjacency matrix *W*. This change can be represented as

$$\left|{}_Tq_{ij}^\varepsilon - {}_Tq_{ij}\right| = \left|\delta \cdot P_1\left(\left\{w_{ks}\right\}\right) + \delta^2 \cdot P_2\left(\left\{w_{ks}\right\}\right) + \dots + \delta^{T-1} \cdot P_{T-1}\left(\left\{w_{ks}\right\}\right)\right| < \varepsilon$$

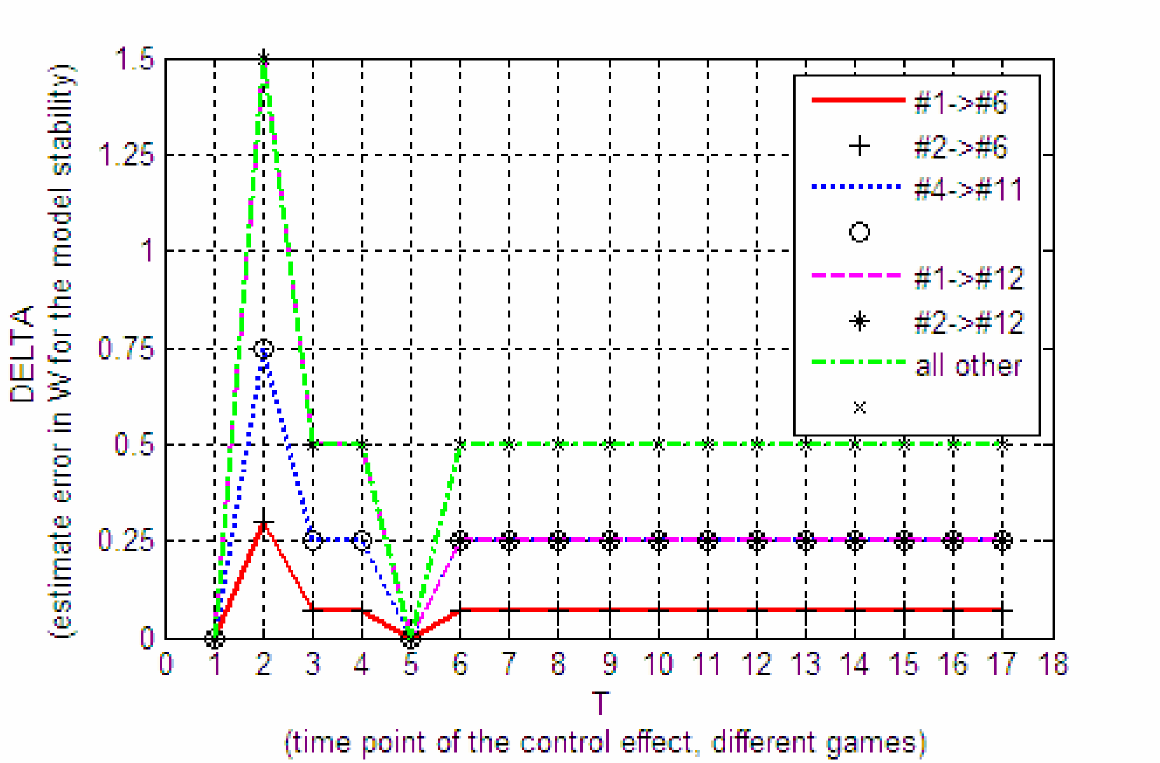
where $P_k(\{w_{ks}\})$ the algebraic sum of products of elements of matrix *W*. To estimate the tolerable error δ in the matrix element w_{rs} it is necessary to estimate the error ε in the value of elements of matrix ${}_TQ$.

Similar to the previous arguments we obtain an estimate of error ε_k^i for the elements of ${}_TQ$

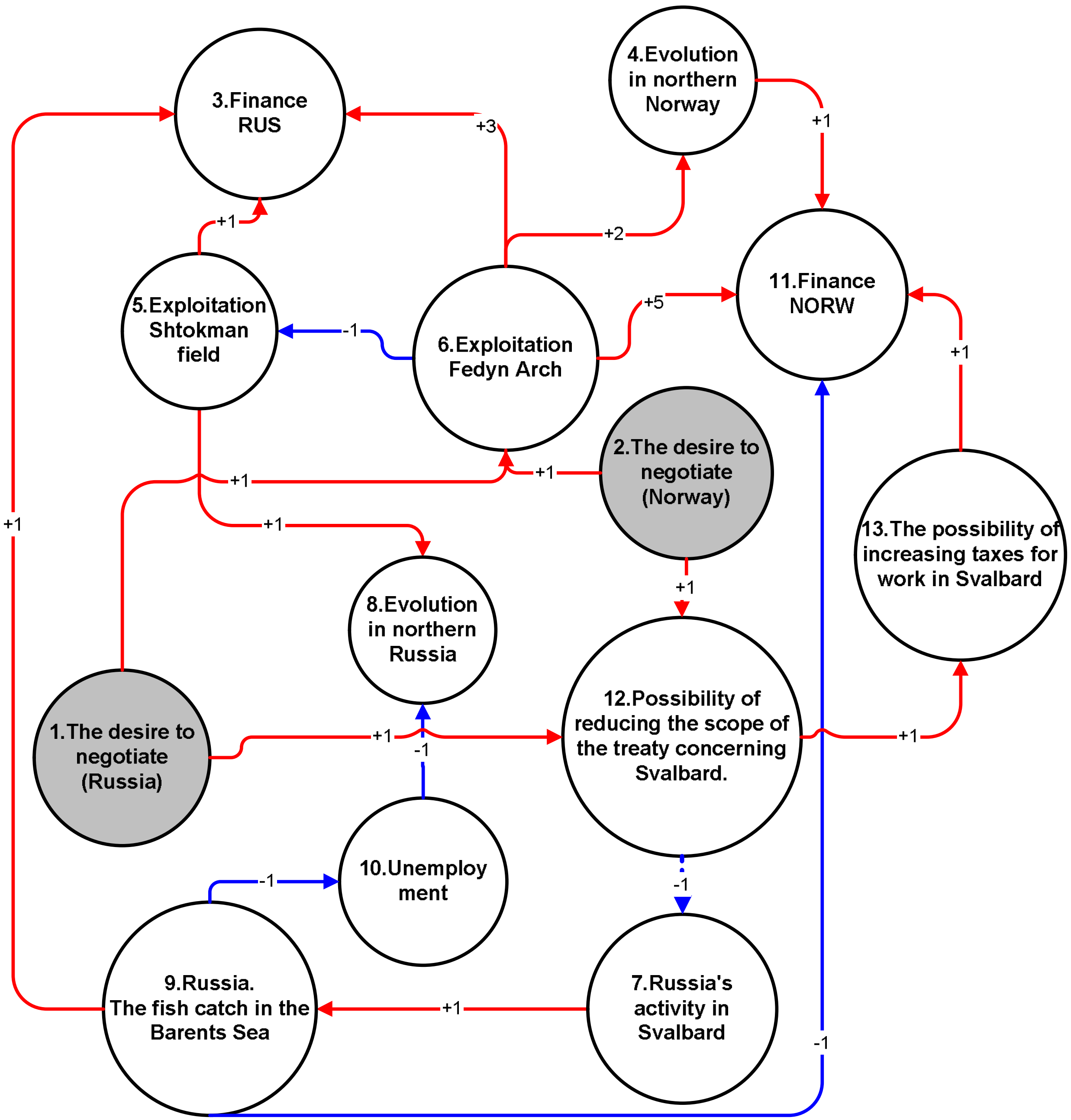
$$_T\alpha_{ik}\ {}_T\alpha_{ik}{}^\varepsilon = {}_T\alpha_{ik}({}_T\alpha_{ik} \pm \varepsilon_k^i \cdot \sum_{s=1}^m \gamma_{is}) > 0 \Rightarrow \varepsilon_k^i < \left|\frac{{}_T\alpha_{ik}}{\sum_{s=1}^m \gamma_{is}}\right|$$

$$\varepsilon = \min_{i \in N} \varepsilon^i = \min_{i \in N} \min_{k \in M_i} \varepsilon_k^i$$

Estimate (lower bound) of tolerance to error in a weight of arc in digraph of cognitive map



Example



Cognitive map that reflects the causal links between concepts in the problem of the disputed territory in the Barents Sea.

The *control concept* for **Russia** is **#1** to **Norway** is **#2**
The *target concepts* for **Russia** is **#3 and #8** to **Norway** is **#4 and #11**

Conclusion

- Concept # 4 should not be selected as the target concept because it is not critical to the invariance result.
- It is shown that the error in the target coefficients γ_{ij} are critical for unchanging of the decision by Russia. This may be an illustration of the unstable situation of Russia in this conflict for a given system of priorities.
- The most critical to errors were links (#1 → #6) and (#2 → #6). This illustrates the importance of concept # 6 for target of the agents.

As a result was evaluated the error in the coefficients of the utility functions of agents as well as in the weights of the arcs of the digraph of cognitive maps in this model. The estimation of errors tolerance made it possible to understand the structural properties of the model and to assess the degree of selecting target concepts feasibility.

Thank you for your attention!