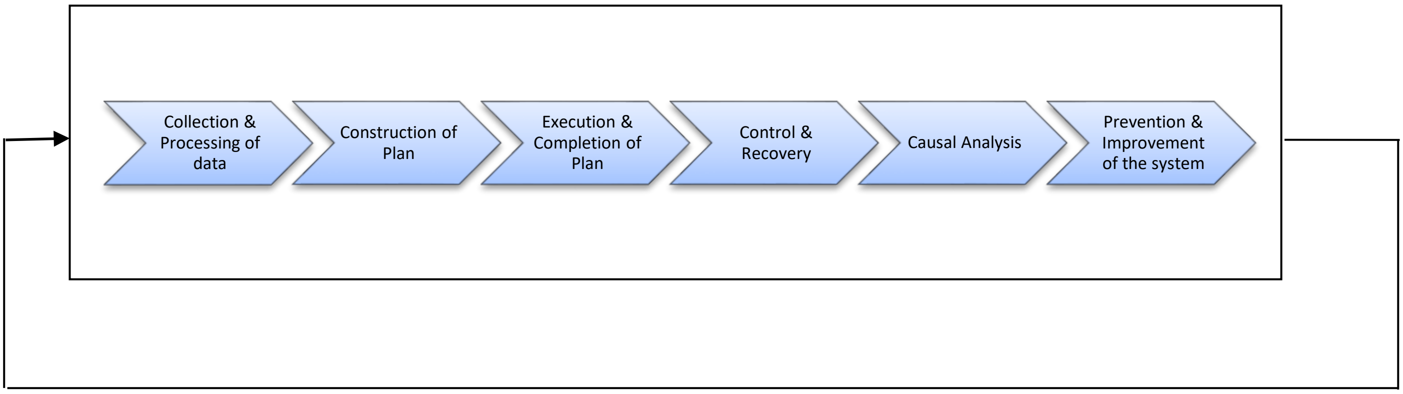
Building an Operational Defense Mechanism

Analysis of the pipeline



Depended on the type of the phenomenon, a backbone configuration for the whole system will be created that will instantiate the system to an instance for dealing with the specific phenomenon.

Collection & Processing of data

In this stage of the pipeline, data is collected and processed in order to establish a comprehensive view of the incident of the phenomenon. Also, data regarding the available resources that are going to build the defense line is collected and processed too.

This stage of the pipeline involves the following processes:

1. Collection of data from a predefined and customizable set of data sources (sensors, geographical data, data from predictions and simulations etc) depended on the type of the phenomenon (e.g for typhoons sensors for air-temperature, humidity, geographical data, simulations etc).
2. Data fusion of the data coming from different data sources.
3. Further data processing depended on the kind of task the dashboard will serve.
4. Creation of dashboard for visualization of info on geographical or other kind of maps.

It is noticed that any registered user will have access to a dashboard, corresponding to the kind of view (role) he/she has to the system that will provide him with information on the incident of the phenomenon, will allow him/her to organize his/her moves for his/her safety, receive directions and stay aligned with the overall safety plan designed for the whole region (this will be better understood as we go through the presentation of the system).

Plan Construction

Based on the data of the dashboard the user e.g. a governmental officer should be able to have a good concept of the extent and strength of the phenomenon in the region, the available defense forces and the areas that need to be secured. In the case of an individual user the information on his/her dashboard will inform him/her correspondingly about the extent of the phenomenon, critical points of aid, where their loved ones are at the moment and so on.

We assume that Plan Construction stage is carried out on a simulation basis, where the phenomenon is not directly threatening the region. The data collected and simulation algorithms can reliably represent the incident of the phenomenon for this region and a near optimal strategy can be constructed for the preparedness of the defense line. However, we can anticipate that this stage can be tuned to work well with runtime construction of a plan, while the phenomenon is directly threatening the region.

The top level of decision-making (level 1) e.g. the government authorities, in this stage of the pipeline, will be able to develop a strategy with the aid of a sophisticated mechanism I am proposing in this work (but another one could be used instead without constraint) and based on the dashboard they have in front of them, they will issue appropriate instructions to the next level of the decision-making hierarchy (level 2), which have to be executed as accurately as possible.

In order to achieve this goal, the second level will come up with a strategy with the aid of the sophisticated mechanism proposed and based on the dashboard of their reference, which is most likely different from that of level 1 and which corresponds to the context and nature of this execution level of the hierarchy, they will finally issue appropriate instructions towards the next execution level.

In this way, the appropriate instructions for each level traverse the execution hierarchy and after these instructions have been executed, the outcome updates accordingly the dashboards of each decision-making level. With informed dashboard, level 1 will issue a new set of instructions based on the constructed strategy, on real time, beginning a new cycle of plan communication across the levels of the execution hierarchy.

This hierarchy, it is reasonable to think that it can reach the level of individual citizens, living in the region, for example, who receive directions from the level above and make their optimal plan too. However, both people and institutions being in zone 1, which we regard as hazardous region, close to inflicted area, will have to directly follow the directions coming from the above level and those belonging to zones 2, 3, 4,.., which are not in as much dangerous situation, will have more freedoms in choosing an optimal plan of rescue for them, using the system for Plan Construction. In either case, both categories are informed about the situation, receive directions and are aligned with the overall safety strategy/plan. We can even consider “drone assistants” as constituent parts receiving instructions and coming up with their own optimal plan.

The sophisticated mechanism proposed ensures, through this cycle, which instructions follow across the decision-making hierarchy, that, firstly, each level comes up with a near optimal policy for fulfilling the instruction that has been given by the above level, based on a necessarily well informed dashboard that they have in front of them, each level is agnostic of the structure and the context of the below level (hence no particular domain knowledge or special skills are required for any decision-making level), but the mechanism indicates through an iterative process how each level will fulfill the instruction been given. In this way all the constituent parts of the execution hierarchy align very well with the decided strategy, which strategy is decided on real time and adapts to the changes of the environment in order to succeed a near optimal outcome. Thus, based on this mechanism, each level of the hierarchy exhibits elasticity, good communication with the environment, good communication with the other parts of the defense line, good communication with the goal and good communication with the time.

The approach described above seems suitable to be also applied to other fields of decision-making.

It is important to notice that, although this approach seems quite promising, we have not completed the experimentation needed in order to have a detailed view of how it works.

Execution & Completion of Plan

In this stage, the plan of actions decided in the previous stage is executed, adapted to the runtime environmental conditions, consulting the Plan Construction mechanism, based on the informed dashboards that have been updated to depict the runtime conditions.

The completion of plan execution is indicated by the condition, where the phenomenon has decreased its power and the situation has been set, for the most part, under control. The region has been protected adequately well under a good strategy and orchestration of actions or not well enough in either case.

It is strongly encouraged, a system for tracking the flow of instructions and outcome to be adopted in this pipeline stage, which will permit to identify defective points in the defense system (a level in the decision-making process, the strategy itself, an execution unit and so on) later in the Causal Analysis stage of the pipeline.

Control & Recovery

Control refers to all those actions that need to be scheduled in order to ensure that there are not any parametric hazardous conditions that have been caused as a result of the phenomenon's blow or out of error during the execution of plan stage that may put in danger people's lives and cause a second round of devastation (such as leaking gases, destroyed cables, danger of explosion, contaminated water/food, landfall and so on).

Also, in this stage, collection of clues and elements from the inflicted region is taking place in order to feed the next stage of the pipeline, Causal Analysis. Furthermore, a first estimation of the size of the damage and its distribution across the region is made in this stage, which is a useful input to the functionality that supports the Recovery phase.

Recovery refers to the orchestration of First Aid and Support for the people to overcome the adverse situation and for the services to be brought back to normal operation as quickly as possible.

An analogous communication and orchestration mechanism described above roughly (Plan Construction stage) could be applied in this part of the system too in order to orchestrate the actions for distributing aid and support, while ensuring that the most appropriate distribution is taking place in the places of need.

At this point, we can identify that solutions for tracking the distribution of aid and support can find a good place in this stage of the pipeline.

For the Recovery we realize that healthcare unit's ecosystem plays an important role and it is expected that they can respond well in their duty under these new conditions.

Support can be realized by throughout all the community of users who want to contribute. By keeping record of the needs rising out of people in inflicted regions and "booking" contributions by users capable of delivering end-to-end support.

NOTE

At this point of the analysis, we recognize the benefit of developing and maintaining a mode of preparedness and contribution to address adverse situations, in, first of all, any institution directly connected to the operation of the pipeline stages, such as institutions of the healthcare ecosystem, First Aid and Support institutions and so on. This mode is essentially a functional mechanism designed to meet this purpose. Other institutions and companies are also encouraged to develop a binding in their platforms to the system, to serve the effective operation of any of the pipeline stages.

Causal Analysis

This might be the most important stage of the pipeline. Based on clues and elements having been collected during the Control phase and records kept during the tracking mechanisms of the previous stages of the pipeline, carrying out analytics and research, we try to answer the questions "why this phenomenon has made its appearance to such an extent at this particular time and place?", "what was good and what was bad in the system's response, which is the cause of the bad response?".

The outcome of the research will be fed to the next stage of the pipeline, Prevention & Improvement of the system.

Prevention & Improvement of the system

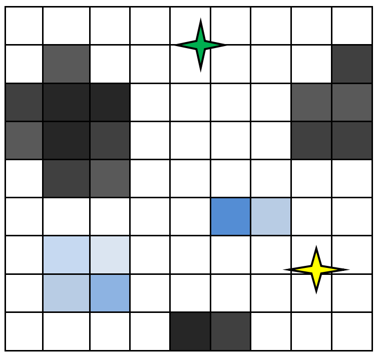
In this stage, we implement the actions suggested by the previous stage of the pipeline in order to control and eliminate the causes that were highlighted during the research phase of that stage as leading causes of the catastrophic phenomenon and system's defective points.

Regarding the improvement of the system, all stages of the pipeline are potential receptors of modifications.

Below i present the proposed mechanism for the “communication” of a strategy across the levels of a decision-making hierarchy for the successful construction of plan, for securing a region against natural catastrophic phenomena. The way the mechanism works is also depicted in summary in the description of the Plan Construction stage of the suggested pipeline.

For better clarification of the mechanism, the presentation is made based on a simple case study and includes the corresponding notation and background, the formulation of the problem of plan construction based on the specifications we have mentioned previously in the analysis, the algorithmic steps of the mechanism and explanatory notes.

In the beginning, we assume that Level 1 encounters a dashboard similar to the one below in our simplistic case study.



The dashboard depicts the conditions that exist in the region of our interest.

To each cell of the grid we correspond a value that represents the estimated energy distribution as for the effect of that cell towards enforcing or threatening the defense. This value is calculated based on geographical data, on data coming from sensors, data from domain databases, from simulations and so on with the aid of scales for stratifying conditions, during the first stage of the pipeline, Collection & Processing of data. For example, if the incident of the phenomenon is wildfires, the dark regions on the dashboard represent the inflicted places on the geographical map, the blue the defense forces, the stars resources that are critical (such as an army base, a forest region and so on) and the white cells are neutral regions as for this level of decision-making. The shades of each color represent differences in the energy distribution of each entity, which energy distribution is estimated using algorithms and data relevant to the nature of the corresponding entity. The energy distribution of each entity determines the behavior of the entity in time and space, the velocity of the entity on the board and its power towards enhancing or deteriorating defense. So, in the above example, the very dark regions correspond to a place where wildfire is of significant extent maybe because of dense and flammable vegetation existing in that place. The difference shades of blue represent the size and the equipment that each emergency center possesses, based on the domain records they keep in their datastores. For the purpose of constructing a strategy at this level of decision-making, we will represent places with critical resources as blue cells of very high energy.

The dashboards are represented for each level as arrays of 3 dimensions and are denoted as s (Level 1), ρ (Level 2), f (Level 3) and d (Level 4) in our example. With 'a' we denote the action decided by a level. 'p' represents the policy decided to be followed (or strategy for Level 1), which for a given configuration of the dashboard depends on the action been chosen at each turn.

In the below diagram (Fig1), it is displayed how the "communication" of the plan and the alignment with the strategy adopted by Level 1 could be achieved. (The diagram shows the basic idea behind the mechanism for achieving these specifications for this simple case. In real decision making scenarios we expect more parameters to come in and refinement of the algorithm that will be applied)

According to this approach, in Level 1 a model trained with a Reinforcement Learning (RL) technique accepts as input the configuration of the dashboard s and based on the Goal of the strategy, which is to increase the area that defense forces control, produces a policy, which indicates, what next action, level 1 should choose. The nature of this action is the placement of defense force of specific energy distribution on the board. The energy distribution on the specific coordinates on the board, E(X, Y) is going to be the instruction towards the next level of decision making.

The instruction issued on Level 1 is going to be the Goal for the next level, Level 2, and the Goal of the Reinforcement Learning model used by this level of decision making. The RL model of this level is preconfigured using elements from ρ and sets as Goal the minimization of the divergence between E_1 and the sum of E_{2a}, E_{2b}, E_{2c} for a specific area on board. E_{2a}, E_{2b}, E_{2c} represent instructions to the next level of decision making, which in this case is composed of three emergency centers. The emergency centers have to produce and deliver in tandem the required energy distribution at the region of interest. However, in reality these centers are not of equivalent potential, each one may have strong points and lacks at the same time, so the energy distribution that each one can produce may vary. On the other hand, the Emergency operations' director may be agnostic of the exact condition of each center nor is it an easy task to decide the most cost-effective combination of forces. So, based on the RL model of Level 2, E_{2a}, E_{2b}, E_{2c} maybe initialized at random values which add up to E_1 and after a number of iterations is capable to stabilize to a combination of E_{2a}, E_{2b}, E_{2c} that approximates E_1 adequately well.

At this point of time, the emergency centers of Level 3 are facing each one a goal that have to achieve, which in fact it is feasible based on the potential of their unit, hence likely to be achieved. Each center separately possesses a dashboard which is different in many aspects than those of the levels above it. It is important to notice, that different mechanisms can be deployed in order a center to be facilitated in the achievement of their goal, which any of them may use a different kind of input. However, we have assumed that the dashboard contains all necessary data that the mechanism can use and hence to produce a near optimal plan. For a unit to achieve its goal, if it has not any levels of decision making to direct, means to produce a plan, which may involve the selection of the appropriate personnel, selection of equipment, complying with guidelines and follow best practices, finding cost-effective access to the region of interest, deciding dynamically on an appropriate plan in the exact geographic region and so on. Taking advantage of these parameters one can think of various goal-driven approaches to deploy. In the case that more levels exist after Level 3 the same approach can be followed as in Level 2, in order to determine an appropriate instruction for energy contribution from the below execution level, which best suits to the potential of the level above and approaches the quantity of energy in need. We also clarify here that the situation, where part of the instruction is executed in one level whereas another by a unit in a lower level can also be modeled by the algorithm. For example, this might be the case where a center in Level 3 cooperates and controls a unit of volunteers. In this case, some part of the energy is produced in Level 3 and the volunteers' unit of Level 4 contributes in its best possible way with the remaining quantity of energy.

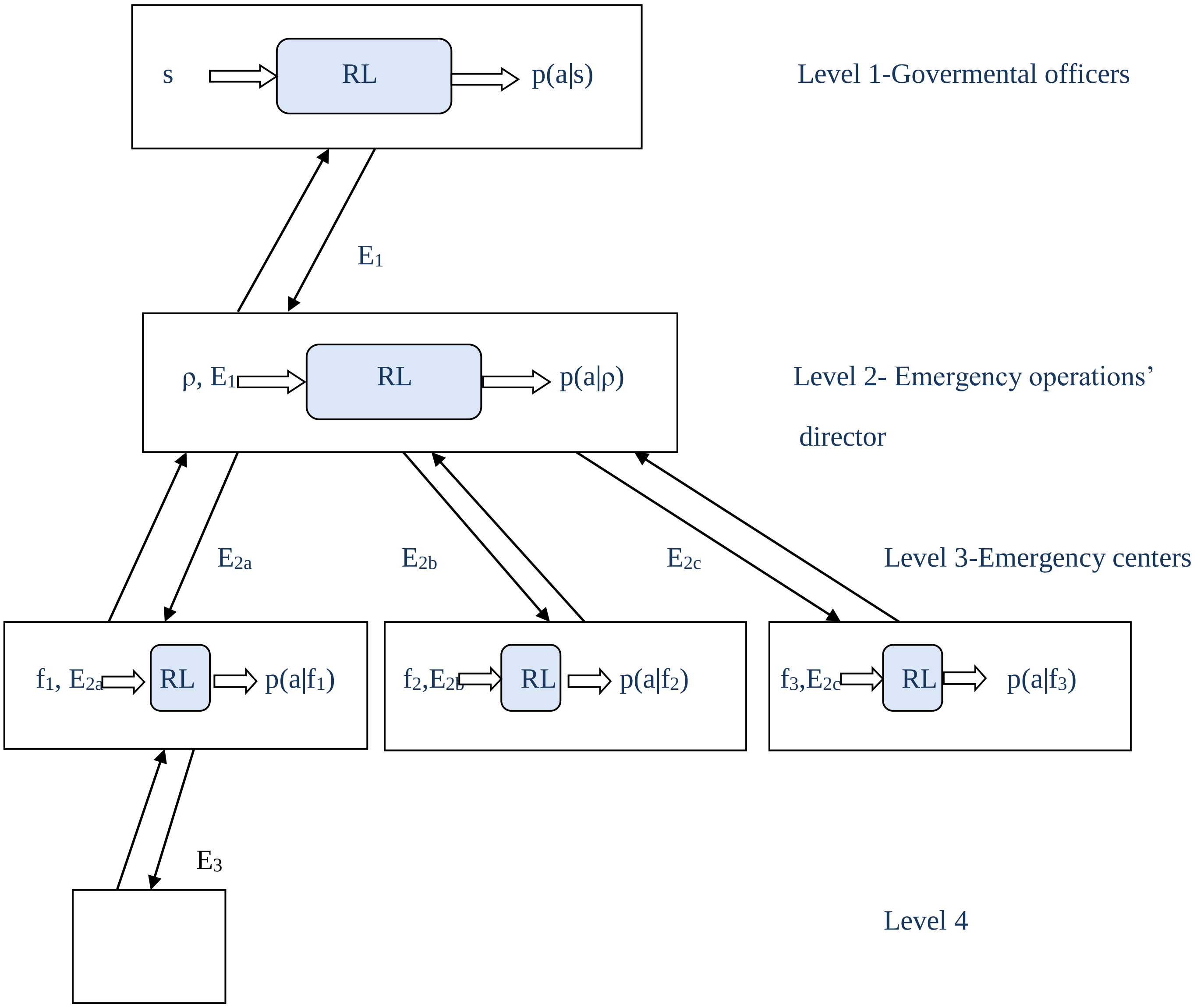
At any case, after the time interval Δt elapses, which corresponds to a reasonable time interval set by Level 2, within which it is expected the instructions issued to the lower levels to have been executed or to be close to completion, the dashboards of all levels are updated by the commitments of energy of all levels of execution and the dashboard of Level 1. At this point of time, Level 1 is in front of a new configuration of the dashboard, based on which decides a new action with the aid of the RL model it uses and hence a new cycle of instruction flow and execution begins.

If, however, by the end of Δt was not possible for the bellow levels to commit the corresponding energy distribution, in case of any unexpected obstacles or because Δt is not tuned properly or because the initial request for the specific energy distribution is beyond the capacity of the constitution of the execution hierarchy, then the execution hierarchy commits with what has achieved by the end of Δt and Level 1 with the aid of the RL mechanism decides the next action that most probably will drive to the success of the strategy.

Algorithm1:

#Level 1

1. Action a indicated by RL of Level 1 based on s.
2. Request defense power of E(X, Y) distribution for specific regions on board.
3. i=1
4. E_i= E(X, Y)
5. Set Δt_i within which the lower level should commit its contribution in energy distribution.
6. Assembly the right combination of requests for energy distribution using RL techniques, so that their aggregation approaches the requested E_i.
7. If i=N-1, commit the energy distribution produced, E_i'. // N is the # levels
8. Recursive step: Execute 5 & 6 for i=i+1
9. Update all the dashboards with the new energy distribution E_i' at the region of interest. s=s' ρ=ρ' f=f' d=d'
10. If the goal has been reached or if time constraints have been exceeded or the strategy has failed, stop. Otherwise go to step 1.



s
RL
p(a|s)
Level 1-Govermental officers
E1
ρ, E1
RL
p(a|ρ)
Level 2- Emergency operations' director
E2a
E2b
E2c
Level 3-Emergency centers
f1, E2a
RL
p(a|f1)
f2,E2b
RL
p(a|f2)
f3,E2c
RL
p(a|f3)
E3
Level 4

Fig 1

From the analysis of the algorithm we conclude that the best selection of an appropriate RL approach at each level of decision-making leads to an even better overall plan. For the Plan Construction stage of the pipeline, as well as for other parts of the system where "communication" and "orchestration" need to be effectively regulated, the appropriate RL algorithms could be chosen to be deployed.

Level 1

In the case of Plan Construction, for Level 1 it seems to be a good choice to experiment with AlphaGo algorithm as described in the paper "Mastering the game of Go with deep neural networks and tree search-Silver D. et al." developed by DeepMind Inc for building a strategy in playing the old Chinese game Go. One may chose an approach for choosing an optimal policy from a great variety of other approaches for dealing with adversarial games, however AlphaGo achieves excellent performance and new variations of it, developed lately by the same team promise even better features.

There is direct correspondence between our setting in Level 1 and that used in AlphaGo. Our goal, as in AlphaGo, is until terminal time T to have managed to reach a configuration on the dashboard, where the defensive forces control most of the region on the board. However, there is one slight difference in our setting, which we can address either by modifying the algorithm and by using an AlphaGoVariation, or by applying a transformation, which can give us an almost equal setting. The difference lies to the fact that we assume that each cell has a certain value corresponding to each potential, whereas in the game of Go all stones are of equal value and are place in cross-sections. We can adopt a one-to-one correspondence between cells and cross-sections, without any problem and for dealing with the stones of different value we could, without any problem, to assume that one action of placement one stone/force on the dashboard of energy E corresponds to placing E stones following appropriate distribution on the dashboard, changing rapidly in one step the configuration of the dashboard.

After applying these changes, we could experiment with finding a near optimal strategy for Level 1, by setting a game with the opponent, where defense plays using AlphaGo (or the modification of it) and the opponent/phenomenon follows the moves that real time simulations indicate, amplified by the value we assign to each cell as a result of context-factors we also consider in the representation of the threatening power of the phenomenon.

Since AlphaGo progresses by placing one stone on the board at each turn, we could force the opponent to pass for some turns (maybe E times) in order to suggest an energy distribution E(X, Y) on the board as an instruction to the next level.

Level 2

In Level 2, we could apply feed forward of the input energy distribution into a neural net, with Level 3 as hidden layer, initial random weights that add up to the input distribution (only 3 parameters), bias vector for level 3 with the potential of each center corresponding to the components of the bias vector and as an objective the minimization of a cost function for approaching the input energy distribution. We could otherwise apply this approach at once for all the levels of the hierarchy towards finding the combination of instructions that achieve approximation of the input energy distribution, while taking for granted (bias vectors) the productivity (potential) of each center at each level.

Level 3

For this level among the various options of mechanisms we can choose for achieving production of E_i distribution by a unit in Level 3 across a region of interest on the dashboard, we could suggest applying a policy gradient method for assembling different combinations of resources that might suggest alternate successful configuration plans at the real geographical space.

Another promising option is to deploy an RL mechanism that facilitates best rout navigation towards the region of interest using detailed geographic representation of the region on the dashboard (some inspiration can be drawn from this paper “Actions Combination Method for Reinforcement Learning-Karanik M. et al.”).

References:

1) *"Mastering the game of Go with deep neural networks and tree search" ,Silver D. et al., Nature*
2) *"Actions Combination Method for Reinforcement Learning", Karanik M. et al.*
3) *"Endocrine Regulation", Science Direct,*