# PRINCIPLE OF LIMITING GENERALIZATIONS

**Iurii Prokopchuk**

Institute of Technical Mechanics of NASU, Dnipropetrovsk, Ukraine
*itk3@ukr.net*

**Method of limiting generalizations (MLG).** The work deals with an efficient method for solution of intellectual logical and computing tasks. The method is based on construction of full knowledge model of the multilevel description of the reality with limiting characteristics. When estimating, the current situation is generalized within the limits proper to the complete model of knowledge. The method corresponds to basic principles of operation of natural intelligence [1-3].

Represent models of the application domain as the tuple $<O, k>$ where $O$ is the model of ontology of that application domain, and $k$ is the model of an adequate system of knowledge. Adequacy of the model of the application domain implies that the model of reality $A(<O, k>)$ coincides with a set of models for every situation including in reality of that application domain.

Represent the model of knowledge $k$ in a developed view as follows:

$$k = \{f/\mu\text{: } k^1 \to k^2\} \cup P_k, \tag{1}$$

where $f/\mu$ is data mapping realizing mathematical models in one way or another;
$\mu$ are the *d*istinct mechanisms of realization of data mapping;
$k^1$ are the input data of the task (description of information environment and job);
$k^2$ are the output data of the task;
$P_k$ are the rules of composition of tasks schemas, i.e. the rules describing modes for unifying local tasks.

Consider specifications of tasks for some classes of knowledge models ($\underline{\tau}/T$ are the results of tests; $d/D$ are the conclusions, diagnoses; $h/H$ are the prediction hypotheses; $r/R$ are the control programs; $T, D, H, R$ are the sorts or the domains) [2]:

$F_1 = \{f/\mu\text{: } \{\underline{\tau}/T\}_1 \to \{\underline{\tau}/T\}_2\}$ is the class of models for computing knowledge;

$F_2 = \{f/\mu\text{: } \{\underline{\tau}/T\} \to d/D\}$ is the class of models of diagnostic knowledge;

$F_3 = \{f/\mu\text{: } \{\underline{\tau}/T\} \to \neg d/D\}$ is the class of knowledge models describing the domain of prohibitions;

$F_4 = \{f/\mu\text{: } \{\underline{\tau}/T\}, \{d/D\} \to \{h/H\}\}$ is the class of models for prediction knowledge;

$F_5 = \{f/\mu\text{: } \{\underline{\tau}/T\}, \{d/D\}, \{h/H\} \to \{r/R\}\}$ is the class of knowledge models for optimization of control;

$F_6 = \{f/\mu\text{: } \{\underline{\tau}/T\} \to \{\underline{\tau}/T\}'\}$ is the class of knowledge models for description of the structure and the dynamics of complicated systems represented as collection of causal and consequent relations (both structural ones and time ones).

The general knowledge model $k$ includes all the above-mentioned classes of models, namely: $F_1 \cup F_2 \cup F_3 \cup F_4 \cup F_5 \cup F_6 \subseteq k$.

The closure of the set of data mapping $F^+/P_k$ is built by means of the rules of composition $P_k$ in solving a specific task [2].

Assume that $R^+ = \{\alpha_1, \ldots, \alpha_m\}$ is the sample of examples with complete information. Suppose that there is a finite set of elementary tests $\{\tau\}$ when any situation of reality $\alpha$ is uniquely re-established from $R^+$ by values of tests $\{\underline{\tau}\}$.

Assume that one of the tests takes values from finite and alternative sets $D = \{d_1,\ldots, d_n\}$. Denote that test by $\tau_d$.

Introduce the condition of separability of real situations based on sets of tests $\{\tau/T\}\backslash\tau_d$ and some transitive metric $\rho$:

$\forall\{\underline{\tau}\}, \{\underline{\tau}\}'$ where $\{\tau\}\subseteq\{\tau/T\}\backslash\tau_d$ and $\exists\alpha, \alpha' \in R^+\text{: } \alpha = \alpha(\{\underline{\tau}\},d), \alpha' = \alpha'(\{\underline{\tau}\}', d')$ the condition should be met: $\rho(\{\underline{\tau}\}, \{\underline{\tau}\}') = 0 \Rightarrow d = d'$.

If a common applied task (the class of applied tasks) may resolved by using in the case a more limited number of notions and statements (theorems), such model of the solution (the model of knowledge) will be considered to be more efficient. Conception of building models of knowledge with a minimal number of objects is dominant for the method under consideration.

Consider the following task.

**A task**. Assume that a representative sample of real situations $R^+$ with complete information is given at a particular level of abstraction (the level is determined by domains). Assume that the metric $\rho$ is given in such a manner that the condition of separability is performed on the set $R^+$. It is required to build a minimal remainder–free model of knowledge on sets $R^+$ from the point of view of an efficiency function $\gamma$: "the classification of the conclusions from $D$"

The paper deals with algorithms for solving the task for different classes of models of knowledge (in the context of a fixed combination of domains).

Notation $\tau/T$ implies that the results of the test $\tau$ can take the values of different domains $T$. Domains can represent a distinct level of generality. Consider the examples.

Assume that T1 – T4 are distinct domains for description of the human temperature:

T1 = [34, 42] degrees;

T2 = {[34, 35], (35, 36.5], (36.5, 36.8], (36.8, 37.4], (37.4, 42]};

T3 = [decreased; normal; elevated; high] temperature;

T4 = [normal; abnormal] temperature.

The above-mentioned groups of domains have the desired property that if the value of the test is given on one domain, values of the test may be determined on domains with a greater number by using the fixed (single) rules of recalculation.

In other words, by using the domains cited an improper order can be given by the criterion of generality, namely: T1 ≤ T2 ≤ T3 ≤ T4.

In the general case an oriented graph of domains can be determined for each test (factor). Any path on the graph implies a possibility of a unique recalculation of values from one domain to other one. One can set some graphs for any test.

In searching through all possible combinations of domains for diverse tests we derive a complete set of descriptions of reality with a variety of levels. A set of optimal models of knowledge for every description forms a complete model of multilevel description of reality. We name such descriptions which cannot be generalized by one test without breaking the condition of separability as critical ones.

We name the model which allows of solving a target task for any presented situation of reality as true one. In so doing, a new situation of reality is generalized at most in the context of the complete model of multilevel description resulting in a simplifying of the solution.

Thus the principle of the method of limiting generalizations lies in the following:

1. The most scanned graph of domains (or some graphs) is built for each test involved in description of the task. Experts in the application domain play a large role in the construction of graphs.
2. An optimal model of knowledge is built for each combination of domains defining the level of generality of description. A set of all optimal models of knowledge defines the complete model of a multilevel description of reality.
3. In searching the solution for a new situation a given situation is generalized at most to one of descriptions including a true model of knowledge (it is desirable to generalize to a critical description). The solution is situated at a new level of description. If the solution is not available, it is necessary to correct models of knowledge. It is important to keep in mind that the availability or the lack of the solution is dependent on a subjective estimate of truth of knowledge models (representativeness of samples at one or another level).

**Model of Multilevel Pattern Sketches.** A nonlinear mathematical model of (visual) pattern perception is presented, which describes stepwise transitions of the internal perception model from one pattern sketch to another in the course of the solution of purpose-oriented problems. A series of sketches of different generality for an arbitrary pattern are uniquely formed on the basis of an inductive recurrent scheme of limiting generalizations, which is based on a complete system of iterating contractive mappings. The limit of generalization is finite sketches consisting of a set of singular points (attractors), i.e. points that vanish on application of any of the iterating mappings.

The representation of any pattern has the form of multilayer hierarchic structure, which reflects, like annual rings, the stages of cognitive formation of the internal forms of the pattern.

A series of pattern sketches is characterized by several integral constants, which are used in one of the pattern comparison schemes. One of the most important constants is the number of layers; other constants are the total number of sketches in all layers and the total number of unique sketches. For an input pattern, all integral constants or part thereof are calculated and compared with the pattern constants from the memory.

In the second scheme, different patterns are compared starting with the maximum generality level, i.e. with finite sketches made up of singular points. If the sketches at the last level of the patterns under comparison coincide, the comparison is continued at the last but one level, and so on, until a solution is obtained. A program that illustrates this approach is developed.

Inductive recurrent scheme for constructing a series of pattern sketches**.**

Suppose we have an infinite field made up of identical squares. Each square has a definite color. Let two colors be distinguished: a background color and an indefinite color. Without a pattern, all squares are of the background color. An arbitrary graphic pattern $O$ is a finite set of squares whose color is other than the background. Let us define a grid that consists of macromeshes of size $M \times L$ squares each. The total number of squares in a mesh is $N = M \times L$. There exist precisely $N$ independent arrangements of the grid. The $i$-th arrangement, $i = 1,\ldots,N$, corresponds to a mapping $T_i$.

*The scheme of iterative generalization-abstracting* Σ[N:P] is as follows.

1. When generalizing in the framework of any $T_i$, the set of the squares of each mesh is transformed into a square of the next-level sketch, and its color is determined as follows: if $P$ squares of the mesh are of like color (which may be the background color), the square of the new sketch takes the same color, otherwise it takes the indefinite color.
2. All mappings $T_i$ ($i$=1,…,$N$) are applied to the original pattern. As a result, $N$ sketches of a first layer are formed. We leave only different sketches in the first layer and apply all mappings $T_i$ ($i$=1,…,$N$) to each of them to obtain sketches of a second layer. Then we leave only unique sketches in the second layer. The iterative process is repeated until no sketch is found in some step.

## References

1. Prokopchuk Iu. Method of Limiting Generalizations for Solving Logical and Computing Tasks // Proceedings of 2nd International Conference on Inductive Modelling (September 15 – 19, 2008, Kyiv, Ukraine). – Kyiv : IRTC ITS, 2008. – PP. 62 – 65.
2. Prokopchuk Yu. A. Intellectual Medical Systems: Formal and Logic Level. - Dnepropetrovsk: ITM of NASU & NSAU, 2007, 259 p. (In Russian)
3. Prokopchuk Iu, Kostra V. Representation of Algorithmic Knowledge in Medical Information Systems // 7th International Conference "Knowledge-Based Intelligent Information and Engineering Systems" (KES 2003) (Oxford, UK, September 3-5, 2003): Proceedings, Part II, Series: Lecture Notes in Artificial Intelligence, Vol. 2773. – Berlin: Springer-Verlag Berlin Heidelberg, 2003. - PP. 1069-1074.