Информационные процессы, Том 21, № 1, 2021, стр. 65–81

Combinatorial planning framework for geological exploration

Mark Sh. Levin

Institute for Information Transmission Problems, RAS, Moscow, Russia
email: mslevin@acm.org

Received 17.03.2021

Abstract—The paper describes a combinatorial framework for planning of geological exploration for oil-gas fields. The suggested solving scheme consists of the stages: (1) building of special 4-layer tree-like model of geological exploration; (2) generations of local design (exploration) alternatives for each low-layer geological object: conservation, additional search, independent utilization, joint utilization; (3) multicriteria assessment of the exploration alternatives and their interrelation (compatibility); (4) hierarchical design of composite exploration plans; (5) integration of the plans into region exploration plans (versions); and (6) aggregation of the region plans (versions) into a general (resultant) exploration plan. Stages 2, 3, 4, and 5 are based on hierarchical multicriteria morphological design (HMMD) method. The composition problem is based on morphological clique model. Aggregation of the obtained modular alternatives (stage 6) is based on multiple choice model. The usage of multiset estimates for alternatives is described as well. The alternative estimates are based on expert judgment. The suggested combinatorial planning methodology is illustrated by realistic numerical examples for geological exploration of Yamal peninsula.

Key words: planning, geological exploration, oil-gas field, combinatorial optimization, morphological analysis, heuristic

1. INTRODUCTION

In recent decades significance of mineral resources and their geological exploring has been increased. This paper addresses combinatorial planning of geological exploration for oil and gas fields. The suggested solving framework is the following:

1. A system analysis of the initial applied problem and its structuring (e.g., partitioning the problem, generation of system requirements/criteria).

2. Design of a special four-layer tree-like model of geological objects: (i) productive stratum (reservoir), (ii) group of productive stratums (reservoirs)), (iii) oil and gas field, and (iv) group of oil and gas fields (region).

3. Generation of local design alternatives DAs (geological exploration) for each bottom-layer geological object: (a) conservation, (b) appraisal work, (c) independent production time, and (d) joint independent production time.

4. Multicriteria assessment of DAs and their interconnection (compatibility IC) and mapping the obtained vector estimates into an ordinal scale.

5. Hierarchical composition of exploration plans for each field.

6. Integration of the obtained exploration plans for the fields into region exploration plans (solution versions).

7. Aggregation of the obtained region plans (solution versions) to obtain a total solution.

Multicriteria morphological design (HMMD) method [10–12] is used for stages 2, 3, 4, 5, 6. Aggregation of the obtained modular solutions (stage 7 above) is based on the strategy with using multiple choice problem [12] An example of combinatorial solution based on multiset estimates of DAs is described as well. All stages above are based on expert judgment.

Note, combinatorial approach for selection of optimal geological actions based on knapsack-like model was described in [9]. The suggested combinatorial framework is illustrated by a realistic numerical example for Yamal peninsula. The preliminary simplified planning example was published in [10, 14]. Mainly, initial information for the example is based on handbook [16] and expert judgment [10, 14]. A preliminary version of the paper was published as electronic preprint [13].

2. SOLVING FRAMEWORK

This section contains basic materials (i.e., the brief descriptions of framework parts [10–12]) for usefulness of readers.

2.1. Multicriteria Ranking (Sorting)

Multicriteria ranking (sorting) is used as important part of many solving frameworks [11, 12, 21]. Let $A = \{1, ..., i, ..., n\}$ be a set of alternatives (items) which are evaluated upon criteria $K = \{1, ..., j, ..., \lambda\}$, and $z(i, j)$ is an estimate (quantitative, ordinal) of alternative i upon criterion j. The matrix $Z = \{z(i, j)\}$ may be mapped into a poset on A. The following resultant kind of the poset as a partition with ordered subsets (a layered structure) is searched for (Fig. 1): $Y = \{Y_1, ..., Y_k, ..., Y_\pi\}$, $Y_{k_1} \& Y_{k_2} = \emptyset$ if $k_1 \neq k_2$, and each alternative from Y_{k_1} (layer k_1) dominates each alternative from Y_{k_2} (layer k_2), if $k_1 \leq k_2$. Thus, each alternative has a priority which equals the number of the corresponding layer. The basic techniques for the multicriteria ranking (sorting) problems are [5, 12, 18, 19, 21]: (1) multi-attribute utility analysis, (2) analytic hierarchy process and its modifications, (3) outranking techniques and their modifications, (4) expert judgment procedures.

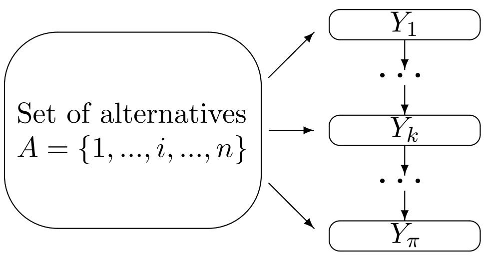

Fig. 1. Multicriteria ranking (sorting)

2.2. Knapsack-like Models

The basic knapsack problem is [3, 6, 15]:

$$\max \sum_{i=1}^{m} c_i x_i \quad s.t. \quad \sum_{i=1}^{m} a_i x_i \ \leq b, \quad x_i \in \{0, 1\}, \ i = \overline{1, m},$$

where $x_i = 1$ if item (element) i is selected, c_i is a value ("utility") of item i, and a_i is a weight of item i (or required resource). Often nonnegative coefficients are assumed. The problem is NP-hard [3]. In multiple choice knapsack problem, the items are divided into groups and it is necessary to select element(s) from each group while taking into account a total resource constraint (or constraints). The basic multiple choice knapsack problem is (Fig. 2):

$$\max \sum_{i=1}^{m} \sum_{j=1}^{q_i} c_{ij} x_{ij}$$

$$s.t. \quad \sum_{i=1}^{m}\sum_{j=1}^{q_i} a_{ij}x_{ij} \leq b, \ \sum_{j=1}^{q_i} x_{ij} = 1, \ i = \overline{1,m}, \quad x_{ij} \in \{0,1\}.$$

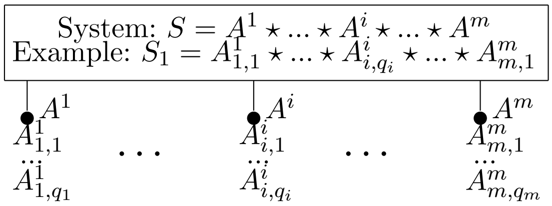

Fig. 2. Illustration for multiple choice problem

Recently, multiple criteria description is often used $\{c_{i,j}\}$ $\forall(i,j)$ (i.e., multi-objective multiple choice knapsack problem). Thus, the vector objective function $\overline{f} = (f^1, ..., f^r)$ is [7,12]:

$$(\max \sum_{i=1}^{m}\sum_{j=1}^{q_i} c^1_{ij}x_{ij} \ , \ ... \ , \ \max \sum_{i=1}^{m}\sum_{j=1}^{q_i} c^r_{ij}x_{ij} \ ).$$

Evidently, here it is necessary to search for the Pareto-efficient (e.g., by the vector objective function above) solutions. Here the following solving schemes can be used (e.g., [2,3,7,12,15]):

1. Heuristics.

2. Dynamic programming (e.g., [2,3,7,15]).

3. Two-stage heuristic based on reducing the initial multicriteria problem to one-objective one: (i) multicriteria ranking of elements $\{(i,j)\}$ (i.e., by vector $(c^1_{ij}, ..., c^p_{ij}, ..., c^r_{ij})$) to get the ordinal priority for each element above (i.e., $\forall(i,j)$), (ii) examination of the obtained multiple choice knapsack problem in which the priorities are used instead of c_{ij}; the solving process of this problem may be based on well-known methods (e.g., greedy heuristic, dynamic programming).

2.3. Morphological approaches

The morphological approach is targeted to combination of alternatives for different system components while taking into account the compatibility of the selected alternatives. Table 1 contains the list of basic morphological methods.

Table 1. Morphological methods

No.	Approach	Source(s)
1.	Basic morphological analysis	[22]
2.	Morphological methods for design	[4]
3.	Morphological methods for planning and forecasting	[1]
4.	General morphological analysis	[17]
5.	Hierarchical morphological design	[10–12]
6.	Hierarchical morphological design with multiset estimates	[12]

2.4. Hierarchical morphological design

Hierarchical morphological multicriteria design (HMMD) method has been described in several publications [10–12]. In HMMD a special hierarchical (tree-like) model of the analyzed system 'morphological tree' is used: (i) tree-like system model, (ii) set of leaf node as bottom-layer components (parts) of the systems, (iii) set of design alternatives (DAs) for each bottom system component, (iv) ranking of DAs for each each bottom system component (to obtain an ordinal estimate/priority for each DA), and (v) ordinal estimates of compatibility between DAs of neighbor system components.

In the basic version of HMMD (ordinal estimated of DAs), the following is assumed: (1) system quality is considered as a two-component estimate: quality of components and quality of their compatibility; (2) monotone criteria for the system and its parts are considered; (3) an ordinal scale is used for quality of system components (i.e., local solutions); and (4) an ordinal scale is used for quality of system component compatibility. The following designations are used: (a) priorities (ordinal estimates) for design alternatives (DAs): $\iota = 1, ..., l$, 1 corresponds the the best quality level; (b) ordinal compatibility for pair of DAs: $w = 0, ..., \nu$, 0 corresponds to impossible (the worst) quality level.

The synthesis problem of a composite solution (DA) consisting of m parts is based on morphological clique model:

Find composite system $S = S(1) \star ... \star S(i) \star ... \star S(m)$, consisting of parts/components (i.e., local DAs) (one representative $S(i)$ for i-th system component $i = 1, ..., m$) with non-zero compatibility estimates between the selected pair of DAs.

The poset of the system quality for composite solution S is based on vector $N(S) = (w(S); e(S))$:

(a) $w(S)$ corresponds to minimum of compatibility estimates for DAs pair in S,

(b) $e(S) = (\eta_1, ..., \eta_\iota, ..., \eta_l)$, where η_ι corresponds to the number of local DAs at quality level ι in S.

Two-criteria optimization model is:

$$\max\ e(S), \quad \max\ w(S), \quad s.t.\ \ w(S) \geq 1.$$

As a result, non-dominated by $N(S)$ (Pareto-efficient) composite solutions are searched for. The model belongs to the class of NP-hard problems. The evident solving scheme involves two stages:

Stage 1. Building of all admissible solutions (composite DAs).

Stage 2. Selection of Pareto-efficient solutions.

Two algorithms can be used for the problem [10,11]: (1) directed enumeration of solutions (start solution(s) corresponds to the best quality estimate(s)), (2) dynamic programming based method (series construction of admissible solutions for system parts). Note, in the case of a small degree of the system tree-based model (for example, [3...7]) algorithmic complexity of the first algorithm is sufficiently small. An illustrative example of tree-component system $S = H \star B \star U$ is depicted in Fig. 3 (priorities of DAs are shown in parentheses). The following solutions can be considered:

(a) $S_1 = B_1 \star H_1 \star U_2$, $N(S_1) = (3; 1, 1, 1)$; (b) $S_2 = B_3 \star H_2 \star U_2$, $N(S_2) = (2; 2, 1, 0)$; and

(c) $S_3 = B_2 \star H_3 \star U_1$, $N(S_3) = (1; 3, 0, 0)$.

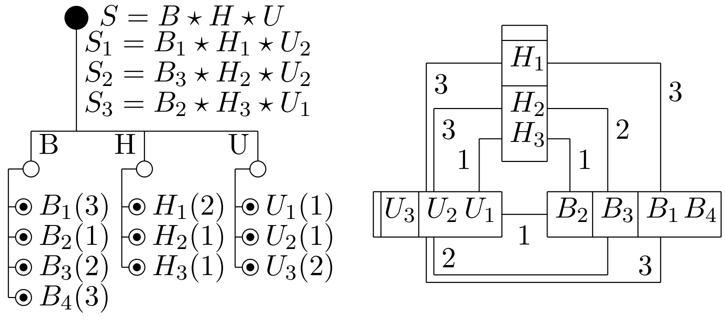

Fig. 3. Illustrative example of combinatorial synthesis

Fig. 4 illustrates the system quality poset without taking into account component compatibility. Here poset parameters are: $m = 3$, $l = 3$. The general system quality poset based on $N(S)$ $(w = 1, 2, 3)$ is depicted in Fig. 5. This poset consists of three posets from Fig. 4.

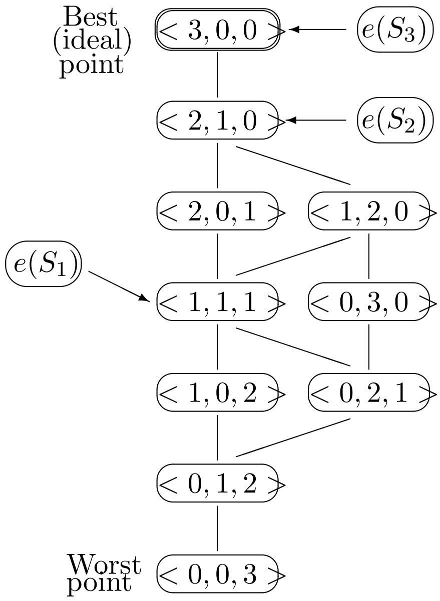

Fig. 4. Poset $e(S) = (\eta_1, \eta_2, \eta_3)$

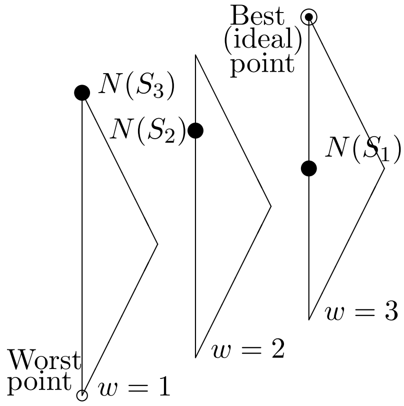

Fig. 5. General system quality poset

Fundamentals of multiset theory are described in [8, 20]. Interval estimates based on multisets and their applications in combinatorial synthesis have been suggested in [12]. In this section, the description of the interval multiset estimates corresponds to material in [12]. The following basic scale is used: $[1, 2, ..., l]$ $(1 \succ 2 \succ 3 \succ ...)$. Interval estimate e for object (alternative) A by scale $[1, l]$ is (position representation): $e(A) = (\eta_1, ..., \eta_\iota, ..., \eta_l)$, where η_ι corresponds to the number of elements at the quality level ι $(\iota = \overline{1, l})$. The assumptions are:

Condition 1: $\sum_{\iota=1}^{l} \eta_\iota = \eta$ (or $|e(A)| = \eta$).

Condition 2: $(\eta_\iota > 0)$ & $(\eta_{\iota+2} > 0)$ $\Longrightarrow$ $\eta_{\iota+1} > 0$ $(\iota = \overline{1, l-2})$.

Presentation of the estimate as multiset is:

$$e(A) = \{\overbrace{1, ..., 1}^{\eta_1}, \overbrace{2, ...2}^{\eta_2}, \overbrace{3, ..., 3}^{\eta_3}, ..., \overbrace{l, ..., l}^{\eta_l}\}.$$

The number of multisets for fixed value of element numbers η is called coefficient of multiset or multiset number:

$$\mu^{l,\eta} = \frac{l(l+1)(l+2)...(l+\eta-1)}{\eta!}.$$

This number corresponds to possible number of estimates or cardinality (without taking into account condition 2). In the case of condition 2, the number of the estimates is decreased. In [12], the following designations for assessment problems based on the interval multiset estimates are suggested: $P^{l,\eta}$.

In the numerical example the following assessment problem is used $P^{3,4}$. Clearly, the basic version of HMMD is based on assessment problem $P^{l,1}$.

Now the integrated multiset estimate is described. There are n initial estimates:

$$e^1 = (\eta_1^1, ..., \eta_\iota^1, ..., \eta_l^1), ..., e^\kappa = (\eta_1^\kappa, ..., \eta_\iota^\kappa, ..., \eta_l^\kappa), ..., e^n = (\eta_1^n, ..., \eta_\iota^n, ..., \eta_l^n).$$

The integrated multiset estimate is:

$$e^I = (\eta^I_1, ..., \eta^I_\iota, ..., \eta^I_l), \quad \eta^I_\iota = \sum_{\kappa=1}^{n} \eta^\kappa_\iota \;\; \forall \iota = \overline{1,l}.$$

The following basic operation is used: $\uplus$: $e^I = e^1 \uplus ... \uplus e^\kappa \uplus ... \uplus e^n$.

The vector proximity between two multiset estimates $e(A_1)$, $e(A_2)$ is:

$$\delta(e(A_1), e(A_2)) = (\delta^-(A_1, A_2), \delta^+(A_1, A_2)),$$

where (i) δ^- corresponds to the number of one-step changes (modifications) of quality element $\iota+1$ into quality element ι ($\iota = \overline{1, l-1}$) (this is an improvement); (ii) δ^+ corresponds to the number of one-step changes (modifications) of quality element ι into quality element $\iota + 1$ ($\iota = \overline{1, l-1}$) (this is decreasing of quality). This description corresponds to modification as editing of object (alternative) A_1 into alternative A_2. In addition, the following is assumed: $|\delta(e(A_1), e(A_2))| = \max\{|\delta^-(A_1, A_2)|, |\delta^+(A_1, A_2))|\}$.

Further, aggregation of estimate (as searching for a median) is examined. There are a set of estimates (as a set of objects/alternatives):

$$\widehat{E} = \{e_1, ..., e_\kappa, ..., e_n\},$$

the set of possible estimates is $\widehat{D}$ ($\widehat{E} \subseteq \widehat{D}$). The aggregation estimate as generalized median is [12]:

$$M^g = \arg min_{X \in \widehat{D}} \biguplus_{\kappa=1}^{n} \mid \delta(e(X), e_\kappa) \mid .$$

Thus, combinatorial synthesis problem based on multiset estimates of DAs is the following:

$$\max \;\; e(S) = M^g = \arg min_{X \in \widehat{D}} \biguplus_{\kappa=1}^{n} \mid \delta(e(X), e_\kappa) \mid ,$$

$$\max \;\; w(S),$$

$$s.t. \;\; w(S) \geq 1.$$

2.5. Aggregation of modular solutions

In [12], basic aggregation strategies for modular solutions are considered. Let $\overline{S} = \{S^1, ..., S^n\}$ be a set of initial modular solutions. A general aggregation strategy is targeted to searching for consensus/median solution S^M (this is generalized median) for the initial solutions $\overline{S} = \{S^1, ..., S^n\}$:

$$S^M = \arg \;\; min_{X \in \overline{S}} \, (\sum_{i=1}^{n} \; \rho(X, S^i)),$$

where $\rho(X, Y)$ is a proximity between two solutions $X, Y \in \overline{S}$. This problem (searching for the generalized median) is often NP-hard. It may be reasonable to use simplified (approximate) strategies, for example: (a) selection of solution from the set of initial solutions (i.e., set median), (b) extension strategy, (c) compressed strategy. The last two strategies are as follows:

1. *Extension strategy*: *1.1.* design of a 'kernel' for the initial solutions (substructure or an extended substructure), *1.2.* generation of some additional elements for possible inclusion into the 'kernel', *1.3.* selection of the additional elements while taking into account their 'profit' and resource requirements (e.g., cost) (here basic knapsack problem can be used).

2. *Compression strategy*: *2.1.* design a super structure for the initial solutions, *2.2.* generation of the superstructure elements as possible candidates for deletion, *2.3.* selection of the elements for deletion from the superstructure while taking into account their 'profit' resource requirements (e.g., cost) (here knapsack problem with minimization of the objective function can be used).

In the geological example, the extension aggregation strategy is used.

2.6. General framework

The suggested general framework is the following (Fig. 6):

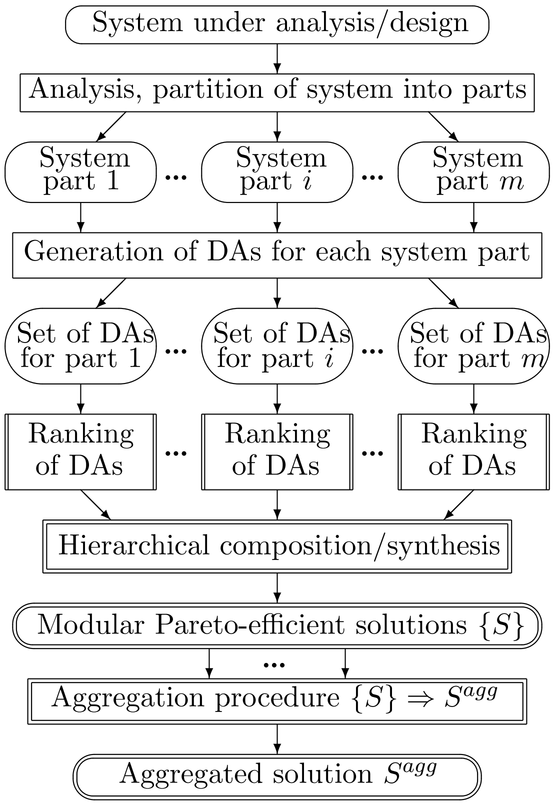

Fig. 6. General combinatorial framework

1. An analysis of the initial applied problem and a preliminary its structuring (for example, partitioning the problem into parts, generation of basic requirements/criteria).

2. Designing a special four-layer tree-like model (as a multi-layer model of geological objects): (i) productive stratum (reservoir), (ii) group of productive stratums (reservoirs)), (iii) oil and gas field, and (iv) group of oil and gas fields (region).

3. Generating a set of local design alternatives DAs (for geological exploration) for each bottom-layer geological object: (a) conservation, (b) appraisal work, (c) independent production time, (d) joint independent production time.

4. Multicriteria assessment of DAs and their interconnection (compatibility IC) and mapping the obtained vector estimates into an ordinal scale.

5. Hierarchical (bottom-up) composition of composite exploration plans for each field.

6. Integration of the obtained plans for the fields into region plans (solution versions).

7. Aggregation of the obtained region plans (solution versions) to obtain a total solution.

Stages 2, 3, 4, 5, 6 are based on hierarchical morphological multicriteria design (HMMD) method (ranking of DAs, selection and composition of DAs) [10–12]. Stage 7 consists in aggregation of the obtained modular solutions (detection of a 'kernel' of the obtained solutions and its extension by some additional solution elements) [12]. The stages above are based on expert judgment.

3. EXAMPLE OF GEOLOGICAL EXPLORATION

Combinatorial planning of geological exploration is examined for oil and gas fields of Yamal peninsula [16] (Fig. 7). The general plan involves five parts: $S = A^1 \star A^2 \star A^3 \star A^4 \star A^5$, where A^1 corresponds to field Kharosovey, A^2 corresponds to field Arkticheskoe, A^3 corresponds to field Neitinskoe, A^4 corresponds to Kruzensternskoe, A^5 corresponds to field Bovanenkovskoe.

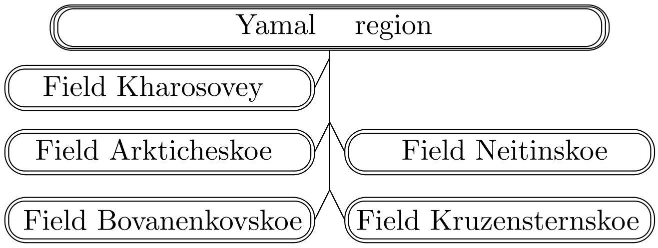

Fig. 7. Oil-gas fields (Yamal region)

The solving scheme consists of two stages:

Stage 1. Hierarchical combinatorial construction of the exploration plan for oil and gas fields (here only two oil and gas fields are described).

Stage 2. Composition of the general exploration plan for region.

3.1. Problem formulation

The following four-layer hierarchy of geological objects is considered: (a) productive stratum (reservoir) (bottom hierarchical level); (b) bore hole as a group of productive stratums (reservoirs); (c) oil and gas field; and (d) group of oil and gas fields (region). The following assessment parameters (attributes) are used:

1. parameter of reservoir existence ('3' corresponds to existence of reservoir, '2' corresponds to prospective geological position (horizon) in the field, '1' corresponds to prospective geological position (horizon) in the traprock);

2. cover of thickness, m;

3. type of fluid, i.e., classification factor: (i) gas, (ii) gas and condensate (condensed fluid), (iii) oil;

4. volume of geological reserves or resources (gas - million cubic meters, oil - thousand tonnes);

5. production rate of work wellsite (cubic metes per 24 hours);

6. complexity of geological situation ('1' - simple, '2' - complex, '3' - very complex);

7. reliability (risk) to obtain the results ([0...100]);

8. validity (adequacy) of assessment of geological reserves (i.e., oil/gas/ condensate in place, probable reserves) (C_1 - 20%, C_2 - 50%, C_3 - 80%, etc.);

9. proximity to technological base (gas-oil pipeline, km).

Eight DAs are examined for each geological object (as stratum) (Table 2) (the corresponding bottom index is used for the designation).

Table 2. Design alternatives (DAs) for geological exploration

	Notation	Content of geological exploration
1.	X_1	conservation
2.	X_2	appraisal work
3.	X_3	independent production time (gas)
4.	X_4	independent production time (oil)
5.	X_5	independent production time (oil and gas)
6.	X_6	joint independent production time (gas)
7.	X_7	joint independent production time (oil)
8.	X_8	joint independent production time (oil and gas)

Further, a subset of design alternative/actions (DAs) for each geological object (stratum) at bottom layer of the system model (i.e., productive stratum) is selected (from initial eight basic DAs) (expert judgment). This is a preliminary selection at the bottom layer of the problem. At the next step, the selected DAs are used as a basis for composition of composite DAs for more higher layer of the problem (i.e., for group of geological objects as bore holes, and for fields)

Each strategy component (geological object, group of objects, strategy) is noted by symbol (the level of effectiveness of priority ι is pointed out for each components in parentheses). It is assumed that experts have their skills for the following: (1) selection of DAs for each geological object, (2) ranking of DAs for each geological object, (3) assessment of compatibility among DAs (by an ordinal scale).

The illustrative hierarchical model of oil and gas field is depicted in Fig. 8.

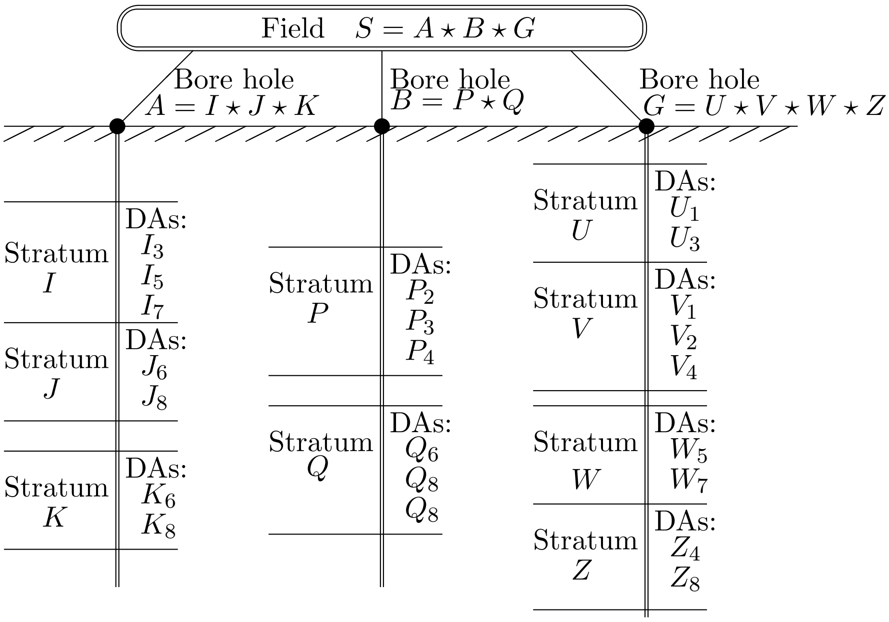

Fig. 8. Illustration for hierarchical model of field

3.2. Examples for oil-gas fields

The modular exploration strategy for field Arkticheskoe is shown in Fig. 9. Table 3 contains compatibility factors for the strategy elements (here 'C5+v' corresponds to level of hydrocarbon in gas as 'C5' and more).

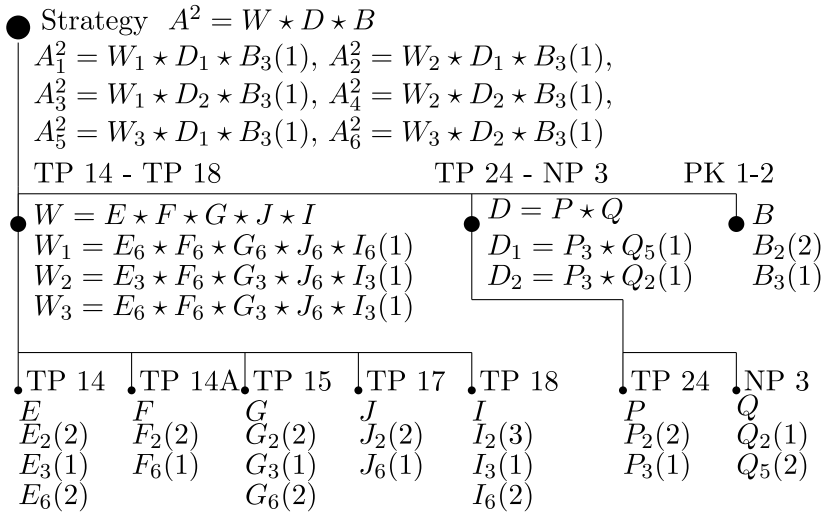

Fig. 9. Strategy for field Arkticheskoe

Table 3. Compatibility factors for DAs pair

	DA & DA	Factors
1.	TP 14 E & TP 14A F	Geological reserves, proximity, 'C5+v'
2.	TP 14 E & TP 15 G	Geological reserves, proximity, 'C5+v'
3.	TP 14 E & TP 17 J	Geological reserves, proximity
4.	TP 14 E & TP 18 I	Geological reserves, proximity
5.	TP 14A F & TP 15 G	Geological reserves, proximity
6.	TP 14A F & TP 17 J	Geological reserves, proximity
7.	TP 14A F & T 18 I	Proximity
8.	TP 15 G & TP 17 J	Geological reserves, proximity, 'C5+v'
9.	TP 15 G & TP 18 I	Geological reserves, proximity, 'C5+v'
10.	TP 17 J & TP 18 I	Geological reserves, proximity, 'C5+v'
11.	TP 24 P & NP 3 Q	Proximity, 'C5+v'

Compatibility estimates between DAs (expert judgment) are contained in Table 4 and Table 5.

Further, intermediate composite DAs for group of geological objects are obtained for TP 14 - TP 18 (W), TP 24 - NP3 (D) (Fig. 9, Table 6):

1. $W_1 = E_6 \star F_6 \star G_6 \star J_6 \star J_6$, $N(W_1) = (4; 2, 3, 0)$; $W_2 = E_3 \star F_6 \star G_3 \star J_6 \star J_3$, $N(W_2) = (2; 5, 0, 0)$; $W_3 = E_6 \star F_6 \star G_3 \star J_6 \star J_3$; $N(W_3) = (3; 4, 1, 0)$.

2. $D_1 = P_3 \star Q_5$, $N(D_1) = (4; 1, 1, 0)$; $D_2 = P_3 \star Q_2$, $N(D_2) = (3; 2, 0, 0)$.

Fig. 10 depicts quality of composite DAs for component W.

Thus, 6 versions of exploration strategy (field Arkticheskoe) are obtained:

$A^2_1 = W_1 \star D_1 \star B_3(1)$, $A^2_2 = W_2 \star D_1 \star B_3(1)$, $A^2_3 = W_1 \star D_2 \star B_3(1)$,

$A^2_4 = W_2 \star D_2 \star B_3(1)$, $A^2_5 = W_3 \star D_1 \star B_3(1)$, $A^2_6 = W_3 \star D_2 \star B_3(1)$.

Table 7 contains some examples of bottlenecks and possible improvement operations for intermediate composite DAs.

The exploration strategy for field Kruzensternskoe is shown in Fig. 11. Table 8 contains compatibility factors for strategy elements. The compatibility estimates among DAs (expert judgment) are presented in Table 9, Table 10. Composite DAs for B and H are presented in Table 11. The obtained two solutions for field Kruzensternskoe are: $A^4_1 = B_1 \star H_1$, $A^4_2 = B_2 \star H_2$. Fig. 12 illustrates quality of composite solutions for component H.

Table 4. Compatibility for DAs (groups TP 14 - TP 18, part W)

	F_2	F_6	G_2	G_3	G_6	J_2	J_6	I_2	I_3	I_6
E_2	2	3	2	3	4	1	0	2	3	3
E_3	3	4	2	4	4	1	2	3	2	2
E_6	3	4	3	4	4	1	4	3	3	4
F_2			2	3	4	2	3	2	3	2
F_6			1	3	4	3	3	2	3	1
G_2						2	3	2	1	3
G_3						2	4	2	3	1
G_6						2	4	2	3	1
J_2								2	1	1
J_6								1	3	4

Table 5. Compatibility for DAs (groups TP 24 - NP 3, part D)

	Q_2	Q_5
P_2	2	3
P_3	3	4

Table 6. Composite DAs

Intermediate composite DAs	N
$W_1 = E_6 \star F_6 \star G_6 \star J_6 \star I_6$	$4; 2, 3, 0$
$W_2 = E_3 \star F_6 \star G_3 \star J_6 \star I_3$	$2; 5, 0, 0$
$W_3 = E_6 \star F_6 \star G_3 \star J_6 \star I_3$	$3; 4, 1, 0$
$D_1 = P_3 \star Q_5$	$4; 1, 1, 0$
$D_2 = P_3 \star Q_2$	$3; 2, 0, 0$

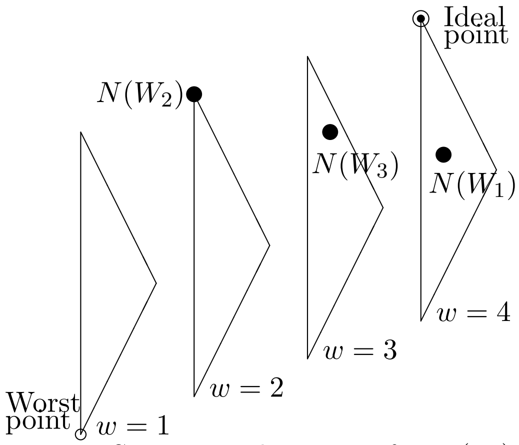

Fig. 10. System quality poset for $N(W)$

Table 7. Bottlenecks and improvement operations

	Intermediate composite DAs	Bottlenecks: DAs	IC	Improvement operation w/r
1.	$D_1 = P_3 \star Q_5$	Q_5		$2 \Rightarrow 1$
2.	$D_2 = P_3 \star Q_2$		(P_3, Q_2)	$3 \Rightarrow 4$
3.	$W_1 = E_6 \star F_6 \star G_6 \star J_6 \star I_6$	E_6		$2 \Rightarrow 1$
4.	$W_1 = E_6 \star F_6 \star G_6 \star J_6 \star I_6$	G_6		$2 \Rightarrow 1$
5.	$W_1 = E_6 \star F_6 \star G_6 \star J_6 \star I_6$	I_6		$2 \Rightarrow 1$
6.	$W_3 = E_6 \star F_6 \star G_3 \star J_6 \star I_3$	E_6		$2 \Rightarrow 1$

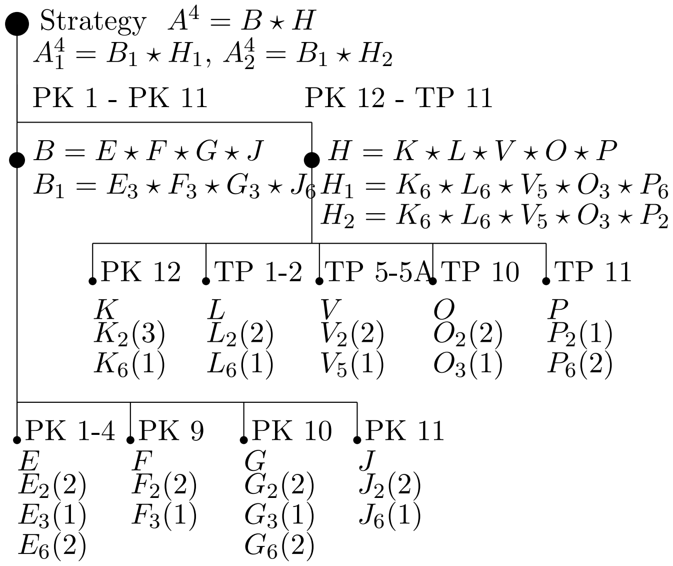

Fig. 11. Strategy for field Kruzensternskoe

Table 8. Compatibility factors for DAs pair (part B)

	DA & DA	Factors
1.	PK 1-4 E & PK 9 F	Geological reserves, proximity
2.	PK 1-4 E & PK 10 G	Geological reserves, proximity
3.	PK 1-4 E& PK 11 J	Geological reserves, proximity
4.	PK 9 F & PK 10 G	Geological reserves, proximity
5.	PK 9 F & PK 11 J	Geological reserves, proximity
6.	PK 10 G & PK 11 J	Geological reserves, proximity
7.	PK 12 K & TP 1-2 L	Geological reserves, proximity
8.	PK 12 K & TP 5-5A V	Geological reserves, proximity
9.	PK 12 K & TP 10 O	Geological reserves, proximity
10.	PK 12 K & TP 11 P	Proximity, 'C5+v'
11.	TP 1-2 L & TP 5-5A V	Geological reserves, proximity, 'C5+v'
12.	TP 1-2 L & TP 10 O	Geological reserves, proximity, 'C5+v'
13.	TP 1-2 L & TP 11 P	Geological reserves, proximity, 'C5+v'
14.	TP 5-5A V & TP 10 O	Geological reserves, proximity, 'C5+v'
15.	TP 5-5A V & TP 11 P	Geological reserves, proximity, 'C5+v'
16.	TP 10 O & TP 11 P	Geological reserves, proximity, 'C5+v'

Table 9. Compatibility for DAs (groups PK 1 - PK 11, part B)

	F_2	F_3	G_2	G_3	G_6	J_2	J_3	J_6
E_2	2	1	2	1	2	2	1	2
E_3	4	3	4	3	1	4	3	1
E_6	1	4	1	4	2	1	3	4
F_2			3	4	2	3	4	2
F_3			3	4	4	3	4	4
G_2						3	4	4
G_3						4	4	4
G_6						3	4	4

Table 10. Compatibility for DAs (groups PK 12 - TP 11, part H)

	L_2	L_6	V_2	V_5	O_2	O_3	P_2	P_6
K_2	4	3	2	4	3	1	4	3
K_6	1	4	3	4	3	4	3	4
L_2			2	3	4	2	3	4
L_6			2	4	4	4	3	4
V_2					4	4	2	3
V_5					3	4	3	4
O_2							4	4
O_3							3	4

Table 11. Intermediate composite DAs

Composite DAs	N
$B_1 = E_3 \star F_3 \star G_3 \star J_3$	$3; 4, 0, 0$
$H_1 = K_6 \star L_6 \star V_5 \star O_3 \star P_6$	$4; 4, 1, 0$
$H_2 = K_6 \star L_6 \star V_5 \star O_3 \star P_2$	$3; 5, 0, 0$

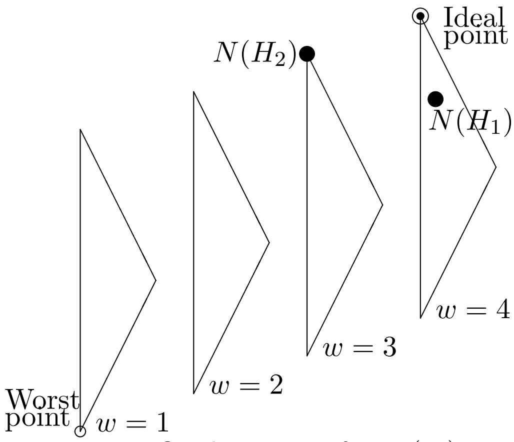

Fig. 12. Quality poset for $N(H)$

3.3. Exploration plan for region

Thus, the following composite exploration strategy for region is obtained (Fig. 13):

0. General composite strategy $S = A^1 \star A^2 \star A^3 \star A^4 \star A^5$

- **1.** Strategy for oil-gas field Kharosovey: A^1_1.
- **2.** Strategies for oil-gas field Arkticheskoe: A^2_1, A^2_2, A^2_3, A^2_4, A^2_5, A^2_6.
- **3.** Strategy for oil-gas field Neitinskoe: A^3_1.
- **4.** Strategies for oil-gas field Kruzensternskoe: A^4_1, A^4_2.
- **5.** Strategies for oil-gas field Bovanenkovskoe: A^5_1, A^5_2.

Finally, 24 composite exploration strategies for the region are (without compatibility analysis):
$S_1 = A^1_1 \star A^2_1 \star A^3_1 \star A^4_1 \star A^5_1$, $S_2 = A^1_1 \star A^2_1 \star A^3_1 \star A^4_2 \star A^5_1$,
$S_3 = A^1_1 \star A^2_1 \star A^3_2 \star A^4_1 \star A^5_1$, $S_4 = A^1_1 \star A^2_1 \star A^3_2 \star A^4_2 \star A^5_1$,
$S_5 = A^1_1 \star A^2_2 \star A^3_1 \star A^4_1 \star A^5_1$, $S_6 = A^1_1 \star A^2_2 \star A^3_1 \star A^4_2 \star A^5_1$,
$S_7 = A^1_1 \star A^2_2 \star A^3_2 \star A^4_1 \star A^5_1$, $S_8 = A^1_1 \star A^2_2 \star A^3_2 \star A^4_2 \star A^5_1$,
$S_9 = A^1_1 \star A^2_3 \star A^3_1 \star A^4_1 \star A^5_1$, $S_{10} = A^1_1 \star A^2_3 \star A^3_1 \star A^4_2 \star A^5_1$,
$S_{11} = A^1_1 \star A^2_3 \star A^3_2 \star A^4_1 \star A^5_1$, $S_{12} = A^1_1 \star A^2_3 \star A^3_2 \star A^4_2 \star A^5_1$,

$S_{13} = A^1_1 \star A^2_4 \star A^3_1 \star A^4_1 \star A^5_1$, $S_{14} = A^1_1 \star A^2_4 \star A^3_1 \star A^4_2 \star A^5_1$,
$S_{15} = A^1_1 \star A^2_4 \star A^3_2 \star A^4_1 \star A^5_1$, $S_{16} = A^1_1 \star A^2_4 \star A^3_2 \star A^4_2 \star A^5_1$,
$S_{17} = A^1_1 \star A^2_5 \star A^3_1 \star A^4_1 \star A^5_1$, $S_{18} = A^1_1 \star A^2_5 \star A^3_1 \star A^4_2 \star A^5_1$,
$S_{19} = A^1_1 \star A^2_5 \star A^3_2 \star A^4_1 \star A^5_1$, $S_{20} = A^1_1 \star A^2_5 \star A^3_2 \star A^4_2 \star A^5_1$,
$S_{21} = A^1_1 \star A^2_6 \star A^3_1 \star A^4_1 \star A^5_1$, $S_{22} = A^1_1 \star A^2_6 \star A^3_1 \star A^4_2 \star A^5_1$,
$S_{23} = A^1_1 \star A^2_6 \star A^3_2 \star A^4_1 \star A^5_1$, $S_{24} = A^1_1 \star A^2_6 \star A^3_2 \star A^4_2 \star A^5_1$.

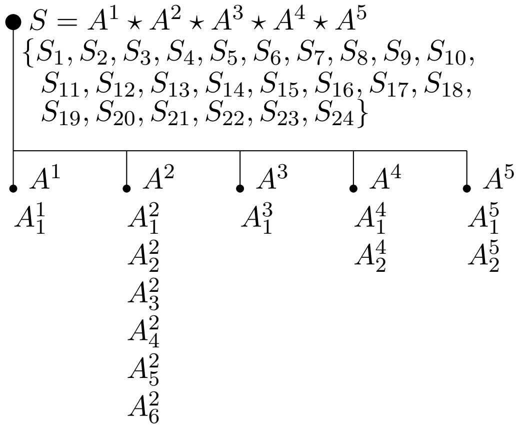

Fig. 13. Composite strategy for region

Now an additional analysis of the obtained strategies can be considered to design the best final strategy (e.g., multicriteria analysis and selection, expert judgment). On the other hand, the final strategy can be build by aggregation of the obtained solutions.

3.4. Aggregation of solutions

In the considered example, there are 24 solutions (previous section): S_1,...,S_{24}. The substructure of the solutions is shown in Fig. 14. This structure is used as a 'kernel' for an extension process. The superstructure is shown in Fig. 15.

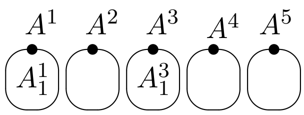

Fig. 14. Substructure ('kernel')

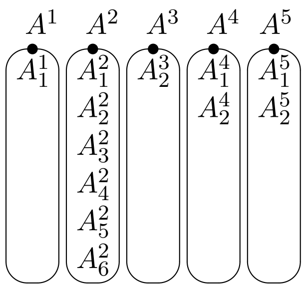

Fig. 15. Superstructure

Table 12 contains design alternatives for extension of the 'kernel' including their estimates (ordinal scales are used, expert judgment).

Table 12. Extension versions

κ	Versions	Binary variable	Cost a_{ij}	Profit c_{ij}
1.	A^2_1	x_{11}	4	4
2.	A^2_2	x_{12}	6	6
3.	A^2_3	x_{13}	3	2
4.	A^2_4	x_{14}	3	3
5.	A^2_5	x_{15}	4	3
6.	A^2_6	x_{16}	5	3
7.	A^4_1	x_{21}	3	4
8.	A^4_2	x_{22}	3	3
9.	A^5_1	x_{31}	3	3
10.	A^5_2	x_{32}	4	4

It is assumed, the DAs are compatible. The aggregation problem (extension strategy) is based on multiple choice problem:

$$\max \sum_{i=1}^{3} \sum_{j=1}^{q_i} c_{ij} x_{ij}$$

$$s.t. \ \sum_{i=1}^{3} \sum_{j=1}^{q_i} a_{ij} x_{ij} \leq b, \quad \sum_{j=1}^{q_i} x_{ij} = 1 \ \ \forall i = \overline{1,3}, \quad x_{ij} \in \{0,1\}.$$

In this model, $q_1 = 6$, $q_2 = 2$, $q_3 = 2$. By the usage of a greedy algorithm (i.e., linear ordering of elements by c_i/a_i) the following solutions are obtained for four versions of constraints:

(1) $b^1 = 9$: $(x_{14} = 10,\ x_{21} = 1,\ x_{31} = 1)$,
$S^{agg}_{b^1} = A^1_1 \star A^2_4 \star A^3_1 \star A^4_1 \star A^5_1 = R_3 \star P_3 \star D_2 \star Q_4 \star U_1 \star Z_1 \star Y_2 \star O_1$;
(2.1.) $b^2 = 10$: $(x_{14} = 1,\ x_{21} = 1,\ x_{32} = 1)$,
$S^{agg1}_{b^2} = A^1_1 \star A^2_4 \star A^3_1 \star A^4_1 \star A^5_2 = R_3 \star P_3 \star D_2 \star Q_4 \star U_1 \star Z_1 \star Y_2 \star O_1$;
(2.2.) $b^2 = 10$: $(x_{11} = 1,\ x_{21} = 1,\ x_{31} = 1)$,
$S^{agg2}_{b^2} = A^1_1 \star A^2_1 \star A^3_1 \star A^4_1 \star A^5_1 = R_3 \star P_3 \star D_2 \star Q_4 \star U_1 \star Z_1 \star Y_2 \star O_1$;
(3) $b^3 = 11$: $(x_{11} = 1,\ x_{21} = 1,\ x_{32} = 1)$,
$S^{agg}_{b^3} = A^1_1 \star A^2_1 \star A^3_1 \star A^4_1 \star A^5_2 = R_4 \star P_3 \star D_2 \star Q_4 \star U_1 \star Z_1 \star Y_2 \star O_1$;
(4) $b^4 = 11$: $(x_{12} = 1,\ x_{22} = 1,\ x_{31} = 1)$,
$S^{agg}_{b^4} = A^1_1 \star A^2_2 \star A^3_1 \star A^4_2 \star A^5_1 = R_4 \star P_3 \star D_2 \star Q_4 \star U_1 \star Z_1 \star Y_2 \star O_1$.

3.5. Example of multiset estimates based synthesis

A scale based on multiset estimates (as a poset) for the used assessment problem $P^{3,4}$ is depicted in Fig. 16. The illustrative numerical example is based on multiset estimates for Arkticheskoe oil-gas field (Fig. 17). Multiset estimates for local DAs are shown in Fig. 17 (in parentheses). Compatibility estimates from Table 4 are used. Two solutions are considered:

(i) $W^M_1 = E_6 \star F_6 \star G_6 \star J_6 \star I_6$, $\quad N(W^M_1) = (w(W^M_1); e(W^M_1)) = (4; 1, 3, 0)$;
(ii) $W^M_2 = E_6 \star F_6 \star G_3 \star J_6 \star I_3$, $\quad N(W^M_2) = (w(W^M_2); e(W^M_2)) = (3; 3, 1, 0)$.

Estimates $e(W^M_1) = (1, 3, 0)$, $e(W^M_2) = (3, 1, 0)$ are medians for estimates of the corresponding components.

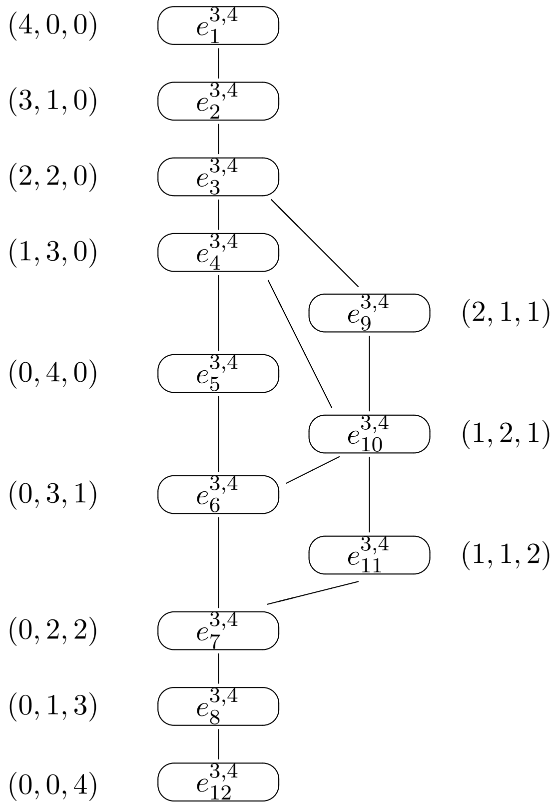

Fig. 16. Estimates for assessment problem $P^{3,4}$

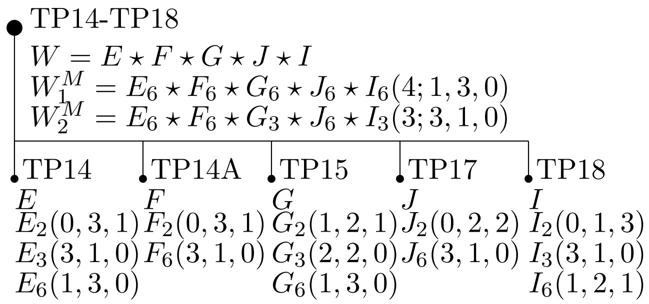

Fig. 17. Arkticheskoe oil-gas field (multiset esitmates)

4. CONCLUSION

In this paper, hierarchical combinatorial planning of geological exploration is described. The approach is based on the following: (a) expert judgment; (b) combinatorial synthesis as bottom-up selection and composition of local solutions (design/exploration alternatives DAs) into composite solutions at the higher layer of the plan hierarchy; and (c) aggregation of the obtained plans (solution versions). The approach is illustrated by a realistic numerical example for oil and gas geological planning (Yamal peninsula). It may be reasonable to consider the following future directions: (1) examination of multistage exploration strategies; (2) study of combinatorial evolution models for oil and gas field(s); (3) consideration of the considered planning approach in other domains, and (4) using the suggested framework in education.

Note the material of the article can be used as a short tutorial.

The author states that there is no conflict of interest.

REFERENCES

1. Ayres R.U., Technological Forecasting and Long-time Planning. New York: McGraw-Hill, 1969.
2. Cormen T.H., Leiserson C.E., Rivest R.L., Introduction to Algorithms. 3rd ed., Cambridge, MA: MIT Press, 2009.
3. Garey M.R., Johnson D.S., Computers and Intractability. San-Francisco: W.H. Freeman, 1979.
4. Jones J.C., Design Methods. 2nd ed., New York: Wiley, 1992.
5. Keeney R.L., Raiffa H., Decisions with Multiple Objectives: Preferences and Value Trade-offs. London: Wiley, 1976.
6. Kellerer H., Pferschy U., Pisinger D., Knapsack Problems. Berlin: Springer, 2004.
7. Klamroth K., Wiecek M., Dynamic programming approaches to the multiple criteria knapsack problem. Nav. Res. Log., 2000, vol. 47, pp. 57–76.
8. Knuth D.E., The Art of Computer Programming. Vol. 2, Seminumerical Algorithms. Addison Wesley, Reading, 1998.
9. Koltun A.A., Pershin O.Y., Ponomarev A.M., Models and algorithms for the selection of optimal set of geotechnical arrangements on the oil fields. Autom. Rem. Contr., 2005, vol. 66, no. 8, pp. 35–45.
10. Levin M.S., Combinatorial Engineering of Decomposable Systems. New York: Springer, 1998.
11. Levin M.S., Composite Systems Decisions. London: Springer, 2006.
12. Levin M.S., Modular System Design and Evaluation. Cham: Springer, 2015.
13. Levin M.S., Combinatorial framework for planning in geological exploration. Electr prepr, 14 p, Jan. 22, 2018. http://arxiv.org/abs/1801.07229 [cs.AI]
14. Levin M.S., Poroskun V.I., Combinatorial planning of oil-gas exploration. In: Geological exploration and utilization of natural resources. Issue 4, Moscow, Ministry of Nature Resources, Private Company 'Geoinformmark', 1997, pp 58–63 (in Russian).
15. Martello S., Toth P., Knapsack Problem: Algorithms and Computer Implementation. New York: Wiley, 1990.
16. Maximov S.P. (ed), Oil-gas Field in USSR. Handbook. Moscow: Nedra Publ. Hause, 1979 (in Russian).
17. Ritchey T., Wicked Problems/Social Messes: Decision Support Modelling with Morphological Analysis. Berlin: Springer, 2012.
18. Roy B., Multicriteria Methodology for Decision Aiding. Dorthrecht: Kluwer Academic Publishers 1996.
19. Saaty T.L., The Analytic Hierarchy Process, New York: MacGraw-Hill, 1988.
20. Yager R.R., On the theory of bags. Int. J. General Systems, 1986, pp. 23–37.
21. Zopounidis C., Doumpos M., Multicriteria classification and sorting methods: A literature review. Eur. J. Oper. Res., 2002, vol. 138, no. 2, pp. 229–246.
22. Zwicky F., Discovery Invention, Research Through the Morphological Approach. New York: McMillan, 1969.