Logic and Logical Philosophy
Volume 20 (2011), 111–137
DOI: 10.12775/LLP.2011.006

Victor K. Finn
Maria A. Mikheyenkova

PLAUSIBLE REASONING FOR THE PROBLEMS OF COGNITIVE SOCIOLOGY

Abstract. The plausible reasoning class (called the JSM-reasoning in honour of John Stuart Mill) is described. It implements interaction of three forms of non-deductive procedures — induction, analogy and abduction. Empirical induction in the JSM-reasoning is the basis for generation of hypotheses on causal relations (determinants of social behaviour). Inference by analogy means that predictions about previously unknown properties of objects (individual's behaviour) are inferred from causal relations. Abductive inference is performed to check on the explanatory adequacy of generated hypotheses. To recognize rationality of respondents' opinion deductive inference is used. Plausible reasoning, semantics of argumentation logic and deductive recognition of opinion rationality represent logical tool for cognitive sociology problems.

Keywords: plausible reasoning, induction, analogy, abduction, knowledge discovery, reasoning about causality.

1. Introduction

The actual problem of knowledge extraction from unordered and ill-formalized data is solved by different methods that form a widely developing branch of Artificial Intelligence — Intelligent Data Analysis (IDA) [23]. Several problems are solved in IDA frames. We need subject domain choice and formalized heuristics for problem solving in these domains. The heuristics formalization is possible if formal languages and logical means for reasoning formalization are constructed. These automated formal reasoning form instruments for IDA problems solving in

Special Issue: *Logic in Cognitive Science*. Guest Editors: J. Malinowski and R. Palczewski
 ISSN: 1425-3305

corresponding domains (models). IDA methods should extract samples and causality (in data mining sense) as far as should form fragments of theories using empirical data. This activity is nothing but a cognition process in different domains. As a result, we have the Intelligent Systems realizing these functions.

It seems perspective to employ logical (logical-combinatorial) methods for sciences where the problem of structurization, ordering and systematization of initial data and regularities extraction from the facts are yet unsolved. Life sciences and sciences on social behaviour are the examples of such disciplines. Statistical analysis and inference, being quantitative and numerical in heart, proved to be insufficient for many theoretical and practical problems in sociology. The situation has led to wide spread of qualitative methods which are often in contrast with statistical ones in sociology.

To objectify particular kinds of qualitative analysis' results some formal instruments have been developed (see [21], for example). An impressive review [14] can be considered as a guide to computer instruments integrating qualitative and quantitative techniques. But most part of the approaches solves local problems of particular researches and is not intended to implement the general scheme of sociologist's cognitive activity "data analysis – forecast – explanation" [10].

Grounded theory [26] — the established methodology of qualitative analysis in sociology — is considered to be the concentrated demonstration of cognitive approach to sociological research. The theory is created on the basis of facts by inductive analysis up to saturation sampling. To objectify this process formalized heuristics are needed. Formalized heuristics allow to discover knowledge from unordered and unformalized empirical data systematically. Such heuristics creation seems to be the approach to the fundamental problem of cognitive sociology solving.

There are two directions in cognitive sociology (in the narrow sense). The first one investigates socially caused features of individual and collective thinking, information perception processes, social aspects of decision making [29], studies thinking processes activated in the appreciation of the question and in answering — in sociological polls [27]. The second one, on the contrary, considers cognitive features of social behaviour, studies cognitive activity influence on the behaviour. We will be interested in both formalization of cognitive process in sociology and studying the social cognitivity effects — behaviour motivation, opinion rationality, concept formation and so on.

2. The JSM-Reasoning as the Synthesis of Cognitive Procedures

The JSM Method of Automated Generation of Hypotheses (JSM-MAGH) is the tool for Intelligent Data Analysis, which formalizes special class of plausible reasoning for knowledge discovery in special Data Bases [11]. The proposed technique represents class of formal heuristics — a synthesis of cognitive procedures: empirical induction (formal generalization of John S. Mill's Methods of Causal Reasoning [17]), causal analogy and Charles S. Peirce's abduction [18] (acceptance or rejection of hypotheses by an explanation of initial data). This synthesis (called JSM-reasoning) is formalized by specially created many-valued logic with two external truth values t and f and a countably infinite set of internal truth values of four types, namely $+1$, -1, 0, and τ (see below) [9]. The JSM-reasoning realizes natural heuristics (which is the tool of formalized qualitative Intelligent Data Analysis) and proves to be an adequate instrument for formalization of reasoning in social behaviour sciences [15] and life sciences [5].

The synthesis of procedures is the operational definition of "cause – effect" relation corresponding to knowledge about some subject domain and its entities and problems. The relation that is defined by positive (+) and negative (−) hypotheses on causal dependence (in contrast with conditional proposition "if p, then q") is generated from initial database ("training sample" in the theory of machine learning). Then the mentioned relation is explained by abduction and the JSM-causality relation is formed. The idea of causality in JSM-reasoning is based upon the principle of structuralism: hypothetical causes (of the properties presence in an object) — as propensity in Karl R. Popper's sense — are the result of structured facts similarities.

We suppose the knowledge in empirical subject domains (that deals with the open world, progressively replenished with new facts) to be represented in the form of quasiaxiomatic theory [9] $QAT = \langle \Sigma, \Sigma', R \rangle$. Here Σ is an open axiom set, describing a subject domain (SD) incompletely. Σ corresponds to knowledge base KB and contains "core axioms" — basic axioms that coincide in all QAT — and "specific axioms", encoding subject domain under investigation. The empirical data about individual SD objects is contained in an open set Σ', corresponding to base of facts and hypotheses generated by R. The rules of inference R of QAT are divided into rules of reliable (that is, deductive) inference R_d and rules of plausible (that is, non-deductive) inference (RPI) R_p,

$R = R_d \cup R_p$. If empirical information about the particular objects of the subject domain has subjective representation (in sociology, for example), QAT is supplemented with argumentation scheme to convert the data into some sort of objective representation.

The general JSM-MAGH features – synthesis of cognitive procedures, knowledge representation as an open theory QAT, constructive generation (by RPI's) of truth values assignment, taking into account both positive and negative causes in hypotheses acceptance for sufficient argumentation, – allow the method to be considered as the adequate formal instrument for qualitative analysis in sociology. This statement has found confirmation in some studying.

3. The Formal Tools of the JSM Method

Let us now characterize language L for JSM-reasoning [2]. The language L has the expressive force of the language of weak second order predicate logic [4]:

- variables for natural numbers: m, n, k, $\ldots$,
- variables for objects and sub-objects: X, Z, V, $\ldots$ (maybe with subscripts); constants C, C_1, C_2, $\ldots$,
- variables for sets of properties: Y, U, W, $\ldots$ (maybe with subscripts); constants A, Q, A_1, Q_1, A_2, Q_2, $\ldots$,
- predicative symbols: $\Rightarrow_1$, $\Rightarrow_2$, ${}_3\!\Leftarrow$, $\subseteq$ (for the Boolean structure of data), $=$,
- functional symbols (for the Boolean structure of data): $\cap$, $\cup$, $-$, $\varnothing$.
- types of truth values: 1, -1, 0, τ (factual truth, factual falsity, factual contradiction ("conflict"), uncertainty);

Truth values of the language L formulae are formed by the types of truth values and numbers $n = 0, 1, 2, \ldots$. Here n, the number of application of plausible inference rules, expresses degree of plausibility of hypotheses generated by JSM-reasoning. Truth values $\bar{\nu} = \langle \nu, n \rangle$, where $\nu \in \{1, -1, 0\}$, and the set of truth values (τ, n) are valuations of facts and hypotheses (if $n = 0$ or $n > 0$, respectively). $\bar{\nu}$ and (τ, n) are "internal" (factual) valuations, and t, f – "external" (logical) valuations: true, false. The set (τ, n) is the set of truth values for uncertain valuation of propositions. It is characterized by recurrent expression $(\tau, n) = \{\langle 1, n+1 \rangle, \langle -1, n+1 \rangle, \langle 0, n+1 \rangle\} \cup (\tau, n+1)$.[1]

[1] This expression is not a definition, it expresses the possibility of uncertainty

So, we use an infinitely-valued logic with finite number of truth value types [1]:

- logical connectives and quantifiers of two-valued logic: $\neg$, $\&$, $\vee$, $\to$, $\forall$, $\exists$;
- unary logical connectives of infinitely-valued logic: $J_{\bar{\nu}}$, where

$$J_{\bar{\nu}}\varphi = \begin{cases} t & \text{if } v[\varphi] = \bar{\nu} \\ f & \text{if } v[\varphi] \neq \bar{\nu} \end{cases} \qquad \nu \in \{1, -1, 0, \tau\},$$

  v is the valuation function, the operator $J_{\bar{\nu}}$ is Rosser-Turquette J-operator [22] (taking into account the difference between "internal" "and external" valuations).

To formalize induction (with initially unknown number of examples) the quantifiers $\forall$, $\exists$ over tuples of variable length have been introduced to L [24], so L is the language of weak second order predicate logic (first order predicate logic for finite models [28]). One can find detailed description of the JSM Method of Automated Generation of Hypotheses in [9] (see brief review, for example, in [7]).

We consider the predicate $X \Rightarrow_1 Y$ — "the object X possesses the set of properties" — to be the primitive predicate of the JSM Method. Initial base of facts contains empirical information — does the object have the property involved or not, or there is no information about the object's property at all. In accordance with this information, the proposition "the object C has the set of properties A" is assigned the truth value $\langle\nu, n\rangle$ or (τ, n). Here $\nu \in \{1, -1, 0\}$, $1, -1, 0$ are the types of "internal" truth values. n is a number of JSM-reasoning step (the process is iterative). In terms of JSM-language the mentioned proposition has the form $J_{\langle\nu,\, n\rangle}\,(C \Rightarrow_1 A)$ or $J_{(\tau,\, n)}\,(C \Rightarrow_1 A)$, where $J_{\langle\nu,\, n\rangle}\varphi$ is the operator (defined above), and $J_{(\nu,\, n)}\varphi \rightleftharpoons \bigvee\limits_{i=1}^{n} J_{\langle\nu,i\rangle}\varphi$.

$\langle\nu, n\rangle$ represents "internal" truth values for empirical facts, t, f are "external" truth values of two-valued logic for representation of facts with valuation and RPI's. After JSM-procedures were used, propositions of the form $J_{\langle\nu,\, n\rangle}(C' \Rightarrow_2 A)$ or $J_{\langle\nu,\, n\rangle}(Q \;{}_3\!\Leftarrow C')$, $(n > 0)$ are generated (depending on the type of inference: direct or inverse, respectively — see below). The first means that proposition "the subobject C' is a cause

reduction during the process of hypotheses (with truth-values $\bar{\nu} = \langle\nu, n+1\rangle$, $\nu \in \{1, -1, 0\}$) generation by JSM-reasoning. Let's note, that the degree of hypotheses plausibility decreases when the number of JSM-reasoning application increases.

of the presence of the set of properties A" has the truth value type ν at the n-th step of JSM-reasoning. So, the predicate $V \Rightarrow_2 W$ represents a causal relation "V causes W" and describes the fragment of the knowledge base. The second means that proposition "the opinion Q is the consequence of differential characteristics C' of the subject" has the truth value $\langle \nu, n \rangle$. So, the predicate $W \ {}_3{\Leftarrow}\ V$ is interpreted as "the opinion W is the consequence of the subject's differential characteristics V".

Propositions $J_{\langle \nu, 0 \rangle}(C \Rightarrow_1 A)$ are facts, $J_{\langle \nu,\, n \rangle}(C \Rightarrow_j A)$ $(j = 1, 2)$ and $J_{\langle \nu,\, n \rangle}(Q \ {}_3{\Leftarrow}\ C)$, $n > 0$, are hypotheses.

Let $\mathbf{U}^{(1)} = \{d_1, ..., d_{r_1}\}$, $\mathbf{U}^{(2)} = \{a_1, ..., a_{r_2}\}$ be two sets. We define two Boolean algebras on them, $B_i = \langle 2^{\mathbf{U}^{(i)}}, \varnothing, -, \cap, \cup \rangle$, $i = 1, 2$, B_1 — algebra of objects and subobjects, B_2 — algebra of properties. Variables and constants of sorts 1 and 2 — objects $X \in 2^{\mathbf{U}^{(1)}}$ and sets of properties $Y \in 2^{\mathbf{U}^{(2)}}$, respectively, — are defined in a standard manner (see, for example, [9]). JSM-semantics for analysis and prediction of social behaviour is represented by algebra of subjects of behaviour (individuals) B_1 and algebra of behavioural actions, dispositions or opinions B_2. According to the our assumption — postulate of behaviour — $\mathbf{U}^{(1)}$ contains differential characteristics describing the individual traits of subjects, their social characteristics and biographical data. Let's underline, that Boolean structure is not believed to be the only possible.

3.1. The Induction

As mentioned above, the first stage of JSM-heuristics realizes the inductive procedure (the method of agreement [17], for example; in JSM-reasoning the term "similarity" is used instead of "agreement") of extracting causal dependencies from facts in J. S. Mill's style. J. S. Mill's inductive methods were in effect expressing the idea of knowledge discovery, which is being used in contemporary Intelligent Systems (IS). Mill's induction does not put forward any claim to formulation and justification of universal theories, but is only a heuristic means of extracting new knowledge from empirical data.

To formalize the inductive reasoning (that is the analysis and comparison of facts $J_{\langle \nu, n \rangle}(C \Rightarrow_1 A)$, $C \in 2^{\mathbf{U}^{(1)}}$ and $A \in 2^{\mathbf{U}^{(2)}}$) the predicates $\widetilde{M}^{\,\sigma}_{a,\,n}(V, W, k)$ $(\sigma \in \{+, -\})$ have been constructed [9]. They are used in the 1-st kind rules RPI-I to search for causes.

The predicates $\widetilde{M}^{\sigma}_{a,n}(V, W, k)$, depending on k (the number of similar examples) parameter, formalize direct JSM-reasoning, establishing causal relation of the type "similarity of the subjects implies similarity of behavioural actions (dispositions) of the subjects" ($\Rightarrow_2$). The method is based on the supposition that informational saturation of data on the subjects is greater than of data on their behaviour. But if the problem of opinion analysis is studied, the contrary supposition — that characterization of opinion exceeds the knowledge about the speaking subject — is actual. In this case, we need formalization of reasoning, establishing causal relation of the type "similarity of the subjects' opinion is a consequence of similarity of the subjects themselves" (${}_3\Leftarrow$). Based on this approach, the predicates $\widetilde{\breve{M}}^{\sigma}_{a,n}(V, W, k)$ ($\sigma \in \{+, -\}$) of inverse JSM Method are formulated. These predicates include subformulae that are analogous to subformulae of direct predicates with corresponding changes. They express the following conditions.

(1) The positive inverse predicate of agreement $\widetilde{\breve{M}}^{+}_{a,n}(V, W, k)$ reveals local similarity on (+)-examples $J_{(1,n)}(X_i \Rightarrow_1 Y_i)$, $i = 1, \ldots, k$, where $k \geqslant 2$ is variable, induction parameter, $X_i \in 2^{\mathbf{U}^{(1)}}$, $Y_i \in 2^{\mathbf{U}^{(2)}}$;

(2) The local similarity of the objects (for the Boolean data structure involved) is expressed by variable length subformula

$$(\bigcap_{i=1}^{k} X_i = V) \,\&\, (V \neq \varnothing) \,\&\, (\bigcap_{i=1}^{k} X_i = W) \,\&\, (W \neq \varnothing);$$

(3) Empirical regularity, that characterizes the predicted (+)-causal relation between V and W, is expressed by subformula

$$\forall X \forall Y((J_{(1,n)}(X \Rightarrow_1 Y) \,\&\, (W \subseteq Y)) \to ((V \subset X) \,\&\, (V \neq \varnothing) \,\&\, (\bigvee_{h=1}^{k}(Y = Y_h))))$$

with exhaustibility condition $\bigvee_{h=1}^{k}(Y = Y_h)$ (it means that we consider all appropriate examples).

$\breve{M}^{+}_{a,n}(V, W) \rightleftharpoons \exists k \widetilde{\breve{M}}^{+}_{a,n}(V, W, k)$. The predicates $\widetilde{\breve{M}}^{-}_{a,n}(V, W, k)$ for negative examples analysis are formulated symmetrically.

The formal definition of the predicate of agreement of inverse JSM Method is the following.

$$\widetilde{\breve{M}}^{+}_{a,n}(V, W, k) \rightleftharpoons \exists X_1 \ldots \exists X_k \exists Y_1 \ldots \exists Y_k((\mathop{\&}_{h=1}^{k} J_{(1,n)}(X_h \Longrightarrow_1 Y_h)) \,\&$$

$$(\bigcap_{h=1}^{k} X_h = V) \mathbin{\&} (V \neq \varnothing) \mathbin{\&} (\bigcap_{h=1}^{k} Y_h = W) \mathbin{\&} (W \neq \varnothing) \mathbin{\&}$$
$$\forall i \forall j ((i \neq j) \& (1 \leqslant i, j \leqslant k) \to (X_i \neq X_j)) \mathbin{\&}$$
$$\forall X \forall Y ((J_{(1,n)}(X \Rightarrow_1 Y) \mathbin{\&} (W \subseteq Y)) \to ((V \subset X) \mathbin{\&}$$
$$(V \neq \varnothing) \mathbin{\&} (\bigvee_{h=1}^{k} (Y = Y_h)))) \mathbin{\&} k \geqslant 2).$$

As a result of the rules of plausible inference of the 1-st type (RPI-I) hypotheses of the form $J_{\langle \nu,\, n\rangle}(Q' \,{}_3{\Leftarrow}\, C')$, $n > 0$, are generated. The rules of plausible inference of the 1-st type look as follows:

$$(I)_n^+ \qquad \frac{J_{(\tau,n)}(W \,{}_3{\Leftarrow}\, V), \quad \breve{M}^+_{a,n}(V,W) \mathbin{\&} \neg\, \breve{M}^-_{a,n}(V,W)}{J_{\langle 1,n+1\rangle}(W \,{}_3{\Leftarrow}\, V)}$$

where $J_{(\tau,n)}(W \,{}_3{\Leftarrow}\, V)$ — premise representing uncertainty $(W \,{}_3{\Leftarrow}\, V)$, $J_{\langle 1,n+1\rangle}(W \,{}_3{\Leftarrow}\, V)$ — consequence, which is the hypothesis on the (+) -cause (W is the consequence of V) with the truth value $\langle 1, n+1\rangle$. The rules $(I)_n^-$, $(I)_n^0$ and $(I)_n^\tau$ are formulated analogously: for consequences $J_{\langle -1,n+1\rangle}(W_3 \Leftarrow V)$, $J_{\langle 0,n+1\rangle}(W \,{}_3{\Leftarrow}\, V)$ and $J_{(\tau,n+1)}(W \,{}_3{\Leftarrow}\, V)$ and premises $\neg\, \breve{M}^+_{a,n}(V,W) \mathbin{\&} \breve{M}^-_{a,n}(V,W)$, $\breve{M}^+_{a,n}(V,W) \mathbin{\&} \breve{M}^-_{a,n}(V,W)$ and $\neg\, \breve{M}^+_{a,n}(V,W) \mathbin{\&} \neg\, \breve{M}^-_{a,n}(V,W)$, respectively.

At this stage of JSM-reasoning, the predicate $X \Rightarrow_1 Y$ represented by initial data in base of facts (BF) generates the predicate $W \,{}_3{\Leftarrow}\, V$ representing causal relation $\overset{*}{{}_3}{\Leftarrow}$. The latter represents in knowledge base (KB) the set of hypotheses H_1.

3.2. The Analogy

The rules of plausible inference of the 2-nd type (RPI-II) are inferences by analogy, which utilize the similarities of objects on the basis of including the generated hypotheses about causes of effects. In this sense one can call these inferences causal analogies. RPI-II bring about forecasts by generating hypotheses of the form $J_{\langle \nu,n+1\rangle}(C \Rightarrow_1 A)$ for cases of uncertainties $J_{(\tau,n)}(C \Rightarrow_1 A)$ on the basis of hypothetical knowledge, which is obtained by means of the rules RPI-I, about (±)-causes $J_{\langle \pm 1,n+1\rangle}(Q \,{}_3{\Leftarrow}\, C')$.

The rules of inference of the 2-nd type are formalized by the predicates $\breve{\Pi}^\sigma_n(V,W)$, where $\sigma \in \{+,-,0,\tau\}$, $\breve{\Pi}^\pm_n(V,W) \rightleftharpoons \exists k \tilde{\breve{\Pi}}^\pm_n(V,W,k)$, k

is a parameter. The predicate $\overset{\smile}{\widetilde{\Pi}}{}^{+}_{n}(V, W, k)$ expresses the following statements. Let $J_{(\tau,\, n)}(V \Rightarrow_1 W)$ and V contains positive (+1) causes X_i for properties Y_i from KB $- X_i \subset V$, $i = 1, \ldots, k$, (the hypotheses $J_{(1,\, n)}(Y_i \ {}_3\!\Leftarrow X_i)$ have been generated at the previous stages of JSM-reasoning), W being covered by these set of properties, $\bigcup\limits_{i=1}^{k} Y_i = W$. At the same time, V does not contain negative $(J_{(-1,n)}(U \ {}_3\!\Leftarrow Z))$ or factual contradictory $(J_{(0,n)}(U \ {}_3\!\Leftarrow Z))$ causes Z for any subset U of properties from W $(U \subseteq W)$.

$$\begin{aligned}
&\overset{\smile}{\widetilde{\Pi}}{}^{+}_{n}(V, W, k) \rightleftharpoons \exists Y_1 \ldots \exists Y_k((\mathop{\&}_{i=1}^{k} \exists X_i(J_{(1,\, n)}(Y_i \ {}_3\!\Leftarrow X_i) \ \& \ (X_i \subset V))) \ \& \\
&\quad (\bigcup_{i=1}^{k} Y_i = W) \ \& \ \forall Y(\exists X(J_{(1,\, n)}(Y \ {}_3\!\Leftarrow X) \ \& \ (X \subset V)) \to \\
&\quad (\bigvee_{i=1}^{k} (Y = Y_i))) \ \& \ \forall U(((U \subseteq W) \ \& \ (U \neq \varnothing)) \to \\
&\quad \neg\exists Z((J_{(-1,n)}(U \ {}_3\!\Leftarrow Z) \vee J_{(0,n)}(U \ {}_3\!\Leftarrow Z)) \ \& \ (Z \subset V)))).
\end{aligned}$$

Then RPI-II is formulated:

$$(II)^{+}_{n} \qquad \frac{J_{(\tau,n)}(V \Rightarrow_1 W), \quad \breve{\Pi}^{+}_{n}(V, W)}{J_{\langle 1,n+1\rangle}(V \Rightarrow_1 W)}.$$

For the predicate $\breve{\Pi}^{-}_{n}(V, W)$ all conditions are formulated symmetrically.

The predicate $\breve{\Pi}^{0}_{n}(V, W)$ describes the cases when V contains both positive (+1) cause for the subset of the set of properties W and negative (−1) cause for another subset of W, or factual contradictory cause for the subset of W. The predicate $\breve{\Pi}^{0}_{n}(V, W)$ is defined as follows.

$$\begin{aligned}
\breve{\Pi}^{0}_{n}(V, W) \rightleftharpoons{} & \exists X_1 \exists Y_1 \exists X_2 \exists Y_2 (J_{(1,\, n)}(Y_1 \ {}_3\!\Leftarrow X_1) \ \& \ J_{(-1,n)}(Y_2 \ {}_3\!\Leftarrow X_2) \ \& \\
& (Y_1 \cap Y_2) \neq \varnothing \ \& \ (X_1 \subset V) \ \& \ (X_2 \subset V) \ \& \ (Y_1 \subseteq W) \ \& \\
& (Y_2 \subseteq W)) \vee \exists X \exists Y (J_{(0,n)}(Y \ {}_3\!\Leftarrow X) \ \& \ (X \subset V) \ \& \ (Y \subseteq W)).
\end{aligned}$$

The following statements hold

1. $\forall V \forall W (\breve{\Pi}^{+}_{n}(V, W) \to \neg\breve{\Pi}^{-}_{n}(V, W))$,
2. $\forall V \forall W (\breve{\Pi}^{0}_{n}(V, W) \to \neg\breve{\Pi}^{\sigma}_{n}(V, W))$, $\sigma = +, -$.

The predicate $\breve{\Pi}^{\tau}_{n}(V, W)$ is defined as

$$\breve{\Pi}^{\tau}_{n}(V, W) = \neg\,(\breve{\Pi}^{+}_{n}(V, W) \vee \breve{\Pi}^{-}_{n}(V, W) \vee \breve{\Pi}^{0}_{n}(V, W)).$$

The rules $(II)_n^\sigma$, $\sigma = -1, 0, \tau$, have analogous representation. As a result of RPI-II hypotheses of the form $J_{\langle\nu,n\rangle}(C \Rightarrow_1 A)$ are generated ($\nu \in \{+1, -1, 0, \tau\}$). These hypotheses form the subset H_2 of KB. Generated hypotheses are assumed to be "analogous" to "parents" of arguments — causal hypotheses from H_1.

Therefore, the relation $\,^*_3\!\Leftarrow$ generates the relation $\Rightarrow_1^*$, extending and specifying initial one from KB. Let's note, that RPI-I and RPI-II is applied in consecutive order until stabilization, $H_{1n} = H_{1(n+2)}$, $n, n+2$ are the numbers of KB states ($n > 0$), and $H_{1n} = H_1$. Underline, that H_1 hypotheses are the arguments for H_2 generations: argumentation is formalized by $\breve{\Pi}_n^\sigma(V, W)$.

3.3. The Abduction

At the last, third, stage of JSM-reasoning abductive acceptance of hypotheses is effected after iterative application of RPI-I (induction) and RPI-II (analogy) to the BF until stabilization: after some step of iteration no new hypotheses are generated. Then abduction (in the sense of C. S. Peirce) is applied. The scheme of abductive inference [13] — explanatory acceptance of hypotheses — in JSM-reasoning can be specified as follows:

D-facts from the BF, representing $X \Rightarrow_1 Y$,
$H = H_1 \cup H_2$ — hypotheses generated by RPI-I and RPI-II,
H explains BF.

$\forall h((h \in H) \to h$ is plausible$)$.

Let's recall, that propositions $J_{\langle\nu,0\rangle}(C \Rightarrow_1 A)$ are facts, $J_{\langle\nu,n\rangle}(C \Rightarrow_1 A)$ and $J_{\langle\nu,n\rangle}(Q \,_3\!\Leftarrow C)$, $n > 0$, are hypotheses. Here n is the number of RPI-application, consequently n corresponds to degree of hypothesis plausibility: the more n, the less plausibility of hypothesis h.

The relation of "explanation" can be formalized by means of axioms of causal completeness $ACC^{(\pm)}$assumed for subject domain (society) $W^{(\pm)}$: every fact (social behaviour Y of the individual X) in $W^{(\pm)}$ has the causes $V_1, \ldots, V_k$. By means of $(\pm)$-causes $(\pm)$-facts from the BF are "explained". If there are "unexplained" facts (i.e. $ACC^{(\pm)}$ are not fulfilled), then the BF is to be extended in an interactive mode, and JSM-reasoning is applied to the new initial state of the BF. If $ACC^{(\pm)}$ is fulfilled, then the generated hypotheses in H are accepted (abductive explanation). Thus, quasiaxiomatic theory QAT has the model — the set of positive and negative facts with corresponding positive and negative causes. $ACC^{(+)}$ means that for every (+)-fact from the initial BF there exists the step n of application of RPI-I and RPI-II, such that at this step a hypothesis about the cause of this (+)-fact is generated.

$$ACC^{(+)}: \quad \forall X \forall Y \exists V_1 \ldots \exists V_k \exists W_1 \ldots \exists W_k \exists k((J_{\langle 1,0\rangle}(X \Rightarrow_1 Y) \to \exists n((\mathop{\&}_{i=1}^{k} (J_{(1,n)}(W_i \ {}_3\!\Leftarrow V_i) \ \& \ (V_i \subset X) \ \& \ (V_i \neq \varnothing) \ \& \ (W_i \neq \varnothing)) \ \& \ (\bigcup_{i=1}^{k} W_i = Y))).$$

$ACC^{(-)}$ is formulated in a similar way because of the following basic principles assumed to be fulfilled in the "world" $W^{(\pm)}$ (society, represented by BF) in order to be correctly studied by JSM Method.

There exist both positive ((+)-facts with $\bar{\nu} = \langle 1, 0\rangle$)) and negative ((−)-facts with $\bar{\nu} = \langle -1, 0\rangle$) examples of behaviour involved in base of facts. Every positive (+) and negative (−) example of the studied phenomenon (the relation "object possesses set of properties") should have positive (+) and negative (−) causes (empirical dependencies of the cause-effect type), respectively.

Similarity of objects is a factor of the recognition of determinations ((±)-causes) in the JSM Method. For this reason, an operation of similarity of the examined objects and events should be algebraically defined preliminarily.

The constructive generation of hypotheses H_1 and H_2 and explanation of initial BF by axioms of causal completeness mean the following: in the frames of QAT both data analysis with forecast of studied effects and the formation of new subject domain knowledge, it's systematization by generation of new relations from initial ones are carried out. Therefore, JSM-reasoning realizes the knowledge discovery and so it is considered to be a cognitive reasoning [3].

4. The Opinion Analysis

Let's consider JSM-semantics for opinions analysis. Let T be the theme of a poll characterized by the set P of statements $p_1, \ldots, p_n$. The aim of the poll is to establish respondents' opinion about theme as well as about $p_1, \ldots, p_n$. We'll call the set $P = \{p_1, \ldots, p_n\}$ "frame of the theme T", elements $p_1, \ldots, p_n$ — "roots" of questions (poll parameters).

We consider the so-called closed m-valued sociological poll when only m variants of answers to every question are suggested to respondents. Logical tools of m-valued poll formalization are m-valued deductive logic J_m [1], equivalent formulae calculus EFC-J_m and EFC*-J_m (see Appendix). The formulae of logic J_m are constructed from elementary formulae $J_{\nu_j} p_j$, where p_j is variable, J_{ν_j} is J-operator, with the use of connectives &, $\vee$, $\to$.

Thus, the set of variables values $V_m = \{0, \frac{1}{m-1}, \ldots, \frac{m-2}{m-1}, 1\}$, atomic valuations $v^{(i)}[p_j] = \nu_j^{(i)}$, $i = 1, \ldots, m^n$, $j = 1, \ldots, n$, valuation function for formulae of J_m, quasiformulae and formulae of EFC-J_m and EFC*-J_m are defined. The question $?p_j$ — "What is the value ν for root of question p_j?" ($\nu \in V_m$, the set of values) corresponds to each element p_j ($j = 1, \ldots, n$), the answer being the statement $J_\nu p_j$. Then the answer of the respondent b_i on theme T is represented as J-maximal conjunction of m-valued logic J_m $C_i \rightleftharpoons J_{\nu_1^{(i)}} p_1 \mathbin{\&} \cdots \mathbin{\&} J_{\nu_n^{(i)}} p_n$ ($\rightleftharpoons$ designates the "graphical equality of formulae" predicate). This conjunction is defined analogously to maximal conjunction of Boolean (two-valued) logic. $J_{\nu_k^{(i)}} p_k$ ($k = 1, \ldots, n$) for every p_k is included in C_i without duplicates, $J_{\nu_k^{(i)}} p_k$ and $J_{\nu_j^{(i)}} p_k$, where ${\nu_k}^{(i)} \neq \nu_j^{(i)}$, are not included in C_i simultaneously. Such answer represents interpretation the theme T by i-th respondent, i.e. his opinion.

Let K be the set of all possible answers on the theme T with the frame P,

$K = \{\varphi_i | \varphi_i \rightleftharpoons J_{\nu_1^{(i)}} p_1 \mathbin{\&} \cdots \mathbin{\&} J_{\nu_n^{(i)}} p_n,\ v^{(i)}[p_j] = \nu_j^{(i)},\ \nu_j \in V_m,\ j = 1, \ldots, n,\ i = 1, \ldots, m^n\}$ (φ is a metasymbol).

The number of this set elements is $|K| = m^n$, because m-valued (n-dimensional) vector $\vec{\sigma}^{(i)} = \langle \sigma_1^{(i)}, \ldots, \sigma_n^{(i)} \rangle$ corresponds to the only maximal conjunction (one-to-one correspondence). Note, that number of respondents can precede m^n because of coincidence of answers. But the number of different answers can be less than m^n.

Truth values $\nu \in V_m$ should be sociologically interpreted, of course. For Boolean values interpretation is obvious. For 3-valued poll ($m = 3$, $V_3 = \{0, \frac{1}{2}, 1\}$) truth values can be interpreted in such a way: 0 (falsity) corresponds to the answer "no", $\frac{1}{2}$ (uncertainty) — to the answer "don't know", 1 (truth) — to the answer "yes". For 6-valued poll: ($m = 6$, $V_6 = \{0, \frac{1}{5}, \frac{2}{5}, \frac{3}{5}, \frac{4}{5}, 1\}$). 0 (falsity) corresponds to the answer "no", $\frac{1}{5}$ (degree of falsity) — to the answer "rather, no", $\frac{2}{5}$ (uncertainty) —

"don't know", $\frac{3}{5}$ (factual contradiction) — "both yes and no", $\frac{4}{5}$ (degree of truth) — "rather, yes", 1 (truth) — "yes". In another words, corresponding values of logics J_m should be sociologically comprehensible.

4.1. On Possible Heuristics

Consider $[\varphi_j] = \{J_{\nu_1^{(j)}} p_1, \ldots, J_{\nu_n^{(j)}} p_n\}$ as the set of elements of answer φ_j — J-maximal conjunction. Let $\mathbf{U}^{(2)} = \{\psi | (\psi \rightleftharpoons J_{\nu_i} p_i) \;\&\; (\nu_i \in V_m), i = 1, \ldots, n\}$ be the set of atomic answers. Then opinions analysis can be described as JSM-reasoning problem. We analyze propositions $J_{\bar{\mu}_j}(C_j \Rightarrow_1 [\varphi_j])$ — "the individual (person, subject) C_j has the opinion φ_j" — by induction in order to generate propositions $J_{\bar{\mu}_j}([\psi_j]\,{}_3\!\Leftarrow C'_j)$ — "the opinion ψ_j is the consequence of subject's differential characteristics C'_j". The generated causal relations then used do forecast opinion of new respondents as described above. Here C_j, C'_j, $[\varphi_j]$, $[\psi_j]$ are constants, $C_j, C'_j \in 2^{U^{(1)}}$ ($\mathbf{U}^{(1)}$ is the set of subject's differential characteristics), $[\varphi_j], [\psi_j] \in 2^{\mathbf{U}^{(2)}}$, $\bar{\mu}_j = \langle \mu_j, m \rangle$ is the value, calculated by JSM-MAGH application, where $\mu_j \in \{\pm 1, 0, \tau\}$ and m is a number of JSM-RPI application (plausibility degree of hypothesis). Note, that in this scheme the valuation of empirical statement "the person C has the opinion φ" ($C \Rightarrow_1 [\varphi]$) is the value of a theme T as a whole. This valuation corresponds to JSM-logic with 4 truth value types, not to m-valued set of possible poll answers.

In the wide sense, the approach described is considered to be the heuristic scheme "similarity – analogy – abduction". Therefore, another way of similarity analysis — similarity of subjects' characteristics and behaviour — – is possible. Apart from induction we used an objective analysis of empirical sociological data — for example, with Boolean algebra [19] or fuzzy logic [20] — can be carried out. Heuristic scheme "similarity – analogy – abduction" is specified by the scheme "algebra of logic – analogy – abduction" in this case.

Consider Boolean poll when only two answers — "yes" (1) and "no" (0) — are possible. Then Boolean vector $\vec{\sigma}^{(i)} = \langle \sigma_1^{(i)}, \ldots, \sigma_n^{(i)} \rangle$, where $\sigma_j^{(i)} = 0, 1$, $j = 1, \ldots, n$, $i = 1, \ldots, k$ (number of different opinions, less or equal to number of respondents; $k \leqslant 2^n$) corresponds to opinion φ_i of respondent X_i with respect to frame $P = \{p_1, \ldots, p_n\}$. If $\vec{\sigma}^{(i)}$ corresponds to atomic value $v^{(i)}$, then $\vec{\sigma}^{(i)} = \langle v^{(i)}[p_1], \ldots, v^{(i)}[p_n] \rangle$. Let the valuation of the i-th respondent with respect to theme T be $\sigma^{(i)}$,

$\sigma^{(i)} = 0, 1$, then $\sigma^{(i)}$ corresponds to opinion φ_i (vector $\vec{\sigma}^{(i)}$). Boolean function is defined characterizing the relation between answers to the questions $p_1, \ldots, p_n$ and the theme as a whole.

The set $\Phi^+ = \{\varphi_1, \ldots, \varphi_s\}$ unites all opinions φ_i with positive perception of theme T, $\sigma^{(i)} = 1$, $\varphi_i \rightleftharpoons p_1^{\sigma_1^{(i)}} \,\&\, \cdots \,\&\, p_n^{\sigma_n^{(i)}}$, $i = 1, \ldots, s$ (here $p^\sigma = p, \sigma = 1; p^\sigma = \neg p, \sigma = 0$, as usual). $B = \{X_1, \ldots, X_m\}$ denotes the set of subjects (respondents) with opinion from Φ^+, $B = \{X \mid J_{\langle 1,0\rangle}(X \Rightarrow_1 [\varphi_i]) \,\&\, (\varphi_i \in \Phi^+)\}$ $(i = 1, \ldots, s)$.

The suggested strategy resembles the strategy of QCA — Qualitative Comparative Analysis [19] — in some details. Let's transform perfect DNF $\varphi_1 \vee \cdots \vee \varphi_s$ to reduced DNF $\partial(\varphi_1 \vee \cdots \vee \varphi_s) \rightleftharpoons \chi_1 \vee \cdots \vee \chi_l$ in a standard manner. Then bring implicants χ_j from the set $[\partial\varphi] = \{\chi_1, \ldots, \chi_l\}$ in correspondence with set Φ_j^+ of opinions φ such that opinion φ (as two-valued maximal conjunction) is covered by implicant χ_j, $\Phi_j^+ = \{\varphi | \chi_j \sqsubset \varphi\}$, $j = 1, \ldots, l$. So, $B^{(j)} = \{X_{j_1}, l \ldots, X_{j_h}\}$ is a set of subjects (respondents) with opinions from Φ_j^+, $B^{(j)} = \{X \mid J_{\langle 1,0\rangle}(X \Rightarrow_1 [\varphi_q]) \,\&\, (\varphi_q \in \Phi_j^+)\}$. The similarity of these subjects is $V_j' = \bigcap\limits_{k=1}^{h} X_{j_k}$. To realize heuristic scheme described above assume causal relation $C(V_j', \chi_j)$ (which is ${}_3\!\Leftarrow$ in JSM-heuristic) to be represented by pairs $\langle V_j', \chi_j\rangle$ ($j = 1, \ldots, l$), i.e. opinion $\varphi_q \in \Phi_j^+(J_{\langle 1,0\rangle}(X \Rightarrow_1 [\varphi_q]))$ is the consequence of subject's characteristics V_j', $V_j' \subseteq X$. This causal relation (causality by implicants) can be used to define the predicate of explanation. We need this predicate to specify idea of explanation of examples from base of facts by hypotheses in the scheme of abduction.

$$E(X,Y) \rightleftharpoons \exists V_1' \ldots \exists V_k'((\bigvee_{i=1}^{k}((V_i' \subseteq X) \,\&\, ([\chi_i] \subseteq Y) \,\&\, C(V_i', \chi_i))) \,\&\, J_{\langle 1,0\rangle}(X \Rightarrow_1 Y)),$$

where $[\chi_i]$ is the set of atoms from χ_i, $[\chi_i] = \{p_{i_1}^{\sigma_{i_1}^{(i)}}, \ldots, p_{i_n}^{\sigma_{i_n}^{(i)}}\}$.

Corresponding predicates and rules of inference by analogy are formulated.

The scheme described can be realized for many-valued logics J_m ($m \geqslant 3$) with corresponding perfect and reduced DNF (see below), for example, for 4-valued JSM-logic [8]. The problem of comparison the results obtained by both heuristics ("induction – analogy – abduction" and "algebra – analogy – abduction") seems to be rather interesting. It can be shown that the second heuristic has substantially smaller cognitive possibilities in comparison with JSM Method [16]. This is due to the peculiarities of its formal instruments allowing to study only the closed world (with the stable knowledge). JSM Method, on the contrary, is constructed for the formalization of reasoning in the open worlds. It is this ability that divides cognitive instruments from data processing procedures.

4.2. Formal Representation of a Poll and Deductive Recognition of Rationality

Let $R = \{b_1, \ldots, b_r\}$ be the set of respondents involved in the poll, $R = \{X \mid \exists\varphi J_{\langle 1,0\rangle}(X \Rightarrow_1 [\varphi])\}$, $[\varphi] = \{J_{\nu_1}p_1, ..., J_{\nu_n}p_n\}$. We suppose the set of these respondents answers (opinions) $K' \subseteq K$ (set of all possible answers) to be stable: K' is not changed when the set of respondents is enlarged. It is important, that stabilization of the answers set is experimentally achieved.

Closed m-valued sociological poll on the theme T with value scale $V_m = \{0, \frac{1}{m-1}, \ldots, \frac{m-2}{m-1}, 1\}$ is characterized by the set of statements $P = \{p_1, \ldots, p_n\}$ describing the theme T and the set of involved respondents $R = \{b_1, \ldots, b_r\}$ with stable set of answers (opinions) K'. Tools of logics J_m give us the possibility to extend this characterization.

$J_\nu p$ are atoms of logic J_m; p is a propositional variable, ν is a truth value of m-valued logic J_m, $\nu \in V_m$. The method of analytic tableaux [25] and equivalent formulae calculus EFC-J_m has been formulated for these logics [12]. The equivalency $\leftrightarrow$ is the main connective in EFC-J_m, φ and ψ in $(\varphi \leftrightarrow \psi)$ are quasiformulae formed by J-atoms and connectives &, $\vee$ and $\to$ (negation is definable connective in EFC*-J_m). EFC-J_m is a distributive lattice with zero element (0) and unit element (1) and with the law of excluded $(m+1)$-th: $(J_0p \vee J_{\frac{1}{m-1}}p \vee \ldots \vee J_{\frac{m-2}{m-1}}p \vee J_1p) \leftrightarrow t$. J-perfect DNF and the algorithm for simple implicants receiving are defined in EFC-J_m. The analytic tableaux are built by use of designated formulae $t\varphi$ and $f\varphi$, t and f are signs for φ. The rules include α- and β-rules of two-valued logic [25], special α-rules

$$\frac{tJ_\nu p}{J_\nu p}, \quad \text{where } \nu \in V_m;$$

and ε-rules

$$\frac{fJ_0p}{J_{\frac{1}{m-1}}p|\ldots|J_{\frac{m-2}{m-1}}p|J_1p}, \ldots, \frac{fJ_1p}{J_0p|J_{\frac{1}{m-1}}p|\ldots|J_{\frac{m-2}{m-1}}p}.$$

Designated formulae of J_m logics $t\varphi$ and $f\varphi$ are contrary pairs so as undesignated formulae $J_\nu p$ and $J_\mu p$ ($\nu \neq \mu$). Analytic tableau branch θ is named "closed" if it contains contrary pair. The formula φ is provable in J_m, if analytic tableau with the root $f\varphi$ is closed. The set of undesignated formulae $\Sigma = \{\varphi_1, \dots, \varphi_s\}$ is consistent if analytic tableau with the root $t(\varphi_1 \,\&\, \cdots \,\&\, \varphi_s)$ is open, and contrariwise. A J_m logic formula φ can be reduced to DNF $\partial\varphi$ by analytic tableau method and $\partial\varphi$ can be reduced to J-perfect DNF in EFC-J_m [12].

Consider the set $\Sigma = \{\psi_1, \dots, \psi_s\}$ of logics J_m formulae $\psi_1, \dots, \psi_s$ representing logical dependencies between statements from P — roots of questions $?p_1, \dots, ?p_n$. This set of "meaning postulates" (by R. Carnap) is defined by researcher (sociologist). The set Σ must be consistent and formula $\psi = (\psi_1 \,\&\, \cdots \,\&\, \psi_s)$ is not tautology of the logic J_m (& is the corresponding conjunction of the logic J_m). Further, this set is used for the calculation of consistency and inconsistency degrees for closed m-valued poll.

Thus, closed m-valued sociological poll O_m ($m \geqslant 3$) on the theme T is realized by means of m-valued logic J_m and can be represented as $O_m = \langle J_m, P, \Sigma, K', R\rangle$ with K' being given by researcher (sociologist) as a result of empirical stabilization of the opinions set. Such K' definition seems to be one of the ways of representative (in logical, not statistical sense) sample formation.

Let $\mathrm{Consis}(\Sigma \cup \{\varphi\})$ be metapredicate for set of formulae consistency $(\Sigma \cup \{\varphi\})$. $K^+ = \{\varphi | \mathrm{Consis}(\Sigma \cup \{\varphi\}) \,\&\, (\varphi \in K)\}$ denotes the set of φ (respondents answers) — J-maximal conjunctions of logics J_m — consistent with Σ; $\Delta = \{\varphi | \neg\,\mathrm{Consis}(\Sigma \cup \{\varphi\}) \,\&\, (\varphi \in K)\}$ denotes the set of φ inconsistent with Σ. For real poll K can be substituted by K'. The set Δ can be defined by analytic tableau with the root $f\psi$, where $\psi = (\psi_1 \,\&\, \cdots \,\&\, \psi_s)$: we construct $\partial\psi$ and reduce it to J-perfect DNF $\overline{\psi}$ in EFC-J_m, $[\overline{\psi}]$ is the set of conjunctions from $\overline{\psi}$ and $\Delta = [\overline{\psi}]$.

Let's define functions $\eta(K', K^+) = |K' \cap K^+|/|K'|$ and $\xi(K', \Delta) = |K' \cap \Delta|/|K'|$ characterizing degrees of consistency and inconsistency of m-valued closed poll concerning the theme T respectively (as usual, $|K'|$ is the number of elements of K', and so on). It is easy to show that $\eta(K', K^+) + \xi(K', \Delta) = 1$, because of $K = K^+ \cup \Delta$, $K^+ \cap \Delta = \varnothing$ and $K' \subseteq K$. So, definition of both functions — $\xi(K', \Delta)$ and $\eta(K', K^+)$ — seems to be unnecessary. But their computing procedures are different, so as the procedures computational complexity for real empirical sample.

To calculate degree of consistency the method of analytic tableaux [25] has been used. According to the method, if analytic tableau $\mathfrak{T}$ for the set $\Sigma \cup \{\varphi\}$ is closed, $\Sigma \cup \{\varphi\}$ is inconsistent and $\varphi \in \Delta$, where Δ is the set of "prohibited" maximal conjunctions. Then the following procedure for computing the degree of poll consistency can be suggested. We check does the opinion φ-maximal conjunction of formulae $J_{\nu_i} p_i$ — from K' belong to Δ (for given set Σ of "meaning postulates") or not, then function $\delta(K', \Delta) = 1 - \xi(K', \Delta) = 1 - |K' \cap \Delta|/|K'|$ is calculated. In this approach neither the set of consistent opinions K^+ nor the set of inconsistent opinions Δ are to be constructed.

We can compute functions $\eta(K', K^+)$ and $\xi(K', \Delta)$ using method of analytic tableaux for m-valued logics [6] in another way. Let's construct completed analytic tableau $\mathfrak{T}$ with the root $t(\psi_1 \ \& \ \cdots \ \& \ \psi_s)$, where $\psi_i \in \Sigma$, $i = 1, \ldots, s$. Disjunction of open branches forms DNF, which is transformed to perfect DNF in EFC-J_m. The set containing all J-maximal conjunctions of constructed perfect DNF represents K^+ — the set of respondent answers consistent with Σ. So, the function $\eta(K', K^+)$ — degree of consistency — can be computed. As $\Delta = K - K^+$, the function $\xi(K', \Delta)$ — degree of inconsistency — can be computed as well. But there is another way of Δ constructing by analytic tableaux. Let's construct completed analytic tableau with the root $f(\psi_1 \ \& \ \cdots \ \& \ \psi_s)$. The next steps are the same as described: disjunction of open branches is transformed to perfect DNF, received maximal conjunctions form Δ.

Consistency defined in such a way is considered to be the kind of rationality understood as argumented decision making. This is especially clear if we use the logic JA_4, the special variant of J_m logics, for poll representation. Let's consider the argumentation logic A_4 [8] with the truth values 1, -1, 0, τ interpreted as "factually true", "factually false", "factual contradictory", and "uncertain", respectively. The semantics of A_4 comprises nonempty set of reasons (possible arguments and counter-arguments) $\mathbf{A}$ and functions $g^{\sigma} \colon P \longrightarrow 2^{\mathbf{A}}$, where $\sigma \in \{+, -\}$ and P is a set of propositional variables p, q, r, s (possibly with the subscripts). For example, P is a set of statements $\{p_1, \ldots, p_n\}$ describing the theme T of the poll. $g^+(p)$ and $g^-(p)$ are the sets of arguments and counterarguments for proposition p, respectively, and $g^+(p) \cap g^-(p) = \varnothing$ for any $p \in P$.

The valuation function for atomic formulae is defined as follows:

- $v[p] = 1$ if and only if $g^+(p) \neq \varnothing$ and $g^-(p) = \varnothing$ (it means that variable p has arguments and doesn't have counterarguments, p is factually true);

- $v[p] = -1$ if and only if $g^+(p) = \varnothing$ and $g^-(p) \neq \varnothing$;
- $v[p] = 0$ if and only if $g^+(p) \neq \varnothing$ and $g^-(p) \neq \varnothing$;
- $v[p] = \tau$ if and only if $g^+(p) = \varnothing$ and $g^-(p) = \varnothing$.

The logic JA_4 contains operators J_ν ($\nu \in \{1, -1, 0, \tau\}$), $J_\nu p = t$, if $v[p] = \nu$, $J_\nu p = f$, if $v[p] \neq \nu$, where t and f are truth values of two-valued logic, "true" and "false", respectively. Here 1, -1, 0, τ are factual (internal) valuations, t and f are external ones. The valuation function for atomic formulae of JA_4 is defined as follows: $v[J_\nu p] = t$, if $v[p] = \nu$. So, $v[J_1 p] = t$ if and only if $g^+(p) \neq \varnothing$ and $g^-(p) = \varnothing$; and so on.

4.3. The Predictive Poll

The suggested description of formal parameters of sociological poll is not connected with the strategy of social data analysis in hand. This description can be specified for poll called "predictive". It means heuristic scheme "similarity – analogy – abduction" that is specified by the scheme "empirical induction – structural analogy – constructive abduction" in JSM Method of Automated Generation of Hypotheses. Let's remind that as a result of JSM-reasoning application initial base of facts is extended. The set of respondents involved $R = \{X \mid \exists\, \varphi J_{\langle 1,0\rangle}(X \Rightarrow_1 [\varphi])\}$, $[\varphi] = \{J_{\nu_1} p_1, ..., J_{\nu_n} p_n\}$, is added by generated hypotheses about possible variants of answers for respondents with previously undefined opinions, $R^* = \{X | \exists\, \varphi J_{\langle 1,n\rangle}(X \Rightarrow_1 [\varphi]) \;\&\; (n > 0)\}$. It is possible that set of answers K' is changed in this case. Accordingly, computed rationality characteristics $\eta(K', K^+)$ and $\xi(K', \Delta)$ are changed as well.

Moreover, the abductive acceptance of hypotheses in JSM Method presupposes interactive extension of facts base if degree of facts explanation is insufficient. It leads to both R and, accordingly, K' changing and, as a result, to recomputation the degrees of consistency and inconsistency of the poll.

Let's reveal the meaning of the degree of facts explanation. Consider BF^+ and BF^--parts of BF containing positive and negative facts of investigated relation, $BF^\sigma \subseteq BF$, $|BF^\sigma|$ denotes the number of (σ)-facts, $\sigma \in \{+, -\}$. Let, further, $\widetilde{BF}^\sigma \subseteq BF^\sigma$, where $\widetilde{BF}^\sigma$ is the subset of BF^σ such that axiom of causal completeness $ACC^{(\sigma)}$ is fulfilled, i.e. generatedcausal hypotheses explain facts from $\widetilde{BF}^\sigma$. Then degree of

(σ)-facts explanation is defined as $\rho^\sigma = \frac{|\widetilde{BF}^\sigma|}{|BF^\sigma|}$ and degree of all facts explanation (degree of causal completeness) is $\rho = \frac{|\widetilde{BF}|}{|BF|}$. This value seems to be additional numerical (objective) parameter of a poll.

Non-trivial procedure of abductive convergence can be added to JSM-heuristics on the base of the defined value. Consider the set of extended bases of facts $BF_1 \subset BF_2 \subset \cdots \subset BF_m$. Then if $\rho_1^\sigma \leqslant \cdots \leqslant \rho_m^\sigma$, where $\rho_i^\sigma = \frac{|\widetilde{BF}_i^\sigma|}{|BF_i^\sigma|}$, $i = 1, \ldots, m$, $\sigma \in \{+, -\}$, abductive convergence takes place. Finite number m can be defined if special threshold ρ^σ (usually $0,8 \leqslant \rho^\sigma \leqslant 1$) is given, $\rho_m^\sigma \geqslant \rho^\sigma$. It is obvious, that abductive convergence can serve for reasoning control.

Thus, some objective characteristics describe predictive polls. These are stable set of opinions K' (new answers do not come when new respondents are added), number n of steps of heuristic scheme application (until the set of generated hypotheses is stable), the threshold of abductive convergence ρ satisfying the experimental researcher.

5. Conclusion

The described new technology of logical analysis of opinions has been employed to analyze and to predict electoral preferences of senior students of Russian State University for Humanities (RSUH; on December elections to the State Duma, 2007). Let's outline some features of the research. The respondents (senior RSUH students) descriptions have been represented by three sets of differential indicators to define their similarity which is the base of behaviour (opinion) determination. The first one describes the social characteristics of the subject (in correspondence with the idea of E. Fromm about social character). The second set describes individual traits that do not depend on the social membership of the individual; and, finally, the third one concerns the biography details that are essential for the topic being studied. Structured description of subjects (individuals) affords the possibility to rely on the so-called "postulate of behavior" in the analysis of behaviour (opinion) — the presupposition that a subset of the union of these three sets of differential signs that characterize the subject determines his behaviour (opinion)[2].

[2] One can say that J.S.Mill himself — being a proponent of "psychologism" in sociology — maintained a similar view on behaviour.

So, our questionnaire included, for example, family and financial status, educational level of relatives (as biographical data); the social characteristics described social and political activity of the students, their fundamental values; psychological tests analyzed consideration for another peoples interests, sociability, the level of self-control and so on.

As we were interested in the students electoral intentions, six Russian political parties have been considered to be the themes $T_1, T_2, \ldots, T_6$. In addition, T_7 stands for "will not participate in the elections". The frame P contains the program purposes of different parties (without party itself) concerning the key problems — Politics, Economics, Army, Media, Personal freedom and rights and so on.

The following problems were solving.

1. Determinants of electoral behaviour generation. Electoral behaviour means (in the sense of our study) both political party and opinion choice. Here opinion consists of the elements of frame answering. So, we have studied the statements $J_{\bar{\mu}_j}(C_j \Rightarrow_1 \langle[\varphi_j], T_i\rangle)$, where T_i is one of the parties.

2. Electoral choice prediction. 27 students (from 231) have been asked only about elements of P (without T). Then their electoral choice has been predicted by generated determinants and has been validated by their real election.

3. Rationality analysis. Here the difference between two kinds of a poll was in the focus. In one of them T-answering precedes P-answering, in other — vice versa.

The following causal hypotheses have been generated (for example).

$J_{\langle 1,1\rangle}(\langle[\psi_1], T_1\rangle \;{}_3{\Leftarrow}\; C'_1)$, $C'_1 = \{$female, parents' financial support, part-time work, State grant for the education, unmarried, middle level of political activity$\}$, T_1 is "United Russia" party, $\psi_1 \rightleftharpoons J_1 p_1 \,\&\, J_1 p_2 \,\&\, J_1 p_3 \,\&\, J_{-1} p_4 \,\&\, J_0 p_5$. Here p_1 – "State property should dominate", p_2 – "natural monopolies should belong to the State", p_3 – "military service should be by contract as well as by call up", p_4 – "pension insurance should be supported by employer as well as by the State", p_5 – "Russian external policy should be oriented to the West" (the numbers of elements of P are not the same as in the questionnaire).

$J_{\langle 1,1\rangle}(\langle[\psi_2], T_2\rangle \;{}_3{\Leftarrow}\; C'_2)$, $C'_2 = \{$fourth year student, low level of authoritarian subordination, parents' financial support, unmarried, middle level of political activity$\}$, T_2 is "Union of rights" or "Yabloko" party, $\psi_2 \rightleftharpoons J_1 p_1 \,\&\, J_{-1} p_2 \,\&\, J_{-1} p_3 \,\&\, J_1 p_6 \,\&\, J_1 p_7$. Here p_1, p_2, p_3 are as above,

p_6 — "media should be independent from the State as well as from any organizations", p_7 – "federalism is to be strengthened".

To analyze opinions rationality the meaningful relations between different parties purposes concerning the same points (Army or Media, for example) were transformed into logical connections. So, for the landownership the following connections have been constructed: $\{((J_1 p_{75} \vee J_1 p_{88}) \to J_{-1} p_{57}), (J_1 p_{67} \to J_{-1} p_{88})\}$ and so on. Here p_{57} – "State ownership of land should be the only form of landownership", p_{67} – "Land market should be strictly limited", p_{75} – "Private landownership is possible as well as State ownership of land", p_{88} – "Free sale of agricultural land is necessary".

For the media, for example, the following connection has been suggested: $(J_1 p_{77} \to J_{-1} p_{59} \;\&\; J_{-1} p_{69})$. Here p_{59} – "Media should be independent both from the State and from monopolies", p_{69} — "Favourable conditions for private media development should be created", p_{77} – "The main TV-channels should belong to the State".

Let's underline, that the construction of such connections is the creative activity depending on sociologists' view on the problem and on the general research tasks.

Experimental results in rationality analysis have shown: if T-answering precedes P-answering, then degree of consistency $\eta(K', K^+)$ is higher than in the contrary answering. We received $\eta(K', K^+) = 0.285$ in the first case (102 students) and $\eta(K', K^+) = 0.169$ in the second (102 another students). It is interesting that the degree of consistency for all students (231; $\eta(K', K^+) = 0.199$) is less than for the students who answered "will not participate in the elections" (26 persons, $\eta(K', K^+) = 0.269$). Low general level of opinion consistency illustrates students' both inattentiveness to programs of parties they prefer and to important social and political problems.

The results obtained complement our earlier investigations (published in Russian) in the conclusion that the JSM Method and Intelligent Systems based on it can be regarded as instruments for formalized qualitative analysis of sociological data. Taking into consideration the person's individuality and generating determinants of behaviour, the proposed approach affords the possibility to study further typology of social communities and to construct the models of social structures. In addition, the important feature of the technology is elaboration of the instruments for rational behaviour (and departure from it) analysis.

In conclusion, let's underline that formal tools of JSM Method allow to specify some unclear ideas, transforming them into concepts with viewable and constructively valued content. The examples of such specifications are "cause – effect" relation, context-depending induction, abduction, open theory (quasiaxiomatic theory), determination of social subject's behaviour, formalized method of qualitative analysis of sociological data. All of them can be formalized and experimentally utilized in the frames of the JSM Method of Automated Generation of Hypotheses and Intelligent Systems based on it. So JSM-reasoning, which is the synthesis of cognitive procedures – induction, analogy and abduction — proves to be an effective tool for cognitive sociology: knowledge discovery in special Data Bases (or Facts Bases) [10].

Appendix

Equivalent formulae calculus EFC-$\boldsymbol{J_m}$

The language:

- propositional variables: p, q, r, ... (maybe with subscripts);
- logical connectives:
  - nullary: 0 and 1,
  - unary: J_0, J_1, J_ν, where $\nu \in \{\frac{1}{m-1}, \ldots, \frac{m-2}{m-1}\}$,
  - binary: &, $\vee$, $\leftrightarrow$ (main connective).

Definition of 'quasiformula':

(i) 0 and 1 are quasiformulae;
(ii) $J_\nu \pi$ is a quasiformula, where π is a variable and $\nu \in V_m = \{0, \frac{1}{m-1}, \ldots, \frac{m-2}{m-1}, 1\}$, $m \geqslant 3$;
(iii) if φ, ψ are quasiformulae, then $(\varphi \mathbin{\&} \psi)$, $(\varphi \vee \psi)$ are quasiformulae;
(iv) the only quasiformulae are those given by (i)–(iii).

Definition of 'formula':

- If φ, ψ are quasiformulae, then $(\varphi \leftrightarrow \psi)$ is a formula.

EFC-$\boldsymbol{J_m}$ axioms:

(1a) $(\varphi \mathbin{\&} \varphi) \leftrightarrow \varphi$
(2a) $(\varphi \mathbin{\&} (\psi \mathbin{\&} \chi)) \leftrightarrow ((\varphi \mathbin{\&} \psi) \mathbin{\&} \chi)$
(3a) $(\varphi \mathbin{\&} \psi) \leftrightarrow (\psi \mathbin{\&} \varphi)$
(4a) $(\varphi \mathbin{\&} (\psi \vee \chi)) \leftrightarrow ((\varphi \mathbin{\&} \psi) \vee (\varphi \mathbin{\&} \chi))$

(5a) $(\varphi \,\&\, (\varphi \vee \psi)) \leftrightarrow \varphi$
(6a) $(J_\nu p \,\&\, J_\mu p) \leftrightarrow 0$, where $\nu \neq \mu$
(7a) $(\varphi \,\&\, 1) \leftrightarrow \varphi$
(8a) $(\varphi \,\&\, 0) \leftrightarrow 0$
(1b) $(\varphi \vee \varphi) \leftrightarrow \varphi$
(2b) $(\varphi \vee (\psi \vee \chi)) \leftrightarrow ((\varphi \vee \psi) \vee \chi)$
(3b) $(\varphi \vee \psi) \leftrightarrow (\psi \vee \varphi)$
(4b) $(\varphi \vee (\psi \,\&\, \chi)) \leftrightarrow ((\varphi \vee \psi) \,\&\, (\varphi \vee \chi))$
(5b) $(\varphi \vee (\varphi \,\&\, \psi)) \leftrightarrow \varphi$
(6b) $(J_0 p \vee J_{\frac{1}{m-1}} p \vee \ldots \vee J_{\frac{m-2}{m-1}} p \vee J_1 p) \leftrightarrow 1$
(7b) $(\varphi \vee 0) \leftrightarrow \varphi$
(8b) $(\varphi \vee 1) \leftrightarrow 1$

EFC-J_m inference rules:

$$R1 \qquad \frac{\Lambda}{(\varphi \leftrightarrow \varphi)},$$

where Λ stands for the empty set of premises.

$$R2 \qquad \frac{(\varphi \leftrightarrow \psi)}{(\psi \leftrightarrow \varphi)}$$

$$R3 \qquad \frac{(\varphi \leftrightarrow \chi), \quad (\chi \leftrightarrow \psi)}{(\varphi \leftrightarrow \psi)}$$

$$R4 \qquad \frac{\varphi(p_1, \ldots, p_{i-1}, p_i, p_{i+1}, \ldots, p_n) \leftrightarrow \psi(p_1, \ldots, p_{i-1}, p_i, p_{i+1}, \ldots, p_n)}{\varphi(p_1, \ldots, p_{i-1}, p_j, p_{i+1}, \ldots, p_n) \leftrightarrow \psi(p_1, \ldots, p_{i-1}, p_j, p_{i+1}, \ldots, p_n)}.$$

$R4$ is the following rule of substitution: every occurrence of the variable p_i in the formulae φ and ψ is replaced by the variable p_j (not by a quasiformula χ).

$$R5 \qquad \frac{(\varphi(\chi_1) \leftrightarrow \psi), \quad (\chi_1 \leftrightarrow \chi)}{(\varphi(\chi) \leftrightarrow \psi)}.$$

$R5$ is the rule of equivalent formulae replacement (for every k occurrences χ_1 in φ some of them are replaced by m occurrences χ, $k \geqslant m$).

Let Ψ be the EFC-J_m formula. Ψ is considered to be proved in EFC-J_m if the formulae sequence $\Phi_1, \ldots, \Phi_n$ exists such that $\Phi_n = \Psi$ and every formula Φ_i is either an axiom or Φ_i is derived from the preceding formulae of the sequence Φ_j, Φ_k, where $j, k < i$, by the rules $R1$–$R5$.

It is easy to see that if we'll replace the equivalency connective by the equivalency relation, equational definition of a distributive lattice with zero element (0) and unit element (1) will be obtained.

Note, that if we'll add the axioms for the distribution of J-operators J_0, J_1 and J_ν ($\nu \in \{\frac{1}{m-1}, \ldots, \frac{m-2}{m-1}\}$) with respect to & and $\vee$ to EFC-J_m and enforce the rule $R4$ by substitution a quasiformula χ for the variable p_i, we'll obtain the calculation EFC*-J_m.

Equivalent formulae calculus EFC*-J_m

Definition of 'quasiformula*':

(i) all EFC-J_m quasiformulae are EFC*-J_m quasiformulae (i.e. quasiformulae*);
(ii) if φ is a quasiformula*, then $J_\nu\varphi$ is a quasiformula*, where $\nu \in V_m = \{0, \frac{1}{m-1}, \ldots, \frac{m-2}{m-1}, 1\}$, $m \geqslant 3$;
(iii) if φ and ψ are quasiformulae*, then $(\varphi \mathbin{\&} \psi)$ and $(\varphi \vee \psi)$ are quasiformulae*;
(iv) the only quasiformulae* are those given by (i)–(iii).

Definition of 'formula*':

- If φ and ψ are quasiformulae*, then $(\varphi \leftrightarrow \psi)$ is a formula*.

EFC*-J_m axioms: all EFC-J_m axioms and the following formulae:

(9a) $J_0(\varphi \mathbin{\&} \psi) \leftrightarrow (J_0\varphi \vee J_0\psi)$
(9b) $J_0(\varphi \vee \psi) \leftrightarrow (J_0\varphi \mathbin{\&} J_0\psi)$
(10) $J_1\varphi \leftrightarrow \varphi$
(11) $J_\nu\varphi \leftrightarrow 0$, where $\nu \in \{\frac{1}{m-1}, \ldots, \frac{m-2}{m-1}\}$
(12) $J_0(J_{\frac{i}{m-1}}\varphi) \leftrightarrow \bigvee\limits_{\substack{0 \leqslant k \leqslant m-1 \\ k \neq i}} J_{\frac{k}{m-1}}\varphi$
(13) $J_0 1 \leftrightarrow 0$
(14) $J_0(J_0\varphi) \leftrightarrow \varphi$

EFC*-J_m inference rules: $R1$–$R3$ and $R5$ from EFC-J_m and the following rule:

$$R4^* \qquad \frac{\varphi(p_1, \ldots, p_{i-1}, p_i, p_{i+1}, \ldots, p_n) \leftrightarrow \psi(p_1, \ldots, p_{i-1}, p_i, p_{i+1}, \ldots, p_n)}{\varphi(p_1, \ldots, p_{i-1}, \chi, p_{i+1}, \ldots, p_n) \leftrightarrow \psi(p_1, \ldots, p_{i-1}, \chi, p_{i+1}, \ldots, p_n)},$$

where χ is a quasiformula*.

The law of excluded third $((J_0\varphi \vee \varphi) \leftrightarrow 1)$ and the law of contradiction $((J_0\varphi \,\&\, \varphi) \leftrightarrow 0)$ are proved in EFC*-J_m. A Boolean algebra is embedded to EFC*-J_m with the operator J_0 being the complement.

Implication $(\varphi \rightarrow \psi)$ is expressible by $(J_0\varphi \vee \psi)$. Thus, the procedure of reduction to J-perfect DNF is defined in EFC*-J_m directly. Then the function $\eta(K', K^+)$, degree of consistency, and the function $\xi(K', \Delta)$, degree of inconsistency, can be computed as described above (see Section 4.2).

References

[1] Anshakov, O. M., Finn, V. K., Skvortsov, D. P., "On axiomatization of many-valued logics associated with formalization of plausible reasoning", *Studia Logica* XLVIII, 4 (1989): 423–448.

[2] Anshakov, O. M., Finn, V. K., Skvortsov, D. P., "Plausible reasoning and its logics", pages 87–91 in: *International Conference on Intelligent Systems and Semiotics: A Learning Perspective*, Gaithersburg, MD, 1997.

[3] Anshakov, O. M., Gergely, T., *Cognitive Reasoning*, Springer, Heidelberg, Dordrecht, London – New York 2010.

[4] Barwise, J. (ed.): *Handbook of Mathematical Logic*, Pt. 1, Theory of Models. North-Holland, Amsterdam – New – York – Oxford (1977).

[5] Blinova, V. G., *et al.*, "Toxicology analysis by means of JSM method", *Bioinformatics* 19, 10 (2003): 1201–1207.

[6] Bolc, L., Borowik, P., *Many-Valued Logics*, vol. 1, "Theoretical Foundations", pp. 123–142, Springer-Verlag, 1992.

[7] Burch, R. W., 'Semeiotic data fusion", pages 67–73 in: *3rd International Conference on Information Fusion*, Paris, 2000.

[8] Finn, V. K., "A form of argumentation logic", *Automatic Documentation and Mathematical Linguistic* 30, 3 (1996): 3–27.

[9] Finn, V. K., "The synthesis of cognitive procedures and the problem of induction", *Automatic Documentation and Mathematical Linguistic* 43, 3 (2009): 149–195.

[10] Finn, V. K.: "Toward structural cognitology: Phenomenology of consciousness form the point of view of artificial intelligence", *Russian Journal of Communication* II, 1/2 (2009): 81–104.

[11] Finn, V. K., Mikheyenkova, M. A., "Quantofrenia by P. Sorokin and formalized qualitative analysis", pages 189–192 in: V. A. Mansurov (ed.), *European Society or European Societies: a View from Russia*, Moscow – Lisbon, 2009.

[12] Finn, V. K., Mikheyenkova, M. A., *Introduction to Logic and Intelligent Data Analysis for Sociologists* (in Russian), preprint, 2010.

[13] Josephson, J. R., Josephson, S. G. (eds.), *Abductive Inference: Computation, Philosophy*, Technology. Cambridge Univ. Press, 1994.

[14] Kelle, U. (ed.), *Computer-aided Qualitative Data Analysis: Theory, Methods and Practice*. London. Thousand Oaks, New Delhi: SAGE Publication, 1998.

[15] Mikheyenkova, M. A., *et al.*, "Application of JSM method of automatic hypotheses generation to some problems of sociology", pages 22–25 in: *Workshop of Applied Semiotics, ECAI-96*, Budapest, August 12–16, 1996.

[16] Mikheyenkova, M. A., "On the logical tools of intelligent sociological data analysis" (in Russian), *Artificial Intelligence and Decision Making* 1 (2010): 20–32.

[17] Mill, J. S., *A System of Logic Ratiocinative and Inductive, Being a Connected View of the Principles of Evidence and the Methods of Scientific Investigation*, London, Parker, Son and Bowin, 1843 (1st edition).

[18] Peirce, C. S., "Abduction and induction", pages 150–156 in: L. Buchler (ed.) *Philosophical Writings of Peirce*, Dover Publications, no 4, 1995.

[19] Ragin, C. C., *The Comparative Method: Moving beyond Qualitative and Quantitative Strategies*, University of California Press, Berkley, Los Angeles and London, 1987.

[20] Ragin, C. C., *Fuzzy-Set Social Science*, University of Chicago Press, Chicago, 2000.

[21] Rihoux, B., "Qualitative comparative analysis and related systematic comparative methods", *International Sociology* 21, 5 (2006).

[22] Rosser, J. B., Turquette, A. R., *Many-Valued Logics*, Amsterdam, North-Holland Publishing Company, 1958

[23] Fayyad, U. M., Piatetsky-Shapiro, G., Smyth, P., Uthurusamy, R. (eds.), *Advanced in Knowledge Discovery and Data Mining*, Menlo Park, California, 1996.

[24] Skvortsov, D. P., "On some methods of constructing the languages with quantifiers over tuples" (in Russian), *Semiotics and Informatics* 20 (1983): 102–126.

[25] Smullyan, R. M., *First Order Logic*, N.Y., Springer-Verlag, 1968.

[26] Strauss, A., Corbin, J., *Basics of Qualitative Research*, Thousand Oaks, London, New Dehli, SAGE Publications, 1998.

[27] Sudman S., Bradburn N., Schwartz N., *Thinking about Answers: The Application of Cognitive Processes to Survey Methodology*, San Francisco, Jossey-Bass Inc. Publishers, 1996.

[28] Vinogradov, D. W., “Formalizing plausible arguments in predicate logic”, *Automatic Documentation and Mathematical Linguistic* 34, 6 (2000): 6–10.
[29] Zerubavel, E., *Social Mindscape. An Invitation to Cognitive Sociology*, Harvard Univ. Press, 1997.

Victor K. Finn and Maria A. Mikheyenkova
All-Russian Institute for Scientific and Technical Information
Usievitcha 20
125190 Moscow, Russia
{finn,mmikh}@viniti.ru