ANALYSIS OF HUMAN COMMONSENSE REASONING PROCESSES IN PATTERN RECOGNITION

Xenia Naidenova

Abstract. Some examples of natural human common sense reasoning both in scientific pattern recognition problems and in solving logical games are given. An analysis of inference structure shows that inductive and deductive rules communicate in reasoning. An automated model for detecting the types of woodland from incomplete descriptions of some evidences is also given in this paper. The important part of this model is a small knowledge base of experts' knowledge about natural woodlands as biological formation.

Keywords: commonsense reasoning machine learning, rule-based inference, good diagnostic test

Conference: The paper is selected from XVth International Conference "Knowledge-Dialogue-Solution" KDS 2009, Varna, Bulgaria, June-July 2009

Introduction

We concentrate our attention on analyzing and modeling natural human reasoning in solving different tasks: pattern recognition in scientific investigations, logical games, and investigation of crimes.

An Example of Reasoning Process in Pattern Recognition

The investigation of human reasoning process in different real situations is an inevitable step foregoing any work on modeling this process on computers. For studying, we have taken the process of visual deciphering forest images. In this case, the features of forest regions are investigated under stereoscope by a decipher (operator or executer), registered on blanks, and analyzed by the use of the decision rules created in advance by the specialists based on previous explorations and experiences.

The attempt to automate the process of deciphering forest images leads, first of all, to investigatiтп algorithms used by the experts under visual deciphering these images (photographs).

So, we have analyzed, the dichotomous diagram (scheme) elaborated by an expert for Lena - Angara forest region under deciphering the types of forest plant conditions. The scheme is a decision tree the nodes of which are associated with the factors or attributes' values to be checked during the deciphering process. The sequence of checking the factors in the nodes of scheme is rigidly assigned and optimum in the sense that for recognizing any of the forest plant types that examined in the diagram it is required the smallest of all possible ways in this diagram.

Our studies show that a strict collection of attributes and a strict sequence of their use do not reflect adequately the processes of specialists' natural reasoning during deciphering of photographs. Decision trees help greatly to increase the productivity of the work of decipherers, but they decrease the number of correctly recognized objects.

Decision tree is easier for inexperienced decipherer and more difficult psychologically for experienced deciphers. We have examined two forms of the experts' knowledge representations: 1) the dichotomous scheme familiar to a decision tree, 2) the table of rules reflecting the links between the factors and the type of forest (the types of forest plant conditions) with the indication of the occurrence frequency of factors.

Two deciphers worked: Executer 1 having experience of work more than 10 years and Executer 2, which did not have an experience of forest deciphering, but he knew how to carry out the preliminary processing of photographs and to work with the simplest stereo - instruments.

Two regions have been chosen: the basin of rivers Lena - Angara (Region 1) and Khentey - Chikoy region (Zabaykalie and Mongolia) (Region 2). The results of deciphering are represented in Tables 1, 2. The data of the ground-based assessment have been taken as true. The trustworthiness of recognition was estimated for 214 parts of Region 1 and for 192 parts of Region 2.

For the first experienced Executor 1, in the familiar Region 1, % of correct answers with the use of a strict algorithm falls. Although the percent of correct answers grows in the unknown Region 2 but it grows insignificantly.

For the second inexperienced Executer 2, the use of strict algorithm both in the familiar and in the unknown region leads to an increase % of the correct answers, but the level of correct recognition remains still very low: 62-64%.

Table 1. The Results of deciphering the types of woodland

Executer	Region 1		Region 2	
	By rules	By decision tree	By rules	By decision tree
	% correct results	% correct results	% correct results	% correct results
1	88	79	66	69
2	43	62	49	64

Table 2. The Productivity of Executer 1 and 2

Executer	Region 1		Region 2	
	By rules	By decision tree	By rules	By decision tree
	The number of parts	The number of parts	The number of parts	The number of parts
1	100	326	92	289
2	65	209	70	238

Thus, in the absence of experience and of knowledge, it is unimportant what to use – table of rules or a strict algorithm, the application of algorithm leads only to an increase in productivity of labor of executer.

But if the executor possesses the experience and the knowledge of region (Executor 1 and Region 1), then he will recognize worse with the use of strict algorithm than without it, he psychologically rejects rigid diagram.

But in the unknown region, the behavior of Executer 1 is the same as the behavior of inexperienced one, but also, in this case, the strict algorithm does not lead to the strong improvement in the results - they are located approximately at the same level, as for inexperienced decipherer. Experienced specialist deteriorates his result in unknown regions. As far as inexperienced specialist it is unimportant, in what region to work.

If we remember that, from the point of view of the quantity of information, the set of rules is equivalent to the form of dichotomous scheme, it will become clear that only knowledge of additional information about the region helps executer to increase the percent of correct recognition and the information in the form of rules contributes to the more complete utilization of knowledge during the deciphering than rigid dichotomous diagram.

The more detail analysis of the reasoning of specialist during the deciphering makes it possible to formulate several properties of the natural reasoning process.

1. It is difficult for specialists to describe the sequence of his considerations (reasoning). Sometimes he can not describe it at all. Experienced specialist realizes that the sequence of his reasoning always different - it depends on the concrete situation, in which the deciphering occurs.

2. The same results can be established by using different collections of factors or attributes. For example, the degree of forest health can be refined, in some cases, according to closeness of forest canopy and the sizes of the projections of crowns, and, in other cases, with the use of the height above sea level, steepness of slopes and admixture of the deciduous species to the conifers.

3. Any factor, whose value is recognized (determined) in the course of reasoning, is involved, in turn, in the process of deciphering and used as a new factor for supporting or rejecting hypotheses. For example, after establishing the type of forest by the use of the landscape features, the specialist can use it for evaluating the degree of forest health and the composition of forest (forest structure).

4. The priority of recognizing characteristics of forest cannot be established previously. First of all, are examined the features with the greatest degree of manifestation on aero-cosmo-photographs. The greatest priority belongs to forbidding features, since they exclude the impossible solutions. Diagnostic features play the important role - they make it possible to divide the hypotheses, which appeared during reasoning.

5. Cause-effect relations are used in reasoning not only in the direction “from the reason to the consequence”, but also in the direction “from the consequence to the reason”. This reasoning generates the statement of the form: “if A is true, then at least B can be true”. Let us name this statement hypothetical one.

A process of reasoning covers, in general, introducing or deleting assumptions (values of attributes or factors), hypotheses (values of goal (sub-goal) attributes, factors, objects (class of objects)), measured or observed values of attributes or factors (they are considered to be established correctly). Assumption, hypotheses and recognized values of forest features as a whole are range by their degree of possibility. Only such signs are selected, which should be verified for confirmation or refutation of hypotheses and only these hypotheses.

As a whole the process of deciphering can be presented as follows. By known (o recognized) features and known casual relations, hypotheses are generated of values of unknown features of forest plant conditions. Hypotheses, by means of known causal relations, generate assumptions of values of new involved features. Assumptions are checked against an observed situation on photographs. Assumption can be supported or rejected. Hypotheses associated with rejected values of attributes or factors are deleted from consideration. A hypothesis is admissible if its description is consistent. So the reasoning process is continued until the values of the necessary characteristics of forest plant conditions and totality of the features connected with them are obtained. In particular case, the decision set is empty or it contains a certain set of decisions, which do not contradict an observed situation, but have the different degree of confidence.

When all hypotheses were rejected, this speaks, that the causal relations do not fully reflect situations in the region and the classification, carried out in the stage of studying region, is not successful.

A situation can occur when it is not possible to recognize necessary features on the photograph. In this case, it is possible of all hypotheses to select those, which coincide with the concrete situation in a maximum quantity of features. It is obviously that several hypotheses can be obtained each of which has the certain degree of probability.It is important that the process of natural reasoning makes it possible to estimate not only authenticity of conclusion, but also meaningfully to explain, as this conclusion was obtained.

Knowledge in this model is a system of coordinated links objects $\leftrightarrow$ classes of objects, classes of objects $\leftrightarrow$ properties, objects $\leftrightarrow$ properties. For instance, “all squares are rhombs”, “square is a rhomb”, “all the angles of rectangle are right”, “square is a rhomb all the angles of which is right”, “if the sun is in the sky and not raining, then the weather is good”, “conifers are pine-tree, fir-tree, cedar”. These connections have causal nature and can be formally expressed with the aid of implications. By commonsense reasoning we understand constructing and using the coordinated classification connections between objects, properties and classes. This understanding goes back to the work of Jean Piaget & Bärvel Inhelder (1959).

The use of these connections is based on the application of syllogisms as deductive reasoning rules. These are rules of everyday reasoning or commonsense reasoning. The construction of these connections is a field of the application of ML algorithms. Reducing these algorithms to the approximations of an assigned classification (partitioning) of a given set of objects’ examples gives the possibility to transform them into a model of reasoning in which inductive inference entails applying deductive commonsense reasoning rules.

The following types of rules are used for commonsense reasoning (Naidenova, 2007a):

INSTANCES (evidences) really observed. Instances serve as a source for inductive inference of generalized rules or implicative assertions.

IMPLICATIVE ASSERTIONS describe regular relationships connecting together objects, properties and classes of objects. We consider the following forms of assertions: implication (a, b, c → d), forbidden rule (a, b, c → false (never), diagnostic rule (x, d → a; x, b → not a; d, b → false), rule of alternatives (a or b → true (always); a, b → false), compatibility (a, b, c → VA, where VA is the occurrence's frequency of rule).

COMMONSENSE REASONING RULES (CRRs) are rules with the help of which implicative assertions are used, updated and inferred from instances. The deductive CRRs infer consequences from observed facts with the use of implicative assertions. These rules are the following ones: **modus ponens**: "if A, then B"; A; hence B; **modus ponendo tollens**: "either A or B" (A, B – alternatives); A; hence not B; **modus tollendo ponens**: "either A or B" (A, B – alternatives); not A; hence B; **modus tollens**: "if A, then B"; not B; hence not A; **generating hypothesis**: "if A, then B"; B; A is possible. The inductive CRRs are the canons formulated by J. S. Mill (1900): the method of agreement, the method of difference, the joint method of agreement and difference, the method of concomitant variations, and the method of residuum. These methods are not rules but they are the processes in which implicative assertions are generated and used immediately.

The Structure of a Small Knowledge Base for Inferring the Type of Woodland via an Analysis of Forest's Aerial Photographs

We describe a very simple structure of a knowledge base that is sufficient for our illustrative goal (see also Naidenova, 2007b). The knowledge base (KB) consists of two parts: the Attribute Base (AtB), containing the relations between problem domain concepts (classifications or ontology), and the Assertion Base (AsB), containing the expert's assertions formulated in terms of the concepts.

For example, let objects be a collection of trees such as asp, oak, fir-tree, cedar, pine-tree, and birch. Each name calls the class or the kind of trees (in a particular case, only one tree). Any set of trees can be partitioned into the separate groups depending on their properties. 'Kind of trees' will be the name of a classification, in which 'asp', 'oak', 'fir-tree', 'cedar', 'pine-tree', and 'birch' are the names of classes. Then, in the KB, 'kind of trees' will be used as the name of an attribute the values of which are 'asp', 'oak', 'fir-tree', 'cedar', 'pine-tree', and 'birch'. The link between the name of an attribute and the names of its values is implicative. It can be expressed by the following way: (<name of value1>, <name of value2>, ... , <name of value k >) → <name of attribute>, where the sign "→" denotes the relation "is a".

In our example (asp, oak, fir-tree, cedar, pine-tree, birch) → kind of trees, and, for each value of 'kind of trees', the assertion of the following type can be created: "asp is a kind of trees".

The set of all attributes' names and the set of all values' names must not intersect. This means that the name of a classification cannot simultaneously be the name of a class. However, this is not the case in natural languages: the name of a class can be used for some classification and vice versa. For example, one can say that 'pine-tree', 'fir-tree', 'cedar' are 'conifers'. But one may also say that 'conifers', 'leaf-bearing' are 'kinds of trees'. Here the word 'conifers' serves both as the name of a classification and as the name of a class. In this setting, class is a particular case of classification like object is a particular case of class. By using names in the way we do in real life we permit the introduction of auxiliary names for the subsets of the set of an attribute's values.

The AsB (Assertion Base) contains the expert's assertions. Each assertion links a collection of values of different attributes with a certain value of a special attribute (SA) that evaluates how often this collection of values appears in practice. The values of a special attribute are: always, never, rarely, and frequently. Assertions have the following form: (<name of value>, <name of value>, … , <value of SA>) = true.

For simplicity, we omit the word 'true', because it appears in any assertion. For example, the assertion "pine-tree and cedar can be found frequently in the meadow type of forest" will be expressed in the following way: (meadow, pine-tree, cedar, frequently). We also omit the sign of conjunction between values of different attributes and the sign of disjunction (separating disjunction) between values of the same attribute. For example, the assertion in the form (meadow, pine-tree, cedar, often) is equivalent to the following expression of formal logic: P((type of forest = meadow) & ((kind of trees = pine-tree) V (kind of trees = cedar)) & (SA = frequently)) = true.

Only one kind of requests to the KB is used: SEARCHING VALUE OF <name of attribute> [,<name of attribute>,...]. IF (<name of value>, <name of value>, ...), where "name of value" is the known value of an attribute, "name of attribute" means that the value of this attribute is unknown. For example, the request "to find the type of forest for a region with plateau, without watercourse, with the prevalence of pine-tree" will be represented as follows: SEARCHING VALUE OF the type of forest IF (plateau, without watercourse, pine-tree).

Inferring All Possible Hypotheses About the Type of Woodland from an Incomplete Description of Some Evidences

Let x be a request to the KB equal to:

SEARCHING VALUE OF type of woodland IF (plateau, without watercourse, pine-tree). Let the content of the Knowledge Base be the following collection of assertions:

AtB:

1. (meadow, bilberry wood, red bilberry wood ....) $\rightarrow$ types of woodland;
2. (pine-tree, spruce, cypress, cedars, birch, larch, asp, fir-tree) $\rightarrow$ dominating kinds of trees;
3. (plateau, without plateau) $\rightarrow$ presence of plateau;
4. (top of slope, middle part of slope, ….) $\rightarrow$ parts of slope;
5. (peak of hill, foot of hill) $\rightarrow$ parts of hill;
6. (height on plateau, without height on plateau) $\rightarrow$ presence of a height on plateau;
7. (head of watercourse, low part of watercourse, ….) $\rightarrow$ parts of water course;
8. (steepness $\geq 4°$, steepness $\leq 3°$, steepness $< 3°$, …) $\rightarrow$ features of slope;
9. (north, south, west, east) $\rightarrow$ the four cardinal points;
10. (watercourse, without watercourse) $\rightarrow$ presence of a watercourse.

AsB:

1. (meadow, pine-tree, larch, frequently);
2. (meadow, pine-tree, steepness $\leq 4°$, never);
3. (meadow, larch, steepness $\geq 4°$, never);
4. (meadow, north, west, south, frequently);
5. (meadow, east, rarely);
6. (meadow, fir-tree, birch, asp, rarely);
7. (meadow, plateau, middle part of slope, frequently);
8. (meadow, peak of hill, watercourse heads, rarely);
9. (plateau, steepness $\leq 3°$, always);
10. (plateau, watercourse, rarely);
11. (red bilberry wood, pine-tree, frequently);
12. (red bilberry wood, larch, rarely);
13. (red bilberry wood, peak of hill, frequently);
14. (red bilberry wood, height on plateau, rarely);
15. (meadow, steepness $< 3°$, frequently).

The process of reasoning evolves according to the following sequence of steps:

Step 1. Take out all the assertions t in AsB containing at least one value from the request, i.e. $t \in$ AsB and $t \cap x \neq \varnothing$, where x is the request. These are assertions 1, 2, 7, 9, 10, 11, and 14.

Step 2. Delete (from the set of selected assertions) all the assertions that contradict the request. Assertion 10 contradicts the request because it contains the value of attribute 'presence of water course' which is different from the value of this attribute in the request. The remaining assertions are 1, 2, 7, 9, 11, and 14.

Step 3. Take out the values of attribute 'type of woodland' appearing in assertions 1, 2, 7, 9, 11, and 14. We have two hypotheses: 'meadow' and 'red bilberry'.

Step 4. An attempt is made to refute one of the hypotheses. For this goal, it is necessary to find an assertion that has the value of SA equal to 'never' and contains one of the hypotheses, some subset of values from the request and does not contain any other value. There is only one assertion with the value of SA equal to 'never'. This is assertion 2: (meadow, pine-tree, steepness $\leq$ 4°, never). However, we cannot use this assertion because it contains the value 'steepness $\leq$ 4°' which is not in the request.

Step 5. An attempt is made to find a value of some attribute that is not in the request (in order to extend the request). For this goal, it is necessary to find an assertion with the value of SA equal to 'always' that contains a subset of values from the request and one and only one value of some new attribute the values of which are not in the request. Only one assertion satisfies this condition. This is assertion 9: (plateau, steepness $\leq$ 3°, always).

Step 6. Forming the extended request:

SEARCHING VALUE OF the type of woodland IF (plateau, without watercourse, pine-tree, steepness $\leq$ 3°).

Steps 1, 2, and 3 are repeated. Assertion 15 is involved in the reasoning.

Step 4 is repeated. Now assertion 2 can be used because the value 'steepness $\leq$ 4° is in accordance with the values of 'feature of slope' in the request. We conclude now that the type of woodland cannot be 'meadow'. The non-refuted hypothesis is "the type of woodland = red bilberry".

The process of pattern recognition can require inferring new rules of the first type from data, when it is impossible to distinguish inferred hypotheses. In general, there exist two main cases to learn rules of the first type from examples in the process of pattern recognition: i) the result of reasoning contains several hypotheses and it is impossible to choose one and only one of them (uncertainty), and ii) there does not exist any hypothesis.

The Analysis of Inference Structure. The Interactive of Deductive and Inductive Reasoning Rules in Solving Pattern Recognition Problems

It is not difficult to see that steps of reasoning in our example realize the deductive CRRs.

Step 1 performs **Introducing Assertions** into reasoning process. This step is an element of common sense reasoning the task of which is the drawing of knowledge into reasoning process. The selected assertions form (constitute) the meaningful context of reasoning or the region reasoning.

Step 2 performs **Deleting Assertions** from reasoning process. This step uses Rule of Alternative. If an assertion contains a value of attribute not equal to the value of the same attribute in the request, then, by Rule of alternative, the value of this attribute and the assertion containing this value must be deleted from consideration. Consequently, step 2 narrows the context of reasoning by Deleting Values of Attributes and Deleting Assertions from considerations.

Step 3 performs **Introducing Hypotheses** about goal attribute values. These hypotheses are all values of goal attributes appearing in the selected assertions. Hence, the source of hypotheses is the context of reasoning.

Step 4 performs **Deleting Hypotheses** by means of using Interdiction (Forbidden) rules. Let H be a hypothesis and FR be forbidden rule 'H,{Y} $\rightarrow$ never', and {X} be a request, where X, Y – collections of attributes values. If {Y} $\subseteq$ {X}, then hypothesis H is disproved.

Step 5 performs **Introducing Assumptions** about values of attributes. Let A be a value of an attribute not contained in the request and IR be the rule 'A,{Y} $\rightarrow$ always', and {X} be a request, where X, Y – collections of attributes values. If {Y} $\subseteq$ {X}, then the request can be extended as follows: {X'} = {X} $\cup$ A.

For extending the request, it is possible to use **Compatibility Rules** and **Diagnostic Rules** (Step 5). Assumptions, introduced by a Compatibility Rule or Diagnostic Rule can be checked against the evidence by means of photographs. Assumptions contradicting the visible image of forest are deleted from considerations.

Step 6 performs **Forming the Extended Request** in accordance with each not disproved hypothesis. With new extended requests for each hypothesis, the steps 1 – 6 are performed untill only one hypothesis remains.

Calculating the estimate VA requires special consideration. In any case, to do this, we need the function which would be monotonous, continuous and bounded above. We introduce some limitations on using compatibility rules: if value v(A) of an attribute A has been determined by a compatibility rule R with VA equal to Z, then value v(A) must be inferred independently with the same or higher value of VA and by means of the rules containing a combination of attributes not intersecting with the combination of attributes associated with compatibility rule R.

Really, the scheme of knowledge base always permits to do so. Usually the following knowledge base scheme is created by the specialists:

Landscape features $\Rightarrow$ Type of woodland (forest plant conditions);

Morphological features of forest $\Rightarrow$ Type of woodland (forest plant conditions);

Landscape features $\Rightarrow$ Predominant type (species) of trees;

Morphological features of forest $\Rightarrow$ Predominant type (species) of trees;

Landscape features $\Rightarrow$ The productivity of forest (the class of quality);

Morphological features of forest $\Rightarrow$ The productivity of forest (the class of quality);

Type of woodland (forest plant conditions $\Leftrightarrow$ Predominant type (species) of trees:

Type of woodland (forest plant conditions) $\Leftrightarrow$ Predominant type (species) of trees; the productivity of forest (the class of quality).

This scheme corresponds to ideas about the forest as about the biological unity, in which climatic conditions, soil, moisture, watercourses, relief, the growing trees and the associated plants are consistent. The sign $\Rightarrow$ means that the attributes in the left parts of rules determine functionally the attribute in right parts of rules. The sign $\Leftrightarrow$ means that attributes in the left and right parts of rules are strongly interconnected. So, if the type of woodland was determined by landscape features and the predominant type of trees was inferred through the type of woodland with certain value Z of VA, then this type of trees must be supported, for example, by morphological features of forest with value of VA not less than Z.

If the number of hypotheses is more than 1 and no one of them can be disproved, then we deal with a difficult situation of the inference and it is necessary to resort to the aid of the diagnostic rules.

Let r be a diagnostic rule such as ‘X, d $\rightarrow$ a; X, b $\rightarrow$ z’, where ‘X’ is true, and ‘a’, ‘z’ are hypotheses or possible values of some attribute, say A. In this rule, X is a combination of attribute values which can not distinguishes hypotheses ‘a’ and ‘z’ ($z \neq a$); d, b are values of an attribute which distinguish these hypotheses on condition that ‘X’ is true. If ‘X’ is included in the request and the pair of considered hypotheses coincides with hypotheses in the diagnostic rule, then this rule is applicable to the situation of the inference. If value of attribute A is not yet determined, then d and b become the assumptions to be inferred or checked against the evidence. Examples 1 and 2 present some diagnostic rules.

Example 1. If, with familiar landscape features, there are two hypotheses ‘bilberry’ and ‘red bilberry’ of the type of woodland, then, with the highest possibility, if the predominant type of trees is cedar, then the type of woodland is red bilberry, and if the predominant type of trees is pine-tree, then the type of woodland is bilberry.

Example 2 (for aero- photo produced by the survey of the small scale). With other equal morphological features, if it is observed the flat structure of curtains, uniform granularity and the equal height of trees, then the species of trees is pine tree; if it is observed the uneven structure of curtains, uneven granularity and different height of trees, then the species of trees is larch.

However if the inferring or observing of indispensable values of diagnostic attributes was not succeed, then it is necessary to address to inductive inference of a new portion of reasoning rules of the first kind for extending the Knowledge Base.

Now consider a situation when the initial context of reasoning does not contain any hypothesis about the value of goal attribute. In this case, it is natural to extend the request by the use of Introducing Assumptions (step 5) taking as a goal any attribute from the reasoning context. Of course, the equality to 0 of hypotheses' number can indicate the need of expanding the very base of the knowledge.

The result of inferring can be considered satisfactory if the number of hypotheses about woodland type is equal 1 and it is consistent with predominant type of trees and the class of quality. If the inference terminates with several hypotheses or the number of hypotheses is equal to 0, then the Knowledge Base is incomplete and it is necessary to expand it. For this goal, the inductive CRRs are used.

Inductive Extension of Incomplete Knowledge Based by Using the Inductive Reasoning Rules

The deductive reasoning rules act during an inference process by means of extending incomplete descriptions of some evidences with disproving the impossible extensions. This extension is based on good knowledge of the forest regions and the interconnections between the main forest characteristics and the natural factor such as climate, soil, relief, watercourses ect. But the conditions of inference depending on the quality of forest images, the type of instruments used for aero-cosmos-survey of earth surface introduce a lot of uncertainties in the inference process. That's why it is indispensable to draw into reasoning the steps of inductive inference of new implicative assertions. Inductive reasoning can requires introducing in reasoning new attributes, new factors and new observation both on the surface of earth and by the use of the instrument surveys of forest region on earth surface.

Two variant of drawing inductive inference in reasoning are thinkable: 1) using a part of existing Knowledge Base which was out of knowledge context of reasoning if this part contains a set of observations potentially applicable as a source of new implicative assertions about difficult situations of the previous reasoning process; 2) to initiate a new investigation of the forest region for collecting observations to enrich the Knowledge Base.

In the first variant, we could do the purpose-directed steps of inductive reasoning, in the second variant; we have to interrupt the reasoning process.

Let A, B be two hypotheses under investigations. The purpose-directed inductive reasoning means that we must choose in KB the instances containing a set of observed attributes' values of request, say X, then, among these instances, we must select instances in which phenomenon A occurs but phenomenon B does not occur. These two sets of instances must be compared. The attributes' values in which the instances of these sets are different are diagnostic ones; they can form the new diagnostic rules for distinguishing hypotheses A and B.

We can find a lot of good examples of natural human deductive and inductive reasoning in the novels of the famous English writer Conan Doyle, who is the real begetter of the detective-fiction genre as we know it. In the novel 'The Adventure of the Second Stain', Sherlock Holmes knows several international spies which could possess the documents stolen from the Foreign Ministry (Office). There were several men under suspicion with equal possibility to steal the documents (Oberstein, La Rothiere, and Eduardo Lucas), but one of these men differed from all the others by the fact that he lived near the Foreign Ministry (Office). Finally, the following reasoning helps Sherlock Holmes to discover the thief. Holmes said: 'There is one obvious point which would, in any case, have turned my suspicions against Lucas. Godolphin Street, Westminster, is only a few minutes' walk from Whitehall Terrace. The other secret agents whom I have named live in the extreme West End. It was easier, therefore, for Lucas than for the others to establish a connection or receive a message from the European Secretary's household".

In the novel 'Murder into Abby-Grange', there are three glasses, from which, supposedly, men drunk vine. In one of the glasses there was sediment, in two others sediment is absent. Holmes searches for the explanation, which would satisfy this difference in the glasses. Possibly two versions of the explanation: 1) in two glasses, they shook vine before using, while, in the third glass, they did not shake up vine; 2) they drunk only from two glasses, they poured off the remainders in the third glass. Mentally (in mind) Holmes constructs usual situations which could explain the difference between the glasses.

In the novel 'The Adventure of the Yellow Face' (Doyle, 1992), there are two hypotheses and the second one is supported by an assumption, that the inmates were warned of Grant Munro's coming. With the second assumption, the way of Holmes' reasoning can be described as follows:

Facts (evidence):

The inmates of the cottage do not want to meet Grant Munro;

Grant Munro returned at home and spoke with the maid;

Grant Munro saw the maid with whom he had been speaking running across the field in the direction of the cottage;

Grant Munro meets his wife and the maid hurrying back together from the cottage;

Grant Munro entered the cottage;

The cottage was absolutely empty.

Assertions:

If one does not want to meet a person and he is warned that this person is going to visit him, then he conceals himself or goes away;

If one only conceals himself, then he must return;

If one goes away, then his house will be permanently deserted;

If one knows something, then he can say it somebody.

Reasoning:

The maid knows that Grant Munro returned at home, then, knowing this, she visited the cottage, hence she warned the inmates and the wife of Grant Munro that he returned at home.

The inmates do not want to meet Grant Munro hence they concealed themselves or went away.

Holmes says to Grant Munro: «If the cottage is permanently deserted we may have some difficulty, if on the other hand, as I fancy is more likely, the inmates were warned of your coming, and left before you entered yesterday, then they may be back now, and we should clear it all up easily».

A dialog between Holmes and Watson is very remarkable with respect to what must be a good reasoning:

Holmes: - What do you think of my theory?

Watson: - It is all surmise.

Holmes: - But at least it covers all the facts. When new facts come to our knowledge, which cannot be covered by it, it will be time enough to reconsider it. This strategy is supported by the novel 'Murder into Abby-Grange'.

Sherlock Holmes begins his investigation from the study of facts. There is an initial hypothesis, but some facts will not be coordinated with this hypothesis and the story of witnesses. Story contradicts the usual and most probable ideas (rules) about the behavior of the robbers. Facts attest to the idea that the robber had to know house and its inhabitants. Holmes returns to the place of crime and he will more thoroughly inspect it. Thus he obtains facts more newly. New facts make it possible to advance new hypotheses about the nature and the physical force of robber and about the fact that he acted alone. But this makes possible for Holmes to conclude that the lady speaks untruth.

An Example of Reasoning Process in the Game 'The Letter's Loto'

In logical game "The letter's Loto" (Bizam, 1978), oracle (one of players) thinks of a word with the fixed number of letters in it. A guesser names a word with the same number of letters. Then oracle says how many letters in this word are correct. A letter is guessed correctly if it is equal to the letter taking up the same position in the thought word. For example, the oracle thought a word of 3 letters and the guesser said the following words in sequence: LAP HAP HAM HAT RAT CUR (Table 3). Numbers in the last line of Table 3 are the estimations of oracle.

A set of reasoning rules are considered for guessing letters (Table 4). Each rule is based on comparing words and differentiating situations of a game. These rules localize positions with correctly and not correctly guessed letters using the oracle's estimations. Since there can be only one correct letter in each position of a word the correctly guessed letters are used in the same positions in the following trials (words) of the guesser, and incorrect letters are replaced by new ones. These steps of guessing can be reduced to classifying letters in each word and in each position of a word into two classes "correct letters" and "incorrect letters".

Table 3. The situation of the game

L	H	H	H	R	C
A	A	A	A	A	U
P	P	M	T	T	B
1	1	1	2	2	1

Table 4. The Reasoning Rules

№	Reasoning rule	The meaning of rule
1	If changing a letter only in one position does not lead to changing the estimation of oracle, then in this position both letters are incorrect (before and after changing)	Diagnostic rule
2	If two words are different in letters only in one position and the estimations of oracle of these words are different by 1, then the letter in this position is correct in the word with greater estimation of oracle and it is incorrect in the word with smaller estimation of oracle	Diagnostic rule
3	If, in a position of word, the letter is identified as correct, then all other letters in this position are incorrect	Letters' classification in a position of words
4	If, in a position of word, the letter is identified as incorrect, then it is incorrect in any word in this positiono	Letters' classification in a position of words
5	If, in a word, the estimation of oracle is equal to the number of letters remaining after deleting all incorrect letters, then these letters are correct in this word	Localization of correct letters in words
6	If the number of correct letters in a word is equal to the estimation of oracle, then all other letter in this word are incorrect	Localization of incorrect letters in words
7	If Let W1 and W2 be two words of lengths H and with estimations H1 and H2 of oracle respectively. If the sum C = H1 + H2 is greater than H than, then there exist at least C-H positions with coinciding letters and these coinciding letters are correct	Localization of correct letters in words
8	If, for two words, the estimations of oracle are different by K and there are T positions with coinciding letters in these words, then the change of the estimation of oracle is associated with H – T remaining positions of words	Localization of correct (incorrect) letters in words
9	If, in two words, the letters in all positions are different, then the sets of positions with correct letters in these words do not intersect	Localization of correct (incorrect) letters in words
10	If changing the estimation of oracle by K is followed by changing K letters in a word, then, in the word with greater estimation, these K letters are correct	Generalization of rule 2
11	If, in some positions, the letters are identified as correct, then a subtask can be considered with the length of words less by the number of positions with correct letters. The estimations of oracles must be recalculated	Reducing a task to the subtask with smaller dimension
12	The previous rules are applicable for subsets of letters of words with recalculating the estimations of oracle	Reducing a task to the subtask with smaller dimension

Return to our example. Since only one letter is changed in words LAP and HAP and the estimation oracle is not changed (rule 1) letters L and H in the first position are not correct. One of letters A or P must be correct (appearing hypotheses).

Compare words HAM and HAP. Since only one letter is changed in these words but the estimation of oracle is not changed (rule 1) letters M and P in the last position are not correct. We know that H in the first position is not correct. Hence letter A in the second position is correct (rule 5).

Compare words HAT and HAM. Since we have change the last letter and the estimation of oracle is change to 2, letter T in the last position is correct (rule 2).

Comparing words HAT and RAT and the fact that the estimation of oracle is not changed imply that letter R in the first position is not correct (rule 1).

Consider word CUB. We know that letters U and B in this word are incorrect, but the estimation of oracle is equal to 1. Hence letter C in the first position is correct and the word thought by oracle is CAT (rule 5). The result is in Table 5.

Table 5. Example 1:

L	H	H	H	R	C	C
A	A	A	A	A	U	A
			T	T	B	T
1	1	1	2	2	1	

A guesser gives words in sequential manner and he can choose a sequent step from a set of possible ones.

The Interaction of Deductive and Inductive Reasoning Rules in Solving Logical Problems

Besides the rules, given in the table 4, the process of reasoning can contain such logical rules and methods as the introduction of assumptions, the selection of versions, proof by means contradiction (the deductive reasoning). Important role play quantitative assessments - word length, number of positions with the accurate or incorrect letters, difference between the number of accurate positions and the estimation of oracle.

There is a certain freedom in selecting position for testing the new letter and in selecting rule, which can be used in the prevailing situation. This fact means that it is necessary to recognize the applicability of rules. For this goal, the description of situations must be well structured, sufficiently complete in order to reflect all intrinsic properties of situations. Specifically, the selection of rules and the subsequent changes of the letters in words distinguish one version of solution of problem from another.

Thus, the deduction is not separated from the inductive steps of reasoning. The inductive methods of reasoning are connected also with the generalization of rules for plays with words of any length or the decomposition of a task into subtasks. As a whole, it is possible to say that the reasoning are well organized, when there is a good decomposition of task to the sub-tasks and a good interrelation between these sub-task is systematically performed.

Conclusion

This chapter examined the problem of human natural reasoning in real world situations. One of the fundamental questions of natural reasoning is using the interaction of deductive and inductive reasoning rules. The analysis of several examples selected from different fields of thinking shows that natural reasoning is based on both deductive and inductive rules of reasoning.

Bibliography

Bizam, G., & Herczeg, J. (1978). Many-colored Logic. Moscow: Publishing House Mir (in Russian).

Doyle, A. C. (1992). The Adventures of Sherlock Holmes. Great Britain: Wordsworth Editions Limited.

Naidenova, X. A. (2007a). Reducing a class of machine learning algorithms to logical commonsense reasoning operations. In G. Felici, & C. Vercellis (Eds), Mathematical Methods for Knowledge Discovery and Data Mining (pp. 41-64). Hershey – New York: ISR Press.

Naidenova, X.A. (2007b). A model of rule-based logical inference. In Proceedings of XIII-th International conference "Knowledge-Dialog-Solution". Sofia – ITHEA. Vol. 2. 2007. P.379-388

Mill, J. S. (1900). The system of logic. Moscow, Russia: Russian Publishing Company "Book Affair".

Piaget, J., & Inhelder, B. (1959). La genèse des structure logiques élémentaires : classifications et sériations. Neuchâtel: Delachaux & Niestlé.

Author's Information:

Naidenova Xenia Alexandrovna - Military medical academy, Saint-Petersburg, Stoikosty street, 26-1-248,
e-mail: naidenovaxen@gmail.com