Application of Risk Theory Approach to Fuzzy Abduction

V.N. Tsypyschev(✉)

Moscow Technological University,
78, Vernadsky Avenue, Moscow 119454, Russian Federation
tsypyschev@yandex.ru

Abstract. In this article, learning under the absence or incompleteness of some facts or premises about the problem domain is considered. This task does not fall under semi-supervised learning in the classical sense, because there it is assumed that the target signals are known and correct. The assumption of incompleteness is, however, natural for pattern recognition, e.g. for medical diagnostics.

In such a situation, it is natural to base a learning process on abductive reasoning instead of induction or transduction. It is then important to have a quality criterion for the state of knowledge on the object to be studied.

Previously, to reconstruct missing training data, a fuzzy logical approach to the application of the abductive reasoning method was studied. Now, fuzzy abduction is considered from a risk-theoretical point of view.

As a result, in addition to the fuzzy abduction method, a general algorithm is suggested for finding the true state of the object to be studied in the case when known hypotheses about its state are mutually far from each other.

Keywords: Fuzzy abduction · Fuzzy systems · Logical deduction · Risk theory

1 Introduction

During learning processes, sometimes cases appear which may be characterized by incompleteness, implausibility or just absence of some necessary information. Not only, as in semi-supervised learning, target signals might be absent. It may be that more generally knowledge of the object to be studied is too incomplete to infer from known facts.

In this case, the task of learning should be seen as a task of elicitation, eduction or establishing of causal relationships. Unfortunately, in this case, an application of any model is impossible because modelling, as usual, demands presence and completeness of information about the object to be modelled.

In memory of T.Y. Morozova.

R. Silhavy et al. (eds.), *Cybernetics and Mathematics Applications in Intelligent Systems*, Advances in Intelligent Systems and Computing 574, DOI 10.1007/978-3-319-57264-2_2

It is then natural to apply abduction as a reasoning method to the considered problem. This method may be viewed as a special kind of inference generating supplements. Supplements consist of auxiliary disjunction forms necessary to construct a disjunction form of additional premises, which is essential for successful deductive inference.

For this, an analytical model of learning explanations can be used, and from an algorithmic point of view, one can use a resolution method as developed in [1–3,7,8,10–12,14,16–19].

Abduction may be considered as reverse deduction. In classical deductions, it is assumed [9] that the facts are true and the inferences rules by which conclusions are drawn are known and therefore the conclusions one draws are true. Unlike an application of deduction, an application of abduction is characterized by incomplete knowledge of facts and the necessity to reconstruct the cause of known output. For this, the deductive inference rule is transformed into the new abduction rule, which can be stated as follows: If the conclusion Q is true and P causes Q then this suggests that P might also be true. Previously, this approach was for example studied in [6] and many other works referenced above.

Let us assume that there are reasons to construct a decision support system, wherein it is necessary to infer online from some incomplete set of not entirely trusted facts by applying a particular a priory rule given by an expert, resulting in a conclusion and an according reasoning supporting the conclusion. To give an example: a decision-making person supervising some process must have plausible explanations in an abnormal case. It is then obvious that an application of the abduction method might be fruitful.

This approach mandates quality criteria for explanations supporting the conclusions. For example, the above decision-making person must have quantized evaluations of plausibility of each of the explanations or, at least, an ordering of the explanations according to their plausibility.

With inductive learning one searches for the state of the object in the search space, and the method consists in an application of inductive inference rules to a fixed set of initial patterns. This however falls short of what one wishes to obtain here.

This suggests to apply abductive reasoning to construct the necessary decision support system. But then the following problem arises: The learning process is not a homogeneous process, and the quality of generated explanations is not constant in time and depends crucially on the quantity of absorbed information via the learning process. The decision support system as part of the learning system must be controlled permanently with respect to the quality of what has been learned. We call the current state of the learning system studying an object the *current state of knowledge regarding the object* or, shortly, *state of knowledge.*

This suggestion is necessary because the method of abductive learning is very complicated, not properly investigated, and may be only roughly represented by inductive learning. On the other hand, the decision-making person has to apply the most plausible explanation. Hence, it is necessary to suggest some additions to classical abduction.

2 Methods

To start with, we consider the case that $m+1$ simple hypotheses $H_j, j=\overline{0,m}$, are suggested to determine the true position of an object under learning in the search space. Each hypothesis consists in that: if observing position of object $\bar{x}=(x_1,\ldots,x_n)$ falls to domain X_k of the search space then decision γ_k is adopted and it means that this position $\bar{x}$ corresponds to state of knowledge $S_k, k=\overline{0,m}$.

To construct the decision making rule, we shall use the criterion of minimal average risk [5,15].

An application of any prior established decision making rule involves the possibility of false decision because of the probabilistic nature of the considered object. The observed sample of explanations $\bar{x}=(x_1,\ldots,x_n)$ may fall into the domain X_k corresponding to decision γ_k that the statement S_k is true though indeed this sample corresponds to other state S_j, $j\neq k$. The presence of not only true but also of false decisions in the sequence of decisions is a price for making decisions under conditions of incomplete information. Consequences of false decisions are accounted for by a function (matrix) of losses, which ties with every false decision, i.e. a pair (S_j,γ_k), $j\neq k$, the payment $\ddot{I}_{j,k}=\ddot{I}(S_k,\gamma_j)>0$, and with right decision the payment $\ddot{I}_{j,j}=\ddot{I}(S_j,\gamma_j)<\ddot{I}_{j,k}, k\neq j$.

An application of a certain decision making rule means nothing less than that the sample space is divided into domains $\{X_k\}$ and corresponding decisions $\{\gamma_k\}$ are given. For a given state S_k, an average value of losses is equal to the average value of losses in the sample space (mathematical expectation):

$$r_j=\sum_{k=0}^{m}\ddot{I}_{j,k}P(\gamma_k/S_j)=\sum_{k=0}^{m}\ddot{I}_{j,k}P(\bar{x}\in S_k/S_j),$$

wherein $P(\gamma_k/S_j)$ is the conditional probability of getting samples to domain X_k under the condition that the true state is S_j. The conditional average value of losses r_j for state S_j is known as a conditional risk [5].

Taking an average conditional risk for all states S_j, we obtain

$$R=\sum_{j=0}^{m}r_jp_j=\sum_{j=0}^{m}\sum_{k=0}^{m}p_j\ddot{I}_{j,k}P(\bar{x}\in X_k/S_j), \tag{1}$$

wherein p_j is a prior probability of S_j.

This value may be seen as a quality criterion of the decision making rule, consisting of partitioning the sample space into m nonintersecting domains and assigning to each of the domains X_k a decision γ_k that the hypothesis H_k is true.

The probability that the observed sample $\bar{x}$ will entail accepting of decision γ_k under condition the hypothesis H_j is true is equal to

$$P(\gamma_k/H_j)=P(\bar{x}\in X_k/S_j)=\int_{X_k}W(\bar{x}/S_j)d\bar{x}. \tag{2}$$

If we substitute (2) into (1), we obtain the value of the average risk

$$R = \sum_{j=0}^{m} r_j p_j = \sum_{j=0}^{m}\sum_{k=0}^{m} p_j \ddot{I}_{j,k} \int_{X_k} W(\bar{x}/S_j)d\bar{x},$$

which depends on the partitioning the sample space into the domains $X_k, k = \overline{0,m}$. This means that the value of R is a quantized measure of quality of the decision rule.

Now a good criterion to determine the optimal selection of rule-making consists in the minimization of the value of the average risk R.

3 Main Results

Before the suggested contribution is given, we explain the method by an example:

Example 1
Consider the matrix of losses for the case of two hypothesis:

$$\ddot{I} = \begin{pmatrix} \ddot{I}_{0,0} & \ddot{I}_{0,1} \\ \ddot{I}_{1,0} & \ddot{I}_{1,1} \end{pmatrix} \tag{3}$$

wherein $\ddot{I}_{1,0} > \ddot{I}_{0,0} \geq 0,\ \ddot{I}_{1,0} > \ddot{I}_{1,1} \geq 0$, the rows correspond to hypothesis H_0, respectively hypothesis H_1 and the columns correspond to decisions $\gamma_k, k = \overline{0,1}$. The right solution costs are located on the main diagonal, losses for the wrong decisions are located on the side diagonal. The average value of losses (average risk) is equal to

$$R = qr_0 + pr_1, \tag{4}$$

wherein

$$r_0 = \ddot{I}_{0,0}P(\gamma_0/H_0) + \ddot{I}_{0,1}P(\gamma_1/H_0) = \ddot{I}_{0,0}(1-\alpha) + \ddot{I}_{0,1}\alpha, \tag{5}$$

$$r_1 = \ddot{I}_{1,0}P(\gamma_0/H_1) + \ddot{I}_{1,1}P(\gamma_1/H_1) = \ddot{I}_{1,0}\beta + \ddot{I}_{1,1}(1-\beta) \tag{6}$$

are conditional risks corresponding to states H_0, H_1 respectively, α is the probability of type I error, i.e. the probability of rejecting a correct hypothesis H_1 and accepting incorrect hypothesis H_0 (false negative), β is a probability to reject a correct hypothesis H_0 and to accept an incorrect hypothesis H_1 (false positive, probability of type II error).

Substituting (5) and (6) into (4), we obtain that

$$R = q\ddot{I}_{0,0} + p\ddot{I}_{1,0} + q(\ddot{I}_{0,1} - \ddot{I}_{0,0})\alpha - p(\ddot{I}_{1,0} - \ddot{I}_{1,1})(1-\beta). \tag{7}$$

The dependence of an average risk on the domain X_1 is expressed via values α and $1-\beta$. Let us substitute them into (1):

$$R = q\ddot{I}_{0,0} + p\ddot{I}_{1,0} - \int\limits_{X_1} \left[p(\ddot{I}_{1,0} - \ddot{I}_{1,1})W(\bar{x}/S_1) - q(\ddot{I}_{0,1} - \ddot{I}_{0,0})W(\bar{x}/S_0)\right] d\bar{x}, \tag{8}$$

where $W(\bar{x}/S_0), W(\bar{x}/S_1)$ are likelihood functions.

Since $q\ddot{I}_{0,0} + p\ddot{I}_{1,0}$ is a constant term, an average risk R gets its minimal value under condition:

$$\forall \bar{x} \in X_1 \; p(\ddot{I}_{1,0} - \ddot{I}_{1,1})W(\bar{x}/S_1) \geq q(\ddot{I}_{0,1} - \ddot{I}_{0,0})W(\bar{x}/S_0),$$

That is, the set X_1 may be determined as

$$X_1 = \left\{ \bar{x} \; : \; \frac{W(\bar{x}/S_1)}{W(\bar{x}/S_0)} \geq \frac{q}{p} \cdot \frac{(\ddot{I}_{0,1} - \ddot{I}_{0,0})}{(\ddot{I}_{1,0} - \ddot{I}_{1,1})} \right\}.$$

The function

$$l(\bar{x}) = \frac{W(\bar{x}/S_1)}{W(\bar{x}/S_0)}$$

is a likelihood ratio and represents a non-negative random variable obtained by transformation $z = l(\bar{x})$, i.e. by transformation mapping points of n-dimensional sample space into $\mathbb{R}_+$.

End of Example 1

Using the previous example as a basis and continuing by induction, one obtains a proof of the following Theorem:

Theorem 1. *Let assume that the positions S_k of the object O in a search space X are determined by the sample $\bar{x} = (x_1, \ldots, x_n)$. Let us also assume that it is possible to generate $m+1$ distinguishable simple hypotheses $H_j, j = \overline{0,m}$ with distribution density functions $W(\bar{x}/S_j), j = \overline{0,m}$ and suppose that the true state S of an object under consideration is determined by appropriate decision making.*

Than the decision making rule based on minimization of the average risk R may be constructed by division of the sample space X into $m+1$ non-intersected domains $X_0, X_1, \ldots, X_m$ according to this rule: the domain $X_k, k = \overline{1,m}$, is the set of solutions of m linear inequalities

$$\sum_{i=0}^{m} (\ddot{I}_{i,j} - \ddot{I}_{i,k}) \frac{p_i \cdot W(\bar{x}/S_i)}{p_0 \cdot W(\bar{x}/S_0)} \geq 0, j = \overline{0,m}, j \neq k,$$

$$X_0 = X \setminus \bigcap_{k=1}^{m} X_k,$$

and the state S_k is adopted as true if and only if $\bar{x} \in X_k$.

Theorem 1 may be simplified and, moreover, the decision making rule may be transformed to operating in a sample space of fixed dimension:

Theorem 2. *Under the conditions of Theorem 1, let*

$$y_i = \frac{p_i}{p_0} l_i(\bar{x}) = \frac{p_i \cdot W(\bar{x}/S_i)}{p_0 \cdot W(\bar{x}/S_0)}.$$

Then the set $\tilde{X}_k, k = \overline{1,m}$ *is determined by intersection of planes in m-dimensional space*

$$\sum_{i=1}^{m}(\ddot{I}_{i,j} - \ddot{I}_{i,k})y_i \geq \ddot{I}_{0,k} - \ddot{I}_{0,j}, j = \overline{0,m}, j \neq k,$$

$$\tilde{X}_0 = \tilde{X} \setminus \bigcap_{k=1}^{m} \tilde{X}_k,$$

and the state S_k *is adopted as true iff* $\bar{y} \in \tilde{X}_k$.

4 Conclusions

The question of obtaining plausible knowledge about the true position of some object in a learning process under the condition of implausible or uncertain information used in learning was considered. Because usual logical-based reasoning is inapplicable (see, e.g. [4]), a method of quasi-abduction reasoning about the true state of the considered object based not on a logical but on a risk-theoretic approach was suggested. Namely, it was suggested to reduce this problem to the problem of minimizing the risk associated with the decision to be made. This problem was then reformulated as a linear programming problem in a space of fixed dimension. An algorithm for decision-making is provided. The next step of investigation is to provide an algorithm for generating hypotheses about the current state of knowledge regarding an object studied with an abductive learning process.

Author used ideas of [13, 20]. Also he's very grateful to Claus Diem for attention paid to this work.

References

1. Belohlavek, R.: Pavelka-style fuzzy logic in retrospect and prospect. Fuzzy Sets Syst. **281**(15), 61–72 (2015)
2. Cheng, C.-L., Lee, R.C.-T.: Symbolic Logic and Mechanical Theorem Proving. Academic Press, New York (1973)
3. Dubois, Didier, Lang, J., Prade, H.: Fuzzy sets in approximate reasoning, Part 2: logical approaches. Fuzzy Sets Syst. **40**(1), 203–244 (1991)
4. Dubois, D., Esteva, F., Godo, L., Prade, H.: Fuzzy-set based logics - an history-oriented presentation of their main developments. In: Handbook of the History of Logic 8. The Many Valued and Nonmonotonic Turn in Logic, pp. 325–449. Elsevier (2007). ISBN 978-0-444-51623-7
5. Roeser, S., Hillerbrand, R., Sandin, P., Peterson, M. (eds.): Handbook of Risk Theory: Epistemology, Decision Theory, Ethics, and Social Implications of Risk. Springer, Dordrecht (2012)
6. Ivanova, I.A., Morozova, T.Yu.: Logical conclusion in indistinct systems and indistinct abduction. In: 8th Open German-Russian Workshop "PATTERN RECOGNITION and IMAGE UNDERSTANDING" OGRW-8, pp. 96–99 (2014)

7. Ivanova, I., Morozova, T., Nikonov, V., Nikolaev, A.: Fuzzy gestures recognition method in development of contact-less interfaces. Int. J. Adv. Stud. **4**(1), 27–31 (2014)
8. Kim, C.S., Kim, D.S., Park, J.S.: A new fuzzy resolution principle based on the antonyms. Fuzzy Sets Syst. **113**, 299–307 (2000)
9. Kleene, S.C.: Mathematical Logic. Wiley, New York (1967)
10. Lee, C.T.: Fuzzy logic and the resolution principle. J. ACM **19**(1), 109–119 (1972)
11. Leonenkov, A.V.: Fuzzy Modeling in MATLAB and fuzzyTECH Environment, 278 p. BHV, Petersburg, Saint-Petersburg (2013)
12. Mendel, J.M., John, R.I., Liu, F.: Interval type-2 fuzzy logic systems made simple. IEEE Trans. Fuzzy Syst. **14**(6), 808–821 (2006)
13. Pankov, V.L.: Stimulation mechanism effectiveness and potential level of satisfaction of the needs of the employee. HERALD of MSTU MIREA **4**(1), 288–291 (2015)
14. Pedrycz, W., Reformat, M.: Evolutionary fuzzy modeling. IEEE Trans. Fuzzy Syst. **11**(5), 652–665 (2003)
15. Sakawa, M., Nishizaki, I., Uemura, Y.: Fuzzy programming and profit and cost allocation for a product and transportation problem. Eur. J. Oper. Res. **131**(1), 1–15 (2001)
16. Zadeh, L.A.: The concept of a linguistic variable and its applications to approximate reasoning, part I. Inf. Sci. **8**, 199–249 (1975)
17. Zadeh, L.A.: The concept of a linguistic variable and its applications to approximate reasoning, part II. Inf. Sci. **8**, 301–357 (1975)
18. Zadeh, L.A.: The concept of a linguistic variable and its applications to approximate reasoning, part III. Inf. Sci. **9**, 43–80 (1975)
19. Zadeh, L.A.: Fuzzy logic and approximate reasoning. Synthese **30**, 407–428 (1975)
20. Tsypyschev, V.N.: Full periodicity of Galois polynomials over nontrivial Galois rings of odd characteristic. J. Math. Sci. **131**(6), 6120–6132 (2005)

http://www.springer.com/978-3-319-57263-5

Cybernetics and Mathematics Applications in Intelligent Systems
Proceedings of the 6th Computer Science On-line Conference 2017 (CSOC2017), Vol 2

Silhavy, R.; Senkerik, R.; Komínková Oplatková, Z.; Prokopova, Z.; Silhavy, P. (Eds.)

2017, XIV, 446 p. 214 illus., Softcover
ISBN: 978-3-319-57263-5