Fuzzy Propositional Logic for the Knowledge Representation

Alexander Savinov
Institute of Mathematics
Academy of Sciences
Academiei 5
277028 Kishinev
Moldova (CIS)
Phone: (373+2) 73-81-30
EMAIL: 23LSII@MATH.MOLDOVA.SU

Abstract: The paper presents a new alternative approach to thinking about such notions as a proposition, interpretation, inference. Fuzzy propositional logic described generalizes the classical propositional logic in two directions: (i) propositions are considered to be fuzzy, and (ii) logical variables are considered to be many-valued. The formulas are interpreted over the non-fuzzy universe, which is the Cartesian product of the sets of the values of all variables. The main inference rule is the fuzzy resolution. The logic described is implemented in the form of the logical kernel EDIP—the library of functions for MS Windows—what allows to use it as an inference engine in wide range of applications.

I. Introduction

Fuzzy propositional logic (FPL) is fully described in [5]. In this paper the problem of the knowledge representation in this logic is considered. The main notion of FPL is the proposition. We define this notion very formally. Namely, any subset of the universe is a proposition. The universe is a matter of the proposition, i.e. it is what the proposition is about. The proposition is our knowledge about the world, whereas the universe is the world itself.

Now knowing what the proposition is, let us think, e.g., how to represent the proposition about the set consisting of $100,000$ members. If all the members from this set are unique, then there is the only possibility—to assign explicitly to each member some value of the membership function. This method is called *extensive*. If the members are characterized by different properties, then they may be used to assign the same membership function value to a whole class of members with the one property. For example, it can be said that all small green members have the value of the membership function 0.5, big red members—1, and the rest of the members—0. Such representation method is called *intensive*. The logical variables in FPL play the role of the properties or attributes.

From the extensive/intensive point of view FPL gives the formal means for representation and manipulation by fuzzy intensive propositions. The universe in FPL is the Cartesian product of the sets of the values of all variables.

The logic described is a generalization of the fuzzy sectional vectors and matrices technique [4] which originates from the non-fuzzy sectional formalism of Zakrevsky [6]. FPL is developed under the project EDIP. The first version of the expert system shell EDIP [1,2] running under DOS is non-fuzzy, while the second one deals with fuzzy data and conclusions [3]. The Logical Kernel EDIP is a library of functions (dynamically and statically linked) which fully realizes possibilities of FPL and in addition many other functions which one needs in building a knowledge based system. The Logical Kernel is used as a fuzzy inference engine within EDIP Knowledge Manager running under MS Windows and providing general interactive access to the knowledge base.

II. Fuzzy Propositions

Let $x_1, x_2, \ldots, x_n$ be elementary logical *variables*, each of them taking its values from the finite sets $X_1, X_2, \ldots, X_n$ called *domains* respectively. The domain $X_i = \{0, 1, \ldots, n_i - 1\}$ consists of n_i values. The Cartesian product of all domains $X_1 \times X_2 \times \ldots \times X_n$ forms the *universe* Ω with the power $n_1 \times n_2 \times \ldots \times n_n$. The member $\omega = \langle x_1, x_2, \ldots, x_n \rangle \in \Omega$ is an ordered n-tuple of the values of all variables.

It will be supposed that all the domains and consequently the universe are sets in the classical meaning, i.e. they are non-fuzzy. Really, this is subtle but important moment. It reflects the fact that the world itself is non-fuzzy, and it is our knowledge about the world what is fuzzy. Such approach also follows from the believe that the world (or any objective process) is always in some state, it cannot be between states, in several states, beyond the states, or somewhere else.

For example, if there is only one variable (extensive case) and $x_1 = COLOR = \Omega = \{Red, Green, Blue\}$, then it means that there are these and only these colors, and all other ones (e.g., mixed) cannot be considered. If it is necessary to have the other color (e.g., Gray), then there is the only possibility—to add it explicitly to the domain as a separate value. Generally, the attributes may be viewed as sensors.

A truth function $\mu_i(x_i) : X_i \to [0,1]$ defined over the domain is said to be the elementary fuzzy proposition about the variable x_i. The number $\mu_i(j_i) \in [0,1]$ ($j_i = 0, 1, \ldots, n_i - 1$) which the truth function assigns to the value j_i of the domain is said to be the *component* of the proposition.

Elementary propositions can be represented only extensively, because the domain members are characterized by the only property—to be distinguishable from each other. Let us consider the following three possible ways to write elementary propositions.

full: In the full form the variable, all its values, and corresponding components are explicitly indicated, for example:

$$COLOR = \{Red : 0, Green : 0.7, Blue : 1\}.$$

This form is used in EDIP as the basic unit of the knowledge representation language.

brief: In the brief form only the variable and an ordered list of the components are indicated, and the variable is written at the right, for example:

$$\{0, 0.7, 1\}(x_i).$$

This notation will be used in some examples.

symbolic: In the symbolic form the proposition is denoted by a Greek symbol:

$$\mu_i(x_i).$$

If the variable is missed, then lower index of the proposition symbol indicates its number.

If the domain X_i of the variable x_i consists of two values 0 and 1, then it is said to be the Boolean variable. If the truth values of the proposition about the Boolean variable are restricted to $\{0,1\}$, then this proposition

$$\mu(x_i) : \{0,1\} \to \{0,1\}$$

is said to be also Boolean. There exist only 4 Boolean propositions, which are denoted as follows:

- $\{0,0\}(x_i) = 1;$
- $\{1,1\}(x_i) = 0;$
- $\{0,1\}(x_i) = x_i = +x_i = \overset{+}{x}_i;$
- $\{1,0\}(x_i) = \neg x_i = -x_i = \bar{x}_i.$

A special kind of proposition is the *constant proposition*. The domain of the constant proposition includes the only element—the empty set $\emptyset$ (or zero). A lower index of the constant proposition symbol is supposed to be equal to zero. Thus the constant proposition $\mu_0 : \{\emptyset\} \to [0,1]$ is fully defined and distinguished from others by the only component $\mu_0(\emptyset) \in [0,1]$.

More complex propositions are obtained with the help of composition and superposition of the elementary and constant propositions. The composition of propositions is constructed traditionally from two propositions and logical connectives $\wedge$ (conjunction), $\vee$ (disjunction), and $\to$ (implication), i.e. if μ and ν are propositions then $\mu \wedge \nu$, $\mu \vee \nu$, and $\mu \to \nu$ are propositions too.

The superposition is a proposition about another proposition, i.e. if μ and ν are propositions then $\nu(\mu)$ is a proposition too. However, the proposition ν has to be formulated with respect to the interval $[0, 1]$:

$$\nu : [0, 1] \to [0, 1].$$

A proposition μ which is a matter of another proposition may be viewed as a variable taking the values from the domain $[0, 1]$. Hence, one is able to make a proposition about μ as it is been an ordinary variable. But the "variable" μ is subjective, while the elementary variables are objective, because they describe the real world.

Let us define two special propositions about the truth interval $[0, 1]$. The proposition $\neg$ or $-$ called the *negation* is defined as the truth function $\neg(x) = 1 - x$, where $x \in [0, 1]$. The proposition $\neg$ is an analogue of the classical negation; it may be applied to any other proposition. The proposition $+$ called the *assertion* is contrary to the negation and is defined as $+(x) = x$, where $x \in [0, 1]$.

III. Interpretation of Fuzzy Propositions

Any proposition is a well formed formula of FPL (or simply a formula). Formulas of FPL are interpreted over the universe Ω. Interpretation rules define the way by which given the formula one is able to calculate the value the truth function takes in any point of the universe. Thus any formula, including elementary ones, with the help of interpretation rules defines a proposition about the universe. The formula itself in such meaning may be viewed as an intensive representation of the proposition about the universe which can be transformed to the explicit extensive form $\{\langle \omega, \mu(\omega) \rangle, \text{ where } \omega \in \Omega, \ \mu(\omega) \in [0, 1]\}$ by interpretation rules. In other words, interpretation rules allow us to determine the meaning of the intensively represented proposition proceeding from the extensive propositions which make it up.

Formally the interpretation rules are defined as follows.

1. The elementary proposition μ_i defines the proposition $\mu_i : \Omega \to [0, 1]$ over the universe such that

$$\mu_i(\omega) = \mu_i(\langle x_1, x_2, \ldots, x_n \rangle) = \mu_i(x_i);$$

2. The constant proposition μ_0 defines the proposition $\mu_i : \Omega \to [0, 1]$ over the universe such that

$$\mu_0(\omega) = \mu_0(\langle x_1, x_2, \ldots, x_n \rangle) = \mu_0(\emptyset);$$

3. Let μ and ν be formulas, then the formulas $\mu \wedge \nu$ and $\mu \vee \nu$ are interpreted as

$$(\mu \wedge \nu)(\omega) = \min(\mu(\omega), \nu(\omega)),$$

$$(\mu \vee \nu)(\omega) = \max(\mu(\omega), \nu(\omega));$$

   The interpretation of the formula $\mu \to \nu$ is the same as for $\neg(\mu) \vee \nu$:

$$(\mu \to \nu)(\omega) = (\neg(\mu) \vee \nu)(\omega);$$

4. The formula $\nu(\mu)$ is interpreted as a proposition $\nu(\mu) : \Omega \to [0, 1]$ over the universe such that

$$(\nu(\mu))(\omega) = \nu(\mu(\omega)).$$

As usual, some formulas of FPL are supposed to be declared by axioms. If $\{\alpha_1, \alpha_2, \ldots\}$ is a set of axioms then their semantics is defined as follows:

$$\{\alpha_1, \alpha_2, \ldots\}(\omega) = \min(\alpha_1(\omega), \alpha_2(\omega), \ldots).$$

Thus we will consider that all axioms are combined with the connective $\wedge$ and all together they define the semantics just like the one big axiom.

The formula $\mu = \mu_0 \vee \mu_1 \vee \ldots \vee \mu_n$ consisting of the disjunction of the constant proposition μ_0 and elementary propositions μ_i $(i = 1, 2, \ldots n)$ about all variables is a fuzzy *disjunct*. The formula $\mu = \mu_0 \wedge \mu_1 \wedge \ldots \wedge \mu_n$ is a fuzzy *conjunct*.

The main semantical property of disjuncts is that it is possible with the help of one disjunct to represent a proposition about the universe which is equal to 1 everywhere except for the one point where it is equal to any desired value. We will say that the disjunct pricks a hole down to a particular depth (and of a particular width) in the unity level surface. Using this property, it is easy to prove that any semantics can be represented with the help of a finite number of disjuncts by pricking holes in necessary places.

We will suppose that the axioms are fuzzy disjuncts. Such representation is called the fuzzy *conjunctive normal form*.

IV. Knowledge Representation

The truth function over the universe may be viewed as a possibility distribution. If there is not any knowledge, then we suppose that everything is possible, i.e. the possibility distribution is equal 1 in all points of the universe. If one knows something, then it means that particular combinations of the values of attributes are not possible (they are disabled), and consequently the possibility distribution is not equal 1 in these points. Thus, and this is very important, the knowledge in our approach is considered to be *fuzzy constraints* on the possible values of attributes (note that, e.g., in Prolog-like languages it is not so). These constraints define logical connections between attributes.

For example, the fact that there are no small green objects in the nature is a knowledge, whereas, formally, the fact that I saw a big red object is not a knowledge because it has another modal semantics. Knowledge in our approach has a negative character—we know only what is not possible, i.e. what we never saw.

For the knowledge representation the following three types of propositions are used:

- *rules*;
- *negative assertions*;
- *positive assertions*.

The rule is a fuzzy implication, what is rather traditional for the artificial intelligence. The rule explicitly represent logical connections between attributes. One rule is translated into one disjunct, which is added to the knowledge base. For example, the rule

$$SIZE = \{Small : 1, Big : 0\} \quad \rightarrow \quad COLOR = \{Red : 0, Green : 0.5, Blue : 1\}$$

(which means that if the size is small, then the color is exactly not red, and apparently not green) is translated into the disjunct

$$SIZE = \{Small : 0, Big : 1\} \quad \vee \quad COLOR = \{Red : 0, Green : 0.5, Blue : 1\}.$$

The negative assertions are simply implications with no right part. They explicitly represent the fact that particular combinations of the values are disabled.

The *negative example* is the most concrete form of the negative assertion. For example, if we knew that there are no small green objects, then it could be expressed as follows:

$$SIZE = \{Small : 1, Big : 0\} \quad \wedge \quad COLOR = \{Red : 0, Green : 1, Blue : 0\} \quad \rightarrow$$

and in the form of disjunct as

$$SIZE = \{Small : 0, Big : 1\} \quad \vee \quad COLOR = \{Red : 1, Green : 0, Blue : 1\}.$$

This negative example disables the only object $\omega = \langle Small, Green \rangle$.

Let us consider the positive assertions. Let several rules and negative assertions be included into the knowledge base, and then we learn that somebody saw the small green object (or this fact is learned from the data base). How to reflect this fact in the knowledge base? Clear that we have to increase the value of the possibility distribution at the point $\langle Small, Green\rangle$ up to a certain level (if the fact is absolute, then up to 1). However, we cannot do it by adding a new disjunct, since it results in only decrease of the values of the possibility distribution. Formally, if the knowledge base consists of the disjuncts $\mu^1, \mu^2, \ldots \mu^m$, and the positive assertion is represented by the conjunct κ, then the resulting knowledge base is written as follows:

$$(\mu^1 \wedge \mu^2 \wedge \ldots \wedge \mu^m) \vee \kappa.$$

This expression can be equivalently (i.e. without semantics change) transformed to

$$(\mu^1 \vee \kappa) \wedge (\mu^2 \vee \kappa) \wedge \ldots \wedge (\mu^m \vee \kappa).$$

Thus when adding the positive assertion to the knowledge base, each disjunct from it may be changed. One disjunct μ^i and conjunct κ are transformed to n disjuncts:

$$\mu^i \vee \kappa = \mu^i \vee (\kappa_1 \wedge \kappa_2 \wedge \ldots \wedge \kappa_n) = (\mu^i \vee \kappa_1) \wedge (\mu^i \vee \kappa_2) \wedge \ldots \wedge (\mu^i \vee \kappa_n),$$

where

$$\mu^i \vee \kappa_j = (\mu^i_1 \vee \ldots \vee \mu^i_j \vee \ldots \vee \mu^i_n) \vee \kappa_j = \mu^i_1 \vee \ldots \vee (\mu^i_j \vee \kappa_j) \vee \ldots \vee \mu^i_n.$$

The most concrete form of the positive assertion is the *positive example*, which enables (makes it possible) the only combination of the values of attributes.

V. Fuzzy Resolution

It is clear that the formula may be moved off a set of axioms without overall semantics change provided that it is a consequence of some other axiom. Such an operation will be referred to as an *absorption*. Application of the absorption does not result in obtaining new axioms, however it enables us to simplify the representation by eliminating the redundancy.

The main inference rule in FPL is a *fuzzy resolution* which is denoted as $\langle x_k\rangle$. This rule is applied to any two disjuncts on the variable x_k and results in a third disjunct called a *resolvent*. Let μ and ν be two disjuncts, then their resolvent $\mu\langle x_k\rangle\nu$ on kth variable is derived as follows:

$$\begin{array}{ccccccccc}
\mu_0 & \vee & \mu_1 & \vee\ldots\vee & \mu_k & \vee\ldots\vee & \mu_n \\
\nu_0 & \vee & \nu_1 & \vee\ldots\vee & \nu_k & \vee\ldots\vee & \nu_n \\
\hline
\{0\} & \vee & (\mu_1 \vee \nu_1) & \vee\ldots\vee & ((\mu_0 \vee \mu_k) \wedge (\nu_0 \vee \nu_k)) & \vee\ldots\vee & (\mu_n \vee \nu_n)
\end{array}$$

For example, if

$$\mu = \{0\} \vee \{1, 0.5, 0.3, 1\}(x_1) \vee \{0, 1\}(x_2),$$
$$\nu = \{0\} \vee \{0.2, 0.5, 0, 1\}(x_1) \vee \{0, 0\}(x_2),$$

then their resolvents on the first and the second variables are equal to

$$\mu\langle x_1\rangle\nu = \{0\} \vee \{0.2, 0.5, 0, 1\}(x_1) \vee \{0, 1\}(x_2),$$

$$\mu\langle x_2\rangle\nu = \{0\} \vee \{1, 0.5, 0.3, 1\}(x_1) \vee \{0, 0\}(x_2).$$

Let us introduce some new notions which are used to formulate the properties of the resolution.

The *consistency* of the proposition is the maximal value of the corresponding truth function. The *constant* of the proposition μ (denoted as constant(μ)) is the minimal value of the corresponding truth function.

A formula ν is said to be a *logical consequence* of a formula μ iff

$$\forall\omega \in \Omega \quad \mu(\omega) \leq \nu(\omega).$$

Disjuncts μ and ν such that

$$\mu \not\models \mu\langle x_k\rangle\nu \qquad \text{and} \qquad \nu \not\models \mu\langle x_k\rangle\nu$$

hold are said to be adjacent on the variable x_k.

The maximal component of the proposition μ_k which is exactly greater than the corresponding component of the proposition ν_k is referred to as the *critical value* of the proposition μ_k with respect to ν_k:

$$\max_{\mu_k(j_k)>\nu_k(j_k)} \mu_k(j_k), \text{ where } j_k = 0, 1, \ldots, n_k.$$

In order to calculate this critical value one at first needs to select in μ_k all components which are exactly greater than the corresponding components of ν_k, and then to choose among them the maximal component. If

$$\forall j_k \quad \mu_k(j_k) \leq \nu_k(j_k)$$

holds, i.e. there is nothing to choose the maximal component from, then the critical value is supposed to be equal to 0.

In fact we will need a so called *mutual* critical value (denoted as critical(μ_k, ν_k)) which is defined as a minimum of two numbers—the critical value μ_k with respect to ν_k and the critical value ν_k with respect to μ_k.

If the mutual critical value is equal to 0, then it means that the propositions μ_k and ν_k are comparable, i.e. one of them is included into the other:

$$\text{critical}(\mu_k, \nu_k) = 0 \qquad \Leftrightarrow \qquad (\mu_k \models \nu_k \quad \text{or} \quad \nu_k \models \mu_k)$$

Now let us formulate two main properties of the resolution.

Theorem 1. $\mu \wedge \nu \models \mu\langle x_k \rangle \nu$.

Theorem 2. Disjuncts μ and ν are adjacent on the variable x_k iff

$$\forall i \neq k \qquad \text{constant}(\mu_i \vee \nu_i) < \text{critical}(\mu_k, \nu_k).$$

The first theorem says that the resolvent is a consequence of its parents, and this is rather obvious fact. The second theorem gives the criteria of the adjacency of two disjuncts. This criteria may be split into two parts, and then the analogy with the traditional resolution will become more obvious.

If critical$(\mu_k, \nu_k) = 0$, then it is not possible to satisfy the theorem condition, and consequently the necessary condition of the adjacency of two disjuncts is an incomparability of propositions which the resolution is to be on. This condition is a direct generalization of the conventional one, which requires that corresponding literals be contrary ones.

If the incomparability condition is satisfied, then in order that the disjuncts be adjacent, it is also necessary that the constant of each resolvent proposition except kth one be exactly less than the critical value. This condition is a generalization of the conventional one which consists in the absence in the disjuncts of the second pair of contrary literals.

VI. Logical Inference

Let us suppose that a set of disjuncts $\mu^1, \mu^2, \ldots \mu^m$ represent known problem domain laws or its invariant, i.e. some properties which remain unchanged during long enough time interval. In addition we will consider that there is some supplementary information about the possible values of attributes, but which is situational, i.e. it is changed from one case to another. This information is represented in the form of conjunct φ. If some component of φ is not equal to 1, then it means that the attribute cannot take the corresponding value with some degree of confidence.

The problem of carrying out the logical inference consists in finding maximal fuzzy constraints on the possible values of attributes which do not contradict to the knowledge and data. Formally, it is required to find the minimal conjunct ψ satisfying the condition

$$\varphi \wedge (\mu^1, \mu^2, \ldots \mu^m) \models \psi \tag{1}$$

The minimality requirement in the given case means that if any component of ψ is decreased, then (1) is not satisfied.

VII. Logical Kernel EDIP

The Logical Kernel (LK) EDIP is a library of functions (statically and dynamically linked) for Windows which is implemented as a general fuzzy inference engine to be used in wide range of applications based on the attribute models of the problem domain. All information in LK EDIP is contained in a special format file (knowledge base) which is devided into two main parts: syntactical and semantical.

The syntactical part consists of descriptions of the attributes and their values which include the following fields:

- name (textual);
- question (textual);
- comment (textual);
- price (numerical);
- importance (numerical);
- hidden (logical);
- global (logical).

Some information is currently used in Logical Kernel while the other is reserved for future versions.

The semantical part includes a set of the source assertions each of them corresponding to one dependence between attributes. Fuzzy components are of integer type and can be ranged from 0 to 100. The source form of the knowledge base is transformed to the equivalent one in compilation time. The compiled knowledge base is used for the logical inference in run time. This preliminary compilation considerably speeds up the logical inference.

Currently the Logical Kernel is used a fuzzy inference engine within EDIP Knowledge Manager running under MS Windows and providing general interactive access to the knowledge base.

References

[1] V.I. Levchenko, The diagnostic system based on the finite predicates, Preprint, Institute of Mathematics and CC AS Moldavia (Kishinev, Shtiintsa, 1990) (in Russian).

[2] V.I. Levchenko, A.A. Savinov, Dialog control and logical inference in finite predicates, in: Applied Systems of Artificial Intelligence, (Kishinev, Shtiintsa, 1991) 40–46 (in Russian).

[3] V.I. Levchenko, A.A. Savinov, The representation of fuzzy knowledge in the diagnostic expert system shell EDIP, Proc. 2nd Int. Conf. on Fuzzy Logic and Neural Networks—IIZUKA'92 (Iizuka, Japan, July 17–22, 1992).

[4] A.A. Savinov, The matrix representation of fuzzy knowledge in the attribute models, Preprint, Institute of Mathematics and CC AS Moldavia (Kishinev, Shtiintsa, 1991) (in Russian).

[5] A.A. Savinov, Fuzzy propositional logic, Fuzzy Sets and Systems (submitted).

[6] A.D. Zakrevsky, The logical inference in finite predicates, Preprint No.6, Institute of Technical Cybernetics AS Belorussia (Minsk, 1989) (in Russian).