PRUF - A LANGUAGE FOR THE REPRESENTATION OF MEANING IN NATURAL LANGUAGES

L.A. Zadeh*
Computer Science Division
University of California
Berkeley, CA 94720

PRUF—an acronym for Possibilistic Relational Universal Fuzzy—is a designation for a novel type of synthetic language which is intended to serve as a target language for the representation of meaning of expressions in a natural language.

PRUF is based on the premise that the uncertainty in the interpretation of a proposition in a natural language is possibilistic rather than probabilistic in nature. This implies that a proposition, p, of the form "X is F" where X is the name of an object and F is a fuzzy subset of a universe of discourse U, translates in PRUF into an expression, P, which defines a procedure whose domain is the set of possibly fuzzy relations in a database. Acting on these relations, P yields a possibility distribution (or a set of possibility distributions) which characterizes the information, I(p), conveyed by p.

The concept of a possibility distribution is defined as follows (Zadeh 1977). Let X be a variable taking values in U, and let F be a fuzzy subset of U which is characterized by a membership function u_F: U->[0,1], with $u_F(u)$, u€U, representing the grade of membership of u in F. Then the proposition "X is F" induces a possibility distribution, TI^, which is given by the possibility association equation

$$\Pi_X = F \qquad (1)$$

Thus, (1) implies that the possibility that X may take the value u is given by

$$\text{Poss}\{X=u\} = \pi_X(u) = \mu_F(u) \qquad (2)$$

and, more generally, if A is a nonfuzzy subset of U,

$$\text{Poss}\{X \in A\} = \sup_{u \in A} \pi_X(u) \; . \qquad (3)$$

For example, the proposition "X is small," where small is a fuzzy set defined by

$$\text{small} = 1/1 + 1/2 + 0.8/3 + 0.6/4 + 0.4/5 + 0.2/6 \qquad (4)$$

implies that the possibility that X is 3 is 0.8, and that X is either 4 or 5 is 0.6.

As a simple illustration, the proposition "John is young" leads to the possibility association equation $\Pi_{\text{Age(John)}} = \text{young}$, where Age(John) is an implied attribute of John and young is a fuzzy subset of the interval [0,100].

A proposition such as "John resides near Boston," translates in PRUF into the expression

$$\text{RESIDES}[\text{subject} = \text{John}; \; \Pi_{\text{location}} = \text{cityl NEAR}[\text{city2} = \text{Boston}]]$$

where NEAR is a fuzzy relation which associates with each pair of cities (cityl,city2) the degree to which they are near one another; RESIDES is a relation which is particularized by (a) assigning the value John to the attribute subject; and (b) assigning the possibility distribution $_{c}i_{t}y\text{lNEAR}[\text{city2} \blacksquare \text{Boston}]$ to the attribute location, in which NEAR[city2 ■ Boston] denotes a particularization of the relation NEAR with city2 set equal to Boston, and $_{\text{city}}\text{NEAR}[\text{city2} = \text{Boston}]$ denotes its projection on the domain of the attribute cityl. Typically, an expression in PRUF contains names of relations, names of attributes, particularized possibility distributions which are associated with attributes, and labels of tests. In appearance, the simpler expressions in PRUF resemble those of the query language SQUARE and may involve an arbitrary number of levels of nesting.

PRUF has a variety of facilities for modeling the logical organization of natural languages. Among these are: attribute modification, truth qualification, probability qualification, possibility qualification, quantification, particularization and composition. Its expressive power is considerably greater than that of semantic networks or first order predicate calculus, and by employing the concept of a possibility distribution PRUF provides an effective framework for dealing with the imprecision of natural languages in a systematic fashion.

The logic underlying PRUF is a fuzzy logic, FL, in which the truth-values are linguistic, i.e., are expressed as true, quite true, very true, more or less true, etc., with the base logic, for FL being Lukasiewicz[1]s LAlephi logic in which the truth-values are points in the unit interval (Bellman & Zadeh 1976). In addition to the use of linguistic truth-values, PRUF allows the use of linguistic quantifiers, e.g., many, few, several, most, almost all, etc.; linguistic probabilities expressed as likely, very likely, more or less likely, etc.; and linguistic possibilities expressed as possible, quite possible, slightly possible, almost impossible, etc.

PRUF has only two basic rules of inference: (a) The projection rule, which asserts that from an n-ary possibility distribution IT(x-i X)onecan infer by projection a marginal possibility distribution $\Pi(X_{i1}, \ldots, X_{ii})$ where $(i^{\wedge}, \ldots, i^{\wedge})$ is a subsequence of tne index sequence $(1, \ldots, n)$; and (b) that from $\Pi(x_{if} \ldots, X_n)$ and $\Pi(X_{i1t} \ldots, X_{ijJ}$ one can infer the intersection of $\Pi(X_1 \ldots X_n)$ with the cylindrical extension of $\Pi(x_{i1}, \ldots . X_{iu})^* \; {}^{*n}$ combination, these two rules lead to the compositional rule of inference which includes the classical modus ponens as a special case.

References

R.E. Bellman and L.A. Zadeh, Local and fuzzy logics, ERL Memo. M-584, University of California, Berkeley (1976).

L.A. Zadeh, Fuzzy sets as a basis for a theory of possibility, ERL Memo. M77/12 (1977). To apppear in the International Journal of Fuzzy Sets and Systems.

* Currently on leave with the Computer Science Department, IBM Research Laboratory, San Jose, CA 95143. Research supported by the National Science Foundation Grant MCS-76-06693.