Conservation of fuzziness in rule interpolation

T.D. Gedeon[*] and L.T. Kóczy

Department of Telecommunications and Telematics
The Technical University of Budapest
Budapest, Sztocek u. 2
H-1111 HUNGARY
{ gedeon | koczy }@ttt-202.ttt.bme.hu

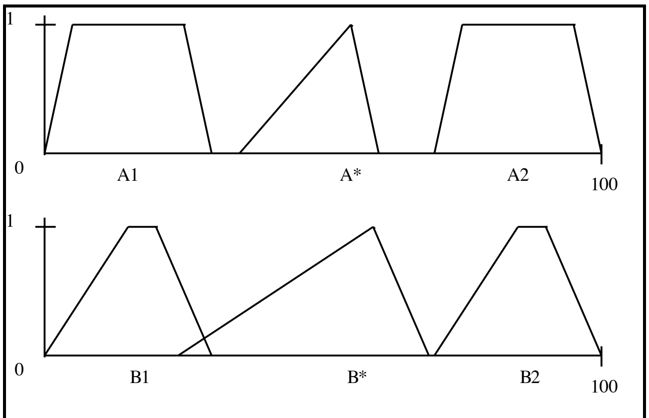
Abstract

Approximate reasoning using fuzzy rule based systems have wide application in for example industrial control and pattern recognition areas. Sparse rule bases which do not contain redundant information can provide computational advantages to comprehensive rule bases. Also, sometimes there are natural gaps in the knowledge base. Fuzzy rule interpolation is used to provide conclusions for observations for which there may be no overlap with even the supports of existing rules in the rule base. All of the methods are descendants of the Kóczy and Hirota (1990) method of linear interpolation, and have various advantages and disadvantages. The criteria for evaluating interpolation methods must include the ability to form conclusions where it is appropriate, the formation of intuitively acceptable conclusions, and a computational complexity which would allow it to be useful in reducing the size of fuzzy rule bases.

In this paper we introduce our method which is conservative with respect to the degree of local fuzziness in the rule base. The notion of intuitively acceptable is important, and is already inherent in the field. That is, the use of linguistic variables which are used as fuzzy rules is in some sense intuitive and qualitative.

Introduction

In this section we discuss the philosophical basis of our approach. Figure 1 illustrates a pair of rules:



A1 >> 0, 5, 25, 30
A2 >> 70, 75, 95,100
B1 >> 0, 15, 20, 30
B2 >> 70, 85, 90,100
A* >> 35, 55, 55, 60
B* >> 24, 59, 59, 69

Figure 1. Rule set p2.

The two rules are *if x is A1 then y is B1*, and *if x is A2 then y is B2*. The observation A* does not overlap with the antecendents of either of the flanking two rules. An interpolated conclusion B* is shown, between the flanking consequents of the two rules. For simplicity of presentation, both the antecedent and consequent universes of discourse have been scaled to the 0 to 100 range, and rules are represented by 2 support and 2 core points (to allow trapezoidal forms).

The initial starting point is that we do assume as little homogeneity in the rule base as possible. Thus, we treat the nearest core points of rules as all that is visible. Thus, A* is in a valley between A1 and A2, and whether A1 or A2 are actually plateaux is not visible, and is not used. The core points of B* are derived by simple linear interpolation between the nearest core points of A1, A2 and B1, B2.

Once the core points of B* are determined, we reduce further the assumption of homogeneity in the rule base. That is, the interpolation of the right of B* is only between the the rightmost core point of B* and

[*] on leave from the School of Computer Science and Engineering, The University of New South Wales.

the leftmost core point of B2. This is consistent with the premise that A* need not be symmetrical, and with the notion that determining the right side of B* should be based on 'nearby' information such as the right side of A* and the left sides of A2 and B2. This is in contrast to the initial method which derived the right side of B* from the right sides of A1, B1, A2, B2 and A*. All of these except the latter are further away from the right side of A* or B* than the 'nearby' information we propose to use.

The use of such 'nearby' information has real world plausibility, while there are few domains where it could be readily proven that there is justification for the assumption of handedness of the the rule shapes.

In view of the claims to intuitive acceptability we make, we will first describe and derive our formulation geometrically before expressing the method in terms of the more usual equations using fuzzy distances.

Geometric Description of Method

Given the subsection of the previous figure shown in Figure 2.

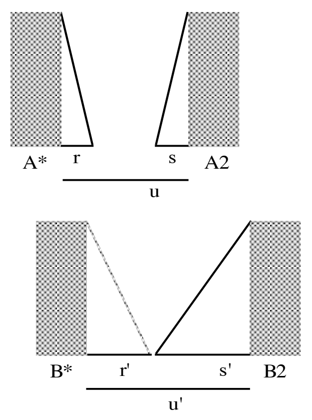


Figure 2. Section of p2.

Note that the grey regions highlight the (assumed) unknown nature of the rest of the rule base. The labels r, s indicate the spread of the observation and the rule antecedent and represent their fuzziness. Clearly A2 is more fuzzy than A*. The labels r', s' indicate the spread of the conclusion (drawn in grey as it is yet to be calculated) and the rule consequent. By observation, our intuition tells us that B2 is more fuzzy than A2, due to the shallower slope. We have not normalised the A*,A2 and the B*,B2 distances, as we prefer to do this explicitly using the values of u, u' as appropriate.

Without using any information from rule 2 other than its core distances, we could calculate a value for r' as shown in Equation (1).

$$r' = r \cdot \frac{u'}{u} \tag{1}$$

This is merely the normalisation of r from the A*,A2 distance metric into the B*,B2 distance. Most likely, the value of r' will be some increase or decrease of the effect of this normalisation. We could similarly calculate a value for the normalisation of s into the B*,B2 distance, called s''.

$$s'' = s \cdot \frac{u'}{u} \tag{2}$$

The relationship between the actual value of s' we can measure from B2, and the calculated value s'' provides an indication of the difference in fuzziness from rule antecedent to consequent. As we have already noted, A2 is clearly fuzzier than B2. This difference can be incorporated into (1) producing:

$$r' = r \cdot \frac{u'}{u} \cdot \left(1 + \frac{s' - s''}{z}\right) \tag{3}$$

Since we are interested in the relative change in fuzziness from rule antecedent to consequent, we divide s' - s'' by z, being either s', or s''. The more conservative choice of divisor is s', which we shall see produces a good compromise between two extremes. With this substitution we produce:

$$r' = r \cdot \frac{u'}{u} \cdot \left(2 - \frac{s}{s'} \cdot \frac{u'}{u}\right) \tag{4}$$

Note that this will not work for crisp rule consequents, so in this case we make the less conservative, and perhaps more obvious, substitution producing Equation 5.

$$r' = r \cdot \frac{s'}{s} \tag{5}$$

Note that r' is no longer dependent on the ratio of the different metrics, and is solely determined by the ratio of rule consequence to antecedent fuzziness. This explains our previous comment regarding Equation (3) being a good compromise, between interpolation solely on the basis of the change in metric, versus solely on the basis of the change in rule fuzziness but ignoring the change in metric.

For completeness, we note that Equation (5) will not work for crisp rule antecendents, hence we resort to Equation (1), which is appropriate as we need to use Equation (1) only in the case where both rule antecedent and conclusion are crisp, where the only information available is the change in metric.

We can now explain the shape of r' in Figure 2. The larger distance B*,B2 versus A*,A2 would be expected to increase the width of r' versus r, and the shallower slope of B2 versus A2 also indicates a icrease of fuzziness. The steep slope of A* indicates low fuzziness in the observation which effect is combined. The actual values are shown in Figure 1, the value of r is 5, and the value of r' is 10.

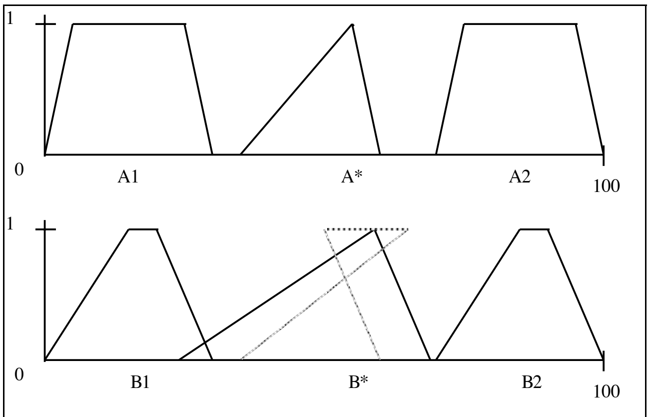
Definitions

Given fuzzy sets $A, B \in P(X)$, where $P(X)$ is the normal and convex fuzzy power set.

The geometric description of the method is readily expressed in the usual notations. The only significant modification we need to make is to note that fuzziness can not be negative. Thus, we need to ensure that the reduction in fuzziness is no larger than the fuzziness present. This incidentally guarantees the normality of the conclusions generated. None of the examples shown in this paper would have been abnormal in this fashion.

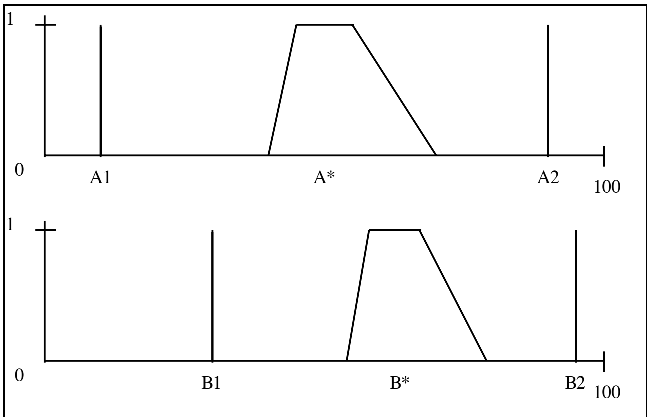
Results

In this section we show a number of examples of rule interpolation using our fuzziness conservation rule interpolation method. The diagrams show as comparison the initial linear interpolation method results in grey. Note that most of the following figures were chosen to highlight the occasionally abnormal results of the linear interpolation method. Our method is effectively a piecewise use of the linear interpolation method and as such does not suffer from the same effects.



A1 >> 0, 5, 25, 30
A2 >> 70, 75, 95,100
B1 >> 0, 15, 20, 30
B2 >> 70, 85, 90,100
A* >> 35, 55, 55, 60
B* >> 24, 59, 59, 69 g&k
B* >> 35, 65, 50, 60 k&h

Figure 3. Rule set p2, methods G&K and K&H.



A1 >> 10, 10, 10, 10
A2 >> 90, 90, 90, 90
B1 >> 30, 30, 30, 30
B2 >> 95, 95, 95, 95
A* >> 40, 45, 55, 70
B* >> 54, 58, 67, 79 g&k
B* >> 54, 58, 67, 79 k&h

Figure 4. Rule set p9, methods G&K and K&H.

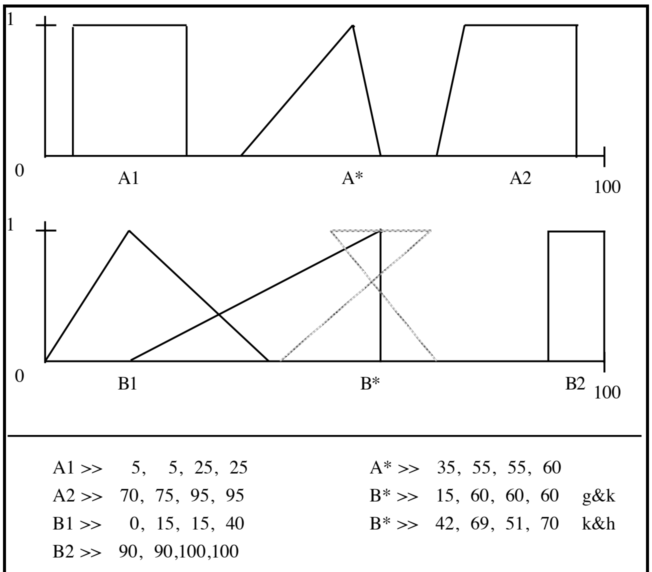


Figure 5. Rule set p11, methods G&K and K&H.

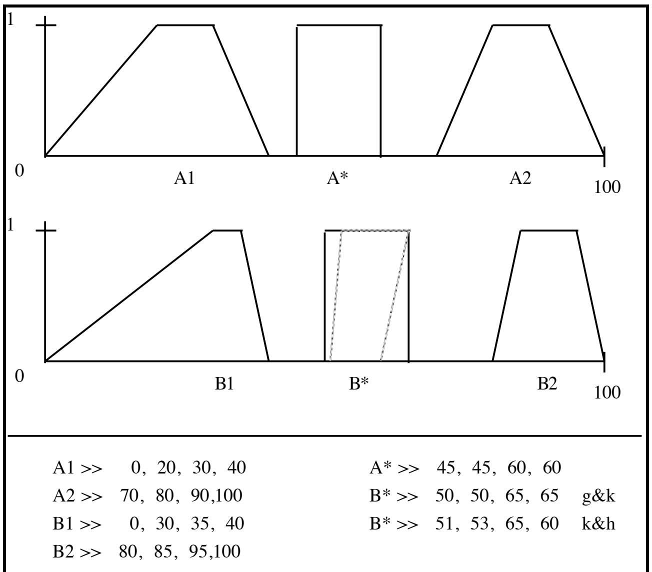


Figure 6. Rule set p12, methods G&K and K&H.

Discussion and Conclusion

The interpolated rules we produce are intuitive in the sense of maintaining the local change in fuzziness in the rule base.

The difference between the geometric interpretation of a rule interpolation transform and the conceptual meaning of the operations implemented by the geometric transform is the root of the occasional problem with abnormality of results using Kóczy and Hirota's rule interpolation technique. In this sense, our approach here is an advance in providing a rationale for clipping the application of the geometrical transform to normal conclusions, as well as the localisation of the rule interpolation.

From an alternative viewpoint, however, we can view the situations in which a rule interpolation technique produces abnormal results as indicating something about the interaction of the application domain, and the rule interpolation method chosen. We examine these implications in another paper.

Reference

Kóczy, LT and Hirota, K, 1990, "Fuzzy inference by compact rules," Proc.of Int. Conference on Fuzzy Logic & Neural Networks IIZUKA '90, Iizuka Fukuoka, pp. 307-310.