A distance measure for the mutation of fuzzy rules

Angelika Krone, Timo Slawinski
Chair of Control Engineering
(Prof. Dr. H. Kiendl)
Faculty of Electrical Engineering
University of Dortmund
D-44221 Dortmund

1 Introduction

A common approach in the field of EA design is to translate the real-world problem into a standard representation. A disadvantage of this mapping between the phenotype and genotype is that the effects of genetic operators on the phenotype space are difficult to determine and that problem knowledge is not considered. Another approach is to use the concept of the Metric Based Evolutionary Algorithm (MBEA) [1]:

- The domain knowledge is expressed by a metric on the phenotype space, such that similar individuals according to the metric have similar fitness values.
- The metric on the phenotype space is preserved by the coding function for the genotype space.
- The genetic operators (mutation, recombination) are designed in consideration of a set of formal requirements (bias free operation, locality, reachability, feature preservation).

In the field of fuzzy modelling an appropriate metric for the distance between fuzzy rules is a difficult problem, especially if rules of different lenghth are used. A first approach has been made by heuristic assumptions about the effect of modifications of a fuzzy rule [5,6]. The disadvantage of this approach is that the spatial position of the rules is neglected and that the defined distance measure is not directly utilized for the mutation operator. In this paper, a distance measure is described that considers the spatial situation and can be much better exploited for the design of the mutation operator. It is developed for the evolutionary rule search in the Fuzzy-ROSA[1] method [2–4].

In Section 2, the elements of the search space are defined and the problem of finding an appropriate distance measure is discussed. In Section 3, the distance measure is defined by a relation vector and a cumulating distance value. Figures illustrate the distance measure. How this distance measure can be used for the mutation operator is explained in Section 4. Section 5 gives a conclusion.

[1] Rule Oriented Statistic Analysis

2 Elements of the search space and their interrelation

The elements of the search space are premises (if-clauses) of fuzzy rules like

$$\begin{array}{ll} \text{IF} & ((X_1 = a_{1,3}) \wedge (X_4 = a_{4,2}))) \\ \text{THEN} & (Y = b_3) \end{array}$$

with P input variables $X_1, X_2, \ldots, X_v, \ldots, X_P$, output variable Y, linguistic input values a_{v,k_v} and linguistic output values b_k. The conclusions of the rules are not included in the search space as it is more efficient to evolve only the premises and then to combine all possible conclusions [4].

Admissible premises are

$$\left(\bigwedge_{\gamma=1}^{G} (X_{v(\gamma)} = a_{v(\gamma),k_{v(\gamma)}}) \right)$$
$$\text{with } X_{v(\alpha)} \neq X_{v(\beta)}$$

for all $\alpha, \beta \in I\!N$ with $1 \leq \alpha, \beta \leq G$ and $G \leq G_{max}$. The value G is called combination depth, the value G_{max} maximum combination depth. The constraint prevents that one variable is considered more than once in a premise.

The premises with $G < P$ (the number of linguistic expressions is smaller than the number of input variables) are called generalizing premises. The variables that are not considered can adopt any value.

We consider linguistic expressions $(X_v = a_{v,k_v})$ that are represented by one-dimensional normal and convex fuzzy sets M_{v,k_v} with membership functions $\mu_{M_{v,k_v}}(x_v)$ and

$$\sum_{k_v=1}^{D_v} \mu_{M_{v,k_v}}(x_v) = 1$$

for all values x_v of the variable X_v. $D_v \geq 2$ is the number of fuzzy sets of the input variable X_v. Then a generalised premise is represented by the following P-dimensional fuzzy set:

$$S((v(1), k_{v(1)}), \ldots, (v(G), k_{v(G)})) :=$$
$$\left\{ ((x_1, x_2, \ldots, x_P); \mu_S(x_1, x_2, \ldots, x_P)) \,|\mu_S = \bigwedge_{\gamma=1}^{G} \left(\mu_{M_{v(\gamma),k_{v(\gamma)}}}(x_{v(\gamma)}) \right) \right\}.$$

Each generalising premise covers several complete input situations. A complete input situation is described by

$$V\left((1, k_1), (2, k_2), \ldots, (P, k_P)\right) :=$$
$$\left\{ ((x_1, x_2, \ldots, x_P); \mu_V(x_1, x_2, \ldots, x_P)) \,|\mu_V = \bigwedge_{v=1}^{P} \left(\mu_{M_{v,k_v}}(x_v) \right) \right\}$$

with

$$\mu_V(x_1, \ldots, x_P) \leq \mu_S(x_1, \ldots, x_P)$$

for all values of $x_1, \ldots, x_P$. Consequently, the interrelation of a set of generalising premises can be very complex. To exploit the whole information, a distance measure must be defined that bases on the calculation of distance and similarity measures of multidimensional fuzzy sets. However, this approach has two essential drawbacks:

- The calculation is very time consuming.
- Such distance and similarity measures can be used if two premises are given. However, they are improper to evolve a new premise from an existing premise by a mutation operator.

To overcome these drawbacks, the relation of two premises is reduced to essential characteristics and the different dimensions are separated.

3 Relation vector

The relation of a premise S_1 to a premise S_2 can be described by a relation vector

$$Rel(S_1, S_2) = (r_1, r_2, \ldots, r_v, \ldots, r_P)$$

with r_v the characterizing value (defined below) for the dimension X_v. In this way, the main spatial relations are represented. The advantage over one aggregated relation value is that starting from one premise, a second premise can be constructed.

The mapping to the single dimensions is illustrated in Figure 1. In Figure 1 (a) the premises are represented by their α-cuts S_α of the associated fuzzy sets S with $\alpha = 0.5$:

$$S_\alpha := \{(x_1, x_2, ..., x_P) | \mu_S(x_1, x_2, \ldots, x_v, \ldots, x_P) \geq \alpha\}.$$

In Figure 1 (b) the premises are represented by their variable-specific α-cuts S_α^v with $\alpha = 0.5$:

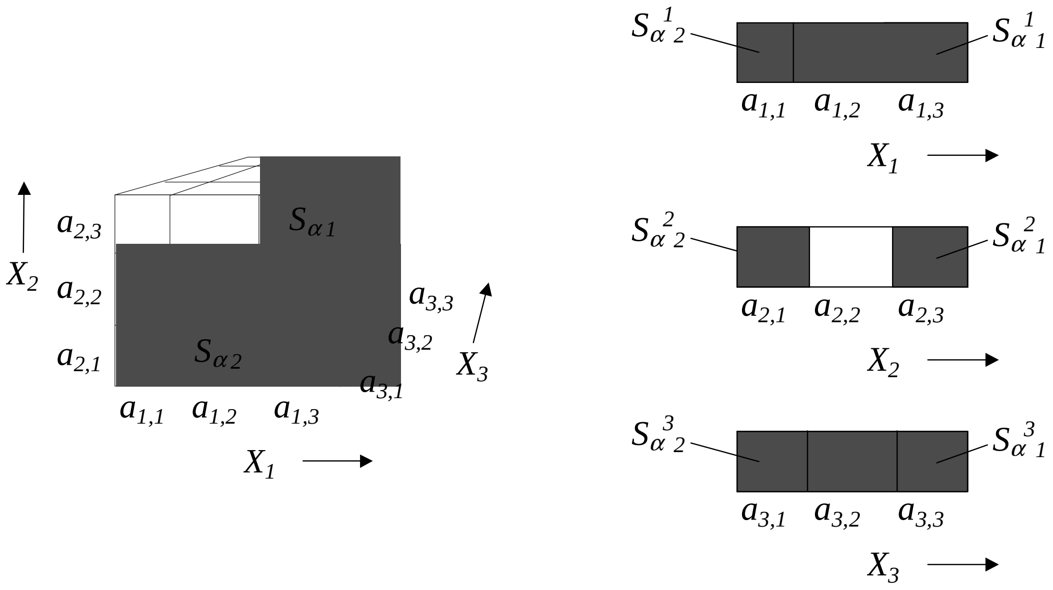
$$S_\alpha^v := \{x_v | \mu_S(x_1, x_2, \ldots, x_v, \ldots, x_P) \geq \alpha\}.$$

Though we usually use the product as AND operator, here, the minimum is used because of the easier presentability of rectangular areas. The analogically defined 0.5-cuts V_α and V_α^v of the complete input situations are represented by enclosing lines.

For determination of the characterizing values r_v of two premises S_1 and S_2, four cases are distinguished:

1. No coverage, no contact:

   The variable-specific α-cuts $S_{\alpha 1}^v$ and $S_{\alpha 2}^v$ cover no joint α-cut V_α^v and are not neighboured (middle diagram of Figure 1 (b)).



(a) Spatial relation. (b) Mapping to single dimensions.

Figure 1: Example for mapping a spatial relation of two premises to the single dimensions.

2. Contact:

   The variable-specific α-cuts $S^v_{\alpha 1}$ and $S^v_{\alpha 2}$ are neighboured. This would be if in the middle diagram of Figure 1 (b) the α-cut $S^1_{\alpha 1}$ was one position further left.

3. Partial coverage:

   The variable-specific α-cuts $S^v_{\alpha 1}$ and $S^v_{\alpha 2}$ cover a joint α-cut V^v_α (top diagram of Figure 1 (b)).

4. Complete coverage:

   The variable-specific α-cuts $S^v_{\alpha 1}$ and $S^v_{\alpha 2}$ cover the identical α-cuts V^v_α (bottom diagram of Figure 1 (b)).

On this basis, the characterizing value r_v is given by[2]

$$r_v = \begin{cases} 0 & : \text{ complete coverage} \\ 1 & : \text{ partial coverage} \\ P+1 & : \text{ contact} \\ (P+1)P+1 & : \text{ no coverage, no contact} \end{cases}$$

Taking only four distinct values has the following advantages:

[2] The explanation for these values demands the definition of the cumulative distance value. Therefore, it is given right behind the definition of the cumulative distance value $\Delta(S_1, S_2)$.

- A further distinction of partial coverages in smaller and larger ones causes that the number of fuzzy sets of the variables play an important role. Following the idea, that small mutations lead to small distances and occur more often than large mutations, the values of variables with a few fuzzy sets are more frequently mutated than those of variables with a higher number of fuzzy sets. This can lead to an undesirable bias.
- Quantifying the case 'no coverage, no contact' is not reasonable as all belonging possibilities $S^v_{\alpha 1}$ have nothing in common with $S^v_{\alpha 2}$ and will anyway reach independent fitness values as there are no common data points supporting the different premises.

The relation vector $(r_1, r_2, \ldots, r_v, \ldots, r_P)$ can be interpreted as a distance vector. The higher the values of the vector are the more differ the positions of S_1 and S_2 with respect to the individual dimensions. The values $(P+1)$ and $((P+1)P+1)$ are chosen to get a well interpretable cumulative distance value:

$$\Delta(S_1, S_2) := \sum_{v=1}^{P} r_v$$

If in all P dimensions there is a partial coverage with $r_v = 1$, then there is $\Delta(S_1, S_2) = P$ which is smaller then the cumulative distance value for one dimension with contact $r_v = P + 1$ and all the other dimensions with $r_v = 0$ ($\Delta(S_1, S_2) = P + 1$). And if in all P dimensions there is a contact with $r_v = P + 1$, then there is $\Delta(S_1, S_2) = P(P + 1)$ which is smaller then the cumulative distance value for one dimension with no coverage and no contact with $r_v = (P+1)P+1$ and all other dimensions with $r_v = 0$ ($\Delta(S_1, S_2) = (P + 1)P + 1$).

In Figure 2 and Figure 3 the relation of one selected premise S_1 to all other possible premises is illustrated. The associated relation vectors and distance values are specified.

4 Distance-based mutation

The distance measure is used to mutate a premise S_1 to a new premise S_{new}. The quantity of mutation is interpreted as distance value $0 \leq \Delta_M \leq \Delta_{Mmax}$ with $\Delta_M \in I\!N$ and $\Delta_{Mmax} = [(P + 1)P + 1)G] + [P - G]$. The first part of the sum in squared brackets represents the cumulated distance values for the G dimensions of the G linguistic expressions of S_1 that all have in an extreme case no coverage no contact. The second part of the sum in squared brackets represents the cumulated distance value for the remaining $P - G$ dimensions that have in an extreme case a partial coverage.

On the basis of this value, a mutation vector $Mut = (m_1, m_2, \ldots, m_P)$ is constructed that is interpreted as relation vector. First the frequencies of the four different elements of the mutation vector are calculated:

$$\begin{aligned}
\#(m_v = (P+1)P+1) &= int(\Delta_M/((P+1)P+1)) = int(Q_1) = N_3 \\
\#(m_v = P+1) &= min[min[P-N_3, int(mod(Q_1)/(P+1)], G-N_3] = N_2 \\
\#(m_v = 1) &= min[P-N_3-N_2, mod(Q_1)-(P+1)N_2] = N_1 \\
\#(m_v = 0) &= P-N_3-N_2-N_1 = N_0
\end{aligned}$$

The minimum functions are necessary, as not for all values of Δ_M a relation vector is existent. For N_2 maximally $P-N_3$ dimensions are left for mutation and for N_1 maximally $P-N_3-N_2$. For N_0 remain $P-N_3-N_2-N_1$ dimensions. For N_2 the combination depth of S_1 must be additionally considered as no more mutations of this kind are possible as linguistic expressions are left in the premise S_1. The function *int* gives the integer value of a quotient and the function *mod* the residual of a division.

The mutation values m_v are allocated to the places of the mutation vector in three steps:

1. The mutation values $m_v = P+1$ and $m_v = (P+1)P+1$ are randomly allocated to places that refer to variables X_v that are considered in the premise S_1.
2. The remaining mutation values $m_v = 1$ are randomly allocated to the remaining places.
3. All not allocated places get the mutation value $m_v = 0$.

An additional difficulty in the allocation process is that the combination depth of the premise must not exceed G_{max}. Thus, the following cases must be distinguished:

- If $\#(m_v = 1) + \#(m_v = P+1) + \#(m_v = (P+1)P+1) \leq G_{max} - G$, then the allocation can be done as described above.
- If $\#(m_v = 1) + \#(m_v = P+1) + \#(m_v = (p+1)P+1) > G_{max} - G$, then step two of the allocation process must be refined.
  - If $\#(m_v = 1) \leq G_{max} + G - 2(\#(m_v = P+1) + \#(m_v = (p+1)P+1))$, then the uprounded value of $(G - G_{max} + \#(m_v = 1))/2$ is the number of mutation values $m_v = 1$ that must be allocated randomly to places that refer to variables X_v that are considered in the premise S_1. The remaining number of mutation values $m_v = 1$ that are allocated randomly to the remaining places.
  - If $\#(m_v = 1) \geq G_{max} + G - 2(\#(m_v = P+1) + \#(m_v = (p+1)P+1))$, then $\#(m_v = 1) - (G_{max} + G - 2(\#(m_v = P+1) + \#(m_v = (p+1)P+1)))$ mutation values $m_v = 1$ must not be allocated.[3]

[3] This number can also be substracted in the calculation of $\#(m_v = 1)$. Then this distinction is superfluent in the allocation process.

In this way, the distance value $\Delta_{M'}$ of the mutation vector *Mut* is either equal to the value of Δ_M or adopts the next smallest possible value. Consequently, for all $\Delta_{M1} < \Delta_{M2}$ there is $\Delta_{M_1'} \leq \Delta_{M_2'}$.

The premise S_{new} is constructed from S_1 along the following mutation rules:

1. $m_v = 0$:

   There is no change with regard to the variable X_v. The definite mutation value is $m_v^{'} = m_v = 0$.

2. $m_v = 1$:

   (a) If X_v is considered in the premise, the associated linguistic expression is deleted.

   (b) If X_v is not considered in the premise, a linguistic expression $(X_v = a_{v,k_v})$ is inserted. The linguistic value a_{v,k_v} is randomly chosen from the possible values with equal probability.

   The definite mutation value is $m_v^{'} = m_v = 1$.

3. $m_v = P + 1$:

   The associated linguistic value a_{v,k_v} of the variable X_v is changed to a neighgboured value. If there are two neighboured values, then one is chosen randomly from the two options with equal probability. The definite mutation value is $m_v^{'} = m_v = P+1$.

4. $m_v = (P + 1)P + 1$:

   The associated linguistic value a_{v,k_v} is changed to another value, but not to a neighgboured value. The definite mutation value is $m_v^{'} = m_v = (P + 1)P + 1$. If there are only neighboured values, then these are accepted. The definite mutation value is $m_v^{'} = P + 1$. The values are chosen randomly with equal probability.

The resulting distance value between the premise S_1 and S_{new} is

$$\Delta(S_1, S_{new}) = \sum_{v=1}^{P} m_v^{'} \leq \Delta_{M'} \leq \Delta_M.$$

In Table 1 an example illustrates the distance-based mutation. The premise

$$((X_1 = a_{1,3}) \wedge (X_4 = a_{4,2})))$$

of a problem with ten input variables, $D_v = 5$ and $G_{max} = 6$ is mutated by different values of Δ_M.

An alternative possibility is to choose directly a mutation vector *Mut* instead of a distance value Δ_M. The advantage is that the allocation time might be lower. However, the disadvantage is that the quantity of mutation is a function of several random functions and thus, the interaction is unclear.

Table 1: Mutations of the premise $S_1 = ((X_1 = a_{1,3}) \wedge (X_4 = a_{4,2}))$ ($D_v = 5$ and $G_{max} = 6$) on the basis of 23 randomly chosen values of Δ_M sorted in ascending order.

Δ_M	*Mut*	$\Delta_{M'}$	S_{new}	$\Delta(S_1, S_{new})$
0	$(0,0,0,0,0,0,0,0,0,0)$	0	$(X_1 = a_{1,3}) \wedge (X_4 = a_{4,2})$	0
0	$(0,0,0,0,0,0,0,0,0,0)$	0	$(X_1 = a_{1,3}) \wedge (X_4 = a_{4,2})$	0
0	$(0,0,0,0,0,0,0,0,0,0)$	0	$(X_1 = a_{1,3}) \wedge (X_4 = a_{4,2})$	0
1	$(0,0,0,0,0,0,0,1,0,0)$	1	$(X_1 = a_{1,3}) \wedge (X_4 = a_{4,2}) \wedge (X_8 = a_{8,1})$	1
1	$(0,0,1,0,0,0,0,0,0,0)$	1	$(X_1 = a_{1,3}) \wedge (X_3 = a_{3,5}) \wedge (X_4 = a_{4,2})$	1
2	$(0,1,0,1,0,0,0,0,0,0)$	2	$(X_1 = a_{1,3}) \wedge (X_2 = a_{2,5})$	2
2	$(0,0,0,0,0,0,0,0,1,1)$	2	$(X_1 = a_{1,3}) \wedge (X_4 = a_{4,2}) \wedge (X_9 = a_{9,5}) \wedge (X_{10} = a_{10,4})$	2
2	$(0,0,1,0,0,0,1,0,0,0)$	2	$(X_1 = a_{1,3}) \wedge (X_3 = a_{3,3}) \wedge (X_4 = a_{4,2}) \wedge (X_7 = a_{7,1})$	2
3	$(0,0,0,1,0,1,0,0,1,0)$	3	$(X_1 = a_{1,3}) \wedge (X_6 = a_{6,1}) \wedge (X_9 = a_{9,4})$	3
3	$(1,0,0,0,1,1,0,0,0,0)$	3	$(X_4 = a_{4,2}) \wedge (X_5 = a_{5,5}) \wedge (X_6 = a_{6,3})$	3
4	$(0,1,1,1,0,0,0,0,1,0)$	4	$(X_1 = a_{1,3}) \wedge (X_2 = a_{2,4}) \wedge (X_3 = a_{3,2}) \wedge (X_9 = a_{9,5})$	4
4	$(0,0,1,0,0,1,0,1,1,0)$	4	$(X_1 = a_{1,3}) \wedge (X_3 = a_{3,1}) \wedge (X_4 = a_{4,2}) \wedge (X_6 = a_{6,4}) \wedge (X_8 = a_{8,1}) \wedge (X_9 = a_{9,1})$	4
6	$(1,0,1,0,0,1,1,1,0,1)$	6	$(X_3 = a_{3,2}) \wedge (X_4 = a_{4,2}) \wedge (X_6 = a_{6,3}) \wedge (X_7 = a_{7,5}) \wedge (X_8 = a_{8,1}) \wedge (X_{10} = a_{10,2})$	1
7	$(1,1,0,1,1,1,0,1,1,0)$	7	$(X_2 = a_{2,1}) \wedge (X_5 = a_{5,4}) \wedge (X_6 = a_{6,5}) \wedge (X_8 = a_{8,1}) \wedge (X_9 = a_{9,3})$	7

10	$(1,1,1,1,0,1,0,1,1,1)$	8	$(X_2 = a_{2,1}) \wedge (X_3 = a_{3,3})$ $\wedge\ (X_6 = a_{6,3}) \wedge (X_8 = a_{8,4})$ $\wedge\ (X_9 = a_{9,3}) \wedge (X_{10} = a_{10,4})$	8
10	$(1,0,0,1,1,1,1,1,1,1)$	8	$(X_5 = a_{5,4}) \wedge (X_6 = a_{6,5})$ $\wedge\ (X_7 = a_{7,5}) \wedge (X_8 = a_{8,1})$ $\wedge\ (X_9 = a_{9,3}) \wedge (X_{10} = a_{10,2})$	8
12	$(11,0,0,0,0,1,0,0,0,0)$	12	$(X_1 = a_{1,4}) \wedge (X_4 = a_{4,2})$ $\wedge\ (X_6 = a_{6,4})$	12
25	$(11,0,1,11,0,0,0,1,1,0)$	25	$(X_1 = a_{1,2}) \wedge (X_3 = a_{3,2})$ $\wedge (X_4 = a_{4,3}) \wedge (X_8 = a_{8,4})$ $\wedge (X_9 = a_{9,1})$	25
31	$(11,1,0,11,0,1,1,0,0,1)$	26	$(X_1 = a_{1,4}) \wedge (X_2 = a_{2,1})$ $\wedge (X_4 = a_{4,1}) \wedge (X_6 = a_{6,2})$ $\wedge (X_7 = a_{7,3}) \wedge (X_{10} = a_{10,2})$	26
52	$(11,1,1,11,1,1,0,0,0,0)$	26	$(X_1 = a_{1,2}) \wedge (X_2 = a_{2,3})$ $\wedge (X_3 = a_{3,1}) \wedge (X_4 = a_{4,1})$ $\wedge (X_5 = a_{5,4}) \wedge (X_6 = a_{6,1})$	26
67	$(11,0,1,11,0,1,0,1,1,0)$	26	$(X_1 = a_{1,4}) \wedge (X_3 = a_{3,1})$ $\wedge (X_4 = a_{4,1}) \wedge (X_6 = a_{6,4})$ $\wedge (X_8 = a_{8,1}) \wedge (X_9 = a_{9,5})$	26
139	$(11,1,0,112,1,0,1,0,1,0)$	127	$(X_1 = a_{1,4}) \wedge (X_2 = a_{2,2})$ $\wedge (X_4 = a_{4,4}) \wedge (X_5 = a_{5,3})$ $\wedge (X_7 = a_{7,3}) \wedge (X_9 = a_{9,1})$	127
172	$(112,1,1,11,0,0,0,1,0,1)$	127	$(X_1 = a_{1,1}) \wedge (X_2 = a_{2,1})$ $\wedge (X_3 = a_{3,2}) \wedge (X_4 = a_{4,3})$ $\wedge (X_8 = a_{8,2}) \wedge (X_{10} = a_{10,2})$	127

5 Conclusions

In this paper, a new distance measure for premises has been developed. It allows to measure the distance between generalizing fuzzy premises according to their spatial interrelation in the space of the input variables. By concentrating to the essential characteristics of this interrelation and separating the different dimensions, this distance measure can be directly used for the mutation of premises. In this way, small mutation quantities cause small distance values and vice versa. The application of this distance measure will further improve the realization of a Metric Based Evolutionary Algorithm (MBEA) in the field of fuzzy modelling. In a next step, simulations are necessary to judge if an improved distance measure will cause improved search results.

Acknowledgement

The research is sponsored by the Deutsche Forschungsgemeinschaft (DFG), as part of the Collaborative Research Center 'Computational Intelligence' (531) of the University of Dortmund.

References

[1] S. Droste and D. Wiesmann. On representation and genetic operators in evolutionary algorithms. In *Reihe Computational Intelligence*. No CI-41/98, ISSN 1433-3325, Collaborative Research Center 531, University of Dortmund, 1998.

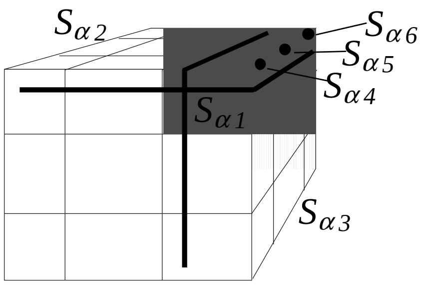
[2] M. Krabs and H. Kiendl. Automatische Generierung von Fuzzy–Regeln mit dem ROSA–Verfahren. *VDI– Berichte*, 1113:29–39, 1994.

[3] A. Krone, T. Baeck, and P. Teuber. Evolutionäres Suchkonzept zum Aufstellen signifikanter Fuzzy–Regeln. *at – Automatisierungstechnik*, 44(8):405–411, 1996.

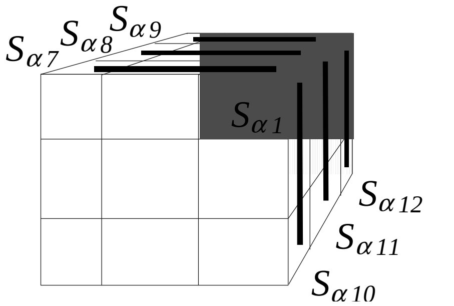
[4] A. Krone and H. Kiendl. Evolutionary concept for generating relevant fuzzy rules from data. *International Journal of Knowledge–based Intelligent Engineering Systems*, 1(4):207–213, 1997.

[5] T. Slawinski, A. Krone, U. Hammel, D. Wiesmann, and P. Krause. A hybrid evolutionary search concept for data-based generation of relevant fuzzy rules in high dimensional spaces. In *Proceedings of Eighth International Conference on Fuzzy Systems, FUZZ-IEEE'99*, Korea, 1999.

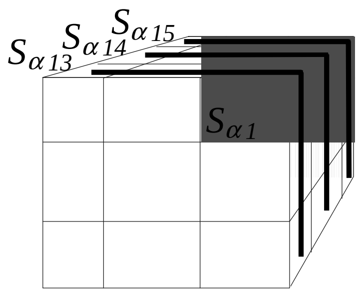
[6] T. Slawinski, A. Krone, U. Hammel, D. Wiesmann, and M. Lindenblatt. A hybrid evolutionary search concept for data–based generation of relevant fuzzy rules in high dimensional spaces (extended version). In *Reihe Computational Intelligence*. CI–50/98, Sonderforschungsbereich 531, Universität Dortmund, 1998.



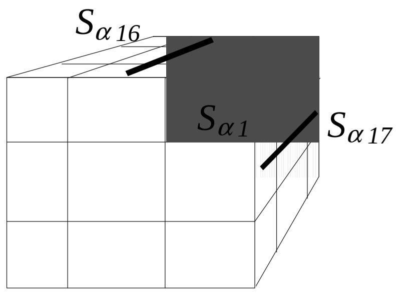
(a) $Rel(S_1, S_2) = (1,0,0)$, $Rel(S_1, S_3) = (0,1,0)$, $Rel(S_1, S_i) = (0,0,1)$ with $i = 4,5,6$ and $\Delta(S_1, S_j) = 1$ with $j = 2,3,\ldots,6$.



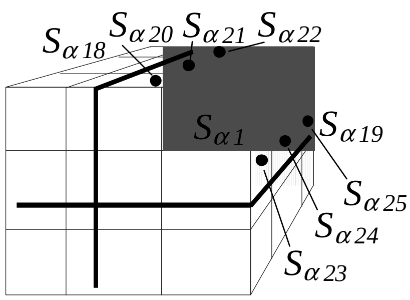
(b) $Rel(S_1, S_i) = (1,0,1)$ with $i = 7,8,9$, $Rel(S_1, S_j) = (0,1,1)$ with $j = 10,11,12$ and $\Delta(S_1, S_k) = 2$ with $k = 7,8,\ldots,12$.



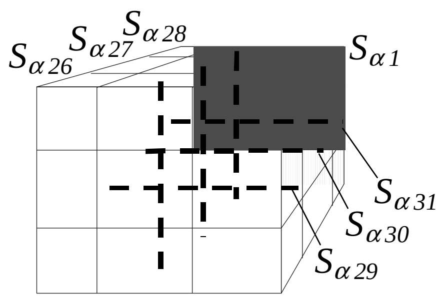
(c) $Rel(S_1, S_i) = (1,1,1)$ and $\Delta(S_1, S_i) = 3$ with $i = 13,14,15$.



(d) $Rel(S_1, S_{16}) = (4,0,0)$, $Rel(S_1, S_{17}) = (0,4,0)$ and $\Delta(S_1, S_i) = 4$ with $i = 16,17$.

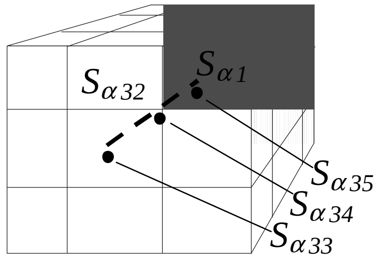


(e) $Rel(S_1, S_{18}) = (4,1,0)$, $Rel(S_1, S_{19}) = (1,4,0)$, $Rel(S_1, S_i) = (4,0,1)$ with $i = 20,21,22$, $Rel(S_1, S_j) = (0,4,1)$ with $i = 23,24,25$ and $\Delta(S_1, S_k) = 5$ with $k = 18,19,\ldots,25$.

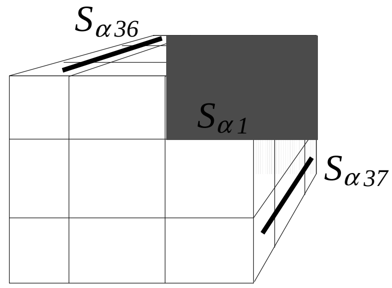


(f) $Rel(S_1, S_i) = (4,1,1)$ with $i = 26,27,28$, $Rel(S_1, S_j) = (1,4,1)$ with $i = 29,30,31$ and $\Delta(S_1, S_k) = 6$ with $k = 26,\ldots,31$.

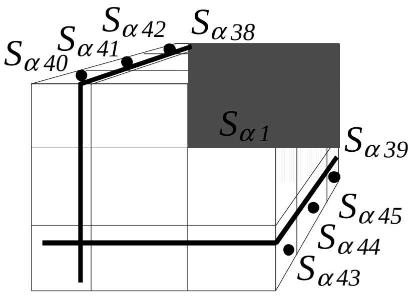
Figure 2: Examples for relation vectors and distance values (part 1). A point represents a premise that covers one complete input situation, a line a premise that covers three complete input situations, and an angle a premise that covers nine complete input situations. The value 4 results from $P+1$ and the value 13 from $(P+1)P+1$ with $P=3$.



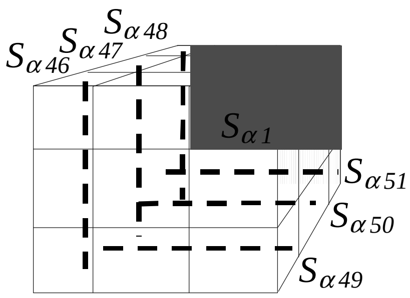
(a) $Rel(S_1, S_{32}) = (4, 4, 0)$, $Rel(S_1, S_i) = (4, 4, 1)$ with $i = 33, 34, 35$ and $\Delta(S_1, S_{32}) = 8$, $\Delta(S_1, S_i) = 9$.



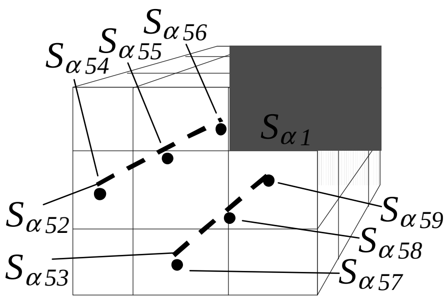
(b) $Rel(S_1, S_{36}) = (13, 0, 0)$, $Rel(S_1, S_{37}) = (0, 13, 0)$ and $\Delta(S_1, S_i) = 13$ with $i = 36, 37$.



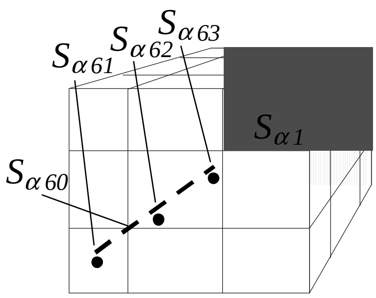
(c) $Rel(S_1, S_{38}) = (13, 1, 0)$, $Rel(S_1, S_{39}) = (1, 13, 0)$, $Rel(S_1, S_i) = (13, 0, 1)$ with $i = 40, 41, 42$, $Rel(S_1, S_j) = (0, 13, 1)$ with $j = 43, 44, 45$ and $\Delta(S_1, S_k) = 14$ with $k = 38, 39, \ldots, 45$.



(d) $Rel(S_1, S_i) = (13, 1, 1)$ with $i = 46, 47, 48$, $Rel(S_1, S_j) = (1, 13, 1)$ with $j = 49, 50, 51$ and $\Delta(S_1, S_k) = 15$ with $k = 46, 47, \ldots, 51$.



(e) $Rel(S_1, S_{52}) = (13, 4, 0)$, $Rel(S_1, S_{53}) = (4, 13, 0)$, $Rel(S_1, S_i) = (13, 4, 1)$ with $i = 54, 55, 56$, $Rel(S_1, S_j) = (4, 13, 1)$ with $i = 57, 58, 59$ and $\Delta(S_1, S_k) = 17$ with $k = 52, 53$, $\Delta(S_1, S_l) = 18$ with $k = 54, 55, \ldots, 59$.



(f) $Rel(S_1, S_{60}) = (13, 13, 0)$, $Rel(S_1, S_i) = (13, 13, 1)$ with $i = 61, 62, 63$ and $\Delta(S_1, S_{60}) = 26$, $\Delta(S_1, S_i) = 27$.

Figure 3: Examples for relation vectors and distance values (part 2).