# Betweenness, Lukasiewicz Rough Inclusions, Euclidean Representations in Information Systems, Hyper-granules, Conflict Resolution

Lech T. Polkowski

Department of Mathematics and Computer Science
University of Warmia and Mazury
Olsztyn, Poland
and
Polish–Japanese Institute of IT. Warszawa, Poland
polkow@pjwstk.edu.pl

**Abstract.** The purpose of this note is to augment information systems with a relation of betweenness among things in the universe of the system and derive from it a geometric representation of granules of things in Euclidean spaces. Next, the notion of betweenness renders a service in introducing a new notion of a hyper–granule. An application to conflict resolution by defining coalitions of agents as granules or hyper–granules and their mixed strategies as elements of convex hulls spanned on things defining the granule is proposed. Finally, hyper–granules offer a new classifying algorithm which exploits neighborhoods of things and in a sense is an improved with respect to similarity variant of nearest neighbor classifier.

**Key words:** rough inclusions, betweenness, Euclidean representations of granules,hyper–granules, coalitions, conflict resolutions, classifier synthesis

## 1 Introduction: Basic Notions

We would like to recall here basic facts about notions mentioned in Abstract, to make this note self–contained. We begin with the Lukasiewicz rough inclusion.

### 1.1 The Lukasiewicz Rough Inclusion

Given a set of things $U$, a rough inclusion is a ternary relation $\mu \subseteq U \times U \times [0, 1]$ which renders the notion of '*to be a part to a degree*'. It is rooted in mereology, see, e.g., [3] whose basic notion is that of a *part* which is a transitive and irreflexive relation $\pi$ on the product $U \times U$ along with its reflexive closure, the *ingredient* $ingr$, i.e., $ingr = part \cup ` =$'. The rough inclusion $\mu$ on the set $U \times U$ should satisfy the requirements

1. I1. $\mu(x, y, 1)$ if and only if $ingr(x, y)$.
2. I2. $\mu(x, y, 1)$ and $\mu(z, x, r)$ imply $\mu(z, y, r)$.
3. I3. $\mu(x, y, r)$ and $s < r$ imply $\mu(x, y, s)$.

A rough inclusion $\mu$ induces a mereological distance $d_\mu$ by means of the formula

$$d_{mu}(x,y) = min\{sup\{r : \mu(x,y,r)\}, sup\{s : \mu(y,x,s)\}\}. \tag{1}$$

Assuming that suprema in the formula (1) are achieved, we have

**Proposition 1.** $d_\mu(x,y) = 1$ *if and only if* $x = y$.

In the case when things in the universe $U$ are described by means of a finite set $A$ of *attributes* so each thing $x \in U$ is represented by its *information set* $\{a(x) : a \in A\}$, we can define the Lukasiewicz rough inclusion $\mu_L$ by taking as its value on a pair $x, y$ in $U$ the quotient of the cardinality of the *indiscernibility set* $IND(x,y) = \{a \in A : a(x) = a(y)\}$ and the cardinality of the set $A$:

$$\mu_L(x,y) = \frac{card(IND(x,y))}{card(A)}, \tag{2}$$

cf., [3], Ch.6.

### 1.2 Betweenness

The notion of betweenness plays an essential role in axiomatization of elementary geometry of Euclidean spaces due to Tarski, see Tarski and Givant [4], it is formalized as a relation $B(x,y,z)$ ('*y is between x and z*'); intuitively, $B(x,y,z)$ means that $y$ lies on the straight line segment with endpoints $x, z$.

Van Benthem [1] proposed an extension of the betweenness relation based on the relation of *nearness* $N(x,y,z)$ ('*x is closer to y than z*') which in terms of the distance $d_\mu$ would be defined by means of

$$N(x,y,z) \text{ if and only if } d_\mu(x,y) > d_\mu(z,y). \tag{3}$$

The relation $N$ thus defined, does satisfy all axioms for nearness in Van Benthem [1], i.e.,

1. N1. $N(z,u,v)$ and $N(v,u,w)$ imply $N(z,u,w)$ (transitivity).
2. N2. $N(z,u,v)$ and $N(u,v,z)$ imply $N(u,z,v)$ (triangle inequality).
3. N3. $N(z,u,z)$ is not true for each pair $u, z$ (irreflexivity).
4. N4. $z = u$ or $N(z,z,u)$ (selfishness).
5. N5. $N(z,u,v)$ implies $N(z,u,w)$ or $N(w,u,v)$ (connectedness).

For the proof, cf., [3] Ch. 6.

Betweenness relation in the sense of Van Benthem $T_B(z,u,v)$ ('*u is between z and v*') introduced in Van Benthem [1] is rendered by a formula

$$T_B(z,u,v) \Leftrightarrow [\text{for each } w \in U\ (u = w) \text{ or } N(u,z,w) \text{ or } N(u,v,w)]. \tag{4}$$

This means that for each thing $w$ distinct from $z$, either u is closer to z than is w or u is closer to v than is w.

## 2 Granules from Indiscernible Things

We assume that things in the set $U$ are described by means of values of $N$ real valued attributes in the set $A$; from this point of view the set $U$ becomes the set $R(U) \subseteq R^N$ of vectors in $R^N$. As said above, each thing $x \in U$ is represented by the vector $[a_i(x)]_{i=1}^N$ where $< a_1, a_2, \ldots, a_N >$ is a fixed ordering of attributes.

We consider a pair $a, b$ of things in $U$ such that $a, b$ have in common attribute values in a set $\Delta \subset A$, where $\Delta$ is non–empty; formally, this means that $IND(a, b) = \Delta$. Let

$$card(\Delta) = \delta \cdot N. \tag{5}$$

*Example 1.* We consider things $a, b$ and a thing $c$ which agrees with $a$ and $b$ on the set $\Delta$ and has $\frac{1}{2} \cdot (N - \delta \cdot N)$ attribute values in common with $a$ and $\frac{1}{2} \cdot (N - \delta \cdot N)$ attribute values in common with $b$; we assume for simplicity that it is possible otherwise we should consider $\lfloor \frac{1}{2} \cdot (N - \delta \cdot N) \rfloor$ and $\lceil \frac{1}{2} \cdot (N - \delta \cdot N) \rceil$, respectively which would only make calculations more cumbersome. Please observe that we do not specify positions of values, i.e., we do not specify particular attributes on which these values are taken, with exception for $\Delta$ only. Therefore, it makes sense to identify all such things into a class $[c]$ in which all things have values on attributes in $\Delta$ same as $a$ and $b$ and share with each of $a, b$ one half of the remaining values.

Our Thesis is

**Proposition 2.** *$c$ is between $a$ and $b$.*

*Proof.* We find the distance $d_{\mu_L}$ between $a, c$ and $b, c$ with respect to the rough inclusion $\mu_L$; by definition (1),

$$d_{\mu_L}(a, c) = \frac{\delta \cdot N + \frac{1}{2} \cdot (N - \delta \cdot N)}{N} = \frac{1 + \delta}{2} = d_{\mu_L}(c, b). \tag{6}$$

We now consider an arbitrary thing $x$ which for some quotient $\alpha$ in $[0, 1]$ has $\alpha \cdot \delta \cdot N$ values of attributes in $\Delta$ in common with $a$ and for some quotient $\beta \in [0, 1]$ has $\beta \cdot (N - \delta \cdot N)$ values of attributes not in $\Delta$ in common with $a$ and at most $(1 - \beta) \cdot (N - \delta \cdot N)$ values in common with $b$. We have

$$d_{\mu_L}(x, a) = \frac{\alpha \cdot \delta \cdot N + \beta \cdot (N - \delta \cdot n)}{N} = \beta + (\alpha - \beta) \cdot \delta, \tag{7}$$

and

$$d_{\mu_L}(x, b) \leq \frac{\alpha \cdot \delta \cdot N + (1 - \beta) \cdot (N - \delta \cdot n)}{N} = 1 - \beta + (\alpha + \beta - 1) \cdot \delta. \tag{8}$$

Let us assume, to the contrary, that (1) $d_{\mu_L}(x, a) > d_{\mu_L}(c, a)$ and (2) $d_{\mu_L}(x, b) > d_{\mu_L}(c, b)$.

Condition (1) means after substitution of values in (6) and (7) that

$$\beta + (\alpha - \beta) \cdot \delta > \frac{1}{2} + \frac{\delta}{2}, \tag{9}$$

i.e.,

$$\alpha \cdot \delta + \beta \cdot (1-\delta) > \frac{1}{2} + \frac{\delta}{2}. \tag{10}$$

Similarly, condition (2) yields after values in (6) and (8) are substituted into it,

$$\alpha \cdot \delta - \beta \cdot (1-\delta) > \frac{3}{2} \cdot \delta - \frac{1}{2}. \tag{11}$$

Adding inequalities (10) and (11) yields

$$2 \cdot \alpha \cdot \delta > 2 \cdot \delta, \tag{12}$$

and, as $\delta > 0$,

$$\alpha > 1 \tag{13}$$

which is impossible. This proves our proposition.

This example has served as a motivation for further generalizations.

We have already observed that we have not specified the attributes selected and actually we have discussed classes of equivalence of things, two things $x, y$ being equivalent if and only if they have had same fractions of attribute values for attributes in $\Delta$ and same fraction of attribute values not in $\Delta$. Hence, for fractions $\gamma$ of values in $\Delta$ and $\varepsilon$ of values in $A \setminus \Delta$ common with $a$, and, at most $1-\varepsilon$ values of attributes in $A \setminus \Delta$ in common with $b$, we denote with the symbol $[\gamma, \varepsilon]$ the class of things satisfying those conditions. We regard, hence, the vector $[\gamma, \varepsilon]$ as the *representation* of that class in the vector space $R^2$, the *representation space*. In particular, the thing $a$ is represented as $[1,1]$, the thing $b$ is represented as $[1,0]$, and, the thing $c$ is in the class $[1, \frac{1}{2}]$. Thus, in the Euclidean plane, $c$ is the midpoint in the segment with endpoints $b$ and $c$, i.e. $c$ is between $a$ and $b$ in the elementary geometry sense.

The proof above does suggest a more general result.

**Proposition 3.** *For* $\alpha \in [0,1]$*, the class* $[1, \alpha]$ *is between classes* $[1,0]$ *of* $b$ *and* $[1,1]$ *of* $a$ *in the representation space.*

*Proof.* This proof goes on similar lines as proof of the previous proposition. Let $d$ denotes the class $[1, \alpha]$ and let $x$ be in the class $[\gamma, \varepsilon]$. Assuming to the contrary that (1)$d_{\mu_L}(x,a) > d_{\mu_L}(d,a)$ and (2) $d_{mu_L}(x,b) > d_{mu_L}(d,b)$, we obtain inequalities

$$\gamma \cdot \delta + \varepsilon \cdot (1-\delta) > \delta + \alpha \cdot (1-\delta), \tag{14}$$

and,

$$\gamma \cdot \delta + (1-\varepsilon) \cdot (1-\delta) > \delta + (1-\alpha) \cdot (1-\delta). \tag{15}$$

Sidewise addition of (14) and (15) yields the inequality

$$2 \cdot \gamma \cdot \delta > 2 \cdot \delta, \tag{16}$$

i.e., $\gamma > 1$, impossible. The proposition is proved.

It turns out that the whole interval from $[1,0]$ to $[1,1]$ consists of classes between $[1,0]$ and $[1,1]$. In the representation space of classes $[\gamma, \varepsilon]$, betweennness in the mereological sense coincides with betweenness in the geometric sense of the Euclidean geometry of the plane.

### 2.1 The Case $\Delta = \emptyset$

In this case $\delta = 0$ and our proofs above are not valid. In this case, given things $a, b$ in $U$, with $IND(a, b) = \emptyset$, for a choice of $\gamma \in [0, 1]$, we form things which have $\gamma \cdot N$ attribute values in common with $a$ and $(1 - \gamma) \cdot N$ attribute values in common with $b$. We represent this class of things as before with the vector $[\gamma, 1 - \gamma]$ in the Euclidean plane. In this representation, $a$ is represented as $[1, 0]$ and $b$ is represented as $[0, 1]$, so $[\gamma, 1 - \gamma]$ is a convex combination of $[1, 0]$ and $[0, 1]$.

**Proposition 4.** *For each choice of $\gamma \in [0, 1]$, the class of things represented as $[\gamma, 1-\gamma]$ is between $a$ and $b$ in the sense of betweenness relation $B$.*

*Proof.* For a point $[\varepsilon, \delta]$ with $\varepsilon, \delta \in [0, 1]$ and $\varepsilon + \delta \leq 1$, if, e.g., $\varepsilon > \gamma$ then $\delta < 1 - \gamma$.

It follows that things and their classes which are between $a$ and $b$ are located in the representation space in the segment with endpoints $[1, 0]$ for $a$ and $[0, 1]$ for $b$, i.e., between these endpoints in the geometrical sense.

This suggests a generalization. We define a more general betweenness relation

$$GB(x, a_1, a_2, \ldots, a_n)$$

('*x is between* $a_1, a_2, \ldots, a_n$') if and only if for each thing $y \neq x$ he thing $x$ is closer than $y$ to some $a_i$ in the mereological sense of (3).

### 2.2 The General Case

We consider a set $V = \{a_1, a_2, \ldots, a_n\}$ of things in $U$. For a choice of $\gamma_1, \gamma_2, \ldots, \gamma_n \in [0, 1]$ with $\sum_i \gamma_i = 1$, which we summarily denote by the vector $\overline{\gamma}$, we denote as $(V, \overline{\gamma})$ the class of things which have the fraction $\gamma_i$ of attribute values in common with the thing $a_i$. As above, the fact is true that

**Proposition 5.** *The class $(V, \overline{\gamma})$ of things represented by the vector $\overline{\gamma} = [\gamma_1, \gamma_2, \ldots, \gamma_n]$ satisfies the relation $GB((V, \overline{\gamma}), a_1, a_2, \ldots, a_n)$.*

Proof of this proposition goes on lines of the preceding proof.

**Proposition 6.** *The relation $GB(x, a_1, a_2, \ldots, a_n)$ holds for a class $x$=*

$$(\{a_1, ..., a_n\}, [\gamma_1, \gamma_2, \ldots, \gamma_n]) \tag{17}$$

*if and only if $[\gamma_1, \gamma_2, \ldots, \gamma_n]$ belongs in the convex hull of vectors $[1, 0, 0..., 0]$, $[0, 1, 0, 0, ..., 0]$, ..., $[0, 0, ..., 0, 1]$ representing, in this order, classes $a_1, a_2, ..., a_n$.*

## 3 A Geometric Representation of Granulation

In the general case, the process of forming of a class $x$=$(V, \overline{\gamma})$ represented by the vector $\overline{\gamma} = [\gamma_1, \gamma_2, ..., \gamma_n]$ can be regarded as forming of a granule of things $gr(V, \overline{\gamma})$. This granulation process is different of previously considered in that it does involve a sort of a *shuffling* map, which takes into $x$ things having specified fractions of attribute values in common with the corresponding classes $a_i$ but over *arbitrary sets* of attributes of cardinality $\gamma_i \cdot N$. This secures a kind of control over values in contradiction to our previous granulation paradigm in which only a fixed part of a given thing attribute values was coming from the granule center, cf., [3], Chs. 5–7.

For a given vector $\overline{\gamma}$=$[\gamma_1, \gamma_2, \ldots, \gamma_n]$ in the representation space, representing the granule $gr(V, \overline{\gamma})$, the size of the granule $gr(V, \overline{\gamma})$ is

$$\prod_{j \leq n} \binom{N - \sum_{k<j} \gamma_k \cdot N}{\gamma_j \cdot N}. \tag{18}$$

The number of granules of type represented by a vector $\overline{\gamma} = [\gamma_1, \gamma_2, \ldots, \gamma_n]$ is the number of sequences $k_1, k_2, \ldots, k_n$ of natural numbers such that $\sum_i k_i = N$, i.e., it equals the number of integer–valued vectors on the hyperplane $\sum_i x_i = N$, given recurrently by the function $\phi(N, n)$:

$$\phi(N, 1) = 1, \tag{19}$$

$$\phi(N, k+1) = \sum_{j=0}^{N} \phi(N-j, k). \tag{20}$$

It follows that $\phi(N, n)$ is of order $\Theta(N^{n-1})$.

Let us observe that those estimates concern all objects generated by the process described above; in reality, only a small fraction of those objects will exist in the set $U$. Hence, we define a *U–granule* $g^U(\overline{\gamma})$ as $g(\overline{\gamma}) \cap U$.

### 3.1 Granular Classifiers

Assume that a *decision partition* is imposed on the things in the set $U$ into *decision classes* $D_1, D_2, \ldots, D_m$. For a U–granule $g = g^U(\overline{\gamma})$, we denote with $Pr[g, D_i]$ the probability that a randomly chosen thing in $g$ is assigned to the class $D_i$. Then the *Bayesian decision on* $g$ is the class $D(g) = D_{i^*}$ such that $Pr[g, D_{i^*}] = max\{Pr[g, D_i] : i = 1, 2, \ldots, m\}$.

The mapping $g \rightarrow D(g)$ from $U$–granules into their Bayesian decisions is the *Bayesian granular classifier*. It is deterministic as the covering into granules is finer than the partition into decision classes.

## 4 An Application Proposal: Conflict Resolution

We propose to apply the afore described results to the problem of conflict resolution. In a conflict, we have a finite number of agents declaring their standpoints on a number of issues; those standpoints are conflicting, i.e., on each issue there are agents having distinct standpoints. A resolution of a conflict means some process leading to a rationally chosen set of non–conflicting standpoints for all agents. Examples can be saddle points in two–person zero–sum games, either for pure or mixed strategies, or, the Nash equilibrium points in continuous convex–concave games. In each of these cases, agents reach a set of strategies which is satisfying for each of them.

Such a rationale is difficult to be obtained in less formalized conflicts, and, our example deals with such a conflict. We base our approach on the example of a conflict in Pawlak [2], for which that author gave an analysis in terms of the rough set–theoretical approach.

*Example 2.* The conflict described in [2] does involve six agents: A1, A2,..., A6 and five issues I1, I2, ..., I5 with the standpoints 0 (disagreement), 1 (agreement) and N (the neutral standpoint equivalent to 'do not care'). Hence, when standpoints N and 0 or 1 are confronted, N means 0 or 1, respectively.

This case is visualized in Table Fig.1 below.

| agent | I1 | I2 | I3 | I4 | I5 |
|---|---|---|---|---|---|
| A1 | 0 | 1 | 1 | 1 | 1 |
| A2 | 1 | N | 0 | 0 | 0 |
| A3 | 1 | 0 | 0 | 0 | N |
| A4 | N | 0 | 0 | N | 0 |
| A5 | 1 | 0 | 0 | 0 | 0 |
| A6 | N | 1 | 0 | N | 1 |

**Fig. 1.** The setting of a conflict

In this case, the set of agents $V = \{A_{i_1}, A_{i_2}, \ldots, A_{i_k}\}$ forming a granule is called a *coalition*, and standpoints of them on issues are their *pure strategies*. Given coefficients $\gamma_1, ..., \gamma_k$ summing up to 1, sequences of issues in the granule $(V, [\gamma_1, ..., \gamma_k])$ are *mixed strategies of agents*. Given a partition of the set of agents into coalitions $C_1, C_2, ..., C_m$, we call a conflict resolution a set $i_1, i_2, ..., i_m\}$ of mixed strategies issued from granules generated by $C_1, C_2, .., C_m$ which are consistent, i.e., conflict–less.

We use the Lukasiewicz rough inclusion as the measure of similarity of agents, and, we include in one coalition agents for which pairwise similarity measures are greater than their similarity degrees to other agents.

From Fig. 1, it follows that agents A1 and A6 are similar to degree of 0.8, agents A2, A3, A4, A5 are similar to each other to degree of 1.0, and, agent A1 is similar to A2, A3, A4, A5 to degree not greater than 0.4. Only A6 is similar to A2, A3, A4 to degree 0.8, and, to A5 to degree 0.6.

We decide to consider coalitions $V_1 = \{A1, A6\}$ and $V_2 = \{A2, A3, A4, A5\}$. Forming all possible granules over $V_1$ and $V_2$ yields possible mixed strategies of both coalitions. Due to their number, we list in Figs.2,3 a selection of six of those strategies for each of the two coalitions.

We find exemplary boldfaced mixed strategies **N1ON1** for the coalition $V1$ and **NN0NN** for the coalition $V2$ which are identical for both coalitions: each of them can provide a conflict resolution; it is true that this requires a ‘gentleman’s agreement’ to accept the results of this procedure; for instance, each of these solutions would require that A1 gives up on its standpoint on issue I1, which in real practice can be impossible due, e.g., to political and religious reasons. The specific bargained for standpoints can be: $01001, 11001, 01011, 11011$.

| | | | | | |
|---|---|---|---|---|---|
| N11N1 | N1011 | **N10N1** | N1011 | 010N1 | 01111 |

**Fig. 2.** Selected mixed strategies for the coalition $V_1$

| | | | | | |
|---|---|---|---|---|---|
| 1N00N | 1N000 | NN0N0 | **NN0NN** | N00N0 | 100NN |

**Fig. 3.** Selected mixed strategies for the coalition $V_2$

## 5 The Thing Point of View: Hyper–Granules

The approach presented above begins with a group $G$ of things and ends with the notion of collection of things between $G$. Among those things are virtual ones not present in the decision/information system. Therefore, we now propose the analysis from the viewpoint of things in the information system.

We consider a maximal set of things $\mathcal{X}$ with the property

$$\forall x \in \mathcal{X}.\exists Y \subseteq \mathcal{X} \setminus \{x\}.Btw(x, Y). \tag{21}$$

We call $\mathcal{X}$ a *hyper–granule*.

Lemma 1. For each $x \in \mathcal{X}$, it is true that $Btw(x, \mathcal{X} \setminus \{x\}$.

Lemma 2. For each $y \notin \mathcal{X}$, it is true that $Btw(y, Y)$ holds for no $Y \subseteq \mathcal{X}$. Moreover, there exists an attribute $a$ such that the value $a(y) \neq a(x)$ for each $x \in \mathcal{X}$. Contrarily, for each $x \in \mathcal{X}$, and, for each attribute $a$, it is true that $a(x) \in \{a(y) : y \in \mathcal{X} \setminus \{x\}$.

The hyper–granule $\mathcal{X}$ is *attribute–self–contained* in the sense of Lemma 2.

For the hyper–granule $\mathcal{X}$, we consider the complementary set of objects $U \setminus \mathcal{X}$;

let $\mathcal{X}'$ be a hyper–granule of objects in $U \setminus \mathcal{X}$. Iterating this procedure, we define in the universe set $U$ a set of hyper–granules $\{\mathcal{X}^{(i)} : i \in I\}$, where $\mathcal{X}^{(j+1)}$ is a hyper–granule in the set $U \setminus \bigcup_{i \leq j} \mathcal{X}^{(i)}$. The remnant of $U$ consist of outliers unable to enter any hyper–granule.

We denote with the symbol $H(x)$ the hyper–granule containing the thing $x$. For a thing $x$, we consider sets of thing in $H(x) \setminus \{x\}$ which we denote with the generic symbol $N(x)$ with the property that $Btw(x, N(x))$ and all coordinates of the vector representing $x$ with respect to $N(x)$ are positive; we will call any such set a *neighborhood of x*; in the example in Fig.4, for instance, $N(4) = \{8, 10\}$.

A neighborhood $N(x)$ of $x$ is *irreducible* if it is of minimal possible cardinality; such is $N(4)$ pointed to above..

# 6 Applications: Hyper–Granules as Coalitions in Conflicts and a Decision Prediction Modal Logic

We will discus first the problem of conflict resolution.

## 6.1 Hyper–Granules as Coalitions in Conflicts

We return to Fig. 1, showing standpoints of agents A1–A6 on issues I1–I5. As before, we regard the standpoint $N$ as *'don't care'* a fortiori $N$ can be 0 or 1. We observe that A1 cannot be considered as a candidate to the hyper–granule of A2–A6 as the issue I1 on A1 takes value 0 not taken by any of A2–A6.

On the other hand, due to our convention about $N$, agents A2–A6 make a hyper–granule $\mathcal{X}$. This hyper–granule can produce a between element $NN0NN$ representing sixteen specific bargaining propositions, from 00000 to 11011. The agent A1 remains as an outlier due to value 0 on I1 which cannot be supplied by any other agent.

The bargaining between $\mathcal{X}$ and A1 can focus on I3 on which A1 takes value of 1 and all other agents adopt the standpoint 0.

## 6.2 A Decision Prediction Logic and a Decision Assignment Algorithm

We will regard the universe $U$ of the informaton system as the training set on which the decision is given, and we consider the additional *test set* on which decision is to be learned on the basis of its values on $U$.

Given a test thing $x$ about which we assume that there exist neighborhoods of it in the set $U$, and for a formula $\phi$, we declare that

$$x \models \phi \Leftrightarrow y \models \phi \text{ for each irreducible } N(x) \text{and each } y \in N(x)\ . \tag{22}$$

Modal operators $L$ of necessity and $M$ of possibility are introduced as follows.

$$x \models L\phi \Leftrightarrow y \models \phi \text{ for each } \ N(x) \text{ and each } y \in N(x)\ . \tag{23}$$

$$x \models M\phi \Leftrightarrow y \models \phi \text{ for some } N(x) \text{ and some } y \in N(x)\ . \quad (24)$$

It follows that

$$L = \neg M \neg. \quad (25)$$

$$x \models L(\phi \Rightarrow \psi) \Rightarrow (x \models L\phi \Rightarrow x \models L\psi). \quad (26)$$

$$x \models L\phi \Rightarrow x \models \phi. \quad (27)$$

Our decision relation formula $\phi$ is $v_d \in A$ where $A \subseteq V_d$, i.e, $A$ is a subset of the set of decision values; the formula $x \models (v_d \in A)$ reads that decision value proposed for $x$ is in the set $A$ of decision values. Necessitation means stressing this hypothesis by conforming it on all neighborhoods of $x$, and, possibility indicates possible sets of values of decision for $x$.

## 7 Appendix: Computational Aspects of Hyper–Granules

We consider an information system $\mathcal{S} = (U, A, V)$ where $U$ is a set of things, $A$ is a set of attributes,and, $V$ is a set of *attribute values*. We will need the notion of a *dual information matrix* $\mathcal{S}^*$ defined as the triple $(A, V, U)$, where for each pair $(a, v)$ the entry in the cell $\mathcal{S}^*(a, v)$ is

$$\{x \in U : a(x) = v\}. \quad (28)$$

**Computing hyper–granules**

The Algorithm proceeds as follows.

**HYPER–GRANULE (U,A,V)**
1. Form the dual information matrix $\mathcal{S}^*$;
2. For each $x$ in $U$ do
3. if there is a cell $(a, v)$ such that $\mathcal{S}^*(a, v) = \{x\}$
4. then remove $x$ from all cells.
6. Repeat steps 2-4 until
7. all cells are either empty or each contains at least two things;
8. Return $\mathcal{X}$=the set of all things that occur in at least one non–empty cell.

Repeating the algorithm with the dual information matrix $(A, V, U \setminus \mathcal{X})$ we may eventually obtain further hyper–granules.

The complexity of the algorithm is $\Theta(|A| \cdot |V| \cdot |U|)$ as we do take into account neither the cost of inserting a symbol into a cell nor deleting it from a cell.

We consider an information system TEST in Fig. 4. The dual information matrix TEST$^*$ is shown in Fig. 5. We have to remove from all cells the things $2, 5, 9$. As each remaining cell is either empty or some at least two–element set, the hyper–granule is $\mathcal{X} = \{1, 3, 4, 6, 7, 8, 10\}$. For the remaining subsystem TEST1=($A, V, \{2, 5, 9\}$), we obtain the dual information matrix TEST1* shown in Fig. 6. After steps 2–4 of the algorithm, all cells are empty so there is no second hyper–granule: things $2, 5, 9$ are outliers.

We would like to observe in addition that it is easy to read off from the dual information matrix the coordinates of a thing in the representation space; e.g., the thing $4$ can be represented as the vector $[0, 0, 0, 0, 0, 0, 0, \frac{1}{3}, 0, \frac{2}{3}]$ in the simplex spanned on unit vectors representing, respectively, things $1, 2, ..., 10$ in the vector space $R^{10}$. Hence, $N(4) = \{8, 10\}$ is an irreducible neighborhood of the thing $4$; another irreducible neighborhood of $4$ is $\{1, 10\}$ with coordinates $[\frac{1}{2}, 0, ..., 0, \frac{1}{2}]$.

### 7.1 A Decision Assigning Algorithm

For a new test thing $y$, we consider in the universe $U$ (the training set) the hyper–granule $H(y) \subseteq U \cup \{y\}$ along with irreducible neighborhoods $N_1(y), \ldots, N_k(y)$ (if there are any). Let $D(y) \subseteq V_d$ be the least set of decision values with the property that

$$\text{If } x \in \bigcup_i N_i(y) \text{ then } d(x) \in D(y). \tag{29}$$

By (22), the hypothetical $d(y)$ belongs in $D(y)$. Now, given $j \leq k$, the neighborhood $N_j(y)$ votes for decision value $d_j(y)$ in the manner as follows. For $z \in N_j(y)$, we denote with $q_j(z)$ the coordinate with which $z$ enters the vector representing $y$ with respect to $N(y)$. Then

$$d_j(y) = argmin_{v \in D(y)} ||v - \sum_z q_j(z) \cdot d(z)||, \tag{30}$$

where $|| - ||$ is a metric chosen for Euclidean space containing $V_d$. Given a parameter $p \leq k$, we select $p$ neighborhoods from among $N_1(y), N_2(y), \ldots, N_k(y)$ with greatest values of $q_j = max_z q_j(z)$ and let

$$d(y) = argmin_{v \in D(y)} ||v - \sum_{j \leq p} \frac{q_j}{\sum_{j \leq p} q_j} \cdot d_j(y)||. \tag{31}$$

This approach may be regarded as a two–step variant of *nearest neighbor classifier*: neighborhoods which are irreducible are at the same time closest to the thing; an important factor not present in usual nearest neighbor classifiers is that neighborhoods guarantee also that attribute values in the thing come from neighborhood members which double stresses the similarity among the thing and neighborhood members.

To give a procedure for computing neighbors of things, we define some useful notions,

$$I(x, y) = \{a \in A : a(x) = a(y)\}, \tag{32}$$

and,

$$f(x, y) = \frac{card(I(x, y))}{card(A)}. \tag{33}$$

| thing | $a_1$ | $a_2$ | $a_3$ | $a_4$ | $a_5$ | $a_6$ |
|---|---|---|---|---|---|---|
| 1 | 1 | 0 | 0 | 1 | 1 | 0 |
| 2 | 0 | 1 | 1 | 0 | 1 | 2 |
| 3 | 1 | 0 | 0 | 1 | 0 | 1 |
| 4 | 0 | 0 | 1 | 1 | 1 | 0 |
| 5 | 0 | 0 | 2 | 1 | 0 | 0 |
| 6 | 1 | 1 | 0 | 0 | 1 | 1 |
| 7 | 0 | 1 | 0 | 1 | 0 | 1 |
| 8 | 1 | 0 | 1 | 0 | 1 | 0 |
| 9 | 1 | 2 | 0 | 0 | 2 | 1 |
| 10 | 0 | 0 | 1 | 1 | 0 | 0 |

**Fig. 4.** Information system TEST

| value | $a_1$ | $a_2$ | $a_3$ | $a_4$ | $a_5$ | $a_6$ |
|---|---|---|---|---|---|---|
| 0 | $\{2,4,5,7,10\}$ | $\{1,3,4,5,8,10\}$ | $\{1,3,6,7,9\}$ | $\{2,6,8,9\}$ | $\{3,5,7,10\}$ | $\{1,4,5,8,10\}$ |
| 1 | $\{1,3,6,8,9\}$ | $\{2,6,7\}$ | $\{2,4,8,10\}$ | $\{1,3,4,5,7,10\}$ | $\{1,2,4,6,8\}$ | $\{3,6,7,9\}$ |
| 2 | $\emptyset$ | $\{9\}$ | $\{5\}$ | $\emptyset$ | $\{9\}$ | $\{2\}$ |

**Fig. 5.** The dual information matrix $TEST^*$

| value | $a_1$ | $a_2$ | $a_3$ | $a_4$ | $a_5$ | $a_6$ |
|---|---|---|---|---|---|---|
| 0 | $\{2,5\}$ | $\{5\}$ | $\{9\}$ | $\{2,9\}$ | $\{5\}$ | $\{5\}$ |
| 1 | $\{9\}$ | $\{2\}$ | $\{2\}$ | $\{5\}$ | $\{2\}$ | $\{9\}$ |
| 2 | $\emptyset$ | $\{9\}$ | $\{5\}$ | $\emptyset$ | $\{9\}$ | $\{2\}$ |

**Fig. 6.** The dual information matrix $TEST1^*$

For $X \subseteq A$ and $x \in U$, we let $V(x,X) =$ the sequence $a(x) : a \in X$ in the assigned order of attributes; the symbol $0^n$ denotes the sequence of length $n$ of 0.

**Procedure Irreducible Neighborhood(x)**

Input: the thing x

Output List of irreducible neighborhoods of x

variable sets $\overline{z}$, $A_{\overline{z}}(x,y)$, $n_{\overline{z}}(x,y)$, $\mathcal{N}(x,y)$, $v_{\overline{z}}(x,y)$

Initialization: $\mathcal{N}(x,y) \leftarrow \emptyset$,
1. order coefficients $f(x,y)$ in descending order

2. in descending order of $f(x,y)$ do

3. while $A_{\overline{z}}(x,y) \neq A$ do

4. $\overline{z} \leftarrow \{y\}$, $A_{\{y\}}(x,y) \leftarrow I(x,y)$, $n_{\{y\}}(x,y) \leftarrow \{y\}$, $v_{\{y\}}(x,y) \leftarrow [V(x, I(x,y)), 0^{|A|-|I(x,y)|}]$

5. in descending order of $f(x,z')$, for each $z' \notin \overline{z}$ such that $(A \setminus A_{\overline{z}}(x,y)) \cap I(x,z') \neq \emptyset$ do

6. $A_{\overline{z}\cup\{z'\}}(x,y) \leftarrow A_{\overline{z}}(x,y) \cup (A \setminus A_{\overline{z}}(x,y)) \cap I(x,z'))$,

$n_{\overline{z}\cup\{z'\}}(x,y) \leftarrow n_{\overline{z}}(x,y) \cup \{z'\}$,

$\overline{z} \leftarrow \overline{z} \cup \{z'\}$,

$v_{\overline{z}\cup\{z'\}}(x,y) = v_{\overline{z}}(x,y) + [0^{|A|-|v_{\overline{z}}(x,y)|}, V(x,(A \setminus A_{\overline{z}}(x,y)) \cap I(x,z'), 0^{|A|-|v_{\overline{z}}(x,y)|-|V(x,(A\setminus A_{\overline{z}}(x,y))\cap I(x,z')|}]$,

$p(z') \leftarrow |V(x,(A \setminus A_{\overline{z}}(x,y)) \cap I(x,z'))|$

7. if $A_{\overline{z}}(x,y) = A$ then

8. $\mathcal{N}(x,y) \leftarrow \mathcal{N}(x,y) \cup \{n_{\overline{z}}(x,y)\}$

9. from $\bigcup_y \mathcal{N}(x,y)$ output the list of sets of minimal cardinality=the list of irreducible neighborhoods of $x$

10. $\frac{|I(x,y)|}{|I(x,y)|+\sum_{z'\in n_{\overline{z}}\setminus\{y\}} p(z')}$ is the coordinate of $y$ in the representation vector of $x$, and, $\frac{p(z')}{|I(x,y)|+\sum_{z'\in n_{\overline{z}}\setminus\{y\}} p(z')}$ is the coordinate of $z'$ in the representation vector of $x$ with respect to the irreducible neighborhood $n_{\overline{z}}(x,y)$.

As an example let us assign decision values to things $1, 2, \ldots, 10$ in TEST, in Fig. 7.1. For simplicity, the thing 4 is regarded as a test thing. The set $D(4)$ is $\{0, 1\}$. Irreducible neighborhoods of 4 are of cardinality 2 and they are: $N_1(4) = \{1, 10\}$ with coordinates $[\frac{1}{6}, \frac{5}{6}]$; $N_2(4) = \{1, 10\}$ with coordinates $[\frac{4}{6}, \frac{2}{6}]$; $N_3(4) = \{7, 8\}$ with coordinates $[\frac{2}{6}, \frac{4}{6}]$; $N_4(4) = \{8, 10\}$ with coordinates $[\frac{1}{6}, \frac{5}{6}]$.

Now, $N_1(4)$ votes for $\frac{1}{3} \cdot 1 + \frac{2}{3} \cdot 1 = 1 = d_1(4)$;
$N_2(4)$ votes for $\frac{4}{6} \cdot 0 + \frac{2}{6} \cdot 1 \rightarrow 0 = d_2(4)$;
$N_3(4)$ votes for $\frac{2}{6} \cdot 1 + \frac{4}{6} \cdot 1 = 1 = d_3(4)$;
$N_4(4)$ votes for $\frac{1}{3} \cdot 1 + \frac{2}{3} \cdot 1 = 1 = d_4(4)$.

The final decision value is the nearest of $0, 1$ to the value $\frac{\frac{2}{3}\cdot 1+\frac{5}{6}\cdot 1+\frac{4}{6}\cdot 0+\frac{4}{6}\cdot 1}{\frac{2}{3}+\frac{5}{6}+\frac{4}{6}+\frac{4}{6}} = \frac{13}{17}$ which is $d(4) = 1$.

| value | $a_1$ | $a_2$ | $a_3$ | $a_4$ | $a_5$ | $a_6$ |
|---|---|---|---|---|---|---|
| 0 | $\{4,7,10\}$ | $\{1,3,4,8,10\}$ | $\{1,3,6,7\}$ | $\{6,8\}$ | $\{3,7,10\}$ | $\{1,4,8,10\}$ |
| 1 | $\{1,3,6,8\}$ | $\{6,7\}$ | $\{4,8,10\}$ | $\{1,3,4,7,10\}$ | $\{1,4,6,8\}$ | $\{3,6,7\}$ |

**Fig. 7.** The dual information matrix $TEST^*$ after removing outliers

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 2 | 1 | *new thing* | 1 | 2 | 1 | 1 | 0 | 1 |

**Fig. 8.** Decision values for TEST; thing 4 regarded as a test thing

## References


1. Van Benthem, J.: The Logic of Time. D. Reidel, Dordrecht, 1983.
2. Pawlak, Z.: An inquiry into anatomy of conflicts. Journal of Information Sciences 109 (1998), pp 65–78.
3. Polkowski, L.: Approximate Reasoning by Parts. An Introduction to Rough Mereology. Springer Verlag, Berlin, ISRL vol. 20, 2011.
4. Tarski, A., Givant, S.: Tarski's system of geometry. Bull. Symbolic Logic 5(2), June 1999, pp 175–214.