Chapter 7

Reasoning with Rough Inclusions

Granular Computing, Granular Logics, Perception Calculus, Cognitive and MAS Reasoning

Rough mereology allows for a plethora of applications in various reasoning schemes due to universality of its primitive predicate of a part to a degree. We have already stressed that by its nature, rough mereology is especially suited to reasoning with collective concepts like geometric figures or solids, or, concepts learned by machine learning methods, i.e., with collective concepts. Those applications are presented in Ch. 8 and Ch. 9. In this chapter, we begin this discussion with a formal approach to the problem of granulation of knowledge and then we examine rough mereological logics: from our results in Ch. 6 it follows that representing implication with a rough inclusion μ leads to logics which extend and generalize fuzzy logics. As an application, we propose a formal rendering of the idea of perception calculus, due to Zadeh [67]. We apply rough mereological schemes to reasoning by multi–agent (MAS) systems, and finally we present a rough mereological variant of cognitive reasoning in neural–like systems.

0.1 On granular reasoning

The creator of Fuzzy Set Theory Lotfi A. Zadeh proposed to compute with granules in Zadeh [66]. The idea was natural, as fuzzy reasonings are carried out in terms of fuzzy membership functions. A fuzzy membership function μ_X maps a universe U of objects into the interval $[0, 1]$ and it represents the membership in a set X as a membership to a degree. The value $\mu_X(x) = r$ is interpreted as the statement that the object x is an element of the set X to the degree of r. The mapping $U \to U/\mu_X$ which sends each object x to its fibre $\mu_X^{-1}(\mu_X(x))$ identifies into the granule $g_\mu(r)$ all objects that belong in X to the degree r.

All fuzzy constructs are then expressed in terms of those granules. In this sense, fuzzy reasoning is in a natural way reasoning with granules. Granules in this reasoning are constructed in a uniform way, i.e., all objects in a granule share the same property of external character: they belong to an "oracle" X to the same degree; changing X produces a variety of granules to reason with. The relation forming any granule is an equivalence R_X: the universe is decomposed into fibres of a fuzzy membership function μ_X in question, and the knowledge base is composed of all relations R_X for subsets $X \subseteq U$ of the universe of objects.

The same conclusion concerns rough set reasoning: elementary objects in reasoning are indiscernibility classes – elementary granules which are elements of a partition of the universe of objects by an indiscernibility relation $IND(B)$ for some set B of attributes, see Ch. 4. These granules are the smallest objects which can be described in terms of attributes and their values, i.e., in terms of descriptors and they are used in forming descriptions of objects, in building decision rules and classifiers as well as control algorithms.

Reasoning by means of aggregating objects, situations, cases, etc., into granules of similar entities is common to all forms of human reasoning. It is therefore important to capture this form of reasoning in its essential and typical facets and render it in mereological environment with help of mereological notions.

In Lin [21], [22], topological character of granules was recognized and the basic notion of a neighborhood system as the meaning of the collection of granules on the universe of objects was brought forth, Lin [22] recognized the import of tolerance relations, see Nieminen [37], cf., Ch. 1, sect. 13, by discussing tolerance induced neighborhoods.

In all hybrid approaches involving fuzzy or rough sets along with neural networks, genetic algorithms, etc., etc., one is therefore bound to compute with granules; this fact testifies to the importance of granular structures.

In search of adequate similarity relations, various forms of granules were proposed and considered as well as experimentally verified as to their effectiveness. In information systems, indiscernibility classes were proposed as granules, or, more generally, *templates* have served that purpose, i.e., meanings of generalized descriptors of the form $(a \in W_a)$ where $W_a \subseteq V_a$ with the meaning $[[(a \in W_a)]] = \{u \in U : a(u) \in W_a\}$, see Nguyen S. H. [36]; clearly, templates are aggregates, in ontological sense, of descriptors, i.e., they form "big" granules. Their usage is motivated by their potentially greater descriptive force than that of descriptors; a judicious choice of sets W_a should allow for constructing of a similarity relation that would reflect satisfactorily well the dependence of decision on conditional attributes.

As means for granule construction, rough inclusions have been considered and applied, in Polkowski [41] – [52]. The idea of granule formation and analysis rests on usage of the mereological class operator in the framework of mereological reasoning.

Granules formed by rough inclusions are used in models of fusion of knowledge, rough–neural computing and in building many–valued logics reflecting the rough set ideology in reasoning, and these forms of reasoning are discussed in further parts of this chapter.

Granulation of knowledge can be considered from a few angles

1. *General purpose of granulation*;

2. *Granules from binary relations*;

3. *Granules in information systems from indiscernibility*;

4. *Granules from generalized descriptors*;

5. *Granules from rough inclusions – mereological approach.*

We briefly examine those facets of granulation.

0.2 On methods for granulation of knowledge

Granulation of knowledge comes into existence for a few reasons; the principal one is founded on the underlying assumption of basically all paradigms, viz., that reality exhibits a fundamental continuity, i.e., objects with identical descriptions in the given paradigm should exhibit the same properties with respect to classification or decision making, in general, in their behavior towards the external world.

For instance, fuzzy set theory assumes that objects with identical membership descriptions should behave identically and rough set theory assumes that objects indiscernible with respect to a group of attributes should behave identically, in particular they should fall into the same decision class.

Hence, granulation is forced by assumptions of the respective paradigm and it is unavoidable once the paradigm is accepted and applied. Granules induced in the given paradigm form the first level of granulation.

In the search for similarities better with respect to applications like classification or decision making, more complex granules are constructed, e.g., as unions of granules of the first level, or more complex functions of them, resulting, e.g., in fusion of various granules from distinct sources.

Among granules of the first two levels, some kinds of them can be exhibited by means of various operators, e.g., the class operator associated with a rough inclusion.

0.2.1 Granules from binary relations

Granulation on the basis of binary general relations as well as their specializations to, e.g., tolerance relations, has been studied by Lin [23] – [28], in particular as an important notion of a neighborhood system; see also Yao

[64], [65]. This approach extends the special case of the approach based on indiscernibility. A general form of this approach according to Yao [64], [65], exploits the classical notion of *Galois connection*, cf., Ch. 1 sect. 4: two mappings $f : X \rightarrow Y, g : Y \rightarrow X$ form a Galois connection between ordered sets $(X, <)$ and $(Y, \prec)$ if and only if the equivalence $x < g(y) \Leftrightarrow f(x) \prec y$ holds.

In an information system (U, A), for a binary relation R on U, one considers the sets $xR = \{y \in U : xRy\}$ and $Rx = \{y \in U : yRx\}$, called, respectively, the *successor neighborhood* and the *predecessor neighborhood* of x. These sets are considered as granules formed by a specific relation of being affine to x in the sense of the relation R. Other forms of granulation can be obtained by comparing objects with identical neighborhoods, e.g., $x \equiv y \Leftrightarrow xR = yR$.

Saturation of sets of objects X, Y with respect to the relation R leads to sets X^*, Y^* such that $X * R = Y* \Leftrightarrow RY^* = X^*$, forming a Galois connection. This approach is closely related to the Formal Concept Analysis of Wille [63].

0.2.2 Granules in information systems from indiscernibility

Granules based on indiscernibility in information/decision systems, see Ch. 4, sect. 2, are constructed as indiscernibility classes: given an information system (U, A) and the collection $IND = \{IND(B) : B \subseteq A\}$ of indiscernibility relations, each elementary granule is of the form of an indiscernibility relation $[u]_B = \{v \in U : (u, v) \in IND(B)\}$ for some B. Among those granules, there are minimal ones: granules of the form $[u]_A$ induced from the set A of all attributes.

Granules $[u]_A$ form the finest partition of the universe U which can be obtained by means of indiscernibility; given the class $[u]_B$ and the class $[u]_{A \setminus B}$, we have $[u]_A$=$[u]_B \cap [u]_{A \setminus B}$ hence $[u]_B = \bigcup_{v \in [u]_B \cap DIS_{A \setminus B}(u)} [v]_A$, where $v \in DIS_{A \setminus B}(u)$ if and only if there exists an attribute $a \in A \setminus B$ such that $a(u) \neq a(v)$. It is manifest that granules $[u]_B$ can be arranged into a tree with the root $[u]_\emptyset = U$ and leaves of the form $[u]_A$.

Granules based on indiscernibility form a complete Boolean algebra generated by atoms of the form $[u]_A$: unions of these atomic granules are closed on intersections and complements and these operations induce into granules the structure of a field of sets.

Atomic granules are made into some important unions by approximation operators, see Ch. 4, sect. 2: given a concept $X \subseteq U$, the lower approximation $\underline{B}X$ to X over the set B of attributes is defined as the union $\bigcup\{[u]_B : [u]_B \subseteq X\}$; the operator $L_B : Concepts \rightarrow \ Granules$ sending X to $\underline{B}X$ is monotone increasing and idempotent: $X \subseteq Y$ implies $LX \subseteq LY$ and $L \circ L = L$. Similarly, the upper approximation $\overline{B}X = \bigcup\{[u]_B : [u]_B \cap X \neq \emptyset\}$ to X over B, makes some elementary granules into the union; the operator U^B

sending concepts into upper approximations is also monotone increasing and idempotent.

0.2.3 Granules from generalized descriptors

Some authors have made use of generalized descriptors called *templates*, see Nguyen S. H. [36]; a template is a formula $T : (a \in W_a)$, where $W_a \subseteq V_a$, with the meaning $g(T) : \{u \in U : a(u) \in W_a\}$. The granule $g(T)$ can be represented as the union $\bigcup_{u \in W_a} [u]_a$; granules of the form $[u]_a$ are also called *blocks*, see Grzymala–Busse [15], Grzymala–Busse and Ming Hu [16].

0.3 Granules from rough inclusions: The mereological approach to granulation of knowledge

Assume that a rough inclusion μ is given along with the associated ingredient relation $ingr$, as in postulate RINC1, in Ch. 6, sect 1.

The granule $g_\mu(u, r)$ of the radius r about the center u is defined as the class of property $\Phi^\mu_{u,r}$

$$\Phi^\mu_{u,r}(v) \Leftrightarrow \mu(v, u, r) \tag{0.1}$$

The granule $g_\mu(u, r)$ is defined by means of

$$g_\mu(u, r) = Cls\Phi^\mu_{u,r} \tag{0.2}$$

Properties of granules depend, obviously, on the type of rough inclusion used in their definitions. We consider separate cases, as some features revealed by granules differ from a rough inclusion to a rough inclusion. The reader is asked to refer to Ch. 5 for description of mereological reasoning, which is going to be used in what follows.

In case of Archimedean t–norm–induced rough inclusions, see Ch. 6, sect. 3, or metric–induced rough inclusions, see Ch. 6, sect. 7, by their transitivity, and symmetry, the important property holds, see Polkowski [51].

Proposition 1. *In case of a symmetric and transitive rough inclusion μ, for each pair u, v of objects , and $r \in [0, 1]$, $ingr(v, g_\mu(u, r))$ if and only if $\mu(v, u, r)$ holds. In effect, the granule $g_\mu(u, r)$ can be represented as the set $\{v : \mu(v, u, r)\}$.*

Proof. Assume that $ingr(v, g_\mu(u, r))$ holds. Thus, there exists z such that $Ov(z, v)$ and $\mu(z, u, r)$. There is x with $ingr(x, v)$, $ingr(x, z)$, hence, by transitivity of μ, also $\mu(x, u, r)$ holds. By symmetry of μ, $ingr(v, x)$, hence, $\mu(v, x, r)$ holds also $\square$

In case of rough inclusions in information systems, induced by residual implications generated by continuous t–norms, see Ch. 6, sect. 2, L, P, or M, we have a positive case, for the minimum t–norm M

Proposition 2. *For the rough inclusion μ induced by the residual implication $\Rightarrow_M$, due to the minimum t–norm M, and $r < 1$, the relation $ingr(v, g_\mu(u, r))$ holds if and only if $\mu(v, u, r)$ holds.*

Proof. The rough inclusion μ has the form $\mu(v, u, r)$ if and only if $\frac{|IND(v,s)|}{|A|} \Rightarrow_M \frac{|IND(u,s)|}{|A|} \geq r$. Assume that $ingr(v, g_\mu(u, r))$ holds, so by the class definition, there exists z such that $Ov(v, z)$ and $\mu(z, u, r)$ hold. Thus, we have w with $ingr(w, v)$ and $\mu(w, u, r)$ by transitivity of μ and the fact that $ingr(w, z)$. By definition of μ, $ingr(w, v)$ means that $|IND(w, s)| \leq |IND(v, s)|$. As $\mu(w, u, r)$ with $r < 1$ means that $|IND(u, s)| \geq r$ because of $|IND(w, s)| \geq |IND(u, s)|$, the condition $|IND(w, s)| \leq |IND(v, s)|$ implies that $\mu(v, u, r)$ holds as well □

The case of the rough inclusion μ induced either by the product t–norm $P(x, y) = x \cdot y$, or by the Łukasiewicz t–norm L, is a bit more intricate. To obtain in this case some positive result, we exploit the averaged t–norm $\vartheta(\mu)$ defined for the rough inclusion μ, induced by a t–norm T, by means of the formula

$$\vartheta(\mu)(v, u, r) \Leftrightarrow \forall z.\exists a, b.\mu(z, v, a), \mu(z, u, b), a \Rightarrow_T b \geq r \tag{0.3}$$

Our proposition for the case of the t–norm P is

Proposition 3. *For $r < 1$, $ingr(v, g_{\vartheta(\mu)}(u, r))$ holds if and only if $\mu(v, u, a \cdot r)$, where $\mu(v, t, a)$ holds for t which obeys conditions $ingr(t, v)$ and $\vartheta(\mu)(t, u, r)$.*

Proof. $ingr(v, g_{\vartheta(\mu)}(u, r))$ implies that there is w such that $Ov(v, w)$ and $\vartheta(\mu)(w, u, r)$, so we can find t with properties, $ingr(t, w)$, $ingr(t, v)$, hence, by transitivity of $\vartheta(\mu)$ also $\vartheta(\mu)(t, u, r)$.

By definition of $\vartheta(\mu)$, there are a, b such that $\mu(v, t, a)$, $\mu(v, u, b)$, and $a \Rightarrow_P b \geq r$, i.e., $\frac{b}{a} \geq r$. Thus, $\mu(v, u, b)$ implies $\mu(v, u, a \cdot r)$ □

An analogous reasoning brings forth in case of the rough inclusion μ induced by residual implication due to the Łukasiewicz implication L, the result that

Proposition 4. *For $r < 1$, $ingr(v, g_{\vartheta(\mu)}(u, r))$ holds if and only if $\mu(v, u, r + a - 1)$ holds, where $\mu(v, t, a)$ holds for t such that $ingr(t, v)$ and $\vartheta(\mu)(t, u, r)$.*

The two last propositions can be recorded jointly in the form

Proposition 5. *For $r < 1$, and μ induced by residual implications either $\Rightarrow_P$ or $\Rightarrow_L$, $ingr(v, g_{\vartheta(\mu)}(u, r))$ holds if and only if $\mu(v, u, T(r, a))$ holds, where $\mu(v, t, a)$ holds for t such that $ingr(t, v)$ and $\vartheta(\mu)(t, u, r)$.*

Granules as collective concepts can be objects for rough mereological calculi.

0.4 Rough inclusions on granules

Due to the feature of mereology that it operates (due to the class operator) only on level of individuals, one can extend rough inclusions from objects to granules; the formula for extending a rough inclusion μ to a rough inclusion $\overline{\mu}$ on granules is a modification of mereological axiom M3 of Ch. 5

$$\overline{\mu}(g,h,r) \Leftrightarrow \forall z.ingr(z,g) \Rightarrow \exists w.ingr(w,h), \mu(z,w,r). \tag{0.4}$$

Proposition 6. *The predicate $\overline{\mu}(g,h,r)$ is a rough inclusion on granules.*

Proof. $\mu(g,h,1)$ means that for each object z with $ingr(z,g)$ there exists an object w with $ingr(w,h)$ such that $\mu(z,w,1)$, i.e., $ingr(z,w)$, which, by the inference rule implies that $ingr(g,h)$. This proves RINC1. For RINC2, assume that $\mu(g,h,1)$ and $\mu(k,g,r)$ so for each $ingr(x,k)$ there is $ingr(y,g)$ with $\mu(x,y,r)$. For y there is z such that $ingr(z,h)$ and $\mu(y,z,1)$, hence, $\mu(x,z,r)$ by property RINC2 of μ. Thus, $\mu(k,h,r)$. RINC2 follows and RINC3 is obviously satisfied □

We now examine rough mereological granules with respect to their properties.

0.5 General properties of rough mereological granules

They are collected below in

Proposition 7. *The following constitute a set of basic properties of rough mereological granules*

1. *If $ingr(y,x)$ then $ingr(y,g_\mu(x,r))$;*

2. *If $ingr(y,g_\mu(x,r))$ and $ingr(z,y)$ then $ingr(z,g_\mu(x,r))$;*

3. *If $\mu(y,x,r)$ then $ingr(y,g_\mu(x,r))$;*

4. *If $s < r$ then $ingr(g_\mu(x,r),g_\mu(x,s))$,*

which follow straightforwardly from properties RINC1–RINC3 of rough inclusions and the fact that *ingr* is a partial order, in particular it is transitive, regardless of the type of the rough inclusion μ.

For T–transitive rough inclusions, we can be more specific, and prove

Proposition 8. *For each T–transitive rough inclusion μ,* *1. If* $ingr(y, g_\mu(x, r)$ *then* $ingr(g_\mu(y, s), g_\mu(x, T(r, s)))$*;*

2. If $\mu(y, x, s)$ *with* $1 > s > r$*, then there exists* $\alpha < 1$ *with the property that* $ingr(g_\mu(y, \alpha), g_\mu(x, r)$*.*

Proof. Property 1 follows by transitivity of μ with the t–norm T. Property 2 results from the fact that the inequality $T(s, \alpha) \geq r$ has a solution in α, e.g., for $T = P$, $\alpha \geq \frac{r}{s}$, and, for $T = L$, $\alpha \geq 1 - s + r$ □

It is natural to regard granule system $\{g_r^{\mu_t}(x) : x \in U; r \in (0, 1)\}$ as a neighborhood system for a topology on U that may be called the *granular topology.*

In order to make this idea explicit, we define classes of the form

$$N^T(x, r) = Cls(\psi_{r,x}^{\mu_T}) \tag{0.5}$$

where

$$\psi_{r,x}^{\mu_T}(y) \Leftrightarrow \exists s > r.\mu_T(y, x, s) \tag{0.6}$$

We declare the system $\{N^T(x, r) : x \in U; r \in (0, 1)\}$ to be a neighborhood basis for a topology θ_μ. This is justified by the following

Proposition 9. *Properties of the system* $\{N^T(x, r) : x \in U; r \in (0, 1)\}$ *are as follows*

1. $y\ ingr\ N^T(x, r) \Rightarrow \exists\delta > 0.N^T(y, \delta)\ ingr\ N^T(x, r)$*;*

2. $s > r \Rightarrow N^T(x, s)\ ingr\ N^T(x, r)$*;*

3. $z\ ingr\ N^T(x, r) \wedge z\ ingr\ N^T(y, s) \Rightarrow \exists\delta > 0\ N^T(z, \delta)\ ingr\ N^T(x, r) \wedge N^T(z, \delta)\ ingr\ N^T(y, s)$*.*

Proof. For Property 1, $y\ ingr\ N^t(x, r)$ implies that there exists an $s > r$ such that $\mu_t(y, x, s)$. Let $\delta < 1$ be such that $t(u, s) > r$ whenever $u > \delta$; δ exists by continuity of t and the identity $t(1, s) = s$. Thus, if $z\ ingr\ N^t(y, \delta)$, then $\mu_t(z, y, \eta)$ with $\eta > \delta$ and $\mu_t(z, x, t(\eta, s))$ hence $z\ ingr\ N^t(x, r)$.

Property 2 follows by RINC3 and Property 3 is a corollary to properties 1 and 2. This concludes the argument □

Granule systems defined above form a basis for applications, where approximate reasoning is a crucial ingredient.

We begin with a basic application in which approximate reasoning itself is codified as a many–world (intensional) logic, where granules serve as possible worlds.

0.6 Reasoning by granular rough mereological logics

The idea of a granular rough mereological logic, see Polkowski [42], Polkowski and Semeniuk–Polkowska [53], consists in measuring the meaning of a unary predicate in the model which is a universe of an information system against a granule defined by means of a rough inclusion. The result can be regarded as the degree of truth (the logical value) of the predicate with respect to the given granule. The obtained logics are intensional as they can be regarded as mappings from the set of granules (possible worlds) to the set of logical values in the interval $[0, 1]$, the value at a given granule regarded as the extension at that granule of the generally defined intension.

The problem of meaning in intensional contexts, pointed to by usage of, e.g., linguistic contexts like 'he believes that ...', 'he knows that...', 'he thinks that....' for long has attracted the attention of philosophers and logicians, it suffices to mention Immanuel Kant and John Stuart Mill.

Problems that arise here are distinctly illustrated with the well–known example of two phrases: 'the morning star' and 'the evening star'. Both describe the planet Venus in two distinct appearances; once, as the star appearing in the sky at the sunrise, and secondly, as the star that appears in the sky at the sunset. One says that both phrases have the same *denotation*, i.e., the planet Venus, but distinct *meanings*. Terms 'denotation', 'meaning', call for a precise definition and a formal mechanism of inference with them.

The problem was addressed by Gottlob Frege [12] in '*Ü*ber Sinn und Bedeutung', terms usually rendered as *sense* (Sinn) and *reference* (Bedeutung). Thus, phrases 'the morning star' and 'the evening star' have different senses but the common reference – the planet Venus.

Among later approaches to the problem, we would like to point to that proposed by Rudolph Carnap [6]. Carnap introduces an important notion of a *state description* understood as a logical status of a state and determined by assignment of truth or falsity to each atomic object of logic. Carnap calls senses *intensions* and formalizes them as functions on state descriptions whereas references, or, denotations, are defined as values of intensions at particular state descriptions and are called *extensions*.

In this way, intensions are given a functional character, e.g., 'the morning star' and 'the evening star' as intensions describe some various astronomic objects at various solar systems, but at the Sun solar system, at the planet Earth, their extension is the same: the planet Venus.

Functional character of intensions was exploited ingeniously in Montague [34], [35] most notably culminating in His formal model of natural language grammar, the *Montague Grammar*, in which hierarchies of functional intensions model the structure of a sentence.

The idea of functionality of intensions has been exploited most famously in Kripke [19] who produced a semantics for modal logics, cf., Ch. 3, sect. 9: state descriptions have been replaced with *possible worlds* collected in the set of possible worlds W and intensions are valuations at possible worlds on logical variables, the extension at a possible world being the value of a given variable at that world. Imposing on possible worlds an *accessibility relation* R allows for formal definitions of modal notions of *necessity* and *possibility* in the *frame* (W, R).

For a given world w, and a formula ϕ, one declares ϕ *necessarily true* at w if and only if ϕ is true at each world w' accessible from w, i.e., such that $(w, w') \in R$; similarly, ϕ is *possibly true* at w if and only if there exists w' such that ϕ is true at w' and $(w, w') \in R$.

Gallin [13] proposed an axiomatics for intensional logic along with completeness proof; expositions of intensional and modal logics can be found in Van Benthem [3], Hughes and Creswell [18] and Fitting [11].

Our approach to rough mereological logics has an intentional tint, as we assume that possible worlds are granules obtained by means of a rough inclusion μ, and, against these worlds, unary predicates interpreted in the collection of objects, on which μ is defined, are evaluated as to their degree of truth. Thus, intensions are functions on granule collection, and extensions are values of truth at particular granules for particular predicates.

Any attempt at assigning various degrees of truth to logical statements places one in the realm of many–valued logic. These logics describe formally logical functors as mappings on the set of truth values/states into itself hence they operate a fortiori with values of statements typically as fractions or reals in the unit interval $[0, 1]$, see in this respect, e.g., Łukasiewicz [29], [30], Łukasiewicz and Tarski [32], and Hájek [17], see also Ch. 3.

In logics based on implication given by residua of t–norms, negation is defined usually as $\neg x = x \Rightarrow_T 0$. Thus, the Łukasiewicz negation is $\neg_L x = 1 - x$ whereas Goguen's as well as Gödel's negation is $\neg_G x = 1$ for x=0 and is 0 for $x > 0$, see Ch. 3. Other connectives are defined with usage of the t–norm itself as semantics for the strong conjunction and ordinary conjunction and disjunction are interpreted semantically as, respectively, *min, max.*

In this approach a rule $\alpha \Rightarrow \beta$ is evaluated by evaluating the truth state $[[\alpha]]$ as well as the truth state $[[\beta]]$ and then computing the values of $[[\alpha]] \Rightarrow_T [[\beta]]$ for a chosen t–norm T. Similarly other connectives are evaluated.

In the rough set context, this approach would pose the problem of evaluating the truth state of a conjunct α of descriptors; to this end, one can invoke the idea of Łukasiewicz [29] and assign to α a value $[[\alpha]]_L = \frac{|\{u \in U : u \models \alpha\}|}{|U|}$.

Clearly, this approach does not take into account the logical containment or its lack between α and β, and this fact makes the many–valued approach of a small use when data mining tasks are involved.

For this reason, we propose an approach to logic of decision rules which is based on the idea of measuring the state of truth of a formula against a concept constructed as a granule of knowledge; concepts can be regarded as "worlds" and our logic becomes intensional: logical evaluations at a given world are extensions of the intension which is the mapping on worlds valued in the set of logical values of truth.

For our purpose it is essential to extend rough inclusions to sets, cf., Ch. 6 sect. 5; we use the t–norm L along with the representation $L(r,s) = g(f(r)+f(s))$ already introduced in Ch. 4, sect. 9. We denote rough inclusions on sets with the generic symbol ν.

For finite sets X, Y, we let,

$$\nu_L(X,Y,r) \Leftrightarrow g(\frac{|X \setminus Y|}{|X|}) \geq r \tag{0.7}$$

As $g(x) = 1 - x$, we have that $\nu_L(X,Y,r)$ holds if and only if $\frac{|X \cap Y|}{|X|} \geq r$.

Let us observe that ν_L is *regular*, i.e., $\nu_L(X,Y,1)$ if and only if $X \subseteq Y$ and $\nu_L(X,Y,r)$ only with $r = 0$ if and only if $X \cap Y = \emptyset$.

Thus, the ingredient relation associated with a regular rough inclusion is the improper containment $\subseteq$ whereas the underlying part relation is the strict containment $\subset$.

Other rough inclusion on sets which we will use is the 3–valued rough inclusion ν_3 defined via the formula,

$$\nu_3(X,Y,r) \Leftrightarrow \begin{cases} X \subseteq Y \ and \ r = 1 \\ X \cap Y = \emptyset \ and \ r = 0 \\ r = \frac{1}{2} \ otherwise \end{cases} \tag{0.8}$$

We now proceed with a construction of the rough mereological logic.

0.7 A logic for information systems

We assume that an information/decision system (U, A, d) is given, along with a rough inclusion ν on the subsets of the universe U; for a collection of unary predicates Pr, interpreted in the universe U (meaning that for each predicate $\phi \in Pr$ the meaning $[[\phi]]$ is a subset of U), we define the intensional logic GRM_ν by assigning to each predicate ϕ in Pr its intension $I_\nu(\phi)$ defined by its extension $I_\nu^\vee(g)$ at each particular granule g, as

$$I_\nu^\vee(g)(\phi) \geq r \Leftrightarrow \nu(g, [[\phi]], r) \tag{0.9}$$

With respect to the rough inclusion ν_L, the formula (0.9) becomes

$$I_{\nu_L}^\vee(g)(\phi) \geq r \Leftrightarrow \frac{|g \cap [[\phi]]|}{|g|} \geq r \tag{0.10}$$

The counterpart for ν_3 is specified by definition (0.8), and it comes down to the following

$$I^{\vee}_{\nu_3}(g)(\phi) \geq r \ \Leftrightarrow \begin{cases} g \subseteq [[\phi]] \text{ and } r = 1 \\ g \cap [[\phi]] \neq \emptyset \text{ and } r \geq \frac{1}{2} \\ g \cap [[\phi]] = \emptyset \text{ and } r = 0 \end{cases} \tag{0.11}$$

We say that a formula ϕ interpreted in the universe U of an information system (U, A) is *true* at a granule g with respect to a rough inclusion ν if and only if $I^{\vee}_{\nu}(g)(\phi) = 1$.

Hence, for every regular rough inclusion ν, a formula ϕ interpreted in the universe U, with the meaning $[[\phi]] = \{u \in U : u \models \phi\}$, is true at a granule g with respect to ν if and only if $g \subseteq [[\phi]]$.

In particular, for a decision rule $r : p \Rightarrow q$ in the descriptor logic, see Ch. 4, the rule r is true at a granule g with respect to a regular rough inclusion ν if and only if $g \cap [[p]] \subseteq [[q]]$. We state these facts in the following

Proposition 10. *For every regular rough inclusion ν, a formula ϕ interpreted in the universe U, with the meaning $[[\phi]]$, is true at a granule g with respect to ν if and only if $g \subseteq [[\phi]]$. In particular, for a decision rule $r : p \Rightarrow q$ in the descriptor logic, the rule r is true at a granule g with respect to a regular rough inclusion ν if and only if $g \cap [[p]] \subseteq [[q]]$.*

Proof. Indeed, truth of ϕ at g means that $\nu(g, [[\phi]], 1)$ which in turn, by regularity of ν is equivalent to the inclusion $g \subseteq [[\phi]]$ $\square$

We will say that a formula ϕ is a *tautology* of our intensional logic if and only if ϕ is true at every world g.

The preceding proposition implies that,

Proposition 11. *For every regular rough inclusion ν, a formula ϕ is a tautology if and only if $Cls(G) \subseteq [\phi]$, where G is the property of being a granule; in the case when granules considered cover the universe U this condition simplifies to $[[\phi]] = U$. This means for a decision rule $p \Rightarrow q$ that it is a tautology if and only if $[[p]] \subseteq [[q]]$.*

Hence, the condition for truth of decision rules in the logic GRM_{ν} is the same as the truth of an implication in descriptor logic, see Ch. 4, under caveat that granules considered cover the universe U of objects.

Our rough mereological intensional logic depends obviously on the chosen rough inclusion μ on sets – concepts in the universe of objects. We are going to examine relationships of this logic, in case of two rough inclusions, viz., the Łukasiewicz t–norm induced ν_L and the 3–valued rough inclusion ν_3, to many–valued logics based on the Łukasiewicz residual implication, studied in Ch. 3. It turns out that in a sense, rough mereological intensional logics are embedded in, respectively, [0,1]–valued and 3–valued Łukasiewicz logics.

We apply to this end the idea of *collapse.* Collapse in this case consists in omitting the variable symbols while preserving connectives and in this way transforming open formulas of unary predicate calculus into propositional formulas.We show that theorems of rough mereological intensional logic are after collapsing them theorems of the respective Łukasiewicz many valued logic. We also apply a kind of inverse collapse by regarding in the coming sections formulas of many valued propositional calculus as formulas of unary predicate calculus with variables omitted.

0.7.1 Relations to many–valued logics

We examine some axiomatic schemes for many–valued logics with respect to their meanings under the stated in introductory section assumption that $[[p \Rightarrow q]] = (U \setminus [[p]]) \cup [[q]]$, $[[\neg p]] = U \setminus [[p]]$.

We examine first axiom schemes for 3–valued Łukasiewicz logic; we recall, see Ch. 3, axiom schemes given in Wajsberg [62].

W1 $q \Rightarrow (p \Rightarrow q)$

W2 $(p \Rightarrow q) \Rightarrow ((q \Rightarrow r) \Rightarrow (p \Rightarrow r))$

W3 $((p \Rightarrow \neg p) \Rightarrow p) \Rightarrow p$

W4 $(\neg q \Rightarrow \neg p) \Rightarrow (p \Rightarrow q)$

We have as meanings of those formulas

$$[[W1]] = (U \setminus [[q]]) \cup (U \setminus [[p]]) \cup [[q]] = U \tag{0.12}$$

$$[[W2]] = ([[p]] \setminus [[q]]) \cup ([[q]] \setminus [[r]]) \cup (U \setminus [[p]]) \cup [[r]] = U \tag{0.13}$$

$$[[W3]] = (U \setminus [[p]]) \cup [[p]] = U \tag{0.14}$$

$$[[W4]] = ([[p]] \setminus [[q]]) \cup [[q]] = U \tag{0.15}$$

It follows that

Proposition 12. *All instances of Wajsberg axiom schemes for 3–valued Łukasiewicz logic are tautologies of our intensional logic in case of regular rough inclusions on sets.*

The deduction rule in 3–valued Łukasiewicz logic is Modus Ponens: $\frac{p, p \Rightarrow q}{q}$.

In our setting this is a valid deduction rule: if $p, p \Rightarrow q$ are tautologies than q is a tautology. Indeed, if $[[p]] = U = [[p \Rightarrow q]]$, then $[[q]] = U$.

We have obtained

Proposition 13. *Each tautology of 3–valued Łukasiewicz logic is a tautology of rough mereological granular logic in case of a regular rough inclusion on sets.*

In an analogous manner, we examine axiom schemes for infinite valued Łukasiewicz logic, proposed by Łukasiewicz, see Łukasiewicz and Tarski [32] in a form modified due to Meredith [33] and Chang [7], see Ch. 3, sect. 4

L1 $q \Rightarrow (p \Rightarrow q)$

L2 $(p \Rightarrow q) \Rightarrow ((q \Rightarrow r) \Rightarrow (p \Rightarrow r))$

L3 $((q \Rightarrow p) \Rightarrow p) \Rightarrow ((p \Rightarrow q) \Rightarrow q)$

L4 $(\neg q \Rightarrow \neg p) \Rightarrow (p \Rightarrow q)$

As L1 is W1, L2 is W2 and L4 is W4, it remains to examine L3.

In this case, we have

$$[[(q \Rightarrow p) \Rightarrow p]] = (U \setminus [[q \Rightarrow p]]) \cup [[p]]) =$$

$$(U \setminus ((U \setminus [[q]]) \cup [[p]])) \cup [[p]] = ([[q]] \setminus [[p]]) \cup [[p]] =$$

$$[[q]] \cup [[p]]$$

Similarly,

$$[[(p \Rightarrow q) \Rightarrow q]]$$

is

$$[[p]] \cup [[q]]$$

by symmetry, and finally, the meaning $[[L3]]$ is

$$(U \setminus ([[q]] \cup [[p]])) \cup [[p]] \cup [[q]] = U$$

It follows that

Proposition 14. *All instances of axiom schemes for infinite–valued Łukasiewicz logic are tautologies of rough mereological granular logic.*

As Modus Ponens remains a valid deduction rule in infinite–valued case, we obtain, analogous to Prop. 13,

Proposition 15. *Each tautology of infinite–valued Łukasiewicz logic is a tautology of rough mereological granular logic in case of a regular rough inclusion on sets.*

It follows from Prop.15 that all tautologies of *Basic logic*, see Hájek [17], cf., Ch. 3, sect. 6, i.e., logic which is intersection of all many–valued logics with implications evaluated semantically by residual implications of continuous t–norms are tautologies of rough mereological granular logic for each regular rough inclusion ν.

The assumption of regularity of a rough inclusion ν is essential: considering the drastic rough inclusion ν_1, we find that an implication $p \Rightarrow q$ is true only at the world $(U \setminus [[p]]) \cup [[q]]$, so it is not any tautology; this concerns all schemes systems W and L above as they are true only at the global world U.

0.8 A graded notion of truth

As already stated, the usual interpretation of functors $\vee, \wedge$ in many–valued logics is $[[p \vee q]] = max\{[[p]], [[q]]\}$ and $[[p \wedge q]] = min\{[[p]], [[q]]\}$, where $[[p]]$ is the state of truth of p. In case of concept–valued meanings, we admit the interpretation which conforms to accepted in many valued logics (especially in the context of fuzzy set theory) interpretation of min as $\cap$ and max as $\cup$.

The formula $\nu(g, [\phi], 1)$ stating the truth of ϕ at g, ν with ν regular can be regarded as a condition of orthogonality type, with the usual consequences.

1. *If ϕ is true at granules g, h, then it is true at $g \cup h$;*

2. *If ϕ is true at granules g, h then it is true at $g \cap h$;*

3. *If ϕ, ψ are true at a granule g, then $\phi \vee \psi$ is true at g;*

4. *If ϕ, ψ are true at a granule g, then $\phi \wedge \psi$ is true at g;*

5. *If ψ is true at a granule g, then $\phi \Rightarrow \psi$ is true at g for every formula ϕ;*

6. *If ϕ is true at a granule g then $\phi \Rightarrow \psi$ is true at g if and only if ψ is true at g.*

The graded relaxation of truth is given obviously by the condition that a formula ϕ is *true to a degree at least r at g, ν* if and only if

$$I_\nu^\vee(g)(\phi) \geq r \tag{0.16}$$

i.e., $\nu(g, [[\phi]], r)$ holds.

In particular, ϕ is *false* at g, ν if and only if $I_\nu^\vee(g)(\phi) \geq r$ implies $r = 0$, i.e., $\nu(g, [\phi], r)$ implies $r = 0$.

Proposition 16. *The following properties hold with the Lukasiewicz residual implication $\Rightarrow_L$*

1. For each regular ν*,* a formula α is true at g, ν if and only if $\neg\alpha$ is false at g, ν*;*

2. For $\nu = \nu_L, \nu_3$, $I^{\vee}_{\nu}(g)(\neg\alpha) \geq r$ if and only if $I^{\vee}_{\nu}(g)(\alpha) \geq s$ implies $s \leq 1-r$;

3. For $\nu = \nu_L, \nu_3$, the implication

$$\alpha \Rightarrow_L \beta$$

is true at g if and only if

$$g \cap [\alpha] \subseteq [\beta]$$

and $\alpha \Rightarrow_L \beta$ is false at g if and only if

$$g \subseteq [\alpha] \setminus [\beta];$$

4. For $\nu = \nu_L$, if

$$I^{\vee}_{\nu}(g)(\alpha \Rightarrow_L \beta) \geq r$$

then $\Rightarrow_L (t, s) \geq r$, *where* $I^{\vee}_{\nu}(g)(\alpha) \geq t$ *and* $I^{\vee}_{\nu}(g)(\beta) \geq s$.

Further analysis should be split into the case of ν_L and the case of ν_3 as the two differ essentially with respect to the form of reasoning they imply.

Property 4 in Proposition 16 shows in principle that the value of $I^{\vee}_{\nu}(g)(\alpha \Rightarrow \beta)$ is bounded from above by the value of $I^{\vee}_{\nu}(g)(\alpha) \Rightarrow_{t_L} I^{\vee}_{\nu}(g)(\beta))$.

This suggests that the idea of collapse attributed to Leśniewski can be applied to formulas of rough mereological logic in the following form: for a formula $q(x)$ we denote by the symbol q^* the formula q regarded as a sentential formula (i.e., with variable symbols removed) subject to relations

1. $(\neg q(x))^*$ *is* $\neg(q(x)^*)$;

2. $(p(x) \Rightarrow q(x))^*$ *is* $p(x)^* \Rightarrow q(x)^*$.

As the value $[[q^*]]_g$ of the formula $q(x)^*$ at a granule g, we take the value of $\frac{|g \cap [[q(x)]]|}{|g|}$, i.e, $argmax_r\{\nu_L(g, [[q^*]]_g, r)\}$. Thus, Property 4 in Proposition 16 can be rewritten in the form

$$I^{\vee}_{\nu}(g)(\alpha \Rightarrow \beta) \leq [\alpha^*]_g \Rightarrow_L [\beta^*]_g \tag{0.17}$$

The statement follows

Proposition 17. *If $\alpha \Rightarrow \beta$ is true at g, then the collapsed formula has the value 1 of truth at the granule g in the Łukasiewicz logic.*

This gives a necessity condition for verification of implications of rough mereological logics

Proposition 18. *If* $[\alpha^*]_g \Rightarrow_L [\beta^*]_g < 1$*, then the implication* $\alpha \Rightarrow \beta$ *is not true at* g.

This concerns in particular decision rules

Proposition 19. *A decision rule* $p(v) \Rightarrow q(v)$*, is true on a granule* g *if and only if* $[[p^*]]_g \leq [[q^*]]_g$.

In case of ν_3, one can check on the basis of definitions that $I_\nu^\vee(g)(\neg\alpha) \geq r$ if and only if $I_\nu^\vee(g)(\alpha) \leq 1 - r$; thus the negation functor in rough mereological logic based on ν_3 is the same as the negation functor in the 3–valued Łukasiewicz logic. For implication, the relations between granular rough mereological logic and 3–valued logic of Łukasiewicz follow from truth tables for respective functors of negation and implication.

Table 0.1 shows truth values for implication in 3–valued logic of Łukasiewicz. We recall that these values obey the implication $x \Rightarrow_L y = min\{1, 1-x+y\}$. Values of x correspond to rows and values of y correspond to columns in the table of Fig. 1.

Fig. 0.1 Truth values for implication in L_3

$\Rightarrow$	0	1	$\frac{1}{2}$
0	1	1	1
1	0	1	$\frac{1}{2}$
$\frac{1}{2}$	$\frac{1}{2}$	1	1

The table of Fig. 0.2 shows values of implication for rough mereological logic based on ν_3. Values are shown for the extension $I_\nu^\vee(g)(p \Rightarrow q)$ of the implication $p \Rightarrow q$. Rows correspond to p, columns correspond to q.

Fig. 0.2 Truth values for implication $p \Rightarrow q$ in logic based on ν_3

$\Rightarrow$	$I_{\nu_3}^\vee(g)(q) = 0$	$I_{\nu_3}^\vee(g)(q) = 1$	$I_{\nu_3}^\vee(g)(q) = \frac{1}{2}$
$I_{\nu_3}^\vee(g)(p) = 0$	1	1	1
$I_{\nu_3}^\vee(g)(p) = 1$	0	1	$\frac{1}{2}$
$I_{\nu_3}^\vee(g)(p) = \frac{1}{2}$	$\frac{1}{2}$	1	1 when $g \cap [\alpha] \subseteq [\beta]$; $\frac{1}{2}$ otherwise

We verify values shown in Fig. 0.2.

First, we consider the case when $I_{\nu_3}^\vee(g)(p) = 0$, i.e., the case when $g \cap [[p]] = \emptyset$. As

$$g \subseteq (U \setminus [[p]]) \cup [[q]]$$

for every value of $[[q]]$, we have only values of 1 in the first row of Table 2.

Assume now that $I^{\vee}_{\nu_3}(g)(p) = 1$, i.e., $g \subseteq [[p]]$. As

$$g \cap (U \setminus [[p]]) = \emptyset$$

the value of $I^{\vee}_{\nu}(g)(p \Rightarrow q)$ depends only on a relation between g and $[[q]]$. In case

$$g \cap [[q]] = \emptyset$$

the value in Table 2 is 0, in case $g \subseteq [[q]]$ the value in Table 2 is 1, and in case

$$I^{\vee}_{\nu_3}(g)(q) = \frac{1}{2}$$

the value in Table 2 is $\frac{1}{2}$.

Finally, we consider the case when

$$I^{\vee}_{\nu_3}(g)(p) = \frac{1}{2}$$

i.e.,

$$g \cap [[p]] \neq \emptyset \neq g \setminus [[p]]$$

In case $g \cap [[q]] = \emptyset$, we have

$$g \cap ((U \setminus [[p]]) \cup [[q]]) \neq \emptyset$$

and it is not true that

$$g \subseteq ((U \setminus [[p]]) \cup [[q]])$$

so the value in table is $\frac{1}{2}$.

In case $g \subseteq [[q]]$, the value in Table is clearly 1. The case when

$$I^{\vee}_{\nu_3}(g)(q) = \frac{1}{2}$$

remains. Clearly, when $g \cap [[p]] \subseteq [[q]]$, we have

$$g \subseteq (U \setminus [[p]]) \cup [[q]]$$

so the value in Table is 1; otherwise, the value is $\frac{1}{2}$.

Thus, negation in both logic is semantically treated in the same way, whereas treatment of implication differs only in case of implication $p \Rightarrow q$ from the value $\frac{1}{2}$ to $\frac{1}{2}$, when $g \cap [[p]]$ is not any subset of $[[q]]$.

It follows from these facts that given a formula α and its collapse α^*, we have,

$$I^{\vee}_{\nu_3}(g)(\neg\alpha) = [[(\neg\alpha)^*]]_{L_3}, I^{\vee}_{\nu_3}(g)(\alpha \Rightarrow \beta) \leq [[(\alpha \Rightarrow \beta)^*]]_{L_3} \qquad (0.18)$$

A more exact description of implication in both logics is as follows.

Proposition 20. *The following statements are true*

1. If $I^{\vee}_{nu_3}(g)(\alpha \Rightarrow \beta) = 1$ *then* $[(\alpha \Rightarrow \beta)^*]_{L_3} = 1$*;*

2. If $I^{\vee}_{nu_3}(g)(\alpha \Rightarrow \beta) = 0$ *then* $[(\alpha \Rightarrow \beta)^*]_{L_3} = 0$*;*

3. If $I^{\vee}_{nu_3}(g)(\alpha \Rightarrow \beta) = \frac{1}{2}$ *then* $[(\alpha \Rightarrow \beta)^*]_{L_3} \geq \frac{1}{2}$ *and this last value may be 1.*

Proof. We offer a simple check–up on Proposition 20. In Case 1, we have

$$g \subseteq ((U \setminus [[\alpha]]) \cup [[\beta]])$$

For the value of $[(\alpha \Rightarrow \beta)^*]$, consider some subcases.

Subcase 1.1 $g \subseteq U \setminus [[\alpha]]$. Then $[[\alpha^*]] = 0$ and

$$[[(\alpha \Rightarrow \beta)^*]] = [[\alpha^*]] \Rightarrow [[\beta^*]]$$

is always 1 regardless of a value of $[\beta^*]$.

Subcase 1.2 $g \cap [[\alpha]] \neq \emptyset \neq g \setminus [[\alpha]]$ so $[\alpha^*] = \frac{1}{2}$. Then $g \cap [\beta] = \emptyset$ is impossible, i.e., $[\beta^*]$is at least $\frac{1}{2}$ and $[(\alpha \Rightarrow \beta)^*]$=1.

Subcase 1.3 $g \subseteq [[\alpha]]$ so $[[\alpha^*]] = 1$; then $g \subseteq [[\beta]]$ must hold, i.e., $[[\beta^*]] = 1$ which means that $[[(\alpha \Rightarrow \beta)^*]]$=1.

For case 2, we have

$$g \cap ((U \setminus [[\alpha]]) \cup [[\beta]]) = \emptyset$$

hence, $g \cap [[\beta]] = \emptyset$ and $g \subseteq [[\alpha]]$, i.e., $[\alpha^*] = 1, [\beta^*] = 0$ so $[\alpha^*] \Rightarrow [\beta^*]$=0.

In case 3, we have

$$g \cap ((U \setminus [[\alpha]]) \cup [[\beta]]) \neq \emptyset$$

and

$$g \cap [[\alpha]] \setminus [[\beta]] \neq \emptyset$$

Can $[\alpha^*] \Rightarrow [\beta^*]$ be necessarily 0? This would mean that $[\alpha^*] = 1$ and $[\beta^*] = 0$, i.e., $g \subseteq [[\alpha]]$ and $g \cap [[\beta]] = \emptyset$ but then

$$g \cap ((U \setminus [[\alpha]]) \cup [[\beta]]) = \emptyset$$

a contradiction. Thus the value $[[\alpha^*]] \Rightarrow [[\beta^*]]$ is at least $\frac{1}{2}$. In the subcase: $g \subseteq [[\alpha]]$, $g \cap [[\beta]] \neq \emptyset \neq g \setminus [[\beta]]$, the value of $[[\alpha^*]] \Rightarrow [[\beta^*]]$ is $0 \Rightarrow_L \frac{1}{2} = 1$, and the subcase is consistent with case 3 □

Let us mention that a 3–valued intensional variant of Montague Grammar is considered in Alves and Guerzoni[1].

0.9 Dependencies and decision rules

It is an important feature of rough set theory that it allows for an elegant formulation of the problem of dependency between two sets of attributes, see Ch. 4, sect. 4, cf., Pawlak [39], [40], in terms of indiscernibility relations.

We recall, see Ch. 4, sect. 4, that for two sets $C, D \subseteq A$ of attributes, one says that D *depends functionally on* C when $IND(C) \subseteq IND(D)$, symbolically denoted $C \mapsto D$. Functional dependence can be represented locally by means of functional dependency rules of the form

$$\phi_C(\{v_a : a \in C\}) \Rightarrow \phi_D(\{w_a : a \in D\}) \tag{0.19}$$

where $\phi_C(\{v_a : a \in C\})$ is the formula $\bigwedge_{a \in C}(a = v_a)$, and $[\phi_C] \subseteq [\phi_D]$.

A proposition holds

Proposition 21. *If $\alpha : \phi_C \Rightarrow \phi_D$ is a functional dependency rule, then α is a tautology of logic induced by ν_3.*

Proof. For each granule g, we have $g \cap [[\phi_C]] \subseteq [[\phi_D]]$ □

Let us observe that the converse statement is also true, i.e., if a formula $\alpha : \phi_C \Rightarrow \phi_D$ is a tautology of logic induced by ν_3, then this formula is a functional dependency rule in the sense of (0.19). Indeed, assume that α is not any functional dependency rule, i.e., $[[\phi_C]] \setminus [[\phi_D]] \neq \emptyset$. Taking $[[\phi_C]]$ as the witness granule g, we have that g is not any subset of $[[\alpha]]$, i.e, $I^{\vee}_{\nu_3}(g)(\alpha) \leq \frac{1}{2}$, so α is not true at g, a fortiori it is not any tautology.

Let us observe that these characterizations are valid for each regular rough inclusion on sets ν.

A more general and also important notion is that of a local proper dependency: a formula $\phi_C \Rightarrow \phi_D$ where $\phi_C(\{v_a : a \in C\})$ is the formula $\bigwedge_{a \in C}(a = v_a)$, similarly for ϕ_D, is a local proper dependency when $[\phi_C] \cap [\phi_D] \neq \emptyset$.

We will say that a formula α is *acceptable with respect to a collection M of worlds* when

$$I^{\vee}_{\nu_3}(g)(\alpha) \geq \frac{1}{2} \tag{0.20}$$

for each world $g \in M$, i.e, when α is false at no world $g \in M$. Then,

Proposition 22. *If a formula $\alpha : \phi_C \Rightarrow \phi_D$ is a local proper dependency rule, then it is acceptable with respect to all C–exact worlds.*

Proof. Indeed, for a C–exact granule g, the case that $I^{\vee}_{\nu_3}(g)(\alpha) = 0$ means that $g \subseteq [[\phi_C]]$ and $g \cap [[\phi_D]] = \emptyset$.

As g is C–exact and $[[\phi_C]]$ is a C–indiscernibility class, either $[[\phi_C]] \subseteq g$ or $[[\phi_C]] \cap g = \emptyset$. When $[[\phi_C]] \subseteq g$, then $[[\phi_C]] = g$ which makes $g \cap [[\phi_D]] = \emptyset$ impossible.

When $[[\phi_C]] \cap g = \emptyset$, then $g \cap [[\phi_D]] = \emptyset$ is impossible. In either case, $I^{\vee}_{\nu_3}(g)(\alpha) = 0$ cannot be satisfied with any C–exact granule g □

Again, the converse is true: when α is not local proper, i.e., $[[\phi_C]] \cap [[\phi_D]] = \emptyset$, then $g = [[\phi_C]]$ does satisfy $I^{\vee}_{\nu_3}(g)(\alpha) = 0$.

A corollary of the same forms follows for *decision rules* in a given decision system (U, A, d), i.e., dependencies of the form $\phi_C \Rightarrow (d = w)$.

0.10 An application to Calculus of Perceptions

Calculus of perceptions, posed as an idea by Zadeh [67], should render formally vague statements and queries, and answer vague questions by giving a semantic interpretation to vague statements . Here, we apply granular mereological logics toward this problem. To this end, we would like to borrow a part of a complex percept in Zadeh [67] and interpret it in terms of granular logic.

The percept is: (i) *Carol has two children: Robert who is in mid–twenties and Helen who is in mid–thirties* with a query (ii) *how old is Carol.* To interpret it, we begin with Table 0.3 in which a decision system *Age* is given with attributes n – the number of children, a_i – the age of the i–th child for $i \leq 3$, and with the decision *age* – the age of the mother.

We define for a fuzzy concept X represented by the fuzzy membership function μ_X on the domain D_X, the *c–cut* where $c \in [0, 1]$ as the concept $X_C = \{x \in D_X : \mu_X(x) \geq c\}$. Concepts *in mid–twenties*, *in mid–thirties* are represented by fuzzy membership functions, μ_{20}, μ_{30}, respectively

$$\mu_{20}(x) = \begin{cases} 0.25(x - 20),\ x \in [20, 24] \\ 1,\ x \in [24, 26] \\ 1 - 0.25(x - 26),\ x \in [26, 30] \end{cases} \tag{0.21}$$

and

Fig. 0.3 Decision system *Age*

object	*n*	age_1	age_2	age_3	*Age*
1	3	15	22	30	58
2	3	10	12	16	42
3	2	6	10	−−	30
4	2	24	33	−−	56
5	2	28	35	−−	62
6	3	22	33	40	67
7	2	18	25	−−	60
8	2	26	35	−−	63
9	2	22	38	−−	70
10	3	8	12	16	38
11	2	22	32	−−	58
12	3	24	36	40	63
13	2	28	34	−−	60
14	3	26	30	35	65
15	3	18	25	35	60
16	3	6	12	16	40
17	3	22	30	35	65
18	2	24	34	−−	60
19	3	22	30	34	58
20	2	24	35	−−	62

$$\mu_{30}(x) = \begin{cases} 0.25(x-30),\ x \in [30, 24] \\ 1,\ x \in [34, 36] \\ 1 - 0.25(x-36),\ x \in [36, 40] \end{cases} \tag{0.22}$$

The concept *old* is interpreted as

$$\mu_{old}(x) = \begin{cases} 0.02(x-30),\ x \in [30, 60] \\ 0.04(x-60) + 0.6,\ x \in [60, 70] \\ 0,\ else \end{cases} \tag{0.23}$$

We interpret our query by a function $q : [0, 1]^3 \to [0, 1]$, where $f(u, v, w) = t$ would mean that for cut levels u, v, t, respectively for *old*, *in mid–twenties*, *in mid thirties*, the truth value of the statement *Carol is at least* $supp_u$ *old to the degree of t with respect to v, w.*

In our example, letting $u = 0.6$, we obtain $old_{0.6} = [60, 70]$; letting $v = 0.5 = w$, we obtain *in mid twenties*$_{0.5} = [23, 27]$ and *in mid thirties*$_{0.5} = [33, 37]$. In order to evaluate the truth degree t, we refer to the world knowledge of Table 0.3 and we find the set of objects $\Lambda_{v,w}$ with two children of ages respectively in the intervals $[23, 27], [33, 37]$ corresponding to values of v, w as well as the set of objects Γ_u having the value of *Age* in old_u. In our case we have $\Lambda_{0.5,0.5} = \{4, 8, 12, 18, 20\}$ and $\Gamma_{0.6} = \{5, 6, 7, 8, 9, 12, 13, 14, 15, 17, 18, 20\}$.

Finally, we evaluate the truth degree of the predicate *in* $old_{0.6}(x)$ represented in the universe of Table 0.3 by the set $\Gamma_{0.6}$ with respect to the granule $\Lambda_{0.5,0.5}$. We obtain by applying the Łukasiewicz rough inclusion μ_L,

$$(I^{\mu_L}_{\Lambda_{0.5,0.5}})^{\vee}(old_{0.6}(x)) = \frac{|\Lambda_{0.5,0.5} \cap \Gamma_{0.6}|}{|\Lambda_{0.5,0.5}|} = 0.8 \quad (0.24)$$

The result is the statement: *Carol is over 60 years old to degree of 0.8 under the assumed interpretation of in mid twenties*$_{0.5}$, *in mid thirties*$_{0.5}$ *with respect to knowledge in* Table 0.3.

0.11 Modal aspects of rough mereological logics

Modal logics are concerned with formal rendering of *modalities* of *necessity* and *possibility.* Possible world semantics (Kripke semantics) is set in terms of a set W of *possible worlds* endowed with an accessibility relation R: in case $(w, w') \in R$, one says that the world w' is *accessible* from the world w. In case of propositional calculus, intensions are valuations on propositional variables: for each world w and a variable p, a valuation *val* assigns the truth value $val(p, g)$ (True or False). This assignment extends by standard logical calculations to formulas, giving for each formula ϕ and a world w, the value $val(\phi, w)$, see Ch. 3, sect. 9.

Necessity is introduced by the requirement that a formula ϕ is *necessarily true*at a world w if and only if ϕ is true at every world w' accessible from w via R. Possibility is defined as the requirement that ϕ is *possibly true* at w if and only if there is w' accessible from w via R with ϕ true at w'. A formula ϕ is *true at the frame* (W, R) if and only if it is true at each world $w \in W$, and ϕ is *true* if and only if it is true at each frame.

A hierarchy of modal logics is built by demanding that they satisfy some postulates about the nature of necessity and possibility, see Hughes and Creswell [18], or, Lemmon and Scott [20].

Necessity is denoted with L and $L\phi$ reads *necessarily* ϕ, possibility is denoted with M and $M\phi$ reads *possibly* ϕ.

A basic postulate going back to Aristotle, cf., Łukasiewicz [31] is that necessity of an implication along with the necessity of the premise implies necessity of the conclusion, expressed by means of a formula (K),

$$(K)\ L(\phi \Rightarrow \psi) \Rightarrow (L\phi \Rightarrow L\psi) \quad (0.25)$$

It is well–known, see Ch. 3, sect. 9, that (K) is true at every frame regardless of the relation R.

The system in which (K) is satisfied is denoted as the system K.

The next postulate concerns the nature of necessity L and is expressed with a formula (T)

$$(T)\ L\phi \Rightarrow \phi \tag{0.26}$$

The formula (T) is satisfied in every frame in which the relation R is reflexive; the system K with (T) added is denoted T.

As a next step usually the formula ($S4$) is added to the system T

$$(S4)\ L\phi \Rightarrow LL\phi \tag{0.27}$$

(S4) is true in every frame in which the relation R is transitive; the system $S4$ in which $K, T, S4$ hold is satisfied in every frame with the relation R reflexive and transitive.

The postulate ($S5$) establishes a relation between L and M

$$(S5)\ M\phi \Rightarrow LM\phi \tag{0.28}$$

The system $S5$ resulting from adding ($S5$) to the system $S4$ is valid in all frames in which the relation R is an equivalence.

0.11.1 A modal logic with ingredient accessibility

In an attempt at expressing modalities afforded by the rough mereological context, we reach to the relation of an ingredient as an accessibility relation. Thus, we adopt the collection of granules induced by a rough inclusion μ as the set W of possible worlds, and the relation $R(g,h) \Leftrightarrow ingr(h,g)$, i.e., a world (granule) h is accessible via R from a world (granule) g if and only if h is an ingredient of g, where the relation of an ingredient on granules is induced from the relation of ingredient on objects by means of (0.4).

For a predicate α, and a granule g, it follows that α is necessarily true at g if and only if α is true at every ingredient h of g, which means that, for a regular rough inclusion ν on sets, that $[[h]] \subseteq [[\alpha]]$ for each ingredient h of g. Possibility of α at g means that there exists a granule h with $ingr(h,g)$ such that α is true at h, i.e., $[[h]] \subseteq [[g]]$.

As ingredient relation is reflexive and transitive, we have

Proposition 23. *The modal logic obtained by taking the ingredient relation on granules as accessibility relation r and granules of knowledge as possible worlds W satisfies requirements for the system $S4$ in the frame (W,R).*

0.11.2 Modal logic of rough set approximations

Possibility and necessity are introduced in rough set theory by means of approximations: the upper and the lower, respectively. A logical rendering of these modalities in rough mereological logics exploits the approximations. We define two modal operators: M (possibility) and L (necessity).

To this end, we let

$$\begin{cases} I_{\nu}^{\vee}(g)(M\alpha) \geq r \;\Leftrightarrow \nu_L(g, \overline{[\alpha]}, r) \\ I_{\nu}^{\vee}(g)(L\alpha) \geq r \;\Leftrightarrow \nu_L(g, \underline{[\alpha]}, r) \end{cases} \tag{0.29}$$

Then we have the following criteria for necessarily or possibly true formulas.

A formula α is *necessarily true at a granule* g if and only if $g \subseteq \underline{[\alpha]}$; α is *possibly true at* g if and only if $g \subseteq \overline{[[\alpha]]}$.
This semantics of modal operators M, L can be applied to show that rough set structures carry the semantics of S5 modal logics, i.e., the following relations hold at each granule g.

1. $L(\alpha \Rightarrow \beta) \Rightarrow [(L\alpha) \Rightarrow L(\beta)]$;

2. $L\alpha \Rightarrow \alpha$;

3. $L\alpha \Rightarrow LL\alpha$;

4. $M\alpha \Rightarrow LM\alpha$.

We need to show that the meaning of each of formulas 1–4 is U.

Concerning formula 1 (modal formula (K)), we have $[[L(\alpha \Rightarrow \beta) \Rightarrow (L\alpha) \Rightarrow L(\beta)]]{=}(U \setminus \underline{(U \setminus [[\alpha]]) \cup [[\beta]]}) \cup (U \setminus [[\alpha]]) \cup [[\beta]]$.

Assume that $u \notin (U \setminus [[\alpha]]) \cup [[\beta]]$; thus, (i) $[u]_A \subseteq [[\alpha]]$ and (ii) $[u]_A \cap [[\beta]] = \emptyset$. If it were $u \in \underline{(U \setminus [[\alpha]]) \cup [[\beta]]}$ then we would have $[u]_A \subseteq (U \setminus [[\alpha]]) \cup [[\beta]]$, a contradiction with (i), (ii). Thus, the meaning of (K) is U.

For formula 2, modal formula (T), we have $[[L\alpha \Rightarrow \alpha]]{=}((U \setminus \underline{[[\alpha]]}) \cup [[\alpha]])$; as $\underline{[[\alpha]]} \subseteq [[\alpha]]$, it follows that the meaning of (T) is U.

In case of formula 3, modal formula (S4), the meaning is $(U \setminus \underline{[[\alpha]]}) \cup \underline{\underline{[[\alpha]]}}{=}(U \setminus \underline{[[\alpha]]}) \cup \underline{[[\alpha]]}{=}\ U$.

The meaning of formula 4, modal formula (S5), is $(U \setminus \overline{[[\alpha]]}) \cup \underline{\overline{[[\alpha]]}}{=}(U \setminus \overline{[[\alpha]]}) \cup \overline{[[\alpha]]}{=}U$.

It follows that the logic S5 is satisfied within logic induced by ν_L and more generally in logic induced by any regular rough inclusion ν.

0.12 Reasoning in multi–agent and distributed systems

Approximate reasoning is often concerned with 'complex cases' like, e.g., robotic soccer, in which performing successfully tasks requires participation of a number of 'agents' bound to cooperate, and in which a task is performed with a number of steps, see, e.g., Stone [60]; other areas where such approach seems necessary concern assembling and design, see Amarel [2], fusion of knowledge, e.g., in robotics, fusion of information from sensors, see, e.g., Canny [5], Choset et al. [8], or, Stone [60], as well as in machine learning and fusion of classifiers, see, e.g., Dietterich [10].

Rough mereological approach to these problems was initiated with Polkowski and Skowron [54] – [57] and here we give these topics a logical touch.

Rough inclusions and granular intensional logics based on them can be applied in describing workings of a collection of intelligent agents which are called here *granular agents*.

A granular agent a will be represented as a tuple

$$(U_a, \mu_a, L_a, prop_a, synt_a, aggr_a)$$

where

1. *U_a is a collection of objects available to the agent a*;
2. *μ_a is a rough inclusion on objects in U_a*;
3. *L_a is a set of unary predicates in first–order open calculus, interpretable in U_a*;
4. *$prop_a$ is the propagation function that describes how uncertainty expressed by rough inclusions at agents connected to a propagates to a*;
5. *$synt_a$ is the logic propagation functor which expresses how formulas of logics at agents connected to the agent a are made into a formula at a*;
6. *$aggr_a$ is the synthesis function which describes how objects at agents connected to a are made into an object at a.*

We assume for simplicity that agents are arranged into a rooted tree; for each agent a distinct from any leaf agent, we denote by B_a the children of a, understood as agents connected to a and directly sending to a objects, logical formulas describing them, and uncertainty coefficients like values of rough inclusions.

For $b \in B_a$, the symbol x_b will denote an object in U_b; similarly, ϕ_b will denote a formula of L_b, and μ_b will be a rough inclusion at b with values r_b. The same convention will be obeyed by objects at a.
Our scheme should obey some natural postulates stemming from an assumption of regularity of reasoning.

MA1 *If* $ingr_b(x'_b, x_b)$ *for each* $b \in B_a$, *then* $ingr_a(aggr(\{x'_b\}), aggr(\{x_b\})$

This postulate does assure that ingredient relations are in agreement with aggregate operator of forming complex objects: ingredients of composed objects form an ingredient of a complex object. We can say that $aggr \circ ingr = ingr \circ aggr$, i.e, the resulting diagram commutes.

MA2 *If* $x_b \models \phi_b$, *then* $aggr(\{x_b\}) \models synt(\{\phi_b\})$

This postulate is about agreement between aggregation of objects and their logical descriptions: descriptions of composed objects merge into a description of the resulting complex object.

MA3 *If* $\mu_b(x_b, y_b, r_b)$ *for* $b \in B_a$, *then* $\mu_a(aggr(\{x_b\}), aggr(\{y_b\}), prop\{r_b\})$

This postulate introduces the propagation function, which does express how uncertainty at connected agents is propagated to the agent a. One may observe the uniformity of *prop*, which in the setting of MA3 depends only on values of r_b's; this is undoubtedly a simplifying assumption, but we want to avoid unnecessary and obscuring the general view complications, which of course can be multiplied at will.
For Archimedean or metric induced rough inclusions μ, in whose cases $g_\mu(u, r) = \{v : \mu(v, u, r)\}$, see Proposition 1, in this chapter, MA3 induces an aggregation operator on granules

MA4 *For* $b \in B_a$, $ingr_b(x_b, g_{mu_r}(u_b, r_b))$ *implies*

$$ingr_a(aggr(\{x_b\}), g_{\mu_a}(aggr(\{u_b\}), prop(\{r_b\}))$$

Admitting MA4, we may also postulate that in case agents have at their disposal variants of rough mereological granular logics, intensions propagate according to the *prop* functor

MA5 *If* $I^{\vee}_{\nu_b}(g_b)(\phi_b) \geq r_b$ *for each* $b \in B_a$, *then*

$$I^{\vee}_{\nu_a}(aggr(\{g_b\}))(synt(\{\phi_b\})) \geq prop(\{r_b\}))$$

Here, we abuse language a bit, as we write *prop* in the same form as in MA3, again, it is done for simplicity of exposition.

We examine, for example's sake, a simple case of knowledge fusion, cf., Polkowski [51].

Proposition 24. *We consider an agent $a \in Ag$ and – for simplicity reasons – we assume that a has two incoming connections from agents b, c; the number of outgoing connections is of no importance as a sends along each of them the same information. Thus, $B_a = \{b, c\}$.*

We assume that each agent is applying the rough inclusion $\mu = \mu_L$ induced by the Lukasiewicz t–norm L in the frame of the respective information system (U_a, A_a), (U_b, A_b), (U_c, A_c). Each agent is also applying the rough inclusion on sets of the form (0.7) in evaluations related to extensions of formulae intensions.

We consider a simple fusion scheme in which information systems at b, c are combined object–wise to make the information system at a; thus, $aggr_a(x, y) = (x, y)$. Such case may happen, e.g., when an object is described with help of a camera image by some features and at the same time it is perceived and recognized with range sensors like infrared or laser sensors and some localization means like GPS.

Then: uncertainty propagation and granule propagation are described by the Lukasiewicz t–norm L and extensions of logical intensions propagate according to the product t–norm P.

Proof. The set A_a of attributes at a equals then $A_b \times A_c$, i.e., attributes in A_a are pairs (a_1, a_2) with $a_1 \in A_b$ and $a_2 \in A_c$, and a fortiori, the value of this attribute is defined as

$$(a_1, a_2)(x, y) = (a_1(x), a_2(y))$$

It follows that the condition holds

$$IND_a(aggr_a(x, y), aggr_a(x', y')) \Leftrightarrow IND_b(x, x') \text{ and } IND_c(y, y') \quad (0.30)$$

Concerning the function $prop_a$, we consider objects x, x', y, y'; clearly,

$$DIS_a(aggr_a(x, y), aggr_a(x', y')) \quad (0.31)$$

is contained (as a subset) in

$$DIS_b(x, x') \times A_c \cup A_b \times DIS_c(y, y') \quad (0.32)$$

It follows by (0.31), (0.32) that

$$|DIS_a(aggr_a(x, y), aggr_a(x', y'))| \quad (0.33)$$

is less or at most equal to

$$|DIS_b(x,x')| \cdot |A_c| + |A_b| \cdot |DIS_c(y,y')| \tag{0.34}$$

As we know

$$\mu_a(aggr_a(x,y), aggr_a(x',y'), t) \tag{0.35}$$

is satisfied with the maximal value of t equal to

$$1 - \frac{|DIS_a(aggr_a(x,y), aggr_a(x',y'))|}{|A_b| \cdot |A_c|} \tag{0.36}$$

As the value of

$$\frac{|DIS_a(aggr_a(x,y), aggr_a(x',y'))|}{|A_b| \cdot |A_c|} \tag{0.37}$$

in the left–side of (0.36) is not less than

$$\frac{|DIS_b(x,x')| \cdot |A_c| + |A_b| \cdot |DIS_c(y,y')|}{|A_b| \cdot |A_c|} \tag{0.38}$$

which in turn is not less than

$$(\frac{|DIS_b(x,x')|}{|A_b|} + \frac{|DIS_c(y,y')|}{|A_c|} + 1 - 1 \tag{0.39}$$

it follows that

$$1 - \frac{|DIS_a(aggr_a(x,y), aggr_a(x',y'))|}{|A_b| \cdot |A_c|} \tag{0.40}$$

is greater or equal to

$$(1 - \frac{|DIS_b(x,x')|}{|A_b|}) + (1 - \frac{|DIS_c(y,y')|}{|A_c|}) - 1 \tag{0.41}$$

which is equivalent to

$$L(max\{r : \mu_b(x,x',r\}, max\{s : \mu_c(y,y',s\}) \tag{0.42}$$

We have
Claim
em If $\mu_b(x,x',r)$ *and* $\mu_c(y,y',s)$, *then*

$$\mu_a(aggr_a(x,y), aggr_a(x',y'), L(r,s))$$

It follows that the propagation function prop is defined by the Łukasiewicz *t–norm*: $prop(r,s) = L(r,s) = max\{0, r+s-1\}$.

In consequence, the granule propagation functor $prop_a^g$ can be defined in our example as

$$prop_a(g_{\mu_b}(x_b, r_b), g_{\mu_c}(y_c, s_c)) = (g_{\mu_a}(aggr_a(x_b, y_c), L(r_b, s_c)) \quad (0.43)$$

The logic synthesizer $synt_a$ is defined by our assumptions as

$$synt_a(\phi_b, \phi_c) = \phi_b \wedge \phi_c \quad (0.44)$$

Finally, we consider extensions of our logical operators of intensional logic.

We have for the extension

$$I(\mu_a)^{\vee}_{prop^g_a(g_b,g_c)}(synt_a(\phi_b, \phi_c)) \quad (0.45)$$

that it is equal to

$$I(\mu_b)^{\vee}_{g_b}(\phi_b) \cdot I(\mu_c)^{\vee}_{g_c}(\phi_c) \quad (0.46)$$

i.e., it is the defined by the product t–norm P □

In such schemes one can discuss *synthesis* processes: when agents are arranged in a tree T, synthesis over T consists in outputting at the root agent $root(T)$ an object x_T. A characterization of x_T can be given by a postulate that it falls into a granule about a specified *standard* s_T within a radius of r_T.

Aggregation operators of agents $aggr_a$ for non–leaf agents a,compose into a global aggregate operator $aggr_T$, which sends objects chosen at leaf agents, at most one for each agent, into an object at the root of T. In a similar vein, propagation operators $prop_a$ compose into the global propagation operator $prop_T$.

The synthesis procedure could be presented as

Input: $g_{\mu_T}(s_t, r_T)$

1. *Find recursively going down the tree to leaf agents: in each leaf agent* b_l, *an object* s_l *such that*

$$aggr_T(\{s_l\}) = s_T \quad (0.47)$$

2. *Find recursively radii* r_l *at leaf agents which satisfy the condition that,*

$$prop_T(\{r_l\}) \geq r_T \quad (0.48)$$

3. *For each leaf agent* b_l, *find an object* x_l *with the property that,*

$$\mu_l(x_l, s_l, r_l) \ for\ each\ leaf\ agent\ b_l \quad (0.49)$$

4. *Output the object* $aggr_T(\{x_l\})$.

This reasoning scheme seems fairly universal also for networking schemes of reasoning like in neural networks. However, as learning in neural networks is usually based on the idea of gradient search, and in consequence it requires differentiable rough inclusions, we devote to neural network reasoning a separate paragraph.

0.13 Reasoning in cognitive schemes

Neural networks are motivated by the functioning of neural system. Their inception was possible due to achievements of neurophysiology, notably, discovery of the structure of the biological neuron by Ramón y Cajal [58].

On this discovery, McCulloch and Pitts [9] built their model of a computing machine called now *McCulloch–Pitts neuron*. In its simplest form, it consisted of a cell endowed with a threshold Θ, and inputs $x_1, x_2, ..., x_n$ through which binary signals of 0 or 1 could be sent. The one output y could issue a binary signal of either 0 or 1. Inputs are counterparts to the biological *synapses*, and the output models the biological *axon*.

Computation by the neuron is governed by the rule

$$y = \begin{cases} 1 \ in\ case\ \sum_i x_i \geq \Theta \\ 0 \ otherwise \end{cases} \tag{0.50}$$

Thus, the neuron *fires* when the summary input exceeds the threshold, otherwise the neuron remains inactive.

The neuron of McCulloch–Pitts can produce a binary *linear classifier*: the separating hyper–plane, which is to divide the training set into positive and negative examples, is of the form

$$\sum_i x_i = \Theta \tag{0.51}$$

The idea of networks of neurons was advocated by Alan Turing [61] who proposed a learning scheme for networks of neurons connected through *modifiers*.

Increasing classification possibilities were given a neuron with the idea of a *perceptron* due to Rosenblatt [59]. A simplified perceptron adds to McCulloch–Pitts neuron *weights* on inputs, and an additional input with constant value of 1 and a weight b, called *bias*. Thus, the classifying hyper–plane has the form

$$\sum_i w_i \cdot x_i + b = \Theta \tag{0.52}$$

Parameters w_i, b, Θ which can be varied, allow for *learning*. As proved by Novikoff [38], if a binary concept can at all be classified with the hyperplane of the form (0.52), then it is possible to find in a finite number of steps of the *learning procedure* a proper set of parameters starting from, e.g., a randomly chosen set.

The idea of a network of perceptrons returned with Grossberg [14]; it was shown that a network of perceptrons can classify, hence learn, any Boolean function.

A further step in improving efficiency of computation and learning of perceptron networks was due to Bryson and Ho [4]: replacement of the Heaviside–type threshold function with a sigmoid–type differentiable function allowed for *supervised learning* by means of the gradient search, called the *backpropagation*, based on computing the gradient of the *error function* $E(w) = \sum_{x_j}(y_j - t_j)^2$, where w is the vector of weights on all connections in the network, x_j are consecutive input examples, y_j are corresponding responses of the network, and t_j are desired outputs to inputs x_j.

We would like to endow neurons with intelligence based on rough mereological reasoning, see Polkowski [44]. In neural models of computation, an essential feature of neurons is differentiability of transfer functions; hence, we introduced a special type of rough inclusions, called *gaussian* in Polkowski [44] because of their form, by letting,

$$\mu_G(x, y, r) \text{ iff } e^{-|\sum_{a \in DIS(x,y)} w_a|^2} \geq r \tag{0.53}$$

where $w_a \in (0, +\infty)$ is a weight associated with the attribute a for each attribute $a \in A$.

Let us observe in passing that μ_G can be factored through the indiscernibility relation $IND(A)$, and thus its arguments can be objects as well as indiscernibility classes; we will freely use this fact.

Properties of gaussian rough inclusions are, see Polkowski [44]

Proposition 25. *The following hold for each gaussian rough inclusion μ_G, where ingr is the ingredient relation associated with μ_G by postulate RINC1, Ch. 5, sect. 2,*

1. *$ingr(x, y)$ if and only if $DIS(x, y) = \emptyset$;*

2. *There exists a function $\eta(r, s)$ such that $\mu_G(x, y, r), \mu_G(y, z, s)$ imply $\mu_G(x, z, t)$ with $t \leq \eta(r, s))$;*

3. *If $ingr(x, g_{\mu_G}(y, r)$ and $ingr(x, g_{\mu_G}(z, s)$, then $ingr(g_{\mu_G}(x, t), g_{\mu_G}(y, r))$, $ingr(g_{\mu_G}(x, t), g_{\mu_G}(z, s))$ for $t \geq max\{r^4, s^4\}$.*

Proof. Property 1 follows by definition 25.

To verify Property 2, we observe that by definition (25), we have

$$\sum_{a \in DIS(x,y)} w_a \leq (-logr)^{\frac{1}{2}} \tag{0.54}$$

and

$$\sum_{a\in DIS(y,z)} w_a \leq (-logs)^{\frac{1}{2}} \tag{0.55}$$

As $DIS(x,z) \subseteq DIS(x,y) \cup DIS(y,z)$, we denote by t the maximum value of q such that $\mu_G(x,z,q)$, and we obtain that

$$(-logt)^{\frac{1}{2}} \leq (-logr)^{\frac{1}{2}} + (-logs)^{\frac{1}{2}} \tag{0.56}$$

from which by taking squares of both sides, we get at

$$logt \leq log(r \cdot s) - 2(logr \cdot logs)^{\frac{1}{2}} \tag{0.57}$$

so finally

$$t \leq r \cdot s \cdot e^{-2(logr \cdot logs)^{\frac{1}{2}}} \tag{0.58}$$

Thus, it suffices to take $\eta(r,s) = r \cdot s \cdot e^{-2(logr \cdot logs)^{\frac{1}{2}}}$.

In case of Property 3, we observe that as μ_G is symmetric by definition and transitive by Property 2, we have $ingr(x, g_{mu_G}(y,r)) \Leftrightarrow \mu_G(x,y,r)$. Hence, in order to verify Property 3, we have to find t such that $\eta(r,t) \geq r, \eta(s,t) \geq s$. As $\eta(r,t) \geq r$ implies $t \geq r^4$, it suffices to have $t \geq max\{r^4, s^4\}$ to satisfy both conditions □

0.13.1 Rough mereological perceptron

The rough mereological perceptron is modeled on the perceptron, and it consists of an intelligent agent a, endowed with a gaussian rough inclusion μ_a on the information system $I_a = (U_a, A_a)$ of the agent a.

The input to a is in the form of a finite tuple $\overline{x} = (x_1, ..., x_k)$ of objects, and the input $\overline{x}$ is converted at a into an object $x = aggr_a(\overline{x}) \in U_a$ by means of an operator $aggr_a$.

The rough mereological perceptron is endowed with a set of *target concepts* $T_a \subseteq U_a/IND(A_a)$, each target concept a class of the indiscernibility IND_a.

Formally, a rough mereological perceptron is thus a tuple

$$RMP = (a, I_a, \mu_a, aggr_a, T_a)$$

Computing by a network of RMP's, is directed by the gradient of the error function, which in this case has the form of the gradient of the function

$$f(x,y) = e^{-|\sum_{a\in DIS(x,y)} w_a|^2} \tag{0.59}$$

which is

$$\frac{\partial f}{\partial w} = f \cdot (-2 \cdot \sum w_a) \tag{0.60}$$

It follows from the last equation that gradient search would go in direction of minimizing the value of $\sum_a w_a$.

The result of computation with a target $g_{\mu_G}(t,r)$ for a a sample $x_1, ..., x_k$ is a granule $g = g_{\mu_G}(aggr_a(x_1, .., x_k), r(res))$ such that $ingr(g, g_{\mu_G}(t,r))$.

During computation, weights are incremented according to the ideology of backpropagation, by the recipe,

$$w_a \leftarrow w_a + \Delta \cdot \frac{\partial E}{\partial w_a} \tag{0.61}$$

where Δ is the *learning rate*.
At a stage *current* of computing, where $\gamma = r_{current} - r$, for a natural number k, the value of $\Delta_{current}$ which should be taken at the step *current* in order to achieve the target in at most k steps should be taken as, see Polkowski [44]

$$\Delta_{current} \simeq \frac{\gamma}{2 \cdot k \cdot f^2 \cdot (\sum_a w_a)^2} \tag{0.62}$$

References

1. Alves E. H., Guerzoni J. A. D. (1990) Extending Montague's system: A three–valued intensional logic. Studia Logica 49, pp 127–132
2. Amarel S. (1991) Panel on AI and Design. Proceedings of 12th Intern. Conf. on AI, Sydney, pp 563–565
3. Van Benthem J. (1988) A Manual of Intensional Logic. CSLI Stanford University, Stanford, CA
4. Bryson A. E., Ho Yu–Chi (1969) Applied Optimal Control: Optimization, Estimation and Control. Blaisdell Publishing Company, Waltham, MA
5. Canny J. F. (1988) The Complexity of Robot Motion Planning. MIT Press, Cambridge, MA
6. Carnap R. (1947) Necessity and Meaning. Chicago University Press, Chicago, IL
7. Chang C. C. (1958) Proof of an axiom of Łukasiewicz. Trans. Amer. Math. Soc. 87, pp 55–56
8. Choset H., Lynch K.M., Hutchinson S., Kantor G., Burgard W., Kavraki L.E., Thrun S. (2005) Principles of Robot Motion. Theory, Algorithms, and Implementations. MIT Press, Cambridge, MA
9. McCulloch W., Pitts W. (1943) A logical calculus of the ideas immanent in nervous activity. Bulletin of Mathematical Biophysics 7, pp 115–133
10. Dietterich T. G. (2000) Ensemble methods in machine learning. 1st Intern. Workshop on Multiple Classifier Systems. In: Lecture Notes in Computer Science 1857, Springer Verlag, Berlin, pp 1–15
11. Fitting M. C. (2004) First–order intensional logic. Annals of Pure and Applied Logic 127, pp 171–193
12. Frege G. (1892) *Ü*ber Sinn und Bedeutung. Zeitschrift für Philosophie und Philosophische Kritik NF 100, pp 25-50
13. Gallin D. (1975) Intensional and higher–order modal logic. North Holland, Amsterdam
14. Grossberg S. (1973) Contour enhancement, short–term memory, and constancies in reverberating neural networks. Studies in Applied Mathematics 52, pp 213–257
15. Grzymala–Busse J. W. (2004) Data with missing attribute values: Generalization of indiscernibility relation and rule induction. Transactions on Rough Sets I. Lecture Notes in Computer Science 3100, Springer Verlag, Berlin, pp 78–95
16. Grzymala–Busse J. W., Ming Hu (2004) A comparison of several approaches to missing attribute values in data mining. Lecture Notes in Artificial Intelligence 2005, Springer Verlag, Berlin, pp 378–385
17. H*á*jek P. (2001) Metamathematics of Fuzzy Logic. Kluwer, Dordrecht
18. Hughes G. E., Creswell M. J. (1996) A New Introduction to Modal Logic. Routledge, London

19. Kripke S. (1963) Semantical considerations on modal logics. Acta Philosophica Fennica. Modal and Many–Valued Logics pp 83–94
20. Lemmon E. J., Scott D. S. (1963) Sederberg, K. (Ed.). The Lemmon Notes. An Introduction to Modal Logic. Basil Blackwell, Oxford, UK
21. Lin T. Y. (1988) Neighborhood systems and relational database. Abstract. In: Proceedings of CSC'88, pp 725
22. Lin T. Y. (1989) Neighborhood systems and approximation in Database and Knowledge Based Systems. In: Proceedings of the 4th International Symposium on Methodologies for Intelligent Systems (ISMIS), pp 75–86
23. Lin T. Y. (1992) Topological and fuzzy rough sets. In: Słowiński R. (Ed.)(1992) Intelligent Decision Support. Handbook of Applications and Advances of the Rough Sets Theory. Kluwer, Dordrecht, pp 287–304
24. Lin T. Y. (1994) From rough sets and neighborhood systems to information granulation and computing with words. In: Proceedings of the European Congress on Intelligent Techniques and Soft Computing, pp 1602–06
25. Lin T. Y. (1999) Granular Computing: Fuzzy logic and rough sets. In: Zadeh L. A., Kacprzyk J, (Eds.) (1999) Computing with Words in Information/Intelligent Systems 1. Physica Verlag, Heidelberg, pp 183–200
26. Lin T. Y. (2003) Granular computing. Lecture Notes in Computer Science 2639, Springer Verlag, Berlin, pp 16–24
27. Lin T. Y. (2005) Granular computing: Examples, intuitions, and modeling. In: Proceedings of IEEE 2005 Conference on Granular Computing GrC05, Beijing, China. IEEE Press, pp 40–44
28. Lin T. Y. (2006) A roadmap from rough set theory to granular computing. In: Proceedings RSKT 2006, 1st International Conference on Rough Sets and Knowledge Technology, Chongqing, China. Lecture Notes in Artificial Intelligence 4062, Springer Verlag, Berlin, pp 33–41
29. Łukasiewicz J. (1913) Die Logischen grundlagen der Wahrscheinlichtkeitsrechnung. Cracow
30. Łukasiewicz, J. (1920) O logice trójwartościowej (On three–valued logic, in Polish). Ruch Filozoficzny 5, pp 170–171
31. Łukasiewicz, J. (1957) Aristotle's Syllogistic from the Standpoint of Modern Formal Logic. 2nd ed. Oxford University Press, Oxford, UK
32. Łukasiewicz J., Tarski A. (1930) Untersuchungen ueber den Aussagenkalkuels. C.R. Soc. Sci. Lettr. Varsovie 23, pp 39–50
33. Meredith C. A. (1958) The dependence of an axiom of Łukasiewicz. Trans. Amer. Math. Soc. 87, p 54
34. Montague R. (1970) Pragmatics and intensional logic. Synthese 22, pp 68–94
35. Montague R. (1974) Thomason R. (Ed.). Formal Philosophy. Yale University Press, New Haven, CN
36. Nguyen S. H. (2000) Regularity analysis and its applications in Data Mining. In: Polkowski L., Tsumoto S., Lin T. Y. (Eds.) (2000) Rough Set Methods and Applications. New Developments in Knowledge Discovery in Information Systems. Physica Verlag, Heidelberg, pp 289–378
37. Nieminen J. (1988) Rough tolerance equality and tolerance black boxes. Fundamenta Informaticae 11, pp 289–296
38. Novikoff A. B. (1962) On convergence proofs on perceptrons. Symposium on the Mathematical Theory of Automata 12, pp 615–622. Brooklyn Polytechnic Institute, New York
39. Pawlak Z (1982) Rough sets. Intern. J. Comp. Inform. Sci. 11, pp 341–366
40. Pawlak Z. (1991) Rough Sets: Theoretical Aspects of Reasoning about Data. Kluwer, Dordrecht
41. Polkowski L. (2003) A rough set paradigm for unifying rough set theory and fuzzy set theory. Fundamenta Informaticae 54, pp 67–88; and in: Proceedings RSFDGrC03, Chongqing, China, 2003. Lecture Notes in Artificial Intelligence 2639, Springer Verlag, Berlin, pp 70–78

42. Polkowski L. (2004) A note on 3–valued rough logic accepting decision rules. Fundamenta Informaticae 61, pp 37–45
43. Polkowski L. (2004) Toward rough set foundations. Mereological approach. In: Proceedings RSCTC04, Uppsala, Sweden. Lecture Notes in Artificial Intelligence 3066, Springer Verlag, Berlin, pp 8–25
44. Polkowski L. (2004) A rough–neural computation model based on rough mereology. In: Pal S. K., Polkowski L., Skowron A.(Eds.) (2004) Rough–Neural Computing. Techniques for Computing with Words. Springer Verlag, Berlin, pp 85–108
45. Polkowski L. (2005) Formal granular calculi based on rough inclusions. In: Proceedings of IEEE 2005 Conference on Granular Computing GrC05, Beijing, China. IEEE Press, pp 57–62
46. Polkowski L. (2005) Rough–fuzzy–neurocomputing based on rough mereological calculus of granules. International Journal of Hybrid Intelligent Systems 2, pp 91–108
47. Polkowski L. (2006) A model of granular computing with applications. In: Proceedings of IEEE 2006 Conference on Granular Computing GrC06, Atlanta, USA. IEEE Press, pp 9–16
48. Polkowski L. (2007) Granulation of knowledge in decision systems: The approach based on rough inclusions. The method and its applications. In: Proceedings RSEiSP'07 (in memory of Z. Pawlak). Lecture Notes in Artificial Intelligence 4585, pp 69–79
49. Polkowski L. (2007) The paradigm of granular rough computing. In: Proceedings ICCI'07. 6th IEEE Intern. Conf. on Cognitive Informatics. IEEE Computer Society, Los Alamitos CA, pp 145–163
50. Polkowski L. (2008) Rough mereology in analysis of vagueness, In: Proceedings RSKT 2008. Lecture Notes in Artificial Intelligence 5009, Springer Verlag, Berlin, pp 197–205
51. Polkowski L. (2008) A unified approach to granulation of knowledge and granular computing based on rough mereology: A survey. In: Pedrycz W., Skowron A., Kreinovich V. (Eds.) (2008) Handbook of Granular Computing. John Wiley and Sons Ltd., Chichester, UK, pp 375–400
52. Polkowski L. (2009) Granulation of Knowledge: Similarity Based Approach in Information and Decision Systems. In: Meyers R. A. (Ed.) (2009) Springer Encyclopedia of Complexity and System Sciences, Springer Verlag, Berlin, article 00 788
53. Polkowski L., Semeniuk–Polkowska M. (2005) On rough set logics based on similarity relations. Fundamenta Informaticae 64, pp 379–390
54. Polkowski L., Skowron A. (1998) Rough mereological foundations for design, analysis, synthesis and control in distributed systems. Information Sciences. An International Journal 104(1–2), pp 129–156
55. Polkowski L., Skowron A. (1999) Grammar systems for distributed synthesis of of approximate solutions extracted from experience. In: Paun Gh., Salomaa A.(Eds.) (1999) Grammatical models of Multi-Agent Systems. Gordon and Breach, Amsterdam, pp 316–333
56. Polkowski L., Skowron A. (1999) Towards adaptive calculus of granules. In: Zadeh L.A., Kacprzyk J. (Eds.) (1999) Computing with Words in Information/Intelligent Systems 1. Foundations, Physica Verlag, Heidelberg
57. Polkowski L., Skowron A. (2001) Rough mereological calculi of granules: A rough set approach to computation. Computational Intelligence. An Intern. Journal 17(3), pp 472–492
58. Ramón y Cajal S. (1889) Sur la morphologie et les connexions des elements de la retine des oiseaux. Anatomisches Anzeiger 4, pp 111–121
59. Rosenblatt F. (1958) The perceptron: A probabilistic model for information storage and organization in the brain. Psychological Review 65(6), pp 386–408
60. Stone P. (2000) Layered Learning in Multiagent Systems: A Winning Approach to Robotic Soccer. MIT Press, Cambridge, MA
61. Turing A. M. (1948) Intelligent Machinery. A report. National Physical Laboratory. Mathematical Division

62. Wajsberg M. (1931) Axiomatization of the 3–valued sentential calculus (in Polish, a summary in German).C. R. Soc. Sci. Lettr. Varsovie 24, pp 126–148
63. Wille R. (1982) Restructuring lattice theory: An approach based on hierarchies of concepts. In: Rival I. (Ed.)(1982) Ordered Sets. D. Reidel, Dordrecht, pp 445–470
64. Yao Y. Y. (2000) Granular computing: Basic issues and possible solutions. In: Proceedings of the 5^{th} Joint Conference on Information Sciences I. Assoc. Intell. Machinery, Atlantic, NJ, pp 186–189
65. Yao Y. Y. (2005) Perspectives of granular computing. In: Proceedings of IEEE 2005 Conference on Granular Computing GrC05, Beijing, China. IEEE Press, pp 85–90
66. Zadeh L. A.(1979) Fuzzy sets and information granularity. In: Gupta M., Ragade R., Yager R.R. (Eds.) (1979) Advances in Fuzzy Set Theory and Applications. North–Holland, Amsterdam, pp 3–18
67. Zadeh L. A. (2004) Toward a unified theory of uncertainty. In: Proceedings of IPMU 2004, 1. Perugia, Italy, pp 3–4