A Graded Applicability of Rules

Anna Gomolińska*

University of Białystok, Department of Mathematics,
Akademicka 2, 15267 Białystok, Poland,
anna.gom@math.uwb.edu.pl

Abstract. We address the problem of rough applicability of rules within the framework of approximation spaces. The graded applicability of a rule for an object of an approximation space, introduced here, is viewed as a fundamental form of rough applicability. It is based on the graded meaning of a set of formulas, defined in our previous works. The notion of graded applicability enjoys a number of interesting properties and it is useful – in our opinion – in modeling of rule-based complex systems like multi-agent systems.

1 Introduction

When thinking of the problem of application of a rule, one can distinguish, among others, three more specific questions: applicability, results of application, and quality of a rule. The first question concerns the premises, the second – the conclusion(s), and the last one – the relationship between the premises and the conclusion of a rule. In this article, applicability of rules is addressed within the framework of approximation spaces. We introduce a notion of graded applicability of a rule for an object, based on the graded meaning of formulas and their sets [2, 3]. The graded applicability of rules, presented here, is viewed as a fundamental form of rough applicability to be further extended and refined.

The concept of graded applicability is interesting not merely for its theoretical properties. In our opinion, it may be useful in modeling of rule-based complex systems like systems of social agents, where rules are often followed in a flexible way. By means of this notion and its extensions, varied aspects concerning applicability of rules may be investigated and explained in theoretical terms. Another area in which one can use appropriate soft concepts of applicability of rules, based on the notion proposed here, is multi-criterial classification and, in particular, decision making. Graded forms of applicability cope with some cases of missing values of attributes and contribute to the greater effectiveness in classifying objects.

Throughout the paper, for a set x and $n > 0$, $\#x$ denotes its cardinality, $\wp x$ – the power set, and x^n – the Cartesian product of x taken n times. Let

* Many thanks to Andrzej Skowron, Marcin Wolski, and the anonymous referees for valuable remarks. The research has been partially supported by the grant 3T11C00226 from Ministry of Scientific Research and Information Technology of the Republic of Poland.

$i = 1, \ldots, n$, $(x_i, \leq_i)$ be partially ordered sets, and $s, t \in x_1 \times \ldots \times x_n$. Then, $\pi_i(t)$ denotes the i-th element of t. By $\geq_i$ we denote the converse relation of $\leq_i$, whereas by $<_i, >_i$ – the strict versions of $\leq_i, \geq_i$, respectively. $\preceq$ is a partial ordering such that $s \preceq t$ iff $\forall i = 1, \ldots, n.\pi_i(s) \leq_i \pi_i(t)$. Degrees of applicability of rules are constructed from numbers of $[0, 1]$ and a constant c, denoting "crisp" as opposite to "vague". By assumption $c \leq c$, where $\leq$ extends the natural ordering on $[0, 1]$ to the set $T \overset{\text{def}}{=} [0, 1] \cup \{c\}$. Additionally, let $T_1 \overset{\text{def}}{=} T \times [0, 1]$. Parentheses and c will be dropped in formulas if no confusion results.

Section 2 is devoted to the notion of an approximation space. In Sect. 3, the concepts of graded satisfiability and graded meaning of formulas and their sets are recalled. The notion of graded applicability of rules is presented in Sect. 4. In the next section, we give an illustrative example. Section 6 contains a concise summary.

2 Approximation Spaces

By an *approximation space* (AS) we understand a triple $\mathcal{M} = (U, \Gamma, \kappa)$, where U is a non-empty set, $\Gamma : U \mapsto \wp U$ is an *uncertainty mapping*, and $\kappa : (\wp U)^2 \mapsto [0, 1]$ is a *rough inclusion function* (RIF) [6].[1] The ASs, studied initially, were those induced by the *Pawlak information systems* [4].

Elements of U, referred to as objects and denoted by u, v with subscripts if needed, are known by their properties only. Some objects may seem to be similar from an observer's perspective. Objects enjoying the same properties are similar in a peculiar way: They are indiscernible. It is assumed that every object is necessarily similar to itself. Thus, the universe U is covered by a family of clusters of objects, called granules of information. In $\mathcal{M}$, the mapping Γ assigns to every object u, an elementary granule $\Gamma(u)$ of objects similar to u. By assumption, $u \in \Gamma(u)$.

The most popular RIFs, called *standard*, are based on the frequency count. For a finite U and $x, y \subseteq U$, the standard RIF κ^{st} is defined by $\kappa^{st}(x, y) = \begin{cases} \frac{\#(x \cap y)}{\#x} & \text{if } x \neq \emptyset \\ 1 & \text{otherwise.} \end{cases}$ A RIF is *quasi-standard* if it is defined as the standard one for finite first arguments. In general, a RIF assigns to every pair (x, y) of subsets of U, a number of the unit interval $[0, 1]$, expressing the degree of inclusion of x in y. Polkowski and Skowron proposed *Rough Mereology*, a formal theory of being-a-part-in-degree, axiomatically characterizing a general notion of RIF [5]. In our approach, every RIF $\kappa : (\wp U)^2 \mapsto [0, 1]$ is supposed to satisfy (A1)–(A3) for any $x, y, z \subseteq U$:

$$\text{(A1) } \kappa(x, y) = 1 \text{ iff } x \subseteq y.$$
$$\text{(A2) If } x \neq \emptyset, \text{ then } \kappa(x, y) = 0 \text{ iff } x \cap y = \emptyset.$$
$$\text{(A3) If } y \subseteq z, \text{ then } \kappa(x, y) \leq \kappa(x, z).$$

[1] Originally, Γ and κ were equipped with lists of tuning parameters, dropped for simplicity here. Such ASs were referred to as *parameterized*.

The essense of ASs is that sets of objects of the universe are approximated by means of uncertainty mappings and RIFs. There are many alternative ways of approximation (see, e.g., [1] for a discussion and references).

3 The Graded Meaning of Formulas

In this section, we briefly recall the notions of graded satisfiability and graded meaning of formulas, studied in [2, 3]. Suppose a language L is given, expressing properties of $\mathcal{M}$. Formulas of L, denoted by α, β, γ with subscripts if needed, form the set FOR. Assume that a commutative conjunction ($\wedge$) occurs in L. Then, the conjunction of all elements of a non-empty finite set of formulas X is denoted by $\bigwedge X$.

Given a relation of (crisp) satisfiability of formulas for objects of U, $\models_c$, where $u \models_c \alpha$ reads as "α is *c-satisfied* for u", the *c-meaning* of α is understood along the standard lines as the set $||\alpha||_c = \{u \mid u \models_c \alpha\}$. These notions are refined by introducing degrees $t \in [0, 1]$:

$$u \models_t \alpha \text{ iff } \kappa(\Gamma(u), ||\alpha||) \geq t \text{ and } ||\alpha||_t = \{u \mid u \models_t \alpha\}. \quad (1)$$

$u \models_t \alpha$ reads as "α is t-satisfied for u", and $||\alpha||_t$ is called the *t-meaning* of α. Where $t \in T$, the set of all formulas, t-satisfied for an object u, is denoted by $|u|_t$:

$$|u|_t = \{\alpha \mid u \models_t \alpha\}. \quad (2)$$

Along the standard lines, a set of formulas X is *c-satisfied* for u, $u \models_c X$, iff $\forall\alpha \in X.u \models_c \alpha$. The (crisp) meaning of X is the set $||X||_c = \{u \mid u \models_c X\}$. Next, let $\kappa^* : (\wp\text{FOR})^2 \mapsto [0, 1]$ be a RIF and $t \in T_1$. Then,

$$u \models_t X \text{ iff } \kappa^*(X, |u|_{\pi_1(t)}) \geq \pi_2(t) \text{ and } ||X||_t = \{u \mid u \models_t X\}. \quad (3)$$

$u \models_t X$ reads as "X is *t-satisfied* for u", and $||X||_t$ is the *t-meaning* of X.

4 A Graded Form of Applicability of Rules

Rules over L, denoted by r with sub/superscripts if needed, are to describe dependencies and properties of objects of U and their sets in $\mathcal{M}$, and they consist of two components: finitely many premises and conclusion(s), all being formulas of L. From our standpoint it is of minor importance whether or not a rule may have one or more conclusions; we omit this question for the time being. The set of premises of r is denoted by $P(r)$ and the set of all rules over L by RUL. In $\mathcal{M}$, a rule r is *applicable* for an object u iff $P(r)$ is satisfied for u, or in other words, $u \in ||P(r)||$.[2] This concept is refined in two ways: by replacing the crisp satisfiability of premises by its graded counterpart and by

[2] Equivalently, r is applicable for u iff all premises of r are satisfied for u, i.e., $\forall\alpha \in P(r).u \in ||\alpha||$.

weakening the condition that all premises be satisfied. Where $t \in T_1$, $\mathrm{apl}_t(u)$ denotes the set of all rules t-applicable for u. Intuitively, r is *t-applicable* for u[3] if u satisfies a sufficiently large part of $P(r)$ in a sufficient degree, where sufficiency is determined by t; otherwise r is *t-inapplicable* for u. Formally,

$$r \in \mathrm{apl}_t(u) \overset{\mathrm{def}}{\leftrightarrow} \kappa^*(P(r), |u|_{\pi_1(t)}) \geq \pi_2(t), \text{i.e., iff } u \in ||P(r)||_t. \tag{4}$$

Theorem 1. *For any objects u, v, formulas α, β, a set of formulas X, a finite set of formulas Y, rules r, r_1, r_2, a finite non-empty set of rules $\{r_i\}_{i \in I}$, $s_1 \in [0,1]$, $s \in T$, $t, t_1, t_2 \in T_1$, and assuming that κ^* is quasi-standard in (j), we have:*

(*a*) If $\Gamma(u) = \Gamma(v)$, then $\mathrm{apl}_t(u) = \mathrm{apl}_t(v)$.

(*b*) $r \in \mathrm{apl}_{(s,1)}(u)$ iff $P(r) \subseteq |u|_s$ iff $u \in \bigcap\{||\alpha||_s \mid \alpha \in P(r)\}$.

(*c*) $\mathrm{apl}_{(s,0)}(u) = \mathrm{RUL}$.

(*d*) If $P(r) = \{\alpha\}$ and $\pi_2(t) > 0$, then $r \in \mathrm{apl}_t(u)$ iff $u \in ||\alpha||_{\pi_1(t)}$.

(*e*) If $t_1 \preceq t_2$, then $\mathrm{apl}_{t_2}(u) \subseteq \mathrm{apl}_{t_1}(u)$.

(*f*) $\mathrm{apl}_{(1,s_1)}(u) \subseteq \mathrm{apl}_{(s,s_1)}(u) \subseteq \mathrm{apl}_{(0,s_1)}(u) = \mathrm{RUL}$.

(*g*) $\bigcap\{\mathrm{apl}_t(u) \mid u \in U \wedge t \in T_1\} = \{r \in \mathrm{RUL} \mid ||P(r)|| = U\}$.

(*h*) If $P(r_1) \subseteq P(r_2)$ and $r_2 \in \mathrm{apl}_{(s,1)}(u)$, then $r_1 \in \mathrm{apl}_{(s,1)}(u)$.

(*i*) If $\exists \alpha \in P(r).||\alpha||_s = \emptyset$, then $\forall u.r \notin \mathrm{apl}_{(s,1)}(u)$.

(*j*) If $P(r_2) = P(r_1) - Y$ and $||Y||_{(\pi_1(t),1)} = U$, then $r_2 \in \mathrm{apl}_t(u)$ implies $r_1 \in \mathrm{apl}_t(u)$.

(*k*) If $P(r_2) \neq \emptyset$, $P(r_1) = \{\bigwedge P(r_2)\}$, and $\pi_2(t) > 0$, then $r_1 \in \mathrm{apl}_t(u)$ implies $r_2 \in \mathrm{apl}_t(u)$.

(*l*) If $P(r) = \bigcup_{i \in I} P(r_i)$ and $\pi_2(t) = 1$, then $r \in \mathrm{apl}_t(u)$ iff $\{r_i\}_{i \in I} \subseteq \mathrm{apl}_t(u)$.

(*m*) If $P(r_2) = P(r_1) - X$, $\pi_2(t) = 1$, and $||P(r_1) \cap X||_t = U$, then $r_1 \in \mathrm{apl}_t(u)$ iff $r_2 \in \mathrm{apl}_t(u)$.

(*n*) If $P(r_2) = P(r_1) \cup X$, $\pi_2(t) = 1$, and $||X - P(r_1)||_t = U$, then $r_1 \in \mathrm{apl}_t(u)$ iff $r_2 \in \mathrm{apl}_t(u)$.

To give a few comments, observe that t-applicability of rules is determined up to granulation by (a). Next, the partial order on the family of parameterized operators of applicability apl_t reverses the order on parameters by (e), (f). $\mathrm{apl}_{(1,1)}$ is the least element and $\mathrm{apl}_{(0,0)}$ – the greatest one. Finally, axiomatic rules, i.e., rules without premises are t-applicable for all $u \in U$ and $t \in T_1$ in virtue of (g).

5 An Illustrative Example

Consider a Pawlak information system $\mathcal{A} = (U, A)$, where $U = \{2, \ldots, 12\}$ and $a_1, a_2, a_3 \in A$. Any attribute $a \in A$ assigns to each object $u \in U$, a value $a(u) \in$

[3] In other words, r is *applicable* for u *in degree* t.

V_a. Values of attributes are denoted by b with subscripts if needed. Let $b_1 \in V_{a_1}$, $b_2, b_3 \in V_{a_2}$, $b_4 \in V_{a_3}$, and $*$ represent other values. A simple logical language, interpreted in $\mathcal{A}$, is defined. Constant symbols are elements of A and $\bigcup_{a \in A} V_a$. Propositional connectives are the classical ones. Atomic formulas have the form (a_i, b). Formulas are obtained from the atomic formulas along the standard lines. For any formulas $(a_i, b), \alpha, \beta$, satisfiability for an object u is defined as follows: $u \models (a_i, b)$ iff $a_i(u) = b$; $u \models \alpha \wedge \beta$ iff $u \models \alpha$ and $u \models \beta$; $u \models \neg\alpha$ iff $u \not\models \alpha$. Let $\alpha = ((a_1, b_1) \wedge (a_2, b_2)) \vee (a_3, b_4)$, $\beta = \neg(a_2, b_2) \wedge \neg(a_2, b_3)$, and $\gamma = \neg\alpha \vee \beta$. Then, $||\alpha|| = \{2, 4, 6, 8, 11, 12\}$, $||\beta|| = \{4, 5, 11\}$, and $||\gamma|| = \{3, 4, 5, 7, 9, 10, 11\}$. Table 1 shows a fragment of $\mathcal{A}$. According to the mapping Γ, objects $2, 6$ are similar to 2, i.e., $\Gamma(2) = \{2, 6\}$. Subsequently, $\Gamma(3) = \{3, 5, 9\}$, $\Gamma(4) = \{4, 11\}$, $\Gamma(5) = \{4, 5\}$, $\Gamma(6) = \{2, 6, 12\}$, $\Gamma(7) = \{4, 7, 8\}$, $\Gamma(8) = \{3, 8\}$, $\Gamma(9) = \{9, 10\}$, $\Gamma(10) = \{3, 9, 10\}$, $\Gamma(11) = \{2, 11\}$, and $\Gamma(12) = \{6, 12\}$. It is assumed that the RIFs considered are quasi-standard. Consider a rule r with premises α, γ (i.e., $P(r) = \{\alpha, \gamma\}$) and with one conclusion β.[4] For $t = (t_1, t_2) \in T_1$, the t-meaning of $P(r)$ is given in Table 2 and the conditions for t-applicability of r for objects of U in Table 3.

Table 1. A fragment of the information system $\mathcal{A}$.

$a \backslash u$	2	3	4	5	6	7	8	9	10	11	12
a_1	b_1	b_1	$*$	$*$	b_1	$*$	b_1	$*$	$*$	$*$	b_1
a_2	b_2	b_3	$*$	$*$	b_2	b_2	b_3	b_3	b_3	$*$	b_2
a_3	b_4	$*$	b_4	$*$	$*$	$*$	b_4	$*$	$*$	b_4	$*$

Table 2. The t-meaning of $P(r)$.

$t_2 \backslash t_1$	0	$(0, \frac{1}{3}]$	$(\frac{1}{3}, \frac{1}{2}]$	$(\frac{1}{2}, \frac{2}{3}]$	$(\frac{2}{3}, 1]$	c
0	U	U	U	U	U	U
$(0, \frac{1}{2}]$	U	U	U	$U - \{8\}$	$U - \{7, 8\}$	U
$(\frac{1}{2}, 1]$	U	$\{4, 5, 7, 8, 11\}$	$\{4, 5, 8, 11\}$	$\{4\}$	$\{4\}$	$\{4, 11\}$

Thus, r is applicable, i.e., $(c, 1)$-applicable for 4 and 11. Let $t_2 = 1$. r may also be applied for 5 and 8 if $\frac{1}{3} < t_1 \leq \frac{1}{2}$. On the other hand, t-applicability is more restrictive than the crisp one for $t_1 > \frac{1}{2}$. In this case, r is t-applicable for 4 only.

[4] Actually, r is an instance of the resolution rule.

Table 3. The conditions for t-applicability of r.

u	Condition
$2, 3, 6, 9, 10, 12$	$t_1 = 0 \vee t_2 \leq \frac{1}{2}$
4	$t \in T_1$
5	$t_1 \leq \frac{1}{2} \vee t_2 \leq \frac{1}{2}$
7	$t_1 \leq \frac{1}{3} \vee t_2 = 0 \vee ((t_1 \leq \frac{2}{3} \vee t_1 = c) \wedge t_2 \leq \frac{1}{2})$
8	$t_1 \leq \frac{1}{2} \vee t_2 = 0 \vee (t_1 = c \wedge t_2 \leq \frac{1}{2})$
11	$t_1 \leq \frac{1}{2} \vee t_1 = c \vee t_2 \leq \frac{1}{2}$

6 Summary

In the paper, intended as an introduction to a larger study of the problem of soft application of rules, we started with a fundamental concept of graded applicability of a rule for an object of an AS. In our opinion, both this notion as well as some of its extensions and refinements may prove useful in modeling of rule-based complex systems, e.g., systems of social agents, where rules are often applied in a soft, flexible way. For scarce space, the illustration has been limited to one example. For the same reason, we had to omit many interesting references to the literature and the proof of the theorem. Also, only a few short comments on the properties of graded applicability are included.

References

1. Gomolińska, A.: A comparative study of some generalized rough approximations. Fundamenta Informaticae **51(1–2)** (2002) 103–119
2. Gomolińska, A.: A graded meaning of formulas and their sets in generalized approximation spaces. In Czaja, L. (ed.) Proc. Int. Workshop on Concurrency, Specification and Programming CS&P'2003, Czarna, Poland, Warsaw University Publ. ISBN 83-88374-71-0 (2003) 157–170
3. Gomolińska, A.: A graded meaning of formulas in approximation spaces. Fundamenta Informaticae **60** (2004) 159–172
4. Pawlak, Z.: Rough Sets. Theoretical Aspects of Reasoning about Data. Kluwer Acad. Publ. Dordrecht (1991)
5. Polkowski, L., Skowron, A.: Rough mereology: A new paradigm for approximate reasoning. Int. J. Approximated Reasoning **15(4)** (1996) 333–365
6. Skowron, A., Stepaniuk, J.: Tolerance approximation spaces. Fundamenta Informaticae **27** (1996) 245–253