Game semantics and vagueness in natural language

Matthias F.J. Hofer*

Vienna University of Technology

Abstract We take up the challenge to define a procedure to systematically evaluate natural language statements involving vagueness, as is the case for “About half days are nice.”, which is quantified with a vague proportional quantifier and applied to a vague predicate in the scope. Our approach is embedded into an analytic game semantic framework, which extends Giles’s game for Łukasiewicz logic by means of a randomization operator and the Baaz-Delta operator.

Keywords game semantics; fuzzy logic; rough sets; natural language processing; vagueness; proportional quantifiers; vague predicates; multi agents

1 Introduction

The modeling of vague proportional quantifier expressions, like “about half”, “almost all”, and “at least about a third”, and of vague predicates, like “tall”, and “nice” is a great challenge taken up by many different researchers and communities, amongst which we have linguists [3,20,29], philosophers [7,32], computer scientists [1,28], and logicians [3,12,26]. There is a huge amount of literature on fuzzy quantification, summarized in the recent survey article [6], while there is also a whole monograph of Glöckner [15] about this topic. Also, recent developments in the field of mathematical fuzzy logic [4,5] contribute to the matter by addressing it using a game semantic framework, extending Giles’s game (G game) for Łukasiewicz logic by a randomization operator [11]. We here intend to further extend the existing analytic game semantic framework, while following the systematic approach of Liu and Kerre [24], who proposed to split the problem of generalized quantification into four steps in the following way\(^1\):

Type I: the quantifier as well as the scope predicate is crisp;
Type II: the quantifier is crisp, but the scope predicate is vague;
Type III: the quantifier is vague, but the scope predicate is crisp;
Type IV: the quantifier as well as the scope predicate are vague.

\(^*\) Supported by FWF projects I1897-N25 and W1255-N23

\(^1\) Crisp here means binary, hence zero or one valued.
A Type I statement may have the form “All mothers are women.”, or “There
is a man being more than 180 cm tall.”, where the quantifiers, ∀ and ∃, are crisp
in the sense that they only take one of the two classical truth values, zero or one.
The same goes for the properties, “being a mother”, “being a man/woman”, or
“being more than 180 cm tall”, which are also taken as crisp, as each object
either completely fulfills them or not. A general Type IV statement may then be
of the more complicated form “Almost all beaches in Thailand are beautiful.”,
where the quantifier “almost all” and the scope predicate “beautiful” are vague.

To be able to model vague proportional quantifiers and vague predicates
adequately, we show how we can define a projection operator, also known as
Baaz-Delta, in our game semantic setting. Then Type I quantification already
becomes a rather easy task, and we continue to show how vague predicates can be
conceived (Type II), following particular ideas from rough set theory or analytic
philosophy [27] respectively, where vague concepts depend on several crisp ones,
as well as on finitely many agents. We then define vague proportional quantifiers,
again following ideas from rough set theory, namely the one of granular levels
[37]. This goes back to Zadeh [38], and has attracted some intensified attention
during the last years [21,35]. We apply this idea in the context of proportional
quantification (Type III) and eventually combine it with our vague predicates,
to arrive at the general level of Type IV proportional quantification.

The rest of the paper is organized as follows: Section 2 illustrates the core
aspects of the used game semantic framework and defines an important new
operator, the so called Baaz-Delta. Section 3 to Section 6 follow exactly the hi-
erarchy of Liu and Kerre, which means each section corresponds to a respective
type of quantification, as introduced just above. Section 7 summarized the con-
tribution and describes what has to be done to further augment the presented
material.
with $n \in \mathbb{N}$ and risk value assignment $\langle \rangle_I$, i.e. for every atomic formula $A$ we let $\langle A \rangle_I$ be its risk and have $v_I(A) = 1 - \langle A \rangle_I$. Hence, the final risk, from $P$’s perspective, of a game is computed as:

$$\langle A_1, \ldots, A_n | B_1, \ldots, B_m \rangle = \sum_{1 \leq i \leq m} \langle B_i \rangle_I - \sum_{1 \leq j \leq n} \langle A_j \rangle_I \quad (1)$$

Note that the truth function corresponding to the previous game rule matches the well known truth function of Lukasiewicz implication:

$$v_I(F \rightarrow G) = \min(1, 1 - v_I(F) + v_I(G))$$

The negation of a formula $F$, defined as $(F \rightarrow \bot)$, introduces role switch of the players, and the following rule for strong conjunction the principle of limited liability [4]:

**Game Rule 2** ($R_{\&}$) If $P$ asserts $F \& G$, then, if $O$ attacks, $P$ has to either assert $F$ as well as $G$, or else $\bot$.

Again, the truth function turns out to correspond to the known one,

$$v_I(F \& G) = \max(0, v_I(F) + v_I(G) - 1),$$

since, as well as $O$ need not attack a by $P$ asserted formula, also $P$ can hedge her/his loss of asserting more than one formula, due to the definition of the game rule. Here, with the strong conjunction rule it is stated explicitly that $P$ can assert $\bot$ instead of both $F$ and $G$, in case they are both wrong. It be understood that this so called principle of limited liability always be in place throughout this paper, although it sometimes remains implicit [4]. We can give a characterization of strong Lukasiewicz logic via $G$-games as follows:

**Theorem 1** ([10]). For every atomic formula $A$ let $\langle A \rangle$ be its risk and let $I$ be the $L$-interpretation given by $v_I(A) = 1 - \langle A \rangle$. Then, if both, $P$ and $O$, play rationally, any game starting in state $[F]$ will end in a state where $P$’s final risk is $1 - v_I(F)$.

To generalize the game for Lukasiewicz logic $L$, the authors of [12] introduce the following randomizing quantifier rule in contrast to the ones for the existential and universal quantifiers, where either the proponent or the opponent can choose a witnessing constant:

**Game Rule 3** ($R_{\Pi}$) If $P$ asserts $\Pi x F(x)$ then $P$ has to assert $F(c)$ for a randomly picked $c$.

\[ \]

\[ ^2 \text{For simplicity we identify objects from the domain with unique constants from a set called $U$ again.} \]
It’s truth function turns out to be the following:

\[ v_I(\Pi x F(x)) = \frac{\sum_{c \in U} v_I(F(c))}{|U|} = \text{Prop}_x F(x) \]

We now also define the so called Baaz-Delta \([5]\), which leaves the truth value of \( F \) unchanged in case it is 1, and projects all others to 0, in the following way\(^3\):

**Game Rule 4 \((R_\Delta)\)** If \( P \) asserts \( \Delta F \) then, \( O \) attacks by choosing \( j \in \mathbb{N} \), obliging \( P \) to assert \( F_j \).

**Theorem 2.** A \( L(R_\Delta) \)-sentence \( \Delta F \), for a \( L \) formula \( F \) is evaluated to \( v_M(\Delta F) = x \) in an interpretation \( I \) iff every \( G \)-game for \( \Delta F \) augmented by rule \((R_\Delta)\) is \((1-x)\)-valued for \( P \) under risk value assignment \( \langle \cdot \rangle_I \), i.e. \( P \) has a strategy to make his/her final risk at most \( 1 - x \), and \( O \) has a strategy to make \( P \)'s final risk at least \( 1 - x \).

Now we define the following Nabla operator, which, complementary to the Baaz-Delta, leaves the truth value of \( F \) unchanged in case it is 0, and projects all others to 1:

**Definition 1.** \( \nabla F := \neg \Delta \neg F \)

The Delta and Nabla Operator can be seen as tools to create discontinuous truth functions, as they represent the limiting case of what can be expressed in ordinary Lukasiewicz logic, where all truth functions are continuous.

3 Type I - the most simple case

As mentioned in the introduction we follow the systematics of Liu and Kerre, and start with defining crisp quantifiers which may only be applied to crisp predicates. We define them in terms of our game semantic connectives, which we have introduced in section 2. The resulting quantifiers behave as expected, in the sense that they are merely a reformulation of respective quantifiers as they are well known from the literature \([15]\). The main difference here, from the more computational approaches there, is that we embed quantifiers into an analytic framework, instead of just giving the computation rule without relating them to a logical theory. The proportional quantifiers which we define in the present section evaluate for a given number \( k \in [0,1] \) to true if and only if the proportion of elements of the domain that fulfill the scope predicate is \( k \), and to false otherwise, and we denote them by \( Q^{[=k]} \):

**Definition 2.** \( \forall k \in [0,1] \) we define \( Q^{[=k]} x \hat{F}(x) = \Delta(\Pi x \hat{F}(x) \leftrightarrow \bar{k}) \)

\( \hat{F} \) denotes a crisp formula, i.e. one composed of crisp (zero or one valued) predicates only. For formulas \( F \) and \( G \), \( F \leftrightarrow G \) is defined as \( (F \rightarrow G) \& (G \rightarrow F) \), and for \( k \in [0,1] \), \( \bar{k} \) denotes the truth constant with the value \( k \).

\(^3\) For a formula \( F \), \( F^j \) means \( F \& \ldots \& F \), \( j \) times.
Note that we can express the universal and existential quantifier as follows:

\[ \forall x F(x) = \Delta (\Pi x F(x)) \]
\[ \exists x F(x) = \nabla (\Pi x F(x)) \]

So far, we can deal with statements like “Exactly half (of the elements of the domain) are students.”, or “Exactly a third (of the elements of the domain) are mothers.”, as being a student, or mother respectively, is supposed to be a crisp property fully possessed or not by each element of the domain. The next section is devoted to augmenting the framework to predicates which need not be of this kind.

4 Type II - adding vague predicates

Vague predicates, such as “tall”, “friendly”, or “nice” are notoriously difficult to model, since it is not objectively determinable what makes competent speakers judge objects to possess such a property. There is a lot discussion to be found in the literature, coming from the linguists side [19] as well as from the side of fuzzy logicians, who also give models for vague predicates [23]. We here relate our approach to the so called rough set theory [31], and show how we can model vague predicates within our game semantic framework, following the convincing ideas of rough sets theorists [27,36], as well as those of linguists and philosophers [19,32], or computer scientists [14], still staying in a neat and uniform logical framework. We understand vague predicates as dependent on a finite number of crisp predicates in the following way:

**Definition 3.** A vague predicate \( P^v \) comes with a set \( \{P^v_1, \ldots, P^v_{m^v}\} \) of crisp predicates, where \( m^v \in \mathbb{N} \), such that for all \( c \in U \) it holds:

\[ v_I(P^v(c)) = \frac{r}{m^v} \text{ iff } v_I(P^v_i(c)) = 1 \text{ for } r \text{ indices } i \in \{1, \ldots, m^v\} \]

As game rule this may be formalized as follows:

**Game Rule 5 (R^v\_P.)** If \( P \) asserts \( P^v_i(c) \), then, if \( O \) attacks, \( i \in \{1, \ldots, m^v\} \) gets picked randomly, and then \( P \) has to assert \( P^v_i(c) \).

The corresponding truth function can be determined as:

\[ v_I(P^v(c)) = \frac{\sum_{i=1}^{m^v} v_I(P^v_i(c))}{m^v} \]

For simplicity we here assume the influence of each crisp predicate to be the same, hence describe an unweighted scenario. The weighted case can be achieved through allowing for multiple occurrences of the same indices in the sample space of the above rule.
Hence, a vague atomic formula evaluates to (completely) true if and only if all crisp atomic formulas relevant to the vague one evaluate to true. In this sense we can talk about lower approximations in the sense of rough set theory. Also, a vague atomic formula evaluates to a truth value greater than zero, if and only if there is a crisp atomic formula relevant to it, which is true. This corresponds to the idea of an upper approximation with regard to rough set theory. Only in case all relevant crisp atomic formulas evaluate to false, the respective vague atomic formula also evaluates to false. We use the following example, to illustrate how this translates into the language of the theory:

4.1 Example

We define the vague predicate “tasty” through the following three crisp predicates: “salami”, “mushrooms”, and “garlic”. Let two universes $U_1, U_2$ consist of 100 meals each. For the first universe we have 50 portions of pasta, and 50 pizzas, all of which have salami, mushrooms, and garlic on top (hence are tasty!). For the second universe we have 50 portions of pasta, 17 pizzas with only salami and garlic on top, and 33 with only mushrooms on top. Now consider the following statement:

“Exactly 50% of all meals are (fully) tasty pizzas.”

With respect to $U_1$ this statement is true, since the defining properties of “tasty” are (completely) fulfilled for exactly half of the elements of the domain. Formally, we express this statement as:

$$\Delta((\Pi x (\Delta P_{tasty}(x) \land pizza(x)) \leftrightarrow \top))$$

With respect to $U_2$ we still have exactly 50% meals with a tasty-value greater than zero, hence we can still evaluate the following statement as true:

“Exactly 50% of all meals are kind of tasty pizzas.”

This “kind of” - hedge gets formally represented in the following way:

$$\Delta((\Pi x (\nabla P_{tasty}(x) \land pizza(x)) \leftrightarrow \top))$$

We are now able to express a much greater range of statements, namely also those that involve vague predicates. Also we have described the vague linguistic hedge “kind of”, which in natural language expresses a certain kind of uncertainty of the speaker regarding the definition of vague predicates present in some utterance. This becomes possible through the nabla operator $\nabla$, defined in section 2, using the delta operator $\Delta$. 
4.2 Definitions and a multi agent extension

Definition 4. crisp quantifier, vague predicates/formulas

\[ Q^{[=k]}_x(fully \cdot F(x)) := \Delta(\Pi x(\Delta F(x)) \leftrightarrow \bar{k}) \]

\[ Q^{[=k]}_x(kind \ of \ F(x)) := \Delta(\Pi x(\nabla F(x)) \leftrightarrow \bar{k}) \]

We can also think of the following generalization of the just described approach to vague predicates. Instead of taking one particular set of crisp predicates, relevant to describe the meaning of a vague predicate, we may think of many different such, reflecting the fact that different agents (competent speakers) may have different reasons to judge an object as “nice”, or the like. This more general setting can be achieved through simply changing the respective game rule in the following way, after we fixed some notation, to be able to refer to different agents:

Definition 5. An agent \(a_j, j \in \{1, \ldots, m_a\}\) comes with a set of crisp properties \(\{P^{v,a_j}_1, \ldots, P^{v,a_j}_{m_{v,a_j}}\}\) for any vague predicate \(P_v, m_{v,a_j} \in \mathbb{N}\), such that for all \(c \in U\) it holds:

\[ v_I(P^{v,a_j}_v(c)) = r \quad \text{iff} \quad v_I(P^{v,a_j}_{m_{v,a_j}}(c)) = 1 \quad \text{for r indices} \ i \in \{1, \ldots, m_{v,a_j}\} \]

The general game rule for vague predicates now is:

Game Rule 6 \( (R^2_{P_v}) \) If \( P \) asserts \( P_v(c) \), then, if \( O \) attacks, \( j \in \{1, \ldots, m_a\} \) gets picked randomly, followed by a random pick of \( i \in \{1, \ldots, m_{v,a_j}\} \), and then \( P \) has to assert \( P^{v,a_j}_i(c) \).

Similarly to the first rule for vague predicates, we can determine the truth function as:

\[ v_I(P_v(c)) = \frac{\sum_{j=1}^{m_a} \sum_{i=1}^{m_{v,a_j}} v_I(P^{v,a_j}_i(c))}{m_am_{v,a_j}} \]

Again, the weighted case, where the influence of crisp predicates to a vague one, or of the different agents to the evaluation respectively, is not uniform, can easily be achieved, as described above.

5 Type III - vague proportional quantifiers, crisp scope

Modeling vague proportional quantifiers is, as argued by Liu and Kerre [24], as well as by Glöckner and others [12,15], best performed in a step by step manner, first focusing on the quantifiers and only later showing how they can then be applied to formulas which may involve vague atomic subformulas. Following this approach, we now develop a way of evaluating vague proportional quantifiers, again being inspired by rough set theory, while staying in our game semantic
framework. Hence, in this section, we assume the scope formulas of quantifiers to be crisp again, and only combine it all together in the next section.

An idea going back to Zadeh [38], being carried out much in recent years, is granular computing [2]. The idea is to attach a level of granularity to certain scenarios, hence making objects indistinguishable with respect to some (equivalence) relation. This idea, applied to vague concepts [37], is here extended to vague quantification in an seemingly obvious way. We apply the simple idea of tolerance intervals around some crisp value. Take the quantifier expression “about half”, which can be associated to several such, e.g. [37, 5%, 62, 5%], [45%, 55%], [49.5%, 50.5%], or others. We can partition the unit interval in many different ways, where each partitioning then corresponds to some level of granularity. Having several such levels, we can talk about a granular hierarchy [21, 35, 37]. However, following everyday experience, we propose the following systematic refinement procedure:

- 3-partitioning: This can be associated to the common classification into three categories, e.g. “small”, “medium”, and “large”
  - partitioning intervals: \([0, \frac{1}{3}), [\frac{1}{3}, \frac{2}{3}), [\frac{2}{3}, 1]\]
- 5-partitioning: Five categories, say “tiny”, “small”, “medium”, “large”, “huge”
  - partitioning intervals: \([0, \frac{1}{5}), \ldots, [\frac{4}{5}, 1]\]
- 7-partitioning: E.g. “almost none”, “few”, “several”, “about half”, “most”, “many”, “almost all”
  - partitioning intervals: \([0, \frac{1}{7}), \ldots, [\frac{6}{7}, 1]\]
- tenner-partitioning: (About) 0%, 10%, 20%, \ldots, 90%, 100%
  - partitioning intervals: \([0, \frac{1}{10}), [\frac{1}{10}, \frac{3}{10}), \ldots, [\frac{17}{20}, \frac{19}{20}), [\frac{19}{20}, 1]\]
- fiver-partitioning: (About) 0%, 5%, 10%, 15%, \ldots, 90%, 95%, 100%
  - partitioning intervals: \([0, \frac{1}{5}), [\frac{1}{5}, \frac{3}{5}), \ldots, [\frac{37}{40}, \frac{39}{40}), [\frac{39}{40}, 1]\]
- oner-partitioning: (About) 0%, 1%, 2%, 3%, \ldots, 98%, 99%, 100%
  - partitioning intervals: \([0, \frac{1}{100}), [\frac{1}{100}, \frac{3}{100}), \ldots, [\frac{197}{200}, \frac{199}{200}), [\frac{199}{200}, 1]\]
- decimal place-partitioning: (About) 0%, 0.1%, 0.2%, 0.3%, \ldots, 99.8%, 99.9%, 100%
  - partitioning intervals: \([0, \frac{1}{2000}), [\frac{1}{2000}, \frac{3}{2000}), \ldots, [\frac{1997}{2000}, \frac{1999}{2000}), [\frac{1999}{2000}, 1]\]
All these classifications are, of course, somehow freely defined, and may hence be changed accordingly. To describe the semantics of some vague proportional quantifier $Q$, we need to fix a finite number of such levels of granularity, say $GL_1, \ldots, GL_{m_Q}$, $m_Q \in \mathbb{N}$, with respect to which we can evaluate respective statements. In the present case, for statements “about half (of the domain elements) fulfill property $\hat{F}$” we then have acceptance intervals\(^4\) as follows:

- $[\frac{1}{3}, \frac{2}{3})$ (3-partitioning)
- $[\frac{2}{5}, \frac{3}{5})$ (5-partitioning)
- $[\frac{3}{7}, \frac{4}{7})$ (7-partitioning)
- $[\frac{45}{100}, \frac{55}{100})$ (tenner-partitioning)
- $[\frac{47}{100}, \frac{52}{100})$ (fiver-partitioning)
- $[\frac{49}{100}, \frac{50}{100})$ (oner-partitioning)
- $[\frac{49}{100}, \frac{95}{100})$ (decimal place-partitioning)

**Definition 6.** A granularity level $GL$ corresponds to a partitioning of the real unit interval $[0,1]$ into finitely many disjoint intervals $z_1, \ldots, z_{m_GL}$, such that $\bigcup_{i=1}^{m_GL} z_i = [0,1]$.

**Definition 7.** A vague proportional quantifier $Q$ comes with a set $\{GL_1, \ldots, GL_{m_Q}\}$ of granularity levels, where $m_Q \in \mathbb{N}$, such that each such level has an unique acceptance interval for $Q$, i.e. for all granularity levels $GL_i$, $i \in \{1, \ldots, m_Q\}$ there is exactly one interval $z_{Q,GL_i}$ of the corresponding partitioning such that it holds:

$$v_I(Q_{GL_i} x \hat{F}(x)) = (\Delta(\Pi x \hat{F}(x) \rightarrow z^+_{Q,GL_i})) \& (\Delta(z^-_{Q,GL_i} \rightarrow \Pi x \hat{F}(x)))$$

If for some fixed vague proportional quantifier $Q$ and granularity level $GL$, $z_{Q,GL}$ is the acceptance interval for $Q$, we set $z^+_{Q,GL}, z^-_{Q,GL}$ the upper, and lower, boundary of the interval. $Q_{GL_i}$ denotes the quantifier $Q$ restricted to one particular granularity level $GL_i$.

As a game rule, we can express this definition in the following way:

**Game Rule 7 ($R_{GL,III}^{[\approx k]}$)**

If $P$ asserts $Q x \hat{F}(x)$, then, if $O$ attacks, $i \in \{1, \ldots, m_Q\}$ gets chosen randomly, and then $P$ has to assert $(\Delta(\Pi x \hat{F}(x) \rightarrow z^+_{Q,GL_i})) \& (\Delta(z^-_{Q,GL_i} \rightarrow \Pi x \hat{F}(x)))$.

The corresponding truth function is the following:

$$v_I(Q x \hat{F}(x)) = \sum_{i=1}^{m_Q} v_I(\Delta(\Pi x \hat{F}(x) \rightarrow z^+_{Q,GL_i})) \& (\Delta(z^-_{Q,GL_i} \rightarrow \Pi x \hat{F}(x)))$$

The range of statements we can express now includes all those that start with quantifier expression like “about halt”, “about a third”, or “almost all”. By means of combining them, using the logical $\lor$ connective, we can also express

---

\(^4\) I.e. the statement is true if $\text{Prop}_i \hat{F}(x)$ is an element of this acceptance interval.
statements like “at least about half”, or “at most about a third”, by simply linking respective statements together. This allows for an even wider range of quantification than Type II, and hence augments the applicability enormously, as many natural language statements in real life are of this form. In a last remaining step we combine Type II and Type III quantifiers, and end up with the final Type IV quantifiers, which are able to systematically evaluate statements that are vaguely quantified and have vague scope formulas at the same time.

6 Type IV - combining it all together

As our game semantic framework analytically decomposes formulas down to atomic subformulas, quantification, with respect to formulas potentially build by vague atomic subformulas, is rather straightforward, as we only need to combine game rule 5 (or 6) with game rule 7 of the present paper. Hence we may restate game rule 7 with the only adjustment of dropping the hat of the formerly crisp scope formula $\hat{F}$.

**Game Rule 8** ($R_{GL,IV}^{[\approx k]}$)

If $P$ asserts $QxF(x)$, then, if $O$ attacks, $i \in \{1, \ldots, m_Q\}$ gets chosen randomly, and then $P$ has to assert $(\Delta(\Pi x F(x) \rightarrow \tilde{z}_{Q,GL_i}^+)) \& (\Delta(z_{Q,GL_i}^- \rightarrow \Pi x F(x)))$.

If we want to determine the truth function, we find the following:

$$v_I(QxF(x)) = \frac{\sum_{i=1}^{m_Q} v_I(\Delta(\Pi x F(x) \rightarrow \tilde{z}_{Q,GL_i}^+)) \& (\Delta(z_{Q,GL_i}^- \rightarrow \Pi x F(x)))}{m_Q}$$

We are now able to evaluate all sort of vague statements, like “At most about a third (of all domain elements) are nice”, or “Almost all (of the domain elements) are friendly or tall”, as long as the involved quantifiers and predicates are well defined. This allows for a great deal of flexibility, and particularly embeds evaluations into a neat logical machinery, which is fully linked to recent developments in the field of mathematical fuzzy logic [4].

7 Conclusion and Outlook

We defined a way to systematically evaluate natural language statement within an analytic game semantic framework. Our approach follows the hierarchy prescribed by Liu and Kerre and focuses first on crisp quantification with respect to crisp scope formulas. In a next step we described how vague predicates can be defined and show how we can quantify over vague formulas. We then introduce granular levels and define vague proportional quantification based on this notion, followed by a final step, where we combine all together.

The presented procedure may be extended into at least two important directions. One of them is the multi arity of quantifiers, as usually natural language statements are at least binary, as is the case with “Almost all children are
friendly.”. It has been pointed out [15] that even higher arities may be of importance, and hence we focus on this aspect in ongoing work. Another important augmentation is introducing non-proportional quantifiers, which may depend on intensional matters, like “many” and “few” [22,25]. These are closely linked to modal logics, but may also be integrated into the present game semantic setting as we will show in future work.

References