Truth Value Intervals, Bets, and Dialogue Games

Christian G. Fermüller*

Fuzzy logics in Zadeh's 'narrow sense' (Zadeh, 1988), i.e., truth functional logics referring to the real closed unit interval $[0,1]$ as set of truth values, are often motivated as logics for 'reasoning with imprecise notions and propositions' (see, e.g., (Hájek, 1998)). However the relation between these logics and theories of vagueness, as discussed in a prolific discourse in analytic philosophy (Keefe & Rosanna, 2000), (Keefe & Smith, 1987), (Shapiro, 2006) is highly contentious. We will not directly engage in this debate here but rather pick out so-called interval based fuzzy logics as an instructive example to study

1. how such logics are usually motivated informally,
2. what problems may arise from these motivations, and
3. how betting and dialogue games may be used to analyze these logics with respect to more general principles and models of reasoning.

The main technical result[1] of this work consists in a characterization of an important interval based logic, considered, e.g., in (Esteva, Garcia-Calvés, & Godo, 1994), in terms of a dialogue *cum* betting game, that generalizes Robin Giles's game based characterization of Łukasiewicz logic (Giles, 1974), (Giles, 1977). However, our aim is to address foundational problems with certain models of reasoning with imprecise information. We hope to show that the traditional paradigm of dialogue games as a possible foundation of logic (going back, at least, to (Lorenzen, 1960)) combined with bets as 'test cases' for rationality in the face of uncertainty might help to sort out some of the relevant conceptual issues. This is intended to highlight a particular meeting place of logic, games, and decision theory at the foundation of a field often called 'approximate reasoning' (see, e.g., (Zadeh, 1975)).

*This work is supported by FWF project I143–G15.

[1] Due to limited space, we state propositions without proofs.

1 T-norm based fuzzy logics and bilattices

Petr Hájek, in the preface of his influential monograph on mathematical fuzzy logic (Hájek, 1998) asserts:

> The aim is to show that fuzzy logic as a logic of imprecise (vague) propositions does have well developed formal foundations and that most things usually named 'fuzzy inference' can be naturally understood as logical deduction. (Hájek, 1998, p. viii)

As the qualification 'vague', added in parenthesis to 'imprecise', betrays, some terminological and, arguably, also conceptional problems may be located already in this introductory statement. These problems relate to the fact that fuzzy logic is often subsumed under the general headings of 'uncertainty' and 'approximate reasoning'. In any case, Hajek goes on to introduce a family of formal logics, based on the following design choices (compare also (Hájek, 2002)):

1. The set of degrees of truth (truth values) is represented by the real unit interval $[0,1]$. The usual order relation $\leq$ models comparison of truth degrees; 0 represents absolute falsity, and 1 represents absolute truth.

2. The truth value of a compound statement shall only depend on the truth values of its subformulas. In other words: the logics are truth functional.

3. The truth function for (strong) conjunction (&) should be a continuous, commutative, associative, and monotonically increasing function $* : [0,1]^2 \to [0,1]$, where $0 * x = 0$ and $1 * x = x$. In other words: $*$ is a continuous t-norm.

4. The residuum $\Rightarrow_*$ of the t-norm $*$ — i.e., the unique function $\Rightarrow_*$: $[0,1]^2 \to [0,1]$ satisfying $x \Rightarrow_* y = \sup\{z \mid x * z \leq y\}$ — serves as the truth function for implication.

5. The truth function for negation is $\lambda x[x \Rightarrow_* 0]$.

Probably the best known logic arising in this way is Łukasiewicz logic **L** (Łukasiewicz, 1920), where the t-norm $*_{\mathbf{L}}$ that serves as truth function for & is defined as $x *_{\mathbf{L}} y = \max(0, x + y - 1)$. Its residuum $\Rightarrow_{\mathbf{L}}$ is given by $x \Rightarrow_{\mathbf{L}} y = \min(1, 1 - x + y)$. A popular alternative choice for conjunction takes the *minimum* as its truth function. Besides 'strong conjunction' (&), also this latter 'weak (min) conjunction' ($\wedge$) can be defined in all t-norm based logics by $A \wedge B \stackrel{\text{def}}{=} A\&(A \to B)$. Maximum as truth function for disjunction ($\vee$) is always definable from $*$ and $\Rightarrow_*$, too.

Other important logics, like Gödel logic **G**, and Product logic **P**, can be obtained in the same way, but we will confine attention to **L**, here. At this point we like to mention that a rich, deep, and still growing subfield of mathematical logic, documented in hundreds of papers and a number of books (beyond (Hájek, 1998)) is triggered by this approach. Consequently it became evident that degree based fuzzy logics are neither a 'poor man's substitute for probabilistic reasoning' nor a trivial generalization of finite-valued logics.

A number of researchers have pointed out that, while modelling degrees of truth by values in $[0,1]$ might be a justifiable choice in principle, it is hardly realistic to assume that there are procedures that allow us to assign concrete values to concrete (interpreted) atomic propositions in a coherent and principled manner. While this problem might be ignored as long as we are only interested in an abstract characterization of logical consequence in contexts of graded truth, it is deemed desirable to refine the model by incorporating 'imprecision due to possible incompleteness of the available information' (Esteva et al., 1994) about truth values. Accordingly, it is suggested to replace single values $x \in [0,1]$ by whole *intervals* $[a,b] \subseteq [0,1]$ of truth values as the basic semantic unit assigned to propositions. The 'natural truth ordering' $\leq$ can be generalized to intervals in different ways. Following (Esteva et al., 1994) we arrive at these definitions:

Weak truth ordering: $[a_1,b_1] \leq^* [a_2,b_2]$ iff $a_1 \leq a_2$ and $b_1 \leq b_2$

Strong truth ordering: $[a_1,b_1] \prec [a_2,b_2]$ iff $b_1 \leq a_2$ or $[a_1,b_1] = [a_2,b_2]$

On the other hand, set inclusion ($\subseteq$) is called *imprecision ordering* in this context. The set of closed subintervals $\mathrm{Int}_{[0,1]}$ of $[0,1]$ is augmented by the empty interval $\varnothing$ to yield so-called *enriched bilattice* structures $\langle \mathrm{Int}_{[0,1]}, \leq^*, 0, 1, \varnothing, L, N^* \rangle$ as well as $\langle \mathrm{Int}_{[0,1]}, \prec, 0, 1, \varnothing, L, N^* \rangle$, where L is the standard lattice on $[0,1]$, with minimum and maximum as operators, and N^* is the extension of the negation operator N to intervals; in our particular case $N^*([a,b]) = [1-b, 1-a]$ and $N^*(\varnothing) = \varnothing$.

Quite a number of papers have been devoted to the study of logics based on such interval generated bilattices. Let us just mention that the Ghent school of Kerre, Deschrijver, Cornelis, and colleagues has produced an impressive amount of work on interval bilattice based logics (see, e.g., (Cornelis, Deschrijver, & Kerre, 2006)).

While it is straightforward to generalize both types of conjunction (t-norm and minimum) as well as disjunction (maximum) from $[0,1]$ to $\mathrm{Int}_{[0,1]}$ by applying the operators point-wise, it seems less clear how the 'right' generalization of the truth function for implication should look like. In (Cornelis, Arieli, Deschrijver, & Kerre, 2007), (Cornelis, Deschrijver, & Kerre, 2004) $[a,b] \Rightarrow^*_C [c,d] \stackrel{\text{def}}{=} [\min(a \Rightarrow c, b \Rightarrow d), b \Rightarrow d]$ is studied, but in (Esteva et al., 1994) the authors suggest $[a,b] \Rightarrow^*_E [c,d] \stackrel{\text{def}}{=} [b \Rightarrow c, a \Rightarrow d]$. As has been pointed out in (Hájek, n.d.) there seems to be a kind of trade off involved here. While $\Rightarrow^*_C$ preserves a lot of algebraic structure — in particular it yields a *residuated* lattice which contains the underlying lattice over $[0,1]$ as a substructure — the function $\Rightarrow^*_E$ is not a residuum, but leads to the following desirable preservation property that is missing for $\Rightarrow^*_C$. If $\mathcal{M}_2$ is a precisiation of $\mathcal{M}_1$ (meaning: for each propositional variable p, $\mathcal{M}_2$ assigns a subinterval of the interval assigned to p by $\mathcal{M}_1$), than any formula satisfied by $\mathcal{M}_1$ is also satisfied by $\mathcal{M}_2$.[2] Below, we will try to show that a game based approach might justify the preference of $\Rightarrow^*_E$ over $\Rightarrow^*_C$ against a background that takes the challenge of deriving formal semantics from first principles about logical reasoning more seriously than the mentioned literature on 'interval logics'.

2 Worries about truth functionality

It is interesting to note that both, (Esteva et al., 1994) and (Cornelis et al., 2007), (Cornelis et al., 2004), refer to Ginsberg (Ginsberg, 1988), who explicitly introduced bilattices following ideas of (Belnap, 1977). Most prominently Ginsberg considers $\mathcal{B} = \langle\{0, \top, \bot, 1\}, \leq_t, \leq_k, \neg\rangle$ as endowed with the following intended meaning:

- 0 and 1 represent (classical) falsity and truth, respectively, $\top$ represents 'inconsistent information' and $\bot$ represents 'no information'. The idea here is that truth values are assigned after receiving relevant information from different sources. Accordingly $\top$ is identified with the information set $\{0,1\}$, $\bot$ with $\emptyset$ and the classical truth values with their singleton sets.
- $\leq_t$, defined by $0 \leq_t \top/\bot \leq 1$, is the '*truth ordering*'.
- $\leq_k$, defined by $\bot \leq_t 0/1 \leq 1$, is the '*knowledge ordering*'.
- Negation is defined by $\neg(0) = 1$, $\neg(1) = 0$, $\neg(\top) = \top$, $\neg(\bot) = \bot$.

While the four 'truth values' of $\mathcal{B}$ may justifiably be understood to represent different states of knowledge about propositions, it is very questionable to try to define corresponding 'truth functions' for connectives other than negation. Indeed, it is surprising to see how many authors[3] followed (Belnap, 1977) in defending a four valued, truth functional logic based on $\mathcal{B}$. It should be clear that, in the underlying classical setting that is taken for granted by Belnap, the formula $A \wedge \neg A$ can only be false (0), independently of the

[2] Here, a formula is defined to be satisfied if it evaluates to the degenerate interval $[1,1]$.
[3] Dozens of papers have been written about Belnap's 4–valued logic.

kind of information, if any, we have about the truth of A. On the other hand, if we neither have information about A nor about B, then $B \wedge \neg A$ could be true as well as false, and therefore $\bot$ should be assigned not only to A, B, and $\neg A$, but also to $B \wedge \neg A$ (in contrast to $A \wedge \neg A$). This simple argument illustrates that knowledge does not propagate compositionally — a well known fact that, however, has been ignored repeatedly in the literature. (For a recent, forceful reminder on the incoherency of the intended semantics for Belnap's logic we refer to (Dubois, n.d.).)

In our context this warning about the limits of truth functionality is relevant at two separate levels. First, it implies that 'degrees of truth' for compound statements cannot be interpreted *epistemically* while upholding truth functionality. Indeed, most fuzzy logicians correctly emphasize that the concept of degrees of truth is *orthogonal* to the concept of degrees of belief. While truth functions for degrees of truth can be motivated and justified in various ways — below we will review a game based approach — degrees of belief simply don't propagate compositionally and call for other types of logical models (e.g., 'possible worlds'). Second, concerning the concept of *intervals* of degrees of truth, one should recognize that it is incoherent to insist on both at the same time:

1. *truth functions* for all connectives, lifted from $[0,1]$ to $\mathrm{Int}_{[0,1]}$, and
2. the interpretation of an interval $[a,b] \subseteq [0,1]$ assigned to a (compound) proposition F as representing a situation where our best *knowledge* about the (definite) degree of truth $d \in [0,1]$ of F is that $a \leq d \leq b$.

Given the mathematical elegance of 1, that results, among other desirable properties, in a low computational complexity of the involved logics[4], one should look for alternatives to 2. Godo and Esteva[5] have pointed out that, if we insist on 2 just for *atomic* propositions, then at least we can assert that the corresponding 'real', but unknown truth degree of any *composite* proposition F cannot lie outside the interval assigned to F according to the truth functions considered in (Esteva et al., 1994) (described above). However, these bounds are not optimal, in general. As we will see in Section 5, taking clues from Giles's game based semantic for **L** (Giles, 1974), (Giles, 1977), a tighter characterization emerges if we dismiss the idea that intervals represent sets of unknown, but *definite* truth degrees.

[4] It is easy to see that coNP-completeness of testing validity for **L** (and many other t-norm based logics) carries over to the interval based logics described above.
[5] Private communication.

3 Revisiting Giles's game for Ł

Giles's analysis (Giles, 1974), (Giles, 1977) of approximate reasoning originally referred to the phenomenon of 'dispersion' in the context of physical theories. Later (Giles, 1976) explicitly applied the same concept to the problem of providing 'tangible meanings' to (composite) fuzzy propositions.[6] For this purpose he introduces a game that consists of two independent components:

3.1 Betting for positive results of experiments.

Two players — say: *me* and *you* — agree to pay 1€ to the opponent player for every false statement they assert. By $[p_1, \dots, p_m \| q_1, \dots, q_n]$ we denote an *elementary state* of the game, where I assert each of the q_i in the multiset $\{q_1, \dots, q_n\}$ of atomic statements (represented by propositional variables), and you assert each atomic statement $p_i \in \{p_1, \dots, p_m\}$.

Each propositional variable q refers to an experiment E_q with binary (yes/no) result. The statement q can be read as 'E_q yields a positive result'. Things get interesting as the experiments may show dispersion; i.e., the same experiment may yield different results when repeated. However, the results are not completely arbitrary: for every run of the game, a fixed *risk value* $\langle q \rangle^r \in [0, 1]$ is associated with q, denoting the probability that E_q yields a negative result. For the special atomic formula $\bot$ (*falsum*) we define $\langle \bot \rangle^r = 1$. The risk associated with a multiset $\{p_1, \dots, p_m\}$ of atomic formulas is defined as $\langle p_1, \dots, p_m \rangle^r = \sum_{i=1}^{m} \langle p_i \rangle^r$. It specifies the expected amount of money (in €) that has to be paid according to the above agreement. The risk $\langle \rangle^r$ associated with the empty multiset is 0. The risk associated with an elementary state $[p_1, \dots, p_m \| q_1, \dots, q_n]$ is calculated from my point of view. Therefore the condition $\langle p_1, \dots, p_m \rangle^r \geq \langle q_1, \dots, q_n \rangle^r$ expresses that I do not expect (in the probability theoretic sense) any loss (but possibly some gain) when we bet on the truth of the involved atomic statements as stipulated above.

[6] E.g., Giles suggests to specify the semantics of the fuzzy predicate 'breakable' by assigning an experiment like 'dropping the relevant object from a certain height to see if it breaks'. The expected dispersiveness of such an experiment represent the 'fuzziness' of the corresponding predicate. An arguably even better example of a dispersive experiment in the intended context might consist in asking an arbitrarily chosen competent speaker for a yes/no answer to questions like 'Is Chris tall?' or 'Is Shakira famous?' for which truth may cogently be taken as a matter of degree.

3.2 A dialogue game for the reduction of composite formulas.

Giles follows ideas of Paul Lorenzen that date back already to the 1950s (see, e.g., (Lorenzen, 1960)) and constrains the meaning of logical connectives by reference to rules of a dialogue game that proceeds by systematically reducing arguments about composite formulas to arguments about their subformulas.

The dialogue rule for implication can be stated as follows:

($R_{\rightarrow}$) If I assert $A \rightarrow B$ then, whenever you choose to attack this statement by asserting A, I have to assert also B. (And vice versa, i.e., for the roles of me and you switched.)

This rule reflects the idea that the meaning of implication is specified by the principle that an assertion of 'if A, then B' ($A \rightarrow B$) obliges one to assert B, if A is granted.[7]

In the following we only state the rules for 'me'; the rules for 'you' are perfectly symmetric. For disjunction we stipulate:

$R_{\vee}$ If I assert $A_1 \vee A_2$ then I have to assert also A_i for some $i \in \{1, 2\}$ that I myself may choose.

The rule for (weak) conjunction is dual:

$R_{\wedge}$ If I assert $A_1 \wedge A_2$ then I have to assert also A_i for any $i \in \{1, 2\}$ that you may choose.

One might ask whether asserting a conjunction shouldn't oblige one to assert *both* disjuncts. Indeed, for strong conjunction[8] we have

$R_{\&}$ If I assert $A_1 \& A_2$ then I have to assert either both A_1 and A_2, or just $\perp$.

The possibility of asserting $\perp$ instead of the attacked conjunction reflects Giles's 'principle of hedged loss': one never has to risk more than 1€ for each assertion. Asserting $\perp$ is equivalent to (certainly) paying 1€.

In contrast to dialogue games for intuitionistic logic (Lorenzen, 1960), (Felscher, 1985) or fragments of linear logic, no special regulation on the succession of moves in a dialogue is required here. Moreover, we assume that each assertion is attacked at most once: this is reflected by the removal of the occurrence of the formula F from the multiset of formulas asserted by a player, as soon as it has been attacked, or whenever the other player has indicated that she will not attack this occurrence of F during the whole

[7] Note that, since $\neg F$ is defined as $F \rightarrow \perp$, according to ($R_{\rightarrow}$) and the above definition of risk, the risk involved in asserting $\neg p$ is $1 - \langle p \rangle^r$.

[8] Giles did not consider strong conjunction. The rule is from (Fermüller & Kosik, 2006).

run of the dialogue game. Every run thus ends in finitely many steps in an elementary state $[p_1, \ldots, p_m \| q_1, \ldots, q_n]$. Given an assignment $\langle\cdot\rangle^r$ of risk values to all p_i and q_i we say that I *win* the corresponding run of the game if I do not have to expect any loss in average, i.e., if $\langle p_1, \ldots, p_m\rangle^r \geq \langle q_1, \ldots, q_n\rangle^r$.

As an almost trivial example consider the game where I initially assert $p \to q$ for some atomic formulas p and q; i.e., the initial state is $[\| p \to q]$. In response, you can either assert p in order to force me to assert q, or explicitly refuse to attack $p \to q$. In the first case, the game ends in the elementary state $[p\|q]$; in the second case it ends in state $[\|]$. If an assignment $\langle\cdot\rangle^r$ of risk values gives $\langle p\rangle^r \geq \langle q\rangle^r$, then I win, whatever move you choose to make. In other words: I have a winning strategy for $p \to q$ in all assignments of risk values where $\langle p\rangle^r \geq \langle q\rangle^r$.

Note that winning, as defined here, does not guarantee that I do not loose money. I have a winning strategy for $p \to p$, resulting either in state $[\|]$ or in state $[p\|p]$ depending on your (the opponents) choice. In the second case, although the winning condition is clearly satisfied, I will actually loose 1€, if the execution of the experiment E_p associated with *your* assertion of p happens to yield a positive result, but the execution of the same experiment associated with *my* assertion of p yields a negative result. It is only guaranteed that my *expected loss* is non-positive. ('Expectation', here, refers to standard probability theory. Under a frequentist interpretation of probability we may think of it as *average loss*, resulting from unlimited repetitions of the corresponding experiments.)

To state Giles's main result, recall that a valuation v for Łukasiewicz logic $\mathbf{L}$ is a function assigning values $\in [0,1]$ to the propositional variables and 0 to $\bot$, extended to composite formulas using the truth functions $*_\mathbf{L}$, max, min, and $\Rightarrow_\mathbf{L}$, for strong and weak conjunction, disjunction and implication, respectively.

Theorem 1 ((Giles, 1974),(Fermüller & Kosik, 2006))**.** *Each assignment $\langle\cdot\rangle^r$ of risk values to atomic formulas occurring in a formula F induces a valuation v for $\mathbf{L}$ such that $v(F) = x$ if and only if my optimal strategy for F results in an expected loss of $(1 - x)$ €.*

Corollary 1. *F is valid in $\mathbf{L}$ if and only if for all assignments of risk values to atomic formulas occurring in F I have a winning strategy for F.*

4 Playing under partial knowledge

It is important to realize that Giles's game model for reasoning about vague (i.e., here, unstable) propositions implies that *each occurrence* of the same atomic proposition in a composite statement may be evaluated differently at the level of results of associated executions of binary experiments. This feature induces truth functionality: the value for $p \vee \neg p$ is *not* the probability that experiment E_p either yields a positive or a negative result, which is 1 by definition; it rather is $1 - x$, where $x = \min(\langle p \rangle^r, 1 - \langle p \rangle^r)$ is my expected loss (in €) after having decided to bet either for a positive or for a negative result of an execution of E_p (whatever carries less risk for me).

The players only know the individual success probabilities[9] of the relevant experiments. Alternatively, one may disregard individual results of binary experiments altogether and simply identify the assigned probabilities with 'degrees of truth'. In this variant the 'pay-off' just corresponds to these truth values, and Giles's game turns into a kind of Hintikka style evaluation game for **L**.

How does all this bear on the mentioned problems of interpretation for interval based fuzzy logics? Remember that both, (Esteva et al., 1994) and (Cornelis et al., 2007, 2006, 2004) seem to suggest that an interval of truth values $[a, b]$ represents 'imprecise knowledge' about the 'real truth value' c, in the sense that only $c \in [a, b]$ is known. For the betting and dialogue game semantic this suggests that the players (or at least player 'I') now have to choose their moves in light of corresponding 'imprecise' (partial) knowledge about the success probabilities of the associated experiments. However, while this may result in an interesting variant of the Giles's game, its relation to the truth functional semantics suggested for logics based on $\mathrm{Int}_{[0,1]}$ and **L**-connectives is dubious.

The following simple example illustrates this issue. Suppose the interval $v^*(p) = [v_1^*(p), v_2^*(p)]$ assigned to the propositional variable p is $[0, 1]$, reflecting that we have no knowledge at all about the 'real truth value' of the proposition represented by p. According to the truth functions presented in Section 1, the formula $p \vee \neg p$ evaluates also to $[0, 1]$, since $v^*(\neg p) = [1 - v_2^*(p), 1 - v_1^*(p)] = [0, 1]$ and hence $v(p \vee \neg p) = [\max(0, 0), \max(1, 1)] = [0, 1]$. Sticking with the 'imprecise knowledge' interpretation, the resulting interval should reflect my knowledge about my expected loss if I play according to an optimal strategy. However, while $1 - v_2^*(p \vee \neg p) = 0$ is the correct lower bound on my expected loss after performing the relevant instance of E_p, to require that $1 - v_1^*(p \vee \neg p) = 1$ is the *best upper bound* for the loss that I have to expect when playing the game is problematic. When playing a mixed strategy that results in my assertion of either p or of $\neg p$ with equal probability, then my resulting *expected* loss is 0.5€, not 1€.

We introduce some notation to assist precise statements about the relation between the interval based semantics of (Esteva et al., 1994) and Giles's game. Let v^* be an interval assignment, i.e., an assignment of closed

[9] These might well be purely *subjective probabilities* that may differ for the two players. To prove Theorem 1 one only has to assume that I can act on assigned probabilities that determine 'my expectation' of loss.

intervals $\subseteq [0,1]$ to the propositional variables PV. Then $v^*_{\mathbf{L}}$ denotes the extension of v^* from PV to arbitrary formulas via the truth functions $\Rightarrow^*_E$ for implication and the point-wise generalizations of max, min, and $*_{\mathbf{L}}$ for disjunction, weak conjunction, and strong conjunction, respectively. Call any assignment v of reals $\in [0,1]$ *compatible* with v^* if $v(p) \in v^*(p)$ for all $p \in \mathsf{PV}$. The corresponding risk value assignment $\langle\cdot\rangle^r_v$, defined by $\langle p\rangle^r_v = 1 - v(p)$, is also called compatible with v^*.

Proposition 1. *If, given an interval assignment v^*, the formula F evaluates to $v^*_{\mathbf{L}}(F) = [a,b]$ then the following holds:*

- $*$ *For the game in Section 3, played on F: All (pure) strategies for me that are optimal with respect to some fixed risk value assignment $\langle\cdot\rangle^r_v$ compatible with v^* result in an expected loss of at most $(1-a)$ €, but at least $(1-b)$ €.*

Note that in the above statement my expected loss refers to a risk value assignment $\langle\cdot\rangle^r_v$ that is fixed before the dialogue game begins. I will play optimally with respect to this assignment. Since the corresponding *expected* pay-off is all that matters here, we technically still have a game of perfect information and therefore no generality is lost by restricting attention to pure strategies. The bounds given by $v^*_{\mathbf{L}}$ for my expected loss are not optimal in general. In other words, the inverse direction of Proposition 1 does not hold. To see this, consider again the interval assignment $v^*(p) = [0,1]$ resulting in $v^*_{\mathbf{L}}(p \vee \neg p) = [0,1]$. Obviously, I cannot loose more than 1€, even if I play badly, but my *expected* loss under any fixed risk value assignment $\langle\cdot\rangle^r_v$ is never greater than 0.5€ if I play *optimally* with respect to $\langle\cdot\rangle^r_v$.

On the other hand, sticking with our example '$p \vee \neg p$', one can observe that the best upper bound for my loss is indeed 1€ if I do *not know* the relevant risk values and I still have to stick with some pure strategy. This is because the chosen strategy might suggest to assert p even if, unknown to me, the experiment E_p always has a negative result. In other words, the bounds 1 and 0 are optimal now and coincide with the limits of $v^*_{\mathbf{L}}(p \vee \neg p)$. However, in general, this scenario — playing a pure strategy referring to risk values that need not coincide with the risk values used to calculate the expected pay-off — may lead to an expected loss outside the interval corresponding to $v^*_{\mathbf{L}}$. For a simple example consider $p \vee q$, where $v^*(p) = [0.4, 0.4]$, i.e., the players know that the expected loss associated with an assertion of p is 0.6€, and $v^*(q) = [0,1]$, i.e., the risk associated with asserting q can be any value between 1 and 0. We have $v^*_{\mathbf{L}}(p \vee q) = [\max(0, 0.4), \max(0.4, 1)] = [0.4, 1]$. Under the assumption that $\langle q\rangle^r_v = 0$, which is compatible with $v^*(q)$, my best strategy calls for asserting q in consequence of asserting $p \vee q$. But if the state $[\,\|q]$ is evaluated using the risk value $\langle q\rangle^r_v = 1$, which is also compatible with $v^*(q)$, then I have to expect a sure loss of 1€, although $1 - 1 = 0$ is outside $[0.4, 1]$.

5 Cautious and bold betting on unstable propositions

We suggest that a more convincing justification of the formal semantics of (Esteva et al., 1994) arises from the following alternative game based model of reasoning under imprecise knowledge. Like above, let v^* be an assignment of intervals $\subseteq [0,1]$ to the propositional variables. Again, we leave the dialogue part of Giles's game unchanged. But in reference to the partial information represented by v^*, we assign two different success probabilities to each experiment E_q corresponding to a propositional variable q, reflecting whether q is asserted by me or by you and consider best case and worst case scenarios (from my point of view) concerning the resulting expected pay-off. More precisely, my expected loss for the final state $[p_1, \ldots, p_m \| q_1, \ldots, q_n]$ when evaluated v^*-*cautiously* is given by $\sum_{i=1}^{n} \langle q_i \rangle_h^r - \sum_{i=1}^{m} \langle p_i \rangle_l^r$, but when evaluated v^*-*boldly* it is given by $\sum_{i=1}^{n} \langle q_i \rangle_l^r - \sum_{i=1}^{m} \langle p_i \rangle_h^r$, where the risk values $\langle q \rangle_h^r$ and $\langle q \rangle_l^r$ are determined by the limits of the interval $v^*(q) = [a,b]$ as follows: $\langle q \rangle_h^r = 1 - a$ and $\langle q \rangle_l^r = 1 - b$.

Proposition 2. *Given an interval assignment v^*, the following statements are equivalent:*

(i) *Formula F evaluates to $v^*_{\mathbf{L}}(F) = [a,b]$.*

(ii) *For the dialogue game in Section 3, played of F: if elementary states are evaluated v^*-cautiously then the minimal expected loss I can achieve by an optimal strategy is $(1-b)$€; if elementary states are evaluated v^*-boldly then my optimal expected loss is $(1-a)$€.*

6 Conclusion

We have been motivated by various problems that arise from insisting on truth functionality for a particular type of fuzzy logic intended to capture reasoning under 'imprecise knowledge'. Most importantly for the current purpose, we have employed a dialogue *cum* betting game approach to model logical inference in a context of 'dispersive experiments' for testing the truth of atomic assertions. This analysis not only leads to different characterizations of an important interval based fuzzy logic, but relates concerns about properties of fuzzy logics to reflections on rationality *qua* playing optimally in adequate games for 'approximate reasoning'.

Christian G. Fermüller
Institut für Computersprachen, TU Wien
Favoritenstraße 9–11, A–1040 Vienna, Austria
chrisf@logic.at
http://www.logic.at/staff/chrisf/

References

Belnap, N. D. (1977). A useful four–valued logic. In G. Epstein & J. M. Dunn (Eds.), *Modern uses of multiple–valued logic* (pp. 8–37).

Ciabattoni, A., Fermüller, C. G., & Metcalfe, G. (2005). Uniform rules and dialogue games for fuzzy logics. In (pp. 496–510). Springer Verlag.

Cornelis, C., Arieli, O., Deschrijver, G., & Kerre, E. E. (2007). Uncertainty modeling by bilattice–based squares and triangles. *IEEE Transactions on fuzzy Systems*, *15*(2), 161–175.

Cornelis, C., Deschrijver, G., & Kerre, E. E. (2004). Implication in intuitionistic and interval–valued fuzzy set theory: construction, classification, application. *Intl. J. of Approximate Reasoning*, *35*, 55–95.

Cornelis, C., Deschrijver, G., & Kerre, E. E. (2006). Advances and challenges in interval–valued fuzzy logic. *Fuzzy Sets and Systems*, *157*(5), 622–627.

Dubois, D. (n.d.). *On ignorance and contradiction considered as truth–values.* (to appear in the *Logic Journal of the IGPL*)

Esteva, F., Garcia-Calvés, P., & Godo, L. (1994, March). Enriched interval bilattices: An approach to deal with uncertainty and imprecision. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, *1*, 37–54.

Felscher, W. (1985). Dialogues, strategies, and intuitionistic provability. *Annals of Pure and Applied Logic*, *28*, 217–254.

Fermüller, C. G. (2003). Theories of vagueness versus fuzzy logic: Can logicians learn from philosophers? *Neural Network World Journal*, *13*(5), 455–466.

Fermüller, C. G. (2004, October). *Revisiting Giles's game: Reconciling fuzzy logic and supervaluation.* (to appear in *Logic, Games and Philosophy: Foundational Perspectives*, Prague Colloquium October 2004)

Fermüller, C. G. (2007). Exploring dialogue games as foundation of fuzzy logic. In M. Štěpnička et al. (Eds.), *New dimensions in fuzzy logic and related technologies. proceedings of the* 5^{th} *eusflat conference* (Vol. I, pp. 437–444).

Fermüller, C. G., & Kosik, R. (2006). *Combining supervaluation and degree based reasoning under vagueness* (No. 4246). Springer Verlag.

Giles, R. (1974). A non-classical logic for physics. *Studia Logica*, *33*(4), 399–417.

Giles, R. (1976). Łukasiewicz logic and fuzzy set theory. *International Journal of Man-Machine Studies*, *8*(3), 313–327.

Giles, R. (1977). A non-classical logic for physics. In R. Wojcicki & G. Malinkowski (Eds.), *Selected papers on łukasiewicz sentential calculi* (pp. 13–51). please, add address of the publisher: Polish Academy of Sciences.

Ginsberg, M. L. (1988). Multivalued logics: a uniform approach to reasoning in artificial intelligence. *Computational Intelligence*, *4*(3), 265–316.

Hájek, P. (n.d.). *On a fuzzy logic with imprecise truth value.* (Unpublished manuscript)

Hájek, P. (1998). *Metamathematics of fuzzy logic.* Dordrecht: Kluwer.

Hájek, P. (2002). Why fuzzy logic? In D. Jacquette (Ed.), *A companion to philosophical logic* (pp. 595–606). please, add address of the publisher: Blackwell.

Keefe, R., & Rosanna. (2000). *Theories of vagueness.* Cambridge: Cambridge University Press.

Keefe, R., & Smith, P. (Eds.). (1987). *Vagueness: A reader.* please, add address of the publisher: MIT Press.

Lorenzen, P. (1960). Logik und Agon. In *Atti congr. internat. di filosofia* (Vol. 4, pp. 187–194). Sansoni, Firenze.

Łukasiewicz, J. (1920). O logice tròjwartościowej. *Ruch Filozoficzny*, *5*, 169–171.

Shapiro, S. (2006). *Vagueness in context.* Oxford: Oxford University Press.

Zadeh, L. A. (1975). Fuzzy logic and approximate reasoning. *Synthese*, *30*(3–4), 407–428.

Zadeh, L. A. (1988). Fuzzy logic. *IEEE: Computer*, *21*(4), 83–93.