Bilattices Are Nice Things

Melvin Fitting

December 28, 2002

Abstract

One approach to the paradoxes of self-referential languages is to allow some sentences to lack a truth value (or to have more than one). Then assigning truth values where possible becomes a fixpoint construction and, following Kripke, this is usually carried out over a partially ordered family of three-valued truth-value assignments. Some years ago Matt Ginsberg introduced the notion of bilattice, with applications to artificial intelligence in mind. Bilattices generalize the structure Kripke used in a very natural way, while making the mathematical machinery simpler and more perspicuous. In addition, work such as that of Yablo fits naturally into the bilattice setting. What I do here is present the general background of bilattices, discuss why they are natural, and show how fixpoint approaches to truth in languages that allow self-reference can be applied. This is not new work, but rather is a summary of research I have done over many years.

1 Introduction

An obvious way out of the problem posed by a sentence like "This sentence is false," is to declare that it lacks a truth value. That is easy. The difficulty comes in saying which sentences should have truth values. That is hard. Kripke and others have applied fixpoint methods to this problem—one introduces a truth revision operator and looks for a partial truth assignment that does not revise away. I will not attempt to analyze or continue the philosophical discussions and insights that began with Kripke's paper on a theory of truth. I have a more technical goal in mind. The question for this paper is, what is the algebraic structure of the space in which Kripke-style truth-revision operators live, and how may it best be organized?

Kripke took a very concrete approach in his influential paper [18]. A partial truth assignment (valuation) is identified with a pair, $\langle \mathcal{T}, \mathcal{F} \rangle$, of disjoint sets of atomic sentences. Think of $\mathcal{T}$ as the set of atoms assigned true and $\mathcal{F}$ as the set assigned false. Disjointness guarantees there is no ambiguity, but $\mathcal{F}$ is not required to be the complement of $\mathcal{T}$, so some atoms may lack truth values. Next, some scheme is adopted for extending a partial valuation to all formulas, not just atomic. The actual machinery for effecting this extension is not unique—Kripke considered three versions (all of which will be generalized in this paper). Once the extension is made to all formulas, a new partial truth assignment is defined—roughly, the new assignment adds "X is true" to either $\mathcal{T}$ or $\mathcal{F}$ depending on the value given to X by the original partial truth assignment, after it has been extended to all formulas. No matter which of the three schemes for extending partial truth assignments to non-atomic formulas is used, this defines an operator that turns partial valuations into partial valuations. Formally, such an operator maps a pair $\langle \mathcal{T}, \mathcal{F} \rangle$ of disjoint sets of atoms to another such pair. Kripke partially ordered his valuations, essentially by subset. That is, $\langle \mathcal{T}_1, \mathcal{F}_1 \rangle$ is considered less than $\langle \mathcal{T}_2, \mathcal{F}_2 \rangle$ if $\mathcal{T}_1 \subseteq \mathcal{T}_2$ and $\mathcal{F}_1 \subseteq \mathcal{F}_2$. In a sense, this ordering has to do with information; an increase does not mean formulas switch truth values, but rather that more formulas acquire truth values. Kripke then made use of lattice-theoretic properties of this ordering to establish the existence of fixpoints for his truth revision operators. A fixed point of a truth revision operator is a plausible candidate for saying which formulas should be what assigned truth values.

In [7] I took a more abstract approach. Algebraically, the ordering sketched above has the following properties: it is a partial ordering (reflexive, anti-symmetric, transitive); there is a smallest element (it is $\langle\emptyset,\emptyset\rangle$); every non-empty set having an upper bound has a least upper bound. It turns out that having these properties is enough to ensure that any monotone operator has fixpoints having the various features Kripke investigated. Although this is very nice, there is still a sense that more is going on than meets the eye. There really are two orderings being used, not just one. As noted above, Kripke's explicit ordering has to do with information. But if we take a partial valuation in his sense, $\langle\mathcal{T},\mathcal{F}\rangle$, and move an atom from $\mathcal{F}$ to $\mathcal{T}$, we get a new valuation that gives us no more and no less information—we still know truth values for the same atoms—but an atom is now "truer" than it was. There is an implicit ordering involving truth, as well as one involving information. And further, there is some kind of interplay between the information and the truth ordering. For instance, the operations $\wedge$ and $\vee$ of, say, the Kleene strong three-valued logic are definable using the ordering involving truth just referred to, but Kripke's fixpoint construction makes use of the information ordering. The algebraic approach of [7], while abstracting away some of the details of [18], still hides the double ordering structure.

In [3, 4], Belnap introduced a four-valued logic that extends Kleene's strong three-valued logic in a natural way, but has two explicit orderings, on information and on truth. Then in [15, 16] the notion of bilattice was introduced, with Belnap's four-valued logic as the simplest example. I realized that bilattices provide exactly the algebraic structure needed to carry out Kripke's construction, and also that of Yablo [23]. Not only did the mathematical structure serve well, but there is an underlying intuition about it that is quite satisfying.

In the rest of this paper I will sketch the basic ideas of bilattices, and show how they apply to languages allowing self-reference. I will omit all proofs, but I will give references to papers in which they can be found. Thus there will be no self-references.

2 Bilattices—the Basics

Terminology concerning bilattices has varied some in the literature. I will use a version I have found handy—it differs some from that originally introduced by Ginsberg, [15, 16]. I will be using bilattices here simply as generalized truth-value spaces. For seeing them as the basis of a logic in their own right, look at [1, 2], and for a general survey of the whole subject, consult [13].

2.1 Pre-bilattices

As I sketched in the introduction, what is needed are two orderings, one having to do with truth and the other with information. The idea now is to make these explicit, in an abstract sort of way.

Convention The notion of a lattice comes up throughout. All lattices I consider will have tops and bottoms—largest and smallest elements. To keep terminology simple, in everything that follows the term *lattice* means lattice with a top and a bottom.

If you are not familiar with the terminology, in a lattice the greatest lower bound and the least upper bound of two-element sets is required to exist (and hence also for any finite set). The greatest-lower-bound operation is usually called *meet* and the least-upper-bound operation is called *join*. Lattice meet and join operations are always commutative and associative. Requiring a top and a bottom amounts to saying there is an element bigger than all others and an element smaller than all others.

Definition 2.1 A *pre-bilattice* is a structure $\mathcal{B} = \langle \mathcal{B}, \leq_t, \leq_k \rangle$ in which $\mathcal{B}$ is a non-empty set and $\leq_t$ and $\leq_k$ are partial orderings each giving $\mathcal{B}$ the structure of a lattice.

Think of the members of $\mathcal{B}$ as pieces of information that are embodied as truth values in some generalized sense. It is to deal with this dual role that we have the two ordering relations.

The ordering $\leq_k$ should be thought of as ranking "degree of information". Thus if $x \leq_k y$, y gives us at least as much information as x (and possibly more). I suppose this really should be written as $\leq_i$, using i for information instead of k for knowledge. In some papers in the literature i is used, but I have always written $\leq_k$, and now I'm stuck with it. The meet and join operations for $\leq_k$ are denoted $\otimes$ and $\oplus$. The $\otimes$ operation is called *consensus*: $x \otimes y$ is the most information that x and y agree on. The $\oplus$ operation is called *gullability*—a person who is gullable will believe anything. Then $x \oplus y$ should be thought of as combining the information in x with that in y, without worrying about whether the pieces fit together or not. The bottom in the $\leq_k$ ordering is denoted by $\perp$ and the top by $\top$. Think of $\perp$ as representing the state of complete ignorance—no information. Likewise $\top$ represents full information, possibly including inconsistencies.

The relation $\leq_t$ is an ordering on the "degree of truth." The bottom in this ordering will be denoted by *false* and the top by *true*. Thus *false* $\leq_t x \leq_t$ *true* for any $x \in \mathcal{B}$. The meet and join operations for $\leq_t$ will be denoted by $\wedge$ and $\vee$. It is easy to check that when restricted to *false* and *true*, these obey the usual truth-table rules. It is also easy to check that when restricted to *false*, $\perp$ and *true* they obey the rules of Kleene's strong three-valued logic [17] (this works equally well if we restrict to *false*, $\top$ and *true*).

In a lattice, meets and joins of finite sets must exist. What is called a completeness assumption extends this to infinite sets as well. Completeness is needed to adequately interpret quantifiers. Here is the bilattice version of completeness.

Definition 2.2 A pre-bilattice $\langle \mathcal{B}, \leq_t, \leq_k \rangle$ is *complete* if all meets and joins exist, with respect to both orderings. I'll denote infinitary meet and join with respect to $\leq_t$ by $\bigwedge$ and $\bigvee$, and by $\prod$ and $\sum$ for the $\leq_k$ ordering.

2.2 Examples

Suppose we have a certain group of people, $\mathcal{P}$, whose opinions we value. If we ask these people about the status of a sentence X, some will call it true, some false. But also, some may decline to express an opinion, and some may be uncertain enough to say they have reasons for calling it both true and false. We can, then, assign X a kind of generalized truth value, $\langle P, N \rangle$, where P is the set of people in $\mathcal{P}$ who say X is true and N is the set who say it is false. As just noted, we do not require that $P \cup N = \mathcal{P}$, nor that $P \cap N = \emptyset$.

Orderings can be introduced into our people-based structure: set $\langle P_1, N_1 \rangle \leq_k \langle P_2, N_2 \rangle$ if $P_1 \subseteq P_2$ and $N_1 \subseteq N_2$, and set $\langle P_1, N_1 \rangle \leq_t \langle P_2, N_2 \rangle$ if $P_1 \subseteq P_2$ and $N_2 \subseteq N_1$ (note the reversal here). Thus, information goes up if more people express a positive or negative opinion, and truth goes up if people drop negative opinions or add positive ones. This gives us the structure of a pre-bilattice. In it, for example, $\langle P_1, N_1 \rangle \wedge \langle P_2, N_2 \rangle = \langle P_1 \cap P_2, N_1 \cup N_2 \rangle$, and $\langle P_1, N_1 \rangle \otimes \langle P_2, N_2 \rangle = \langle P_1 \cap P_2, N_1 \cap N_2 \rangle$. Reflection should convince you that these are quite natural operations. Also, $\perp = \langle \emptyset, \emptyset \rangle$, $\top = \langle \mathcal{P}, \mathcal{P} \rangle$, *false* $= \langle \emptyset, \mathcal{P} \rangle$, and *true* $= \langle \mathcal{P}, \emptyset \rangle$. You should reflect on these too.

As another example, consider a "fuzzy" truth value space, in which truth values are pairs $\langle p, n\rangle$ of real numbers in the interval $[0, 1]$, where p is "degree of belief," and n is "degree of doubt." Appropriate orderings for this example are $\langle p_1, n_1\rangle \le_k \langle p_2, n_2\rangle$ if $p_1 \le p_2$ and $n_1 \le n_2$; and $\langle p_1, n_1\rangle \le_t \langle p_2, n_2\rangle$ if $p_1 \le p_2$ and $n_2 \le n_1$.

The two examples above can be combined if we consider a collection of people, each of whom has "fuzzy" opinions. I won't follow up on this—you probably get the general idea.

Figure 1 shows the simplest non-trivial example of a pre-bilattice: only the four extreme elements exist and are distinct. It can be thought of as a special case of the people pre-bilattice above, in which there is only one person. This is a fundamental example, and originated before bilattices as such arose—it is the four-valued logic due to Belnap, [3, 4], and will be called $\mathcal{FOUR}$ here. Think of the left-right direction as characterizing the $\le_t$ ordering: a move to the right is an increase. The meet operation for the $\le_t$ ordering, $\wedge$, is then characterized by: $x \wedge y$ is the rightmost thing that is left of both x and y. The join operation, $\vee$ is dual to this. In a similar way the up-down direction characterizes the $\le_k$ ordering: a move up is an increase in information. $x \otimes y$ is the uppermost thing below both x and y, and $\oplus$ is dual. Spatial conventions like these will be used throughout.

Figure 2 shows a pre-bilattice in which subtler distinctions can be registered. As is also the case with $\mathcal{FOUR}$, $\perp$ represents a state of complete ignorance, and $\top$ one of information overload—solid evidence has been supplied both for and against some proposition. Likewise *false* represents the situation in which we have convincing evidence against some proposition, and no evidence in its favor, while *true* is just the opposite. But in Figure 2 there are two more states. Think of *fd* as a state in which we have no evidence in favor of a proposition, but we have some weak evidence against—read *fd* as "false with doubts." Think of *td* likewise as "true with doubts."

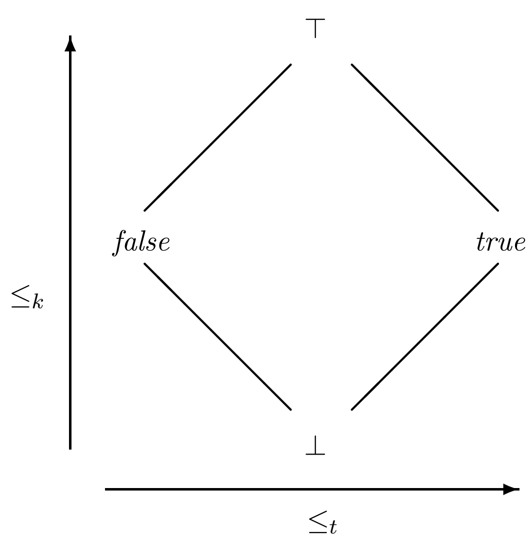

Figure 1: The Bilattice $\mathcal{FOUR}$

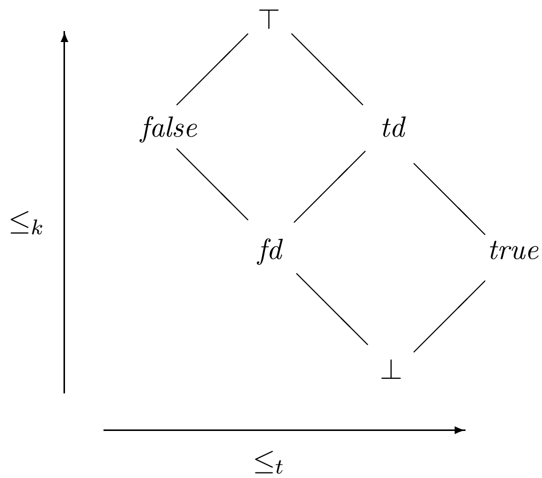

Figure 2: A Six-Valued Bilattice

2.3 Bilattices

A pre-bilattice has two orderings, with no postulated connections between them. I'll reserve the term *bilattice* for pre-bilattices where there are useful connections between orderings. Ginsberg's original definition of bilattice postulated a connection through a negation operation. Here I will use stronger notions that also trace back to Ginsberg.

Definition 2.3 A pre-bilattice $\langle \mathcal{B}, \leq_t, \leq_k \rangle$ is:

1. *an interlaced bilattice* if each of the operations $\wedge$, $\vee$, $\otimes$, and $\oplus$ is monotone with respect to both orderings (the interlacing conditions);

2. *an infinitarily interlaced bilattice* if it is complete and all four infinitary meet and join operations are monotone with respect to both orderings;

3. *a distributive bilattice* if all 12 distributive laws connecting $\wedge$, $\vee$, $\otimes$, and $\oplus$ are valid;

4. *an infinitarily distributive bilattice* if it is complete and infinitary, as well as finitary, distributive laws are valid. Examples of infinitary distributive laws are: $a \wedge \sum_i b_i = \sum_i (a \wedge b_i)$, and $a \otimes \bigwedge_i b_i = \bigwedge_i (a \otimes b_i)$.

A lattice is called distributive if it satisfies distributive laws; for example, a pre-bilattice is a lattice with respect to the $\leq_k$ ordering, and this lattice is distributive if $x \otimes (y \oplus z) = (x \otimes y) \oplus (x \otimes z)$ and $x \oplus (y \otimes z) = (x \oplus y) \otimes (x \oplus z)$ holds. Saying a pre-bilattice is distributive requires that we have distributive lattices with respect to both orderings and, in addition, we have "mixed" distributive laws, such as $x \otimes (y \vee z) = (x \otimes y) \vee (x \otimes z)$. All examples from Section 2.2 are distributive bilattices, and infinitary distributivity is satisfied as well.

In a lattice, meet and join operations are always monotone with respect to the lattice ordering. Thus we always have that $x_1 \leq_t y_1$ and $x_2 \leq_t y_2$ implies $(x_1 \wedge y_1) \leq_t (x_2 \wedge y_2)$. Saying we have an interlaced bilattice adds to this the requirement that monotonicity also work "across" orderings;

for example $x_1 \leq_k y_1$ and $x_2 \leq_k y_2$ implies $(x_1 \wedge y_1) \leq_k (x_2 \wedge y_2)$. It is not hard to show that every (infinitarily) distributive bilattice is also (infinitarily) interlaced, hence the examples of Section 2.2 satisfy the interlacing conditions.

2.4 Negation and conflation

Some bilattices have natural symmetries, and these can be used to characterize interesting subsystems.

Definition 2.4 A bilattice has a *negation* operation if there is a mapping, $\neg$, that reverses the $\leq_t$ ordering, leaves unchanged the $\leq_k$ ordering, and $\neg\neg x = x$. Likewise a bilattice has a *conflation* operation if there is a mapping, $-$, that reverses the $\leq_k$ ordering, leaves unchanged the $\leq_t$ ordering, and $- - x = x$. If a bilattice has both operations, they *commute* if $-\neg x = \neg - x$ for all x.

In the people example of Section 2.2, there are natural notions of negation and conflation. Take $\neg\langle P, N\rangle$ to be $\langle N, P\rangle$—the roles of for and against are switched. And take $-\langle P, N\rangle$ to be $\langle \mathcal{P} - N, \mathcal{P} - P\rangle$, where $\mathcal{P}$ is the set of people. This amounts to a kind of switching to a default position—the people who affirm under a conflation are the people who originally did not deny, for instance. The "fuzzy" example has a similarly defined negation and conflation—I'll leave their formulation to you. For both examples, negation and conflation commute.

In the example of Figure 1, there is a negation operation under which $\neg$*true* = *false*, $\neg$*false* = *true*, and $\perp$ and $\top$ are left unchanged. There is also a conflation under which $-\perp = \top$, $-\top = \perp$ and *true* and *false* are left unchanged. In this example negation and conflation commute. In any bilattice, if a negation or conflation exists the behavior on the extreme elements $\perp$, $\top$, *false*, and *true* will be as it is in $\mathcal{FOUR}$.

The example of Figure 2 does not have either a negation or a conflation. One might, for instance, try introducing a negation by adding to the usual conditions for the extreme elements the requirement that $\neg td = fd$ and $\neg fd = td$, but this will not work. We have $fd \leq_k false$ and negation is required not to affect the $\leq_k$ ordering, so we should have $td \leq_k true$, but in fact we have the opposite. There is a deeper reason for the lack of conflation and negation in this example that will become clear in the next section.

Definition 2.5 Suppose $\mathcal{B}$ is a bilattice with a conflation operation. Call $x \in \mathcal{B}$ *exact* if $x = -x$ and *consistent* if $x \leq_k -x$.

In the bilattice example involving people, Section 2.2, the exact values are those $\langle P, N\rangle$ where N is the complement of P—everyone expresses an unambiguous opinion. The consistent values are those where $P \cap N = \emptyset$, that is, people may be undecided, but they are never contradictory. In the bilattice $\mathcal{FOUR}$ of Figure 1, the exact members are $\{false, true\}$, the classical truth values, and the consistent ones are $\{false, \perp, true\}$, which behave like the values of Kleene's strong three-valued logic, with respect to $\neg$, $\wedge$, and $\vee$. This phenomenon, in fact, is not uncommon. The exact part of a complete bilattice with commuting conflation and negation is always closed under $\neg$, $\wedge$, and $\vee$, and similarly for the consistent part. In addition, the consistent part will always be closed under the infinitary version of $\otimes$, and under the infinitary version of $\oplus$ when applied to a directed set. It is essentially these conditions that were used in [7] for the special case of Kleene's strong three-valued logic, but in fact they obtain much more generally.

3 Where Bilattices Come From

There are standard ways of constructing bilattices that also provide some intuition concerning them. The first is from [16] with extensions of mine, though underlying ideas actually go back somewhat earlier. The second approach is apparently due to me.

3.1 Bilattice product

Suppose we have notions of positive and negative evidence. For instance, positive evidence for a mathematical conjecture might consist of plausibility arguments, computer experiments, almost correct proofs, and so on. Actual proofs would be best possible, of course. Negative evidence might also consist of various informal arguments, with counter-examples as best possible. Let us say we have a way of ranking evidence—this piece is better than that. More formally, say we have two lattices, $\mathcal{L}_1 = \langle \mathbf{L}_1, \leq_1 \rangle$ and $\mathcal{L}_2 = \langle \mathbf{L}_2, \leq_2 \rangle$, where members of $\mathbf{L}_1$ are things that can serve as positive evidence, with $\leq_1$ as a comparison relation, and similarly for $\mathbf{L}_2$ as pieces of negative evidence. The lattices need not be the same.

Definition 3.1 A *bilattice product* $\mathcal{L}_1 \odot \mathcal{L}_2$ is the structure $\langle \mathbf{L}_1 \times \mathbf{L}_2, \leq_t, \leq_k \rangle$ where:

1. $\langle x_1, x_2 \rangle \leq_t \langle y_1, y_2 \rangle$ if $x_1 \leq_1 y_1$ and $y_2 \leq_2 x_2$
2. $\langle x_1, x_2 \rangle \leq_k \langle y_1, y_2 \rangle$ if $x_1 \leq_1 y_1$ and $x_2 \leq_2 y_2$

Think of a member $\langle x, y \rangle$ of $\mathbf{L}_1 \times \mathbf{L}_2$ as encoding evidence about some assertion: evidence for, x, and evidence against, y. Then an increase in information amounts to saying evidence in general goes up. An increase in truth says evidence for increases while evidence against decreases. Earlier examples concerning people and "fuzzyness" are both special cases of this construction.

It is straightforward to show that $\mathcal{L}_1 \odot \mathcal{L}_2$ is always an interlaced bilattice, and is complete if both $\mathcal{L}_1$ and $\mathcal{L}_2$ are complete as lattices. And further, if both $\mathcal{L}_2$ and $\mathcal{L}_2$ are distributive lattices, $\mathcal{L}_1 \odot \mathcal{L}_2$ will be a distributive bilattice.

If $\mathcal{L}_1 = \mathcal{L}_2$ then a negation operaton can be introduced into $\mathcal{L}_1 \odot \mathcal{L}_2$. Set $\neg\langle x, y \rangle = \langle y, x \rangle$. That is, negation switches the roles of positive and negative evidence. Next, suppose $\mathcal{L}_1 = \mathcal{L}_2 = \mathcal{L}$ has what is called a *de Morgen complement* operation, an operation that maps x to $\overline{x}$ such that $x \leq y$ implies $\overline{y} \leq \overline{x}$, and $\overline{\overline{x}} = x$. Then a conflation operation can also be introduced into the bilattice product: set $-\langle x, y \rangle = \langle \overline{y}, \overline{x} \rangle$. Defined these ways, negation and conflation will commute.

The machinery just set forth for constructing various kinds of bilattices is completely general. That is, every distributive bilattice is isomorphic to $\mathcal{L}_1 \odot \mathcal{L}_2$ for some distributive lattices $\mathcal{L}_1$ and $\mathcal{L}_2$, and similarly for the other cases. Proof can be found for the various parts of this family of results in [16, 8, 11, 13]. In a way the result is a descendant of the Polarities Theorem of Dunn, [5].

Consider the familiar lattice $\mathcal{B}$ whose carrier is $\{\mathit{false}, \mathit{true}\}$, with $\mathit{false} < \mathit{true}$. This is a distributive lattice for which the operation $\overline{\mathit{false}} = \mathit{true}$ and $\overline{\mathit{true}} = \mathit{false}$ is a de Morgen complement. Then $\mathcal{B} \odot \mathcal{B}$ is a distributive bilattice with a negation and a conflation. It is, in fact, isomorphically the bilattice $\mathcal{FOUR}$ of Figure 1. Further, let $\mathcal{C}$ be the lattice whose carrier is $\{0, \frac{1}{2}, 1\}$, ordered numerically. Then $\mathcal{B} \odot \mathcal{C}$ is isomorphically the bilattice of Figure 2. Since $\mathcal{B}$ and $\mathcal{C}$ are different, there is no negation or conflation. But since both are distributive lattices, $\mathcal{B} \odot \mathcal{C}$ is a distributive bilattice.

3.2 An interval construction

There is a second approach to the construction of bilattices that provides a somewhat different intuition for them. The idea traces to [16] and was fully developed in [9].

Say we have a lattice L, which we can think of as generalized truth values of some kind, with $\leq_L$ as the ordering relation. As a concrete example, consider the unit interval with the usual ordering. This is linerally ordered, but other more complex examples are easy to come by. For $a, b \in L$ with $a \leq_L b$, by the *interval* $[a, b]$ is meant $\{x \in L \mid a \leq_L x \leq_L b\}$. Let $\mathcal{I}(L)$ be the family of intervals in the lattice L. Uncertain measurement of x may lead to the conclusion that it is between a and b, and so we might use $[a, b]$ as our current information about x. Presumably better measurements will shrink the interval.

Definition 3.2 Let L be a lattice. $\mathcal{K}(L)$ is the structure $\langle \mathcal{I}(L), \leq_t, \leq_k \rangle$ where:

1. $[a, b] \leq_t [c, d]$ if $a \leq_L c$ and $b \leq_L d$;
2. $[a, b] \leq_k [c, d]$ if $[c, d] \subseteq [a, b]$.

Thus an increase in information corresponds to shrinking an interval, and an increase in degree of truth corresponds to shifting the interval rightwards.

If L has a de Morgan complement, we can introduce a notion of negation into $\mathcal{K}(L)$ by setting $\neg[a, b] = [\overline{b}, \overline{a}]$.

Suppose we apply this construction to the simplest non-trivial lattice, $\{\mathit{false}, \mathit{true}\}$, with $\mathit{false} < \mathit{true}$. There are three intervals, $[\mathit{false}, \mathit{true}]$, essentially representing no information, and $[\mathit{false}, \mathit{false}]$ and $[\mathit{true}, \mathit{true}]$, which have narrowed things down as much as possible. The interval structure we get is, in fact, isomorphic to the consistent part of $\mathcal{FOUR}$.

This interval-based construction is also quite general. If we start with any lattice L having a de Morgan complement, $\mathcal{K}(L)$ will be isomorphic to the consistent part of an interlaced bilattice having a negation and a conflation. If L is distributive, the bilattice in question will also be. And conversely, given an interlaced bilatice with a negation and a conflation, its consistent part will always be isomorphic to $\mathcal{K}(L)$ for some lattice L having a de Morgan complement. (And, if the bilattice is distributive, the lattice will also be.)

3.3 Function spaces

There are still other ways of constructing bilattices, though not as general as those discussed above. Among them one stands out as the thing that makes bilattices especially applicable to our present purposes. As was noted in Section 1, Kripke took a very concrete representation for partial truth assignments. Here we can be somewhat more abstract. If we have a bilattice $\mathcal{B}$, we can think of its members as generalized truth values. For instance, as has been noted several times now, in $\mathcal{FOUR}$, the consistent values can be identified with those of Kleene's strong three-valued logic, so we can identify a truth assignment, in Kripke's sense, with a mapping from atomic sentences of a language to the consistent truth values of $\mathcal{FOUR}$. Or we could map to all of $\mathcal{FOUR}$. Or we could use some other bilattice altogether. The key point is that there is a mapping to a bilattice involved.

Definition 3.3 Let $\mathcal{B} = \langle \mathcal{B}, \leq_t, \leq_k \rangle$ be a pre-bilattice and let S be some non-empty set. $\mathcal{B}^S$ is the set of all functions from S to $\mathcal{B}$. Ordering relations are defined on $\mathcal{B}^S$ pointwise, and if $\mathcal{B}$ has a negation or a conflation operation, these are also extended to $\mathcal{B}^S$ in pointwise fashion. The characterizations are as follows.

1. $v \leq_t w$ if $v(s) \leq_t w(s)$ for all $s \in S$;
2. $v \leq_k w$ if $v(s) \leq_k w(s)$ for all $s \in S$;
3. If $\mathcal{B}$ has a negation, a negation operation is defined on $\mathcal{B}^S$ by: $\neg v$ is the mapping such that $(\neg v)(s) = \neg(v(s))$;
4. If $\mathcal{B}$ has a conflation, a conflation operation is defined on $\mathcal{B}^S$ by: $-v$ is the mapping such that $(-v)(s) = -(v(s))$;

It is straightforward that this makes $\mathcal{B}^S$ itself into a pre-bilattice. In fact, $(v \oplus w)(s) = v(s) \oplus w(s)$, and similarly for the other operations. Further, if $\mathcal{B}$ is interlaced, or distributive, $\mathcal{B}^S$ will also be interlaced, or distributive. If $\mathcal{B}$ has a negation operation, the negation operation defined above on $\mathcal{B}^S$ will meet the conditions for being a negation operation, and similarly for conflation.

If we take S to be the set of atomic formulas of some formal language, and $\mathcal{B}$ to be some interesting bilattice of truth values, $\mathcal{B}^S$ is really a space of valuations, and it is guaranteed to have algebraic properties that will make it useful to us.

We can carry this one step further yet. Suppose $\mathcal{B}^S$ is a space of valuations, where $\mathcal{B}$ is a bilattice of some particular kind. As just noted, $\mathcal{B}^S$ will itself be a bilattice of the same kind. Now, suppose we have a modal Kripke model—that is, a relational model for modal logic—with $\mathcal{G}$ as the set of possible worlds. $(\mathcal{B}^S)^{\mathcal{G}}$ is the space of mappings from possible worlds to valuations, and it too must be a bilattice of the same kind as $\mathcal{B}$. This is the natural machinery to use in developing a Kripke-style theory of truth in a modal setting.

4 How To Get Monotone Operators

Kripke defined three different truth revision operators and showed each must have a fixed point. The three differed in the underlying logic used: Kleene's strong three-valued, Kleene's weak three-valued, and supervaluations. As was noted in Section 1, Kripke made use of an ordering on information. All this generalizes quite naturally to a bilattice setting. For starters, we need a language capable of self-reference, and I'll do this in the most familiar way.

From here on *Lang* is the first-order language extending the language of arithmetic with the inclusion of an additional predicate symbol T, intended to be a "truth predicate." The only other relation symbol of *Lang* is $=$. We could, of course, also have relation symbols intended to represent relations in the "real world;" I omit consideration of these here for simplicity. I'll take $\wedge$, $\vee$, and $\neg$ as connectives and $\forall$ and $\exists$ as quantifiers of *Lang*. I'll also assume *Lang* has constant and function symbols for 0, successor, addition, and multiplication, and so, in particular, it has terms that (intuitively) denote exactly the natural numbers. In the usual Gödelian way numbers encode the various syntactic constructs of *Lang*, syntactic manipulations correspond to arithmetic operations on code numbers, and these operations can be represented in *Lang* itself. To keep the details simple, I'll assume the coding is onto from numbers to sentences of *Lang*, and for a closed term t, I'll say t *names* the sentence X if, in the standard model for arithmetic, t designates a code for X.

Since the meaning of the T predicate is what really concerns us now, let $\mathcal{A}$ be the set of all atomic sentences of *Lang* of the form $T(t)$, where t is a closed term (which names a sentence of *Lang*). Let $\mathcal{B}$ be some complete, infinitarily interlaced bilattice with a negation, and a conflation, with negation and conflation commuting. A *valuation* is a mapping from $\mathcal{A}$ to $\mathcal{B}$, and so the space of valuations, $\mathcal{B}^{\mathcal{A}}$, is itself an infinitarily interlaced bilattice with commuting negation and conflation.

From here on the treatment splits depending on how we choose to extend the behavior of valuations beyond the atomic level.

4.1 Kleene's strong three-valued logic generalized

Kleene's strong three-valued logic has probably been the most popular of the various logics used in Kripke's approach to self-reference. In this logic, for instance, $true \vee \perp = true$, informally because if we get enough further information about the second component of the disjunction to assign it a classical truth value, whether we find that component to be *true* or *false* we would still evaluate the disjunction to *true*, so extra information is not really relevant—we can assign *true* right now. Because of the nature of bilattices, this is the simplest partial logic for us to generalize. Let v be a valuation, that is, a member of $\mathcal{B}^{\mathcal{A}}$. Since this assigns values in $\mathcal{B}$ only to *atomic* sentences involving T, our first task is to extend it to a mapping on *all* sentences of *Lang*. I'll denote this extension by v^s; the superscript is for "strong".

1. $v^s(T(t)) = v(T(t))$.
2. If X is an atomic sentence not involving T, it must be a sentence of arithmetic. Set $v^s(X)$ to be *true* or *false* in $\mathcal{B}$ depending on whether X is true or false in the standard model for arithmetic.
3. $v^s(X \wedge Y) = v^s(X) \wedge v^s(Y)$
4. $v^s(X \vee Y) = v^s(X) \vee v^s(Y)$
5. $v^s(\neg X) = \neg v^s(X)$
6. $v^s((\forall x)F(x)) = \bigwedge_t v^s(F(t))$
7. $v^s((\exists x)F(x)) = \bigvee_t v^s(F(t))$

In item 3 the occurrence of $\wedge$ on the left is syntactic—it is a symbol of *Lang*; the occurrence of $\wedge$ on the right denotes the meet operation of $\mathcal{B}$ with respect to the $\leq_t$ ordering. Similar remarks apply to 4 and 5 as well. In item 6 the infinitary meet operation is over the set of all closed terms of *Lang*, and $F(t)$ is the result of substituting t for free occurrences of x in $F(x)$.

A remark that is not needed here, but will be when we come to supervaluations: if v is exact in the bilattice $\mathcal{B}^{\mathcal{A}}$ then for every formula X, $v^s(X)$ is an exact member of $\mathcal{B}$. The argument for this is rather simple. If v is exact in $\mathcal{B}^{\mathcal{A}}$ then exactness also applies pointwise, that is, for every atomic formula $T(t)$, $v(T(t))$ will be exact in $\mathcal{B}$. Now, using the extension v^s to arbitrary formulas, this exactness condition carries over to all formulas. Verifying this uses the fact that conflation does not change the $\leq_t$ ordering (hence conflation preserves $\wedge$, $\vee$, and their infinitary analogs), and the fact that negation and conflation commute.

Now define a mapping Φ^s from valuations to valuations. For a valuation v, set $\Phi^s(v) = v'$ where v' is the valuation such that $v'(T(t)) = v^s(X)$, where the closed term t names the sentence X.

$\mathcal{B}$ is required to be infinitarily interlaced, hence the same applies to $\mathcal{B}^{\mathcal{A}}$, the space of valuations. Given this, it is easy to check that Φ^s is monotone in the following sense (I'll leave details of the verification to you).

$$v \leq_k v' \Longrightarrow \Phi^s(v) \leq_k \Phi^s(v') \tag{1}$$

If we take for $\mathcal{B}$ the bilattice $\mathcal{FOUR}$, Φ^s is Kripke's operator based on Kleene's strong three-valued logic. It maps consistent valuations to consistent valuations, and the consistent subsystem of $\mathcal{FOUR}$ is exactly Kripke's setting.

4.2 Kleene's weak three-valued logic generalized

Kleene's weak three-valued logic assigns a value of $\bot$ to any compound formula in which some part has been assigned $\bot$. Thus, for instance, $\mathit{true} \vee \bot = \bot$, which is a different outcome than we get in Kleene's strong three-valued logic. The weak logic too can be generalized to the bilattice setting—[11] proposes an approach, but here I follow a different one. For motivation, consider once again the bilattice example based on people, from the beginning of Section 2.2. Suppose we have two bilattice values, $A = \langle P_1, N_1\rangle$ and $B = \langle P_2, N_2\rangle$, where the P_i and N_i are sets of people, those expressing opinions for, and against, respectively. Of course $A \wedge B$ was defined earlier, but suppose we want to 'cut this down' by only considering people who have actually expressed an opinion on both propositions A and B. As far as A is concerned, $A \oplus \neg A = \langle P_1 \cup N_1, P_1 \cup N_1\rangle$, and taking the consensus, $\otimes$, of this with an arbitrary member of the people bilattice does, indeed, cut things down to those who have expressed an opinion concerning A. Similarly for B. To keep notational clutter down, suppose I write $\|X\|$ for $X \oplus \neg X$, so what we want for a 'cut down' conjunction is $(A \wedge B) \otimes \|A\| \otimes \|B\|$. We can do a similar thing with disjunction, of course, and with $\forall$ and $\exists$. Negation is somewhat simpler since $\neg A \otimes \|A\| = \neg A$, so we can avoid extra complication in this case.

This suggests we define the following operators for any complete bilattice with negation. The superscript w is for "weak," and in fact, confined to the consistent part of the bilattice $\mathcal{FOUR}$, they are the connectives of Kleene's weak three-valued logic.

1. $X \wedge^w Y = (X \wedge Y) \otimes \|X\| \otimes \|Y\|$
2. $X \vee^w Y = (X \vee Y) \otimes \|X\| \otimes \|Y\|$
3. $\bigwedge^w S = (\bigwedge S) \otimes \prod\{\|X\| \mid X \in S\}$
4. $\bigvee^w S = (\bigvee S) \otimes \prod\{\|X\| \mid X \in S\}$

Once again let v be a valuation in $\mathcal{B}$; I'll extend it to a mapping v^w on all sentences of *Lang* as follows.

1. $v^w(T(t)) = v(T(t))$.
2. If X is a sentence of arithmetic, set $v^w(X)$ to be *true* or *false* in $\mathcal{B}$ depending on whether X is true or false in the standard model for arithmetic.
3. $v^w(X \wedge Y) = v^w(X) \wedge^w v^w(Y)$
4. $v^w(X \vee Y) = v^w(X) \vee^w v^w(Y)$
5. $v^w(\neg X) = \neg v^w(X)$
6. $v^w((\forall x)F(x)) = \bigwedge_t^w v^w(F(t))$
7. $v^w((\exists x)F(x)) = \bigvee_t^w v^w(F(t))$

This time define a mapping Φ^w on valuations by: for a valuation v, $\Phi^w(v) = v'$ where v' is the valuation such that $v'(T(t)) = v^w(X)$, where the closed term t names the sentence X.

Given the various properties of infinitarily interlaced bilattices, it is simple to verify that we again have monotonicity.

$$v \leq_k v' \Longrightarrow \Phi^w(v) \leq_k \Phi^w(v') \tag{2}$$

If we take for $\mathcal{B}$ the bilattice $\mathcal{FOUR}$, Φ^w is Kripke's operator based on Kleene's weak three-valued logic. It maps consistent valuations to consistent valuations.

4.3 Supervaluations generalized

A way of assigning partial truth values that respects tautologies was introduced in [21], and given the name *supervaluation*. The idea is, for a partial truth assignment to atomic formulas, extend it to the non-atomic level by taking a formula X to be true if every total truth assignment that extends the given partial one assigns X true; similarly for false; only in cases of disagreement is there a truth-value gap. Thus $\mathit{true} \vee \perp = \mathit{true}$ as with the strong Kleene logic, but also $P \vee \neg P$ always evaluates to *true* as well. Supervaluation, too, generalizes quite nicely to bilattices.

We need the following properties of infinitarily interlaced bilattices with commuting negation and conflation operations.

- For every x, $x \vee -x$ is exact. Reason: $-(x \vee -x) = -x \vee - - x = -x \vee x = x \vee -x$. (Recall, conflation does not affect the $\leq_t$ ordering, hence preserves $\wedge$ and $\vee$.)
- Every consistent member is below an exact member, in the $\leq_k$ ordering. Reason: Suppose x is consistent; $x \vee -x$ is exact, and x is below it because $x \leq_k -x$ (by consistency), $x \leq_k x$, and we have the interlacing conditions.
- If S is a non-empty set of exact members, $\prod S$ is consistent. Reason: $\prod S \leq_k \sum\{x \mid x \in S\} = \sum\{-x \mid x \in S\} = -\prod\{x \mid x \in S\} = -\prod S$.

Now, suppose v is a *consistent* valuation. A mapping from all formulas to $\mathcal{B}$, called v^{sv}, is defined as follows (the superscript is for "supervaluation").

$$v^{sv}(X) = \prod\{w^s(X) \mid v \leq_k w \text{ and } w \text{ is exact}\}$$

The conditions verified above ensure that this definition is meaningful, and that $v^{sv}(X)$ is a consistent member of $\mathcal{B}$. Now, define an operator *on the consistent part of* $\mathcal{B}^{\mathcal{A}}$ as follows. For a consistent valuation v, $\Phi^{sv}(v) = v'$ where v' is the valuation such that $v'(T(t)) = v^{sv}(X)$, where the closed term t names the sentence X. This operator, applied to a consistent valuation, produces another consistent valuation. And the operator is monotone on consistent valuations, with respect to the $\leq_k$ ordering. In $\mathcal{FOUR}$, the operator Φ^{sv} is exactly Kripke's supervaluation version of a truth revision operator. Finally, although Φ^{sv} is only defined on consistent valuations, it can be artificially extended to the entire of $\mathcal{B}^{\mathcal{A}}$ by setting $\Phi^{sv}(v) = \top$ whenever v is not consistent. This still leaves us with a monotone operator, a fact that will be useful later on.

5 Fixed Points

Three different families of truth-revision operators have now been presented—families because the choice of bilattice is left open. If a truth-revision operator has a fixed point, it must be a valuation that is, in a sense, coherent. Without additional information from the outside, a fixed point cannot be revised away, so it is a plausible way of assigning truth values to formulas while accomodating self-reference. So are there fixed points and, if there are, how do they relate to each other.

5.1 Fixed points exist

Fortunately, showing fixed points exist for each of the three operator families is easy. There is a well-known theorem of Knaster and Tarski [20] that says a monotone mapping on a complete lattice always has a smallest and a biggest fixed point. (Actually it says more, but we won't need the additional information.) Let $\mathcal{B}$ be a infinitarily interlaced bilattice with negation and conflation. Then operators Φ^s, Φ^w, and Φ^{sv} have been defined, and each is monotone with respect to the $\leq_k$ ordering of the bilattice. Since $\mathcal{B}$ is a complete lattice with respect to $\leq_k$, each of the three operators has a smallest and a biggest fixed point.

Carrying things further, it can be shown that smallest fixed points for our three operators must be consistent. So, if we take $\mathcal{B}$ to be the Belnap bilattice $\mathcal{FOUR}$, and restrict our attention to the consistent part, we have duplicated the investigation presented by Kripke in [18]. Some of the mathematics needed to justify Kripke's direct approach is a bit more complicated, essentially because he only considered the consistent part of $\mathcal{FOUR}$, by itself this is not a complete lattice, and so the Knaster-Tarski theorem could not be used. By moving to the bilattice setting, we not only get a simplification in the mathematics, but a considerable generalization—recall some of the bilattice examples discussed earlier, involving populations of people, or fuzzy truth values, or modal models, for instance.

Of course in $\mathcal{FOUR}$, both the sentence asserting its own falsehood and the sentence asserting its own truth are assigned $\perp$ in smallest fixed points, and $\top$ in largest. Since Kripke only worked with the consistent part of $\mathcal{FOUR}$, $\top$ was unavailable, but others have considered analogs of his theory in dual settings, where $\top$ was present. Bilattices combine both developments in a single setting.

5.2 Family structure

Kripke analyized the structure of the family of fixed points, introducing notions of maximal, intrinsic, and so on. I will not try to duplicate that here, except to note that the concept of intrinsic fixed point can be carried over to the general bilattice setting—see Section 7. Instead I want to discuss another topic—what happens if we restrict our formal language. Besides being of interest for its own sake, this also has some relationship to a different fixed point construction that will be presented in Section 6. For the rest of this section, let $\mathcal{B}$ be a complete bilattice—it doesn't matter if it has negation or conflation.

One of the reasons self-reference is such a problem is that the presence of negation in the language makes truth revision operators non-monotonic with respect to the $\leq_t$ ordering, which is the ordering our first impulses direct us toward. So let's get rid of negation. By $Lang^0$ we mean the sublanguage of *Lang* without negation. In $Lang^0$ we can no longer write a formal counterpart of "I am not true," but we still can for "I am true." Truth revision operators Φ^s, Φ^w, and Φ^{sv} can be defined with respect to $Lang^0$ essentially as we did above. If Φ is such an operator, it will be monotone with respect to $\leq_k$, just as before, but now it will also be monotone with respect to $\leq_t$. Since bilattices are lattices with respect to each of their orderings, we can apply the Knaster-Tarski theorem two different ways. There are least and greatest fixed points for Φ with respect to both orderings. The question is, how are these four fixed points related to each other?

Suppose we let p_k and P_k be the least and greatest fixed points of Φ with respect to the $\leq_k$ ordering of $\mathcal{B}$, and let p_t and P_t be the least and greatest fixed points of Φ with respect to $\leq_t$. By definition, all fixed points of Φ must lie between p_k and P_k in the $\leq_k$ ordering, and between p_t and P_t in the $\leq_t$ ordering. Figure 3 shows the configuration. But more, these extreme fixed points are related in the following remarkable way.

$$\begin{array}{rcl} p_k & = & p_t \otimes P_t \\ P_k & = & p_t \oplus P_t \\ p_t & = & p_k \wedge P_k \\ P_t & = & p_k \vee P_k \end{array}$$

Each extreme fixed point with respect to one ordering is a meet or join of the extreme fixed points with respect to the other ordering. As a special case, in $\mathcal{FOUR}$ the truth-teller is assigned $\perp$ and $\top$ by p_k and P_k respectively, and is assigned *false* and *true* by p_t and P_t respectively. And, in any bilattice the following identities hold.

$$\begin{array}{rcl} \perp & = & \mathit{false} \otimes \mathit{true} \\ \top & = & \mathit{false} \oplus \mathit{true} \\ \mathit{false} & = & \perp \wedge \top \\ \mathit{true} & = & \perp \vee \top \end{array}$$

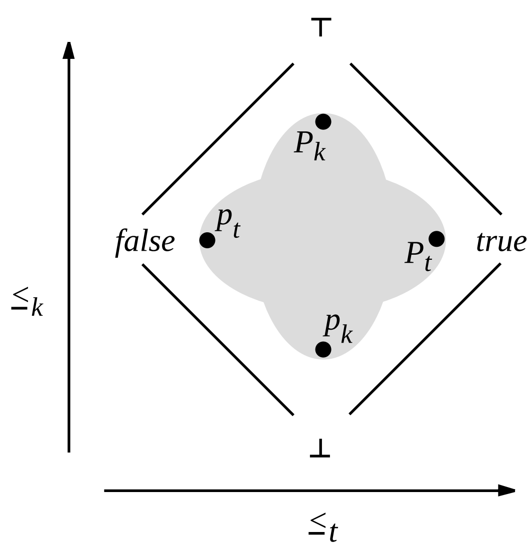

Figure 3: Fixed Points Without Negation

6 Introducing A Bias

The liar sentence can't have a classical truth value, but the truth-teller can—either *true* or *false* is fine. Still it is not implausible to argue that the truth-teller should not be evaluated as *true* because, as Mark Twain once said, "truth is precious, and we should economize it." Dropping negation altogether, as we did above, is not a good solution, because we are then unable to discuss the liar sentence at all. But there is another way—there are fixed points that minimize truth as much as possible. A full presentation can be found in [12], though the ideas grew out of stable model semantics for logic programs, and for non-monotonic reasoning more generally [14, 6, 10]. Here I'll just outline the development, referring to [12] for details.

6.1 Stable Fixed Points

The main idea is to separate the roles of positive and negative in sentences, and then apply monotonic machinery as far as possible. This probably sounds quite mysterious, but bear with me. To make this separation easier, *from now on I assume all formulas are in negation normal form: all occurrences of the negation symbol are at the atomic level.*

For the rest of this section, let $\mathcal{B}$ be an infinitarily interlaced bilattice. With formulas in negation normal form, think of occurrences of $\neg T(x)$ as if they were occurrences of a new atom, a falsehood atom, no longer directly connected with $T(x)$. As enabling machinery, I introduce the notion of a *pseudo-valuation*, a mapping from sentences of the forms $T(t)$ and $\neg T(t)$, independently, to $\mathcal{B}$. Pseudo-valuations are extended to non-atomic sentences inductively, using the bilattice operations $\wedge$, $\vee$, $\bigwedge$, and $\bigvee$. (In essence we are following the strong Kleene scheme; the weak Kleene, or supervaluation versions do not work for what I am about to present.) Pseudo-valuations can be created naturally, starting with valuations.

Definition 6.1 Let v_1 and v_2 be valuations. A pseudo-valuation denoted $v_1 \triangle v_2$ is characterized as follows.

$$(v_1 \triangle v_2)(T(t)) = v_1(T(t))$$
$$(v_1 \triangle v_2)(\neg T(t)) = \neg v_2(T(t))$$

I'll also write $(v_1 \triangle v_2)$ for the extension of this pseudo-valuation to all formulas.

Next, we generalize the truth revision operator to an operator Ψ that uses separate inputs for positive and for negative occurrences of T.

Definition 6.2 Let v_1 and v_2 be valuations. $\Psi(v_1, v_2) = v'$ where v' is the valuation such that $\Psi(v_1, v_2)(T(t)) = (v_1 \triangle v_2)(X)$, where the closed term t names the sentence X.

The operators looked at earlier in this paper, while well-behaved with respect to $\leq_k$, could be rather chaotic with respect to $\leq_t$. This new one is much more orderly, and indeed, all subsequent assertions about what I call the *derived operator* follow from these facts alone.

1. Ψ is monotone in both inputs, under $\leq_k$; if $v_1 \leq_k v_2$ and $w_1 \leq_k w_2$ then $\Psi(v_1, w_1) \leq_k \Psi(v_2, w_2)$.

2. Ψ is monotone in its first input, under $\leq_t$; if $v_1 \leq_t v_2$ then $\Psi(v_1, w) \leq_t \Psi(v_2, w)$.

3. Ψ is anti-monotone in its second input, under $\leq_t$; if $w_1 \leq_t w_2$ then $\Psi(v, w_1) \geq_t \Psi(v, w_2)$.

As noted earlier, in a complete bilattice we have a complete lattice under $\leq_t$ as well as under $\leq_k$. Since Ψ is monotone under $\leq_t$ in its first input, if we hold the second input fixed, and treat the operator as a function of its first input, we can apply the Knaster-Tarski Theorem.

Definition 6.3 The *derived operator* of Ψ is the single input function Ψ' characterized by: $\Psi'(v)$ is the smallest fixed point, in the $\leq_t$ ordering, of the function $(\lambda x)\Psi(x, v)$.

Since I chose to use the *smallest* fixed point, instead of the largest, an explicit bias towards falsehood has been introduced. I could, of course, have gone the other way—that is, everything that follows dualizes.

The mapping Ψ' is another candidate for a truth revision operator. In [12] I gave the following definition: A *GLF-stable* valuation is a fixed point of Ψ'. The 'GLF' stood for *Gelfond-Lifschitz, Fine*, to honor their work on stable model semantics for logic programming, [14, 6]. It can be shown that every GLF-stable valuation is also a fixed point of the operator Φ^s, defined in Section 4.1, so we are talking about a distinguished subclass of something whose investigation began with Kripke. Of course, we don't yet know there are any GLF-stable fixed points, but in fact it can also be shown that Ψ' is monotone with respect to $\leq_k$, and so once again Knaster-Tarski gives us the result we want.

Using Ψ' in $\mathcal{FOUR}$, we get smallest and biggest fixed points that are different from the extreme fixed points given by any of Kripke's operators. In the smallest GLF-stable valuation, the liar sentence has value $\perp$, and in the largest, $\top$. But in both, the truth-teller is simply *false*.

6.2 An Alternating Approach

In [23] a way was introduced for approximating to a fixed point from below and from above. In logic programming a related approach was introduced, called an alternating fixpoint construction, [22]. The idea carries over to the general bilattice setting as well, and is rather easy to describe. To begin, we need a variation on the familiar Knaster-Tarski theorem.

Suppose $\mathcal{L}$ is a complete lattice with $\leq$ as its ordering, and f is a function from $\mathcal{L}$ to itself. Two members, $x, y \in \mathcal{L}$ are called *oscillation points* of f if $f(x) = y$ and $f(y) = x$. They are *extreme* oscillation points if they are comparable (say $x \leq y$) and if a and b are any pair of oscillation points, $x \leq a, b \leq y$. Finally, the mapping f is called *anti-monotone* if $x \leq y$ implies $f(y) \leq f(x)$. The result we need is: an anti-monotonic map on a complete lattice always has a unique pair of extreme oscillation points.

A proof of the result cited above is not difficult. It can be shown by an argument similar to that used to establish Knaster-Tarski, for instance. Or it can be derived directly from Knaster-Tarski by noting that if f is anti-monotone, then f^2 is monotone. Its greatest and smallest fixed points will be extreme oscillation points for f. I omit details here.

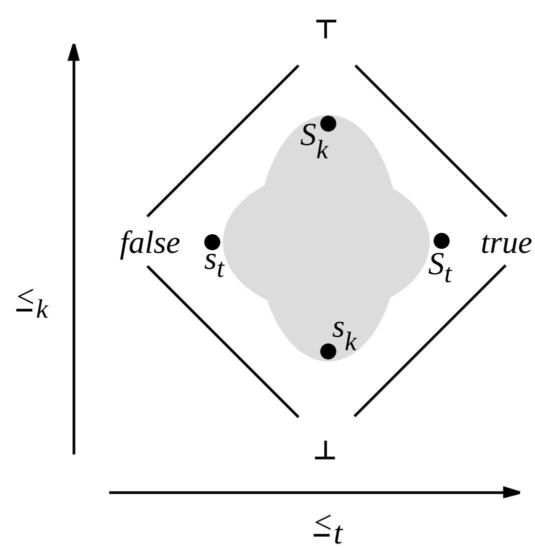

Figure 4: GLF-Stable Fixed Points

Now, let $\mathcal{B}$ be a complete bilattice with negation, define an operator Ψ as we did in Section 6.1, and also its derived operator Ψ'. Earlier I made use of the fact that Ψ' is monotonic with respect to $\leq_k$ to establish there are GLF-stable valuations. But also, Ψ' can be shown to be anti-monotonic with respect to $\leq_t$, and so Ψ' has extreme oscillation points with respect to this ordering. Of course these oscillation points are not GLF-stable valuations themselves, but if p any fixed point of Ψ', p, p is a trivial pair of oscillation points, and so p must be between the extreme oscillation points. Let me be more precise. Suppose we let s_k and S_k be the smallest and biggest fixed points of Ψ' with respect to the $\leq_k$ ordering—these are GLF-stable valuations. And let s_t and S_t be extreme oscillation points of Ψ' with respect to $\leq_t$, with s_t being the smaller—these are not themselves GLF-stable valuations. Then all GLF-stable valuations lie between these four values. The arrangement is shown in Figure 4, with the shaded area representing the GLF-stable valuations. Note that s_t and S_t are not part of the shaded area. What is more remarkable is that the four points are connected in exactly the same way that those of Figure 3 were.

$$
\begin{array}{rcl}
s_k & = & s_t \otimes S_t \\
S_k & = & s_t \oplus S_t \\
s_t & = & s_k \wedge S_k \\
S_t & = & s_k \vee S_k
\end{array}
$$

7 Intrinsic fixed points

Kripke defined a special class of fixed points which he called *intrinsic*, and showed they had a number of interesting features. (An equivalent definition was given independently in [19], using the name *optimal*.) The definition extends to bilattices quite easily, whether we use the strong Kleene, weak Kleene, or supervaluation operator—I'll use Φ for any one of these. In a bilattice (with conflation) a fixed point v of Φ is *intrinsic* if $v \oplus w$ is consistent, for every consistent fixed point w.

Let p_k be the smallest fixed point of Φ with respect to the $\leq_k$ ordering. As I noted before, p_k will itself be consistent. If w is any consistent fixed point, $p_k \oplus w = w$, since p_k is smallest. It follows that p_k is intrinsic. Hence intrinsic fixed points exist—p_k is one. Also every intrinsic fixed point must be consistent, because if w is intrinsic, since p_k is consistent $p_k \oplus w$ must be consistent, but this is just w. This is elementary, but with more work it can be shown that a largest intrinsic fixed point always exists as well—a proof can be found in [12].

In some of the consistent fixed points of Φ a truth-teller is *true*, in some, *false*, so if v is an *intrinsic* fixed point, a truth-teller can be neither *true* nor *false* in v. But a truth-teller will be *false* in every GLF-stable valuation, so no GLF-stable valuation can be intrinsic.

But, the notion of intrinsic can be relativized to GLF-stable valuations. I'll say a GLF-stable valuation v is *GLF-intrinsic* if $v \oplus w$ is consistent, for every consistent GLF-stable valuation w.

It is easy to show that the smallest GLF-stable valuation is GLF-intrinsic, and a largest GLF-intrinsic valuation exists. Very little more is known about the family of GLF-intrinsic valuations.

8 Conclusion

The machinery and results sketched above should provide convincing evidence that Kripke's theory of truth really lives in bilattices. Clearly a bilattice is an elegant place to live. But as the discussion of GLF-intrinsic indicates, the neighborhood needs to be better mapped. I hope I have generated enough interest in some readers of this paper to look further into the matter.

References

[1] O. Arieli and A. Avron. Reasoning with logical bilattices. *Journal of Logic, Language, and Information*, 5(1):25–63, 1996.

[2] O. Arieli and A. Avron. The value of four values. *Artificial Intelligence*, 102:97–141, 1998.

[3] N. D. Belnap, Jr. How a computer should think. In G. Ryle, editor, *Contemporary Aspects of Philosophy*, pages 30–55. Oriel Press Ltd, Stocksfield, 1977.

[4] N. D. Belnap, Jr. A useful four-valued logic. In J. M. Dunn and G. Epstein, editors, *Modern Uses of Multiple-Valued Logic*, pages 8–37. D. Reidel, Dordrecht, 1977.

[5] J. M. Dunn. Relevance logic and entailment. In D. Gabbay and F. Guenthner, editors, *Handbook of Philosophical Logic, Volume III*, chapter 3, pages 117–224. D. Reidel, 1986.

[6] K. Fine. The justification of negation as failure. In J. E. Fenstad, I. T. Frolov, and R. Hilpinen, editors, *Logic, Methodology and Philosophy of Science VIII*, pages 263–301, Amsterdam, 1989. North-Holland.

[7] M. C. Fitting. Notes on the mathematical aspects of Kripke's theory of truth. *Notre Dame Journal of Formal Logic*, 27:75–88, 1986.

[8] M. C. Fitting. Bilattices in logic programming. In G. Epstein, editor, *The Twentieth International Symposium on Multiple-Valued Logic*, pages 238–246. IEEE, 1990.

[9] M. C. Fitting. Kleene's logic, generalized. *Journal of Logic and Computation*, 1:797–810, 1992.

[10] M. C. Fitting. The family of stable models. *Journal of Logic Programming*, 17:197–225, 1993.

[11] M. C. Fitting. Kleene's three-valued logics and their children. *Fundamenta Informaticae*, 20:113–131, 1994.

[12] M. C. Fitting. A theory of truth that prefers falsehood. *Journal of Philosophical Logic*, 26:477–500, 1997.

[13] G. Gargov. Knowledge, uncertainty and ignorance in logic: bilattices and beyond. *Journal of Applied Non-Classical Logics*, 9, 1999.

[14] M. Gelfond and V. Lifschitz. The stable model semantics for logic programming. In R. Kowalski and K. Bowen, editors, *Proc. of the Fifth Logic Programming Symposium*, pages 1070–1080, Cambridge, MA, 1988. MIT Press.

[15] M. Ginsberg. Multi-valued logics. In *Proceedings of AAAI-86, Fifth National Conference on Artificial Intelligence*, pages 243–247, Los Altos, 1986. Morgan Kaufman.

[16] M. Ginsberg. Multivalued logics: a uniform approach to reasoning in AI. *Computational Intelligence*, 4:256–316, 1988.

[17] S. C. Kleene. *Introduction to Metamathematics.* D. Van Nostrand, Princeton, NJ, 1950.

[18] S. Kripke. Outline of a theory of truth. *The Journal of Philosophy*, 72:690–716, 1975. Reprinted in *New Essays on Truth and the Liar Paradox*, R. L. Martin, ed., Oxford (1983).

[19] Z. Manna and A. Shamir. The optimal approach to recursive programs. *Comm. ACM*, 20:824–831, 1977.

[20] A. Tarski. A lattice-theoretical theorem and its applications. *Pacific Journal of Mathematics*, 5:285–309, 1955.

[21] B. Van Fraassen. Singular terms, truth-value gaps, and free logic. *Journal of Philosophy*, 63:481–485, 1966.

[22] A. Van Gelder. The alternating fixpoint of logic programs with negation. In *Proc. 8th ACM Symp. on Principles of Database Systems*, pages 1–10, Philadelphia, 1989. ACM.

[23] S. Yablo. Truth and reflection. *Journal of Philosophical Logic*, 14:297–349, 1985.