MELVIN FITTING

BILATTICES AND THE THEORY OF TRUTH

1. INTRODUCTION

Kripke [13], and independently Martin and Woodruff [14], introduced the fixed point approach as a solution to the problem of assigning truth values to sentences in a language that allowed self-reference, and hence that allowed paradoxical sentences. Crucial to this was the use of a three-valued logic; some sentences must be left lacking a classical truth value, or having the truth value 'underdetermined'. More recently this approach has been extended [17], [18] to use a four-valued logic due to Belnap [2], with a fourth truth value of 'overdetermined'. Moving from a three to a four-valued logic simplifies the mathematics, since we now have a complete lattice instead of a complete semi-lattice to work with. It loses none of the original insights, since Kleene's strong three-valued logic is a natural sublogic of Belnap's. And it makes possible a treatment that has its own intuitive satisfications. After all, a sentence asserting its own falsehood could be taken to be overdetermined as well as underdetermined.

In the fixed point approach, some fixed points are more significant than others. In Kripke's version one is most interested in the 'least' fixed point, where the underlying ordering is one that treats the lack of a classical truth value as 'less than' the possession of one. In the four-valued approach this ordering is extended to count having a classical truth value as 'less than' being overdetermined. It turns out that this ordering meshes well with the usual logical connectives and quantifiers (suitably modified to take the four values into account). What one needs is that the ordering and the connectives interact in ways that 'respect' the order relation, and this is indeed the case.

Ginsberg has introduced into computer science a family of algebraic structures under the general name *bilattice* ([8], [9]). A bilattice can be thought of as a generalized truth value space. In any bilattice there are analogs of the classical logical connectives, and analogs of the

Journal of Philosophical Logic **18**: 225–256, 1989.

ordering relation mentioned above, and the interaction between them is of the kind necessary for a fixed-point theory to be developed. Indeed, the simplest bilattice is just the four-valued logic of Belnap. What is more, a bilattice structure is behind many other naturally occurring examples, such as modal versions based on Kripke models. Finally, there is a general technique for constructing a bilattice, starting with any 'reasonable' space of truth values.

In this paper we first present several concrete examples of bilattices, and discuss them in some detail. Then we present the general notion of bilattice, and establish the basic results concerning it. Next we show that a Kripke-like development can be given uniformly for a whole family of bilattices meeting certain simple conditions. Finally we show an interesting result relating the various extremal fixed points that one obtains in the bilattice setting; and in Belnap's four valued logic as a particular example. This is a result whose very statement would be impossible without the bilattice machinery.

We note that there are close relationships between techniques developed for dealing with the formal concept of truth, and those suitable for the semantics of logic programming in computer science. Heretofore the crossover has been in the direction of computer science; see [4]. This time the movement is the other way. The results here were originally developed in the logic programming setting (see [7]) and then transferred to this area. This relationship holds promise of further enrichment of both fields.

2. THE FOUR VALUED CASE

Let L be a first-order language with connectives $\wedge$, $\vee$ and $\neg$, and quantifiers $\forall$ and $\exists$, and that includes notation for elementary arithmetic. We will always interpret the operators, constant symbols and relation symbols of arithmetic in the standard way. The point of including arithmetic is so that L can code its own syntax. Any other suitable mechanism for this purpose could be used as well. In addition to the technical incorporation of arithmetic we also assume L has atomic sentences that we can think of as being about the real world, such as, it is raining at place a at time b. We use $\mathbf{R}$ for the class of these atomic sentences. Incidentally, we use the term *sentence* to mean *closed formula*. It is only sentences that we will be interested in here. Next, we assume L has 'names' for all objects we wish to talk about, and so we can think of quantification substitutionally. (This is a convenience, not a necessity.) We use **D** for the collection of these names. Finally, we assume L contains a predicate **Tr**, and we will want to think of $\mathbf{Tr}(\ulcorner X \urcorner)$ as asserting the truth of sentence X. (We use $\ulcorner X \urcorner$ for the Gödel number of X, so **Tr** is technically a predicate of numbers.)

If the atomic sentences of **R** are assigned truth values that accord with the real world, and atomic arithmetic sentences are interpreted in the standard way, there remains the problem of interpreting **Tr** so that $\mathbf{Tr}(\ulcorner X \urcorner)$ will be true exactly when the sentence X is true. Since we have sufficient machinery to construct a sentence that asserts its own falsehood, the task is impossible by Tarski's Theorem. But if **Tr** were given a *partial* interpretation (it is *true* on some integers, *false* on some, and underdetermined on some) the problem could be avoided. What Kripke, and Martin and Woodruff, showed was that this could be done in a way that was both clean and plausible. We will not repeat the details here, but simply refer to [13], [14] and incidentally [5].

Now, suppose L is the language of person P. This person, wishing to carry out Kripke's program (blithely ignoring the program's non-constructive aspects) is faced with an immediate problem: how does one ascertain the truth of sentences in **R**, about the real world? Much of what we 'know' is really quite uncertain knowledge. But, for the purposes of Kripke's construction, this is a non-issue. All that matters is that classical truth values be assigned to members of **R**. Whether or not these truth values accurately reflect reality (whatever this means) plays no role. So we might imagine a scenario like this.

Person P is in a cave and cannot see the outside world. But just outside the cave there is a real-world expert who, from time to time, shouts in to person P some sentence about the outside world such as, "It is raining in Detroit, now." Person P accepts these assertions as correct, and assigns truth values to sentences of **R** accordingly. Since **R** may have infinitely many sentences, at any given time there may be members of **R** that have not acquired a truth value by this process. They are given the value underdetermined. At any rate, Kripke's mechanism still applies, and a meaningful partial truth assignment to

Tr, and hence to all sentences of L can still be characterized in terms of a least fixed point model. Instead of 'real-world truth', we are now talking about 'truth as P understands it.' Incidentally, it can be shown that as time passes, and P's expert continues to call out more assertions about the outside world, and the interpretation of **Tr** is recalculated, nothing ever needs to be retracted. That is, if $\mathbf{Tr}(\ulcorner X \urcorner)$ is *true* (or *false*) today, it will still come out *true* (or *false*) tomorrow, no matter what additional information P's expert supplies. This assumes that the expert never takes back or contradicts an earlier assertion, but only adds new ones.

Now consider a situation a little more like the one we generally find ourselves in. P is still in the cave, but this time there are several experts outside, shouting information. It is enough to consider the case of two experts, say A and B; the general case adds no new complications. What if A shouts, "It is raining in Detroit now", while B shouts "It is not raining in Detroit now." Clearly P has a problem. There are two simple courses of action that P could take, assuming P cannot call back to A and B. First, P could decide to insist on a consensus on this issue, and since A and B have disagreed, the sentence "It is raining in Detroit now" is given no truth value (or given the value *undetermined* or $\bot$). Second, P could decide that an expert's opinion should be accepted no matter what, and so "It is raining in Detroit now" is taken to be *both* true and false (or given the truth value *overdetermined* or $\top$). Having $\bot$ as a truth value is sometimes referred to as having a truth value *gap*. In the same way, having $\top$ as a truth value is sometimes referred to as having a truth value *glut*. At any rate, what formal sense can be made of this new truth value?

Belnap, in [2], proposed a four valued logic, which we will call $\mathcal{FOUR}$, that is quite appealing, and we recommend his paper highly. (See also [17].) One can think of Belnap's truth values as sets of truth values in the ordinary sense, an idea originating in [3]. Then the four truth values of $\mathcal{FOUR}$ are: $\{true\}$, which we will write simply as *true*; $\{false\}$, which we will write as *false*; $\{\}$, which we will write as $\bot$ and read as *underdetermined*; and $\{true, false\}$, which we will write as $\top$ and read as *overdetermined.*

In a sense, this four valued logic is quite ancient. Once, the Venerable Malunkyaputta, a disciple of the Buddha, asked for

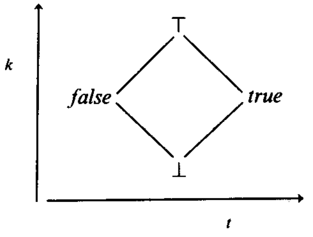

Fig. 1. The logic $\mathcal{FOUR}$.

information on several subjects which he thought to be of importance, and about which the Buddha had not spoken. One of them was the existence of the soul, and the question was: whether it existed; whether it did not exist; whether it neither existed nor did not exist; whether it both existed and did not exist. This is clearly Belnap's logic anticipated and applied. For another pertinent example consider the following, from *Mādhyamika Kārikā*, chapter 18, verse 8, quoted in [15], page 316: "Everything is true; not everything is true; both, everything is true, and not everything is true; or, neither everything is true nor is everything not true. This is the teaching of the Buddha."

The four truth values can be given some natural mathematical structure in a simple way. We follow ideas explicit in [8] and [9] which were implicit in [2]. We wish to give the space $\mathcal{FOUR}$ *two* partial orderings, denoted $\leqslant_k$ and $\leqslant_t$. We represent these in a double Hasse diagram in Figure 1.

Thus $a \leqslant_k b$ if there is an 'uphill' path from a to b. For example, $\bot \leqslant_k$ *false*. And $a \leqslant_t b$ if there is a 'left-right' path from a to b. For example, *false* $\leqslant_t \bot$. Very informally, the intuition is this. $\leqslant_t$ represents an increase in 'truth' or a decrease in 'falseness'. Thus *false* $\leqslant_t \bot$ because in going from $\{$*false*$\}$ to $\{\}$ the degree of falseness has decreased. Likewise $\bot \leqslant_t$ *true* because in going from $\{\}$ to $\{$*true*$\}$ the degree of truth has gone up. $\leqslant_k$ represents an increase in 'knowledge'. Thus $\bot \leqslant_k$ *false* because in going from $\{\}$ to $\{$*false*$\}$ our knowledge has gone from no information to the assignment of the truth-value *false*. And so on.

Under *both* partial orderings $\leqslant_t$ and $\leqslant_k$, we have a lattice (indeed, a complete one) and so meets and joins exist for both orderings.

We use the notation: $\wedge$ and $\vee$ for finite meet and join respectively under the $\leqslant_t$ ordering; $\bigwedge$ and $\bigvee$ for arbitrary meet and join under this ordering; $\otimes$ and $\oplus$ for finite meet and join under the $\leqslant_k$ ordering; and Π and Σ for arbitrary meet and join under this ordering. Finally, we can define negation in a reasonable way: $\neg$*true* = *false*; $\neg$*false* = *true*; $\neg\bot = \bot$; $\neg\top = \top$.

It is easy to check that the operations $\wedge$, $\vee$ and $\neg$, restricted to the two values *true* and *false*, are the usual classical logic operations. Further, when restricted to *true*, *false* and $\bot$, they are the operations of Kleene's strong three valued logic [10].

Finally, we mentioned earlier P's two possible approaches to conflicting information: insist on consensus, or accept anything. It is easy to see that the consensus approach amounts to combining truth values using $\otimes$, while the accept anything version combines truth values by using $\oplus$. Thus we have all the underlying machinery necessary to make formal sense of P's situation.

As starters on the technical machinery we can set up mappings $\mathscr{I}_A$ and $\mathscr{I}_B$ corresponding to experts A and B, as follows. For each sentence X in $\mathbf{R}$, let

$$\mathscr{I}_A(X) = \begin{cases} \textit{true} & \text{if } A \text{ says } X \text{ is true} \\ \textit{false} & \text{if } A \text{ says } X \text{ is false} \\ \bot & \text{otherwise} \end{cases}$$

and similarly for B. Next we can merge these in accordance with P's intentions, as follows. For a sentence X in $\mathbf{R}$, if P has decided to require consensus on X, set $\mathscr{I}(X) = \mathscr{I}_A(X) \otimes \mathscr{I}_B(X)$, and if P is willing to accept any expert opinion, set $\mathscr{I}(X) = \mathscr{I}_A(X) \oplus \mathscr{I}_B(X)$.

By a *valuation* we mean a mapping v from sentences of L to $\mathscr{FOUR}$ that meets the following conditions.

1. v makes the truth functional connectives correspond to the operations $\wedge$, $\vee$ and $\neg$ of $\mathscr{FOUR}$; that is, $v(X \wedge Y) = v(X) \wedge v(Y)$, and so on, and
2. $v((\forall x)\varphi(x)) = \bigwedge_{d\in\mathbf{D}} v(\varphi(d))$; $v((\exists x)\varphi(x)) = \bigvee_{d\in\mathbf{D}} v(\varphi(d))$.

As is the case in classical logic, so too here, valuations are completely determined by their behavior at the atomic level. This is simple to verify.

Suppose we (metaphorically) call a valuation v *realistic* if, for each X in $\mathbf{R}$, $v(X) = \mathscr{I}(X)$, and for each atomic sentence of arithmetic, v assigns truth values in accordance with the standard model. We can partially order these realistic valuations (or for that matter, all valuations) as follows: $v \leqslant_k w$ if $v(X) \leqslant_k w(X)$ for each atomic sentence X. ($\leqslant_t$ extends similarly.) It is not hard to verify that $\leqslant_k$ makes the set of realistic valuations into a complete lattice (also true for $\leqslant_t$), and also if $v \leqslant_k w$ then $v(X) \leqslant_k w(X)$ for *every* sentence of L, not just for atomic ones (not true for $\leqslant_t$ because of the behaviour of negation). The *smallest* realistic valuation in the $\leqslant_k$ ordering, v_0, has the property that $v_0(\mathbf{Tr}(\ulcorner X \urcorner)) = \bot$ for all X.

Following Kripke, we define a mapping Φ from valuations to valuations. For a valuation v, $\Phi(v) = w$ where w is the valuation whose atomic behavior is specified as follows.

1. if X is an arithmetic sentence, $w(X)$ is the truth value of X in the standard model;
2. if $X \in \mathbf{R}$, $w(X) = \mathscr{I}(X)$;
3. $w(\mathbf{Tr}(\ulcorner X \urcorner)) = v(X)$.

Now one can show that Φ maps *realistic* valuations to *realistic* valuations. Also Φ is *monotone* in $\leqslant_k$: $v \leqslant_k w \Rightarrow \Phi(v) \leqslant_k \Phi(w)$. This means the well-known Knaster–Tarski Theorem can be applied. This theorem states that a monotone mapping on a complete lattice has a smallest (and a biggest) fixed point, and indeed the set of all fixed points constitutes another complete lattice [16]. There are many places where proofs of some version of the theorem can be found. One that is readily available is [1] page 197. So, in our case, it follows by the Knaster–Tarski Theorem that Φ has a smallest fixed point, and it is this that we take as supplying the semantics for L. It is the 'intended' model for L, reflecting person P's 'knowledge' about the outside world. Thus we use the ordering $\leqslant_t$ in $\mathscr{FOUR}$ to provide meanings for the logical connectives and quantifiers, but we minimize in the $\leqslant_k$ ordering.

Proofs of our assertions are postponed until Sections 4 and 5. We note however that the following items will be established, and are of interest here. Let v_k be the smallest fixed point of Φ in the $\leqslant_k$ ordering.

1. We have a proper generalization of Kripke's work in the following sense. If persons A and B always agree, then v_k will never take on the value $\top$, and in fact will be identical with the three-valued least fixed point produced by Kripke's original methods.

2. v_k is monotone in time. That is, as time passes, persons A and B may announce truth values for sentences of **R** that they had not previously spoken about. Assuming they never retract or contradict anything they themselves said earlier, the result will be to increase v_k. That is, if v_k^* is the least fixed point of Φ calculated at a later time, then $v_k \leqslant_k v_k^*$. Thus no earlier 'knowledge' is reversed as time passes, provided there is no reversal at the 'real-world' level.

3. A MODAL EXTENSION

Before we turn to the full bilattice generalization, we consider a direct extension of the more-or-less classical situation treated in the previous section. Kripke, in [13], remarked that his construction carried over to models based on modal logic ([11], [12]). Here we look more closely at this assertion. We only consider so-called *constant domain* Kripke models. These are the models in which the Bárcan formula is valid. Models with non-constant domains can also be treated by the methods below, by using the following simple device. Introduce an extra predicate, $E(x)$, whose intended interpretation is: $E(d)$ is true at a world Γ if d is in the domain we want to associate with Γ. That is, E is an existence predicate. Then relativize all quantifiers to E. This amounts to an embedding of variable domain semantics into constant domain semantics. But for simplicity, for the rest of this section we assume L is a first order language, like the one used in the previous section, but with the incorporation of the usual model operators $\Box$ and $\Diamond$; we continue to use **D** for the set of constants of L, and we assume **D** is the domain of quantification for each world of our Kripke modal model.

DEFINITION 3.1. A Kripke model with constant domains is a tuple $\langle \mathscr{G}, \mathscr{R}, \models \rangle$, where $\mathscr{G}$ is a non-empty set (of possible worlds), $\mathscr{R}$ is a binary relation on $\mathscr{G}$ (of accessibility), and $\models$ is a relation between members of $\mathscr{G}$ and sentences of L. We assume the following conditions are met: for every $\Gamma \in \mathscr{G}$

$$\Gamma \models \neg X \Leftrightarrow \Gamma \not\models X$$

$$\Gamma \models X \wedge Y \Leftrightarrow \Gamma \models X \text{ and } \Gamma \models Y$$

$$\Gamma \models X \vee Y \Leftrightarrow \Gamma \models X \text{ or } \Gamma \models Y$$

$$\Gamma \models (\forall x)\varphi(x) \Leftrightarrow \Gamma \models \varphi(d) \text{ for all } d \in \mathbf{D}$$

$$\Gamma \models (\exists x)\varphi(x) \Leftrightarrow \Gamma \models \varphi(d) \text{ for some } d \in \mathbf{D}$$

$$\Gamma \models \Box X \Leftrightarrow \text{for all } \Delta \text{ with } \Gamma\mathscr{R}\Delta,\ \Delta \models X$$

$$\Gamma \models \Diamond X \Leftrightarrow \text{for some } \Delta \text{ with } \Gamma\mathscr{R}\Delta,\ \Delta \models X.$$

In the definition above we intended $\not\models$ to mean the same as not-$\models$. But now suppose we generalize the situation by assuming we have *two independent* relations, $\models$ and $\not\models$, each of which behaves in the expected way. More precisely,

DEFINITION 3.2. By a generalized Kripke model we mean a tuple $\langle \mathscr{G}, \mathscr{R}, \models, \not\models \rangle$ obeying the following conditions.

$$\Gamma \models \neg X \Leftrightarrow \text{not } \Gamma \models X$$

$$\Gamma \models X \wedge Y \Leftrightarrow \Gamma \models X \text{ and } \Gamma \models Y$$

$$\Gamma \models X \vee Y \Leftrightarrow \Gamma \models X \text{ or } \Gamma \models Y$$

$$\Gamma \models (\forall x)\varphi(x) \Leftrightarrow \Gamma \models \varphi(d) \text{ for all } d \in \mathbf{D}$$

$$\Gamma \models (\exists x)\varphi(x) \Leftrightarrow \Gamma \models \varphi(d) \text{ for some } d \in \mathbf{D}$$

$$\Gamma \models \Box X \Leftrightarrow \text{for all } \Delta \text{ with } \Gamma\mathscr{R}\Delta,\ \Delta \models X$$

$$\Gamma \models \Diamond X \Leftrightarrow \text{for some } \Delta \text{ with } \Gamma\mathscr{R}\Delta,\ \Delta \models X$$

$$\Gamma \nvDash \neg X \Leftrightarrow \text{not } \Gamma \nvDash X$$

$$\Gamma \nvDash X \wedge Y \Leftrightarrow \Gamma \nvDash X \text{ or } \Gamma \nvDash Y$$

$$\Gamma \nvDash X \vee Y \Leftrightarrow \Gamma \nvDash X \text{ and } \Gamma \nvDash Y$$

$$\Gamma \nvDash (\forall x)\varphi(x) \Leftrightarrow \Gamma \nvDash \varphi(d) \text{ for some } d \in \mathbf{D}$$

$$\Gamma \nvDash (\exists x)\varphi(x) \Leftrightarrow \Gamma \nvDash \varphi(d) \text{ for all } d \in \mathbf{D}$$

$$\Gamma \nvDash \Box X \Leftrightarrow \text{for some } \Delta \text{ with } \Gamma\mathcal{R}\Delta,\ \Delta \nvDash X$$

$$\Gamma \nvDash \Diamond X \Leftrightarrow \text{for all } \Delta \text{ with } \Gamma\mathcal{R}\Delta,\ \Delta \nvDash X$$

Then, if we have neither $\Gamma \vDash X$ nor $\Gamma \nvDash X$ we can think of X as being underdetermined, or having the truth value $\bot$ at Γ. Likewise if we have both $\Gamma \vDash X$ and $\Gamma \nvDash X$ we can think of X as being overdetermined, or having the truth value $\top$ at Γ.

But there are other ways of bringing truth values into the picture. Often the set of worlds at which a sentence holds is taken to be a kind of generalized truth value. This, in fact, is one way relationships between Kripke-style and algebraic-style semantics for modal logics are established. In our case, however, since we have independent relations $\vDash$ and $\nvDash$, things must be a little more complicated. The following is based on [8] and [9].

By a *truth value* relative to the generalized Kripke model $\mathcal{M} = \langle \mathcal{G}, \mathcal{R}, \vDash, \nvDash \rangle$ we mean a pair, $\langle F, A \rangle$, where both F and A are subsets of $\mathcal{G}$. Informally, we can think of F as the set of worlds in which some sentence holds ($\vDash$), and A as the set of worlds in which the sentence fails ($\nvDash$). More loosely yet, we can think of F as *evidence for*, and A as *evidence against*. We use $\mathbf{Kr}(\mathcal{M})$ for the space of all these truth values.

Just as in Section 2, there are two natural partial orderings that can be put on the space $\mathbf{Kr}(\mathcal{M})$ of truth values. Let $\langle F_1, A_1 \rangle \leqslant_t \langle F_2, A_2 \rangle$ if $F_1 \subseteq F_2$ but $A_2 \subseteq A_1$. Intuitively, $\langle F_2, A_2 \rangle$ is 'more true' than $\langle F_1, A_1 \rangle$ if the evidence for has gone up but the evidence against has gone down. Similarly, let $\langle F_1, A_1 \rangle \leqslant_k \langle F_2, A_2 \rangle$ if $F_1 \subseteq F_2$ and $A_1 \subseteq A_2$. Intuitively, $\langle F_2, A_2 \rangle$ represents 'more knowledge' than $\langle F_1, A_1 \rangle$ if all the evidence, for and against, has gone up. It is not hard to show that under both of these partial orderings the space of truth values is a complete lattice, and so again all meets and joins exist. We continue to use the notation of Section 2: $\wedge$, etc. for $\leqslant_t$ and $\otimes$, etc for $\leqslant_k$.

Since we have complete lattices under both orderings, top and bottom elements exist in $\mathbf{Kr}(\mathscr{M})$ for each. We use *false* and *true* for bottom and top under $\leqslant_t$, and $\bot$ and $\top$ for bottom and top under $\leqslant_k$. Notice that *false* $= \langle \varnothing, \mathscr{G} \rangle$, that is, no evidence for but total evidence against. Likewise $\bot = \langle \varnothing, \varnothing \rangle$, or a complete lack of evidence either way. Similarly for the tops.

We can define a straightforward negation operation: $\neg \langle F, A \rangle = \langle A, F \rangle$. And also, appropriate definitions of $\Box$ and $\Diamond$ operations are not difficult to come by. $\Box(\langle F, A \rangle) = \langle F', A' \rangle$ where $F' = \{\Gamma \mid \forall \Delta\ \Gamma \mathscr{R} \Delta \Rightarrow \Delta \in F\}$ and $A' = \{\Gamma \mid \exists \Delta\ \Gamma \mathscr{R} \Delta \text{ and } \Delta \in A\}$. Likewise, $\Diamond(\langle F, A \rangle) = \langle F', A' \rangle$ where $F' = \{\Gamma \mid \exists \Delta\ \Gamma \mathscr{R} \Delta \text{ and } \Delta \in F\}$ and $A' = \{\Gamma \mid \forall \Delta\ \Gamma \mathscr{R} \Delta \Rightarrow \Delta \in A\}$.

Suppose we define a mapping v from sentences of L to truth values of $\mathbf{Kr}(\mathscr{M})$ by making use of the underlying generalized Kripke model, as follows:

$$v(X) = \langle \{\Gamma \in \mathscr{G} \mid \Gamma \vDash X\}, \{\Gamma \in \mathscr{G} \mid \Gamma \nvDash X\} \rangle$$

It is not hard to show that v has the following properties.

$$\begin{aligned} v(\neg X) &= \neg v(X) \\ v(X \wedge Y) &= v(X) \wedge v(Y) \\ v(X \vee Y) &= v(X) \vee v(Y) \\ v((\forall x)\varphi(x)) &= \bigwedge_{d \in \mathbf{D}} v(\varphi(d)) \\ v((\exists x)\varphi(x)) &= \bigvee_{d \in \mathbf{D}} v(\varphi(d)) \\ v(\Box X) &= \Box(v(X)) \\ v(\Diamond X) &= \Diamond(v(X)) \end{aligned}$$

Suppose, for the rest of this section, we call a mapping v a *valuation* if it has the properties listed above. Again, as in Section 2, we can partially order valuations: $v \leqslant_k w$ if $v(X) \leqslant_k w(X)$ for every atomic X.

This gives us a complete lattice again. Finally, it is not difficult to show that $v \leqslant_k w$ implies $v(X) \leqslant_k w(X)$ for *every* sentence X, not just for atomic ones. (A similar partial ordering based on $\leqslant_t$ is also possible, though again the presence of negation destroys monotonicity with respect to arbitrary sentences).

Suppose we have a mapping $\mathscr{I}$ assigning truth values to atomic sentences about the real world. That is, $\mathscr{I}: \mathbf{R} \to \mathbf{Kr}(\mathscr{M})$. $\mathscr{I}$ can be thought of as arising from the opinions of experts, much as in Section 2, though now if we think in terms of the underlying Kripke model, we will have to associate experts with each possible world.

Again, as in Section 2, we call a valuation v *realistic* if, for each X in $\mathbf{R}$, $v(X) = \mathscr{I}(X)$, and for each atomic sentence X of arithmetic, if X is true in the standard model, $v(X) = true$, and if X is false in the standard model, $v(X) = false$. The ordering $\leqslant_k$ makes the set of realistic valuations into a complete lattice.

Finally, we can define a mapping Φ from valuations to valuations word for word as we did in Section 2. And similar results obtain: Φ has a smallest fixed point (under $\leqslant_k$) v_k which serves as an 'intended' model for the modal language L. Suppose we call a truth value $\langle F, A\rangle$ in $\mathbf{Kr}(\mathscr{M})$ *consistent* if $F \cap A = \varnothing$. If $\mathscr{I}$ only assigns consistent truth values to members of $\mathbf{R}$, the same will be true of v_k. Then, if we translate information about v_k back into information directly involving the possible world model we began with, we have essentially the generalization of Kripke's construction that was mentioned in [13]. But of course the present setting is more general, because it makes allowance for inconsistent truth values as well. And if we begin with a one-world Kripke model (with $\Box$ and $\Diamond$ trivialized) then the $\mathbf{Kr}(\mathscr{M})$ structure of this section is isomorphic to $\mathscr{FOUR}$ from Section 2.

4. INTERLACED BILATTICES

The truth value spaces above, and others as well, can be grouped together and treated uniformly using the notion of *bilattice*, [8], [9]. This section presents such a development, and contains sketches of proofs for many of the results cited earlier. In the next section we apply a Kripke-style development in the general bilattice setting.

A mild warning first, however. The notion of bilattice as currently defined in the literature does not meet our needs; we must require stronger conditions. (In fact, at an earlier stage of development, these conditions were part of the general definition of bilattice). In order to remain compatible with current terminology, we use the technical term *interlaced bilattice* for a structure meeting the conditions we require, so that the term *bilattice* will not wind up with competing definitions.

DEFINITION 4.1. By an interlaced bilattice we mean a set **B** together with two partial orderings, $\leqslant_t$ and $\leqslant_k$, meeting the conditions:

1. each of the two partial orderings gives **B** the structure of a complete lattice (hence arbitrary meets and joins exist with respect to each ordering);

2. the meet and join operation for each partial ordering is monotone with respect to the other ordering.

We call condition (2) above the *interlacing* condition. We will continue using the notation introduced in earlier sections. Arbitrary meets and joins under the $\leqslant_t$ ordering are denoted with $\bigwedge$ and $\bigvee$ respectively; finite meets and joins under this ordering are denoted with $\wedge$ and $\vee$ respectively. Arbitrary meets and joins under the $\leqslant_k$ ordering are denoted with Π and Σ respectively; finite meets and joins under this ordering are denoted with $\otimes$ and $\oplus$ respectively. Also the bottom and top under the $\leqslant_k$ ordering will be denoted $\bot$ and $\top$ respectively, and under the $\leqslant_t$ ordering by *false* and *true* respectively. We assume all interlaced bilattices are non-trivial, $\bot \neq \top$ and *false* $\neq$ *true*.

Since an interlaced bilattice is a lattice under $\leqslant_t$, it follows that $a_1 \leqslant_t a_2$ and $b_1 \leqslant_t b_2$ together imply $a_1 \wedge b_1 \leqslant_t a_2 \wedge b_2$. But by condition (2) of the interlaced bilattice definition, we also have that $a_1 \leqslant a_2$ and $b_1 \leqslant_t b_2$ together imply $a_1 \otimes b_1 \leqslant_t a_2 \otimes b_2$. That is, the meet operation of the $\leqslant_k$ ordering is monotone with respect to the $\leqslant_t$ ordering. Similarly for the other combinations of operation and order. More generally, suppose A and B are two subsets of **B**. We write $A \leqslant_t B$ if: for every $a \in A$ there is some $b \in B$ with $a \leqslant_t b$, and for every $b \in B$ there is some $a \in A$ with $a \leqslant_t b$. Then condition (2) above requires that if $A \leqslant_t B$, then $\Pi A \leqslant_t \Pi B$. And, of course, there are similar implications for the various combinations of orderings and operations. It is this tight interconnection between the partial orderings that makes an interlaced bilattice a coherent mathematical structure and not just a set with two independent lattice orderings on it.

We have required that an interlaced bilattice be a *complete* lattice under both of its partial orderings. In a broader context there may be some value to considering the weaker notion that requires just a lattice structure without insisting on completeness. We have not found any interesting applications however, but if there turn out to be some, the notion above should be rechristened a *complete interlaced bilattice*, reserving the term *interlaced bilattice* for the version that does not insist on completeness. In the interests of simple terminology, we do not do this in the present paper.

The truth value spaces considered in earlier sections, $\mathcal{FOUR}$ and $\mathbf{Kr}(\mathcal{M})$, are all interlaced bilattices. This is not difficult to check, and we do not go through the details. [6] presents a family of examples arising from topological spaces, and indirectly from Kripke intuitionistic logic models. Also any complete lattice trivially yields an interlaced bilattice if we take both the partial orderings to be the same, namely the given lattice ordering. Moreover, there is a general and intuitively appealing method for constructing interlaced bilattices, due to Ginsberg. We describe it briefly.

Suppose $C = \langle C, \leqslant \rangle$ and $D = \langle D, \leqslant \rangle$ are complete lattices. (We use the same notation, $\leqslant$, for both orderings, since context can determine which is meant.) Form the cartesian product $C \times D$, and give it two orderings, $\leqslant_k$ and $\leqslant_t$, as follows.

$$\langle c_1, d_1 \rangle \leqslant_k \langle c_2, d_2 \rangle \text{ if } c_1 \leqslant c_2 \text{ and } d_1 \leqslant d_2$$

$$\langle c_1, d_1 \rangle \leqslant_t \langle c_2, d_2 \rangle \text{ if } c_1 \leqslant c_2 \text{ and } d_2 \leqslant d_1$$

We denote the resulting structure, $\langle C \times D, \leqslant_k, \leqslant_t \rangle$ by $\mathcal{B}(C, D)$. It is easy to check that for complete lattices C and D, $\mathcal{B}(C, D)$ is an interlaced bilattice.

The intuition here is rather nice. Suppose we think of a pair $\langle c, d \rangle$ in $\mathcal{B}(C, D)$ as codifying two independent judgements concerning the 'truth' of some sentence: c represents our degree of belief *in* it, while d

represents our degree of belief *against* it. Since C and D can be different lattices, expressions of belief for and against need not be measured in the same way. If $\langle c_1, d_1\rangle \leqslant_k \langle c_2, d_2\rangle$ then $\langle c_2, d_2\rangle$ embodies more 'knowledge' than $\langle c_1, d_1\rangle$, which is reflected by an increased degree of belief both for and against. On the other hand, if $\langle c_1, d_1\rangle \leqslant_t \langle c_2, d_2\rangle$ then $\langle c_2, d_2\rangle$ embodies more 'truth' than $\langle c_1, d_1\rangle$, which is reflected by an increased degree of belief for, and a decreased degree of belief against.

Suppose we begin with the simplest non-trivial lattice $\mathcal{TR}$: $\{$*false*, *true*$\}$ with *false* $\leqslant$ *true*. Then $\mathcal{B}(\mathcal{TR}, \mathcal{TR})$ is simply an isomorphic copy of $\mathcal{FOUR}$ (Figure 1). In this representation, $\perp$ is $\langle$*false*, *false*$\rangle$, no belief for and no belief against; similarly $\top$ is $\langle$*true*, *true*$\rangle$. Likewise *false* is $\langle$*false*, *true*$\rangle$, no belief for, total belief against; and *true* is $\langle$*true*, *false*$\rangle$.

Another example of importance is a probabilistic one, $\mathcal{B}([0, 1], [0, 1])$, based on the complete lattice $[0, 1]$ with the usual ordering $\leqslant$ of reals.

The earlier discussion of Kripke models for modal logics fits into this paradigm rather nicely. As Ginsberg suggests, we can think of the set of possible worlds in which a formula is true as the *evidence for* a formula; similarly for false and evidence against. Then, given a particular Kripke model with $\mathcal{W}$ as the set of possible worlds, use the power set lattice $P = \langle \mathbf{P}(\mathcal{W}), \subseteq\rangle$ to create the interlaced bilattice $\mathcal{B}(P, P)$. This yields the structure considered in Section 3.

Likewise the family of interlaced bilattices that was considered in [6], based on topological spaces arising out of Kripke Intuitionistic Logic models, is also yielded by this construction technique. Let $\mathcal{T}$ be a topological space. The family $\mathcal{O}$ of open sets is a complete lattice under $\subseteq$. Join is union, while meet is interior of intersection. Likewise the family $\mathcal{C}$ of closed sets is a complete lattice under $\subseteq$. Now consider the interlaced bilattice $\mathcal{B}(\mathcal{O}, \mathcal{C})$.

Thus there are many potential applications for general results on interlaced bilattices. It is to such general results that we devote the rest of this section.

PROPOSITION 4.1. In an interlaced bilattice **B**:

1. $\mathit{true} \oplus \mathit{false} = \top$; $\mathit{true} \otimes \mathit{false} = \perp$;
2. $\top \vee \perp = \mathit{true}$; $\top \wedge \perp = \mathit{false}$.

Proof. Since $\top$ is the largest member of **B** under the $\leqslant_k$ ordering, then $a \oplus \top = \top$ for every a. Also, since *false* is the smallest member under the $\leqslant_t$ ordering, *false* $\leqslant_t \top$ and hence, since we have an interlaced bilattice, $a \oplus false \leqslant_t a \oplus \top$. Now, $true \oplus false \leqslant_t true \oplus \top = \top$. In the other direction, since *true* is the largest member under the $\leqslant_t$ ordering, $\top \leqslant_t true$ and hence $a \oplus \top \leqslant_t a \oplus true$ for any a. Now, $\top = false \oplus \top \leqslant_t false \oplus true$. The other items are also proved by similar methods.

In addition to the lattice operations, interlaced bilattices often have some natural notion of negation. This is the case with the examples of earlier sections.

DEFINITION 4.2. An interlaced bilattice **B** has a weak negation operation if there is a mapping $\neg : \mathbf{B} \to \mathbf{B}$ such that:

1. $$a \leqslant_k b \Rightarrow \neg a \leqslant_k \neg b;$$
2. $$a \leqslant_t b \Rightarrow \neg b \leqslant_t \neg a.$$

Note We have used the term *weak* negation. Generally a bilattice is said to have a negation operation if, in addition to the conditions above, we have $a = \neg\neg a$. In this paper we will not need this extra condition, though it holds in many examples of interest. In the family of interlaced bilattices considered in [6] the negation operation generally did not satisfy this condition, but did satisfy the weaker one: $a \leqslant_t \neg\neg a$, a kind of intuitionistic condition. There are examples of interlaced bilattices with weak negations that do not even satisfy this.

Besides negation, other operators may be of interest, such as $\Box$ and $\Diamond$ from Section 3.

DEFINITION 4.3. A mapping f: $\mathbf{B} \to \mathbf{B}$ is k monotone if $a \leqslant_k b \Rightarrow f(a) \leqslant_k f(b)$. Similarly for t monotone.

In fact, the operations $\Box$ and $\Diamond$ from Section 3 are both k and t monotone. This is easy to verify.

Finally, there is one more operation that many interlaced bilattices admit, though it is less commonly considered than those mentioned above. We call it *conflation*. In an interlaced bilattice $\mathbf{Kr}(\mathscr{M})$, from Section 3, the intended conflation of a truth value $\langle F, A\rangle$ is the truth value $\langle \bar{A}, \bar{F}\rangle$, where the overbar means complement with respect to the set of all possible worlds. The idea here is: the conflation of a truth value is a new truth value which counts as evidence for anything that wasn't counted against originally, and counts as evidence against anything that wasn't counted for originally. In $\mathscr{FOUR}$ we take the conflation of $\top$ to be $\perp$ and conversely, and the conflation of *true* and *false* to be themselves again. And in $\mathscr{B}([0, 1], [0, 1])$ we take the conflation of $\langle a, b\rangle$ to be $\langle 1 - b, 1 - a\rangle$. For the rest of this paper, when we refer to these interlaced bilattices we mean these spaces together with the conflation operations just associated with them.

Not all interlaced bilattices have a conflation operation, though all the ones considered here do. Formally, we require the following (which is related to the notion of *complete selfdual lattice* from [17]).

DEFINITION 4.4. An interlaced bilattice $\mathbf{B}$ has a conflation operation if there is a mapping: $\mathbf{B} \to \mathbf{B}$ such that:

1. $a \leqslant_t b \Rightarrow -a \leqslant_t -b;$
2. $a \leqslant_k b \Rightarrow -b \leqslant_k -a;$
3. $--a = a.$

Not every interlaced bilattice has a conflation operation. For instance $\mathscr{SIX}$ in Figure 2 is an interlaced bilattice without a conflation operation. This interlaced bilattice is, in fact, $\mathscr{B}(A, B)$ where A is the lattice $\{0, 1\}$ and B is the lattice $\{0, 1, 2\}$, where each has the natural ordering. Here a is $\langle 0, 1\rangle$, and b is $\langle 1, 1\rangle$. $\mathscr{SIX}$ also has a weak negation, but not a negation operation.

Conflation is like a dualized version of negation, with the roles of k and t interchanged. It is easily checked that the examples of conflation operations mentioned above do satisfy this definition. We use the obvious extension of notation from elements of an interlaced bilattice to subsets: $-A = \{-a \mid a \in A\}$. (The following has relationships with the notion of *self-duality* from [17].)

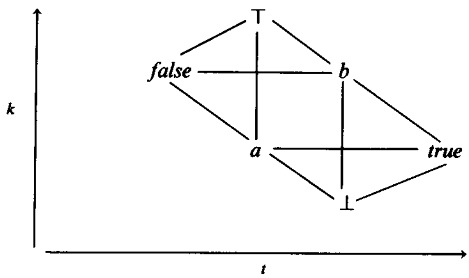

Fig. 2. The interlaced bilattice $\mathcal{SIX}$.

PROPOSITION 4.2. In an interlaced bilattice with a conflation operation:

1. $-\bigwedge A = \bigwedge - A$; $-(a \wedge b) = (-a \wedge -b)$;
2. $-\bigvee A = \bigvee - A$; $-(a \vee b) = (-a \vee -b)$;
3. $-\Pi\, A = \Sigma - A$; $-(a \otimes b) = (-a \oplus -b)$;
4. $-\Sigma\, A = \Pi - A$; $-(a \oplus b) = (-a \otimes -b)$.

Proof. Suppose $a \in A$. Then $\bigwedge A \leqslant_t a$, and so $-\bigwedge A \leqslant_t -a$. Since a is an arbitrary member of A, and hence $-a$ is an arbitrary member of $-A$, this establishes that $-\bigwedge A \leqslant_t \bigwedge - A$.

Conversely, again let $a \in A$, so $-a \in -A$. Then $\bigwedge - A \leqslant_t - a$, so by parts (1) and (3) of the definition, $-\bigwedge - A \leqslant_t a$. Since a is an arbitrary member of A, we have $-\bigwedge - A \leqslant_t \bigwedge A$, and hence $\bigwedge - A \leqslant_t -\bigwedge A$.

Part (1) follows immediately from these two results. The other items are similar.

PROPOSITION 4.3. In any interlaced bilattice **B** with conflation:

1. $-\bot = \top$; $-\top = \bot$;
2. $-$*false* $=$ *false*; $-$*true* $=$ *true*.

Proof. Since $\bot$ is the smallest element in the $\leqslant_k$ ordering, then $\bot \leqslant_k - \top$, and hence $\top \leqslant_k - \bot$. It follows that $\top = -\bot$. This is half of part (1), and the other half follows easily.

For part (2), we first note that, since *false* is the smallest member of **B** under the $\leqslant_t$ ordering, then $(*)$ *false* $\vee\ a = a$ for every a. Now $-$*false* $=$ $-$*false* $\vee$ *false* [by $(*)$] $=$ $-$*false* $\vee$ $--$*false* [by definition of conflation] $=$ $-$(*false* $\vee$ $-$*false*) [by Proposition 4.2 part (2)] $=$ $-(-$*false*$)$ [by $(*)$ again] $=$ *false* [by definition of conflation]. The second half of part (2) is similar.

In general, we are interested in interlaced bilattices with other operations besides the lattice ones, typically $\neg$, and maybe $\Box$ and $\Diamond$, as in the examples of Section 3.

DEFINITION 4.5. If **B** is an interlaced bilattice with a conflation operation, and f: $\mathbf{B} \rightarrow \mathbf{B}$, we say the conflation operation commutes with f provided $-f(a) = f(-a)$.

We leave the reader to check that the conflation operation for the interlaced bilattice $\mathscr{FOUR}$ commutes with $\neg$. Also the conflation operation for the interlaced bilattice $\mathbf{Kr}(\mathscr{M})$ commutes with $\neg$, $\Box$, and $\Diamond$. The definition of *commutes with* extends in an obvious way to binary operations, and beyond. Then Proposition 4.2 says conflation always commutes with $\wedge$, $\bigwedge$, $\vee$ and $\bigvee$.

Next, certain substructures of an interlaced bilattice are often of interest to us. The conflation operation provides a useful mechanism for identifying them. (The notion we call consistent below is analogous to that of *E-under* in [17], while that of exact is analogous to *E–U*.)

DEFINITION 4.6. In an interlaced bilattice **B** with conflation, for $a \in \mathbf{B}$,

1. a *is* exact *if* $a = -a$,
2. a *is* consistent *if* $a \leqslant_k -a$.

In the interlaced bilattice $\mathscr{FOUR}$, the exact truth values are the classical ones, *false* and *true*, and the consistent ones are these together with $\bot$, that is, the values used in Kleene's three valued logic. In $\mathbf{Kr}(\mathscr{M})$, the exact truth values are $\langle F, A\rangle$ where F and A are complementary. The consistent ones are those for which $F \cap A = \varnothing$. In $\mathscr{B}([0, 1],$

[0, 1]) the exact values are pairs $\langle a, b\rangle$ for which $a + b = 1$, while the consistent ones are those for which $a + b \leqslant 1$.

In $\mathcal{FOUR}$, the exact truth values are closed under the operations of classical logic. This is not surprising, since the exact truth values here are just the two values of classical logic. But the observation generalizes, to the following.

PROPOSITION 4.4. In an interlaced bilattice with conflation, the exact truth values include false and true and are closed under $\wedge$, $\vee$, $\bigwedge$ and $\bigvee$. Further, if f is an operation that the conflation operation commutes with, then the exact truth values are closed under f as well. The exact truth values do not include $\bot$ or $\top$ and are not closed under $\oplus$ or $\otimes$.

Proof. false and *true* are exact by Proposition 4.3. If a and b are exact, by Proposition 4.2, $-(a \wedge b) = -a \wedge -b = a \wedge b$, hence the exact truth values are closed under $\wedge$. The other closure cases are similar. Neither $\bot$ nor $\top$ are exact by Proposition 4.3 (and the fact that $\bot \neq \top$). The exact truth values are not closed under $\oplus$ because otherwise *true* $\oplus$ *false* $= \top$ would be exact, and it is not. Similarly for $\otimes$.

Again in $\mathcal{FOUR}$, the consistent truth values are also closed under the operations $\vee$, $\bigvee$, $\wedge$, $\bigwedge$ and $\neg$ because, after all, the consistent truth values of $\mathcal{FOUR}$ are just the values of Kleene's system, and the operations listed are Kleene's. Moreover, under the operations arising from the $\leqslant_k$ ordering, there is closure generally under the meet operation Π, and closure under the join operation Σ whenever it is applied to a *directed* set (a set is directed if any two members have a common upper bound in the set). In other words, under the $\leqslant_k$ ordering we have a *complete semi-lattice*. This plays a key role in Kripke's approach, though the terminology itself was not used (see [5]). And once again, we are dealing with general facts.

PROPOSITION 4.5. In an interlaced bilattice with conflation, the consistent truth values include the exact truth values, and are closed under $\wedge$, $\vee$, $\bigwedge$ and $\bigvee$. Further, if f is a k-monotone operation that the conflation operation commutes with, then the consistent truth values are closed under f as well. The consistent truth values are closed under Π, and under Σ when applied to a directed set.

Proof. Every exact truth value is trivially consistent. Suppose S is a set of consistent truth values. Then it follows that $S \leqslant_k -S$ so, $\bigwedge S \leqslant_k \bigwedge -S = -\bigwedge S$ by Proposition 4.2. Hence the consistent truth values are closed under $\bigwedge$. The other closure cases are similar.

Again, let S be a set of consistent truth values, so that $S \leqslant_k -S$. Then using Proposition 4.2, $\Pi S \leqslant_k \Pi - S = -\Sigma S$. But also, $\Pi S \leqslant_k \Sigma S$, so $-\Sigma S \leqslant_k -\Pi S$. It follows that $\Pi S \leqslant_k -\Pi S$, so ΠS is consistent. Thus the consistent truth values are closed under Π.

Finally suppose S is a set of consistent truth values that is also directed by $\leqslant_k$. To show ΣS is consistent we must show $\Sigma S \leqslant_k -\Sigma S = \Pi - S$. To show this, it is enough to show that, for any a, $b \in S$, $a \leqslant_k -b$. But, since S is directed, and $a, b \in S$, there must be some $c \in S$ with $a \leqslant_k c$ and $b \leqslant_k c$. From the second of these inequalities, $-c \leqslant_k -b$. Also since the members of S are consistent, $c \leqslant_k -c$. Combining these inequalities, we have $a \leqslant_k -b$. (This neat argument is from [17].)

5. BILATTICES APPLIED

In Section 4 we established some general properties of interlaced bilattices, and observed that the truth value structures of Sections 2 and 3 met the conditions necessary for them to have these properties. Now we show how a Kripke-like development can be applied using interlaced bilattices, hence to the structures already discussed.

As in previous sections, let L be a first-order language that includes notation for arithmetic. L may or may not also have modal operators. Let $\mathbf{R}$ be the class of atomic sentences of L that are 'about' the real world, that is, are neither arithmetic sentences nor of the form $\mathbf{Tr}(n)$. And let $\mathbf{B}$ be an interlaced bilattice with weak negation, fixed for the rest of this section. If L has modal operators, then we assume there are corresponding operators defined on $\mathbf{B}$, denoted $\Box$ and $\Diamond$, which are monotone with respect to the two partial orderings of $\mathbf{B}$. Finally, let $\mathscr{I}$ be a mapping from the sentences of $\mathbf{R}$ to $\mathbf{B}$. We think of $\mathscr{I}$ as supplying information about the 'real world', and this information may be incomplete or contradictory. That is, $\mathscr{I}$ may map some sentences of **R** to consistent though not exact truth values, and some to members of **B** that are not even consistent truth values (assuming that **B** has a conflation operation, so that notions of consistent and exact are meaningful).

A *valuation* is a mapping v from sentences of L to **B** meeting the usual conditions:

$$\begin{aligned} v(\neg X) &= \neg v(X) \\ v(X \wedge Y) &= v(X) \wedge v(Y) \\ v(X \vee Y) &= v(X) \vee v(Y) \\ v((\forall x)\varphi(x)) &= \bigwedge_{d \in \mathbf{D}} v(\varphi(d)) \\ v((\exists x)\varphi(x)) &= \bigvee_{d \in \mathbf{D}} v(\varphi(d)). \end{aligned}$$

and if L has modal operators, then also:

$$\begin{aligned} v(\Box X) &= \Box(v(X)) \\ v(\Diamond X) &= \Diamond(v(X)). \end{aligned}$$

We continue to write $v \leqslant_k w$, for valuations v and w, provided $v(A) \leqslant_k w(A)$ for *atomic* A. Similarly for $\leqslant_t$. We have assumed the modal operation symbols, if present, are interpreted by monotone operators. The lattice operators are monotone in both orderings. And the negation operation is monotone in $\leqslant_k$, though not in $\leqslant_t$. Then we immediately have the following.

PROPOSITION 5.1. Let v and w be valuations.

1. $v \leqslant_k w \Rightarrow v(X) \leqslant_k w(X)$ for every sentence X;
2. $v \leqslant_t w \Rightarrow v(X) \leqslant_t w(X)$ for every sentence X that does not contain negations.

We continue to call a valuation *realistic* if, for each X in **R**, $v(X) = \mathscr{I}(X)$, and for each atomic sentence X of arithmetic, if X is true in the standard model, $v(X) = true$, and if X is false in the standard model, $v(X) = false$. The collection of all realistic valuations is a complete lattice under both the $\leqslant_k$ ordering and the $\leqslant_t$ ordering. Under the $\leqslant_k$ ordering, the smallest realistic valuation maps all sentences of the form $\mathbf{Tr}(n)$ to $\perp$, while the smallest realistic valuation under the $\leqslant_t$ ordering maps such sentences to *false*. We will discontinue discussion of the $\leqslant_t$ order until Section 6 because of the limitation on negation in Proposition 5.1. The only role of this ordering, for now, is to supply a natural lattice-theoretic interpretation for the logical connectives $\wedge$, $\vee$, and the quantifiers.

Next, we define a mapping Φ on valuations just as we did in earlier sections. We repeat the definition for convenience. For a valuation v, $\Phi(v) = w$ where w is the valuation whose atomic behavior is:

1. if X is an arithmetic sentence, $w(X)$ is the truth value of X in the standard model;
2. if $X \in \mathbf{R}$, $w(X) = \mathscr{I}(X)$,
3. $w(\mathbf{Tr}(\ulcorner X \urcorner)) = v(X)$.

We have directly from Proposition 5.1, and a small amount of work:

PROPOSITION 5.2. Φ maps realistic valuations to realistic valuations, and $v \leqslant_k w \Rightarrow \Phi(v) \leqslant_k \Phi(w)$.

The Knaster–Tarski Theorem applies, and says that Φ has a smallest fixed point (under the $\leqslant_k$ ordering). Indeed, it says Φ has a largest fixed point, and the collection of all fixed points is a complete lattice. At any rate, the smallest fixed point is the one of most interest: it provides a 'meaning' for L, including the problematic predicate $\mathbf{Tr}$, that is coherent, though partial, and involves no 'unnecessary knowledge'.

The usual proof of the Knaster–Tarski Theorem actually establishes a stronger result that is often considered to be a kind of generalized induction. Suppose $\mathbf{T}$ is a monotone map from a complete lattice to itself. Then not only will $\mathbf{T}$ have a smallest fixed point $v_{\mathbf{T}}$, but if w is any member of the lattice such that $\mathbf{T}(w) \leqslant w$, then $v_{\mathbf{T}} \leqslant w$. This has an important application here. Recall, Φ depends on $\mathscr{I}$, and so we write $\Phi_{\mathscr{I}}$ to emphasize this.

PROPOSITION 5.3. Suppose $\mathscr{I}$ and $\mathscr{J}$ are both mappings from **R** to **B** and $\mathscr{I} \leqslant_k \mathscr{J}$. Let $v_{\mathscr{I}}$ be the least fixed point of $\Phi_{\mathscr{I}}$ (in the $\leqslant_k$ ordering), and $v_{\mathscr{J}}$ be the least fixed point of $\Phi_{\mathscr{J}}$. Then $v_{\mathscr{I}} \leqslant_k v_{\mathscr{J}}$.

Proof. It is quite easy to see that, since $\mathscr{I} \leqslant_k \mathscr{J}$ then for every valuation v, $\Phi_{\mathscr{I}}(v) \leqslant_k \Phi_{\mathscr{J}}(v)$. Then $\Phi_{\mathscr{I}}(v_{\mathscr{J}}) \leqslant_k \Phi_{\mathscr{J}}(v_{\mathscr{J}}) = v_{\mathscr{J}}$, since $v_{\mathscr{J}}$ is a fixed point of $\Phi_{\mathscr{J}}$. It follows that the *least* fixed point of $\Phi_{\mathscr{I}}$ must be below $v_{\mathscr{J}}$, that is, $v_{\mathscr{I}} \leqslant_k v_{\mathscr{J}}$.

It is this proposition that substantiates earlier claims that, as time goes on, and more information is gained about the 'real world', no knowledge is lost, where by knowledge we mean that supplied by the least fixed point model.

Now suppose that **B** is also an interlaced bilattice *with conflation.* Then notions of exact and consistent truth values are applicable. We also assume that negation and the modal operators (if present) commute with conflation. It is now an easy consequence of Proposition 4.5 that, if $\mathscr{I}$ maps all members of **R** to consistent members of **B**, then the least fixed point of $\Phi_{\mathscr{I}}$ will also take on only consistent truth values. We omit the proof of this, but it accounts for why, in most earlier treatments, truth value 'gluts' could be safely ignored. They do not come up unless they are forced to. On the other hand, one can easily create a language more general than L, which contains propositional connectives corresponding to the bilattice operations $\oplus$ and $\otimes$. If we were working with such a language, dealing with inconsistent truth values would be common, because of results like Proposition 4.1.

6. EXTREMAL FIXED POINTS

We have been considering fixed points of operators assigning truth values; fixed points that are smallest in the $\leqslant_k$ ordering. According to the Knaster–Tarski Theorem, monotone operators on a complete lattice have *biggest* as well as smallest fixed points. Moreover, there is another lattice ordering, $\leqslant_t$ present. It has supplied us with meanings for logical connectives and quantifiers, but it could not play a role like that of $\leqslant_k$ because the presence of negations destroys monotonicity with respect to this ordering. However, Proposition 5.1 implies that monotonicity still obtains under $\leqslant_t$ provided no negations are present. With no negations allowed, liar sentences are no longer possible, though sentences asserting their own truth are.

> For the rest of this section we assume L is a first order language without negation.

Under this assumption the mapping Φ defined in the previous section is monotone under both the $\leqslant_k$ and the $\leqslant_t$ orderings, and hence *four* extremal fixed points exist. It turns out that, under certain restrictions on the underlying interlaced bilattice of truth values, there are rather elegant relationships between them. We devote the rest of this section to stating and proving them. We begin by introducing the appropriate interlaced bilattice restrictions.

A bilattice has four basic binary operations, $\otimes$, $\oplus$, $\wedge$ and $\vee$, and so there are twelve possible distributive laws:

$$a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c)$$

$$a \wedge (b \oplus c) = (a \wedge b) \oplus (a \wedge c)$$

etc.

DEFINITION 6.1. We call an interlaced bilattice distributive if all twelve distributive laws hold.

Distributive interlaced bilattices are fairly common; $\mathcal{FOUR}$ for instance is distributive. We have the following simple general result whose proof we omit.

PROPOSITION 6.1. If C and D are distributive lattices then $\mathcal{B}(C, D)$ is a distributive interlaced bilattice.

In fact, distributivity implies the interlacing condition. Suppose B has two partial orderings, $\leqslant_k$ and $\leqslant_t$, each of which makes B into a complete lattice. And suppose all twelve distributive laws hold. Then we can argue that $a \leqslant_k b$ implies $a \wedge c \leqslant_k b \wedge c$ as follows.

$a \leqslant_k b$ (assumption)

$a \oplus b = b$

$$(a \oplus b) \wedge c = b \wedge c$$

$$(a \wedge c) \oplus (b \wedge c) = b \wedge c \text{ (distributivity)}$$

$$a \wedge c \leqslant_k b \wedge c.$$

Similarly for other order/operation combinations.

DEFINITION 6.2. For the mapping Φ, the least and greatest fixed points under the $\leqslant_k$ ordering are denoted v_k and V_k; the least and greatest fixed points under the $\leqslant_t$ ordering are v_t and V_t.

Now the main result of this section is easily stated.

THEOREM 6.2. If the interlaced bilattice **B** of truth values is distributive and finite. Then:

1. $v_k = v_t \otimes V_t$
2. $V_k = v_t \oplus V_t$
3. $v_t = v_k \wedge V_k$
4. $V_t = v_k \vee V_k$

These interconnections between the extremal fixed points place severe restrictions on what is possible, but also provide interesting side information. We discuss this briefly with respect to item (1); the other three items have similar consequences.

Item (1) can be read as saying that v_k is a consensus model between those that are least and greatest in the $\leqslant_t$ ordering. Suppose X is a sentence asserting its own truth. In Kripke's least fixed point model, whose construction we can carry our in $\mathcal{FOUR}$, X is underdetermined, $\perp$, although there are models in which X is *true* and others in which X is *false*. In fact, it is easy to see that $v_t(X) = \textit{false}$ and $V_t(X) = \textit{true}$, and Theorem 6.2, part (1), says $v_k(X) = \textit{false} \otimes \textit{true} = \perp$.

Finally, suppose we continue using $\mathcal{FOUR}$, and in addition that $\mathcal{I}$ maps all members of **R** to consistent truth values. As we observed at the end of Section 5, then v_k can only take on consistent truth values, and we are essentially in the setting Kripke used, except for the absence of negations. Under these circumstances, item (1) above implies the following for atomic **X**:

$$v_k(X) = \textit{true} \text{ if and only if } v_t(X) = \textit{true}$$

$$v_k(X) = \textit{false} \text{ if and only if } V_t(X) = \textit{false}.$$

For instance, suppose $v_t(X) = \textit{true}$. Since $v_t \leqslant_t V_t$, then $V_t(X) = \textit{true}$ and so $v_k(X) = \textit{true} \otimes \textit{true} = \textit{true}$. Conversely, suppose $v_t(X) \neq \textit{true}$. Then $v_t(X)$ must be $\bot$ or *false*, and either way, $v_t(X) \otimes V_t(X)$ cannot be *true*.

The rest of this section is devoted to a proof of the result above.

LEMMA 6.3. Suppose that in a distributive interlaced bilattice, $a_1 \leqslant_t A_1$ and $a_2 \leqslant_t A_2$ (or $a_1 \leqslant_k A_1$ and $a_2 \leqslant_k A_2$). Then:

$$(a_1 \otimes A_1) \wedge (a_2 \otimes A_2) = (a_1 \wedge a_2) \otimes (A_1 \wedge A_2),$$

$$(a_1 \oplus A_1) \wedge (a_2 \oplus A_2) = (a_1 \wedge a_2) \oplus (A_1 \wedge A_2),$$

$$(a_1 \wedge A_1) \otimes (a_2 \wedge A_2) = (a_1 \otimes a_2) \wedge (A_1 \otimes A_2),$$

etc.

Proof. Using a distributive law:

$$(a_1 \otimes A_1) \wedge (a_2 \otimes A_2) = (a_1 \wedge a_2) \otimes (a_1 \wedge A_2) \otimes (A_1 \wedge a_2) \otimes (A_1 \wedge A_2) = *$$

If $a_1 \leqslant_t A_1$ and $a_2 \leqslant_t A_2$ then using the basic bilattice monotonicity conditions:

$$* \leqslant_t (a_1 \wedge a_2) \otimes (A_1 \wedge A_2) \otimes (A_1 \wedge A_2) \otimes (A_1 \wedge A_2) = (a_1 \wedge a_2) \otimes (A_1 \wedge A_2)$$

Similarly

$$* \geqslant_t (a_1 \wedge a_2) \otimes (a_1 \wedge a_2) \otimes (a_1 \wedge a_2) \otimes (A_1 \wedge A_2) = (a_1 \wedge a_2) \otimes (A_1 \wedge A_2).$$

The other parts of the Lemma are proved in a similar way.

LEMMA 6.4. Suppose that in a finite, distributive interlaced bilattice, $a_i \leqslant_t A_i$ for each $i \in I$ (or $a_i \leqslant_k A_i$ for each $i \in I$). Then

$$\bigwedge_{i\in I} (a_i \otimes A_i) = \bigwedge_{i\in I} a_i \otimes \bigwedge_{i\in I} A_i$$

$$\bigwedge_{i\in I} (a_i \oplus A_i) = \bigwedge_{i\in I} a_i \oplus \bigwedge_{i\in I} A_i$$

$$\prod_{i\in I} (a_i \wedge A_i) = \prod_{i\in I} a_i \wedge \prod_{i\in I} A_i$$

etc.

Proof. Since the interlaced bilattice is assumed to be finite, 'infinitary' operations are really 'finitary' ones, so the proof is like that of the previous lemma.

Next, we can define operations $\oplus$, etc. on the space of valuations, in a pointwise way.

DEFINITION 6.3. For valuations v and w, $v \oplus w$ is the valuation given by: $(v \oplus w)(A) = v(A) \oplus w(A)$ for atomic A. Similarly for $\otimes$, $\vee$ and $\wedge$.

The definition above contains a restriction to the *atomic* level. The extension to arbitrary formulas does not always follow, but the following result gives circumstances under which it does.

PROPOSITION 6.5. Suppose **B** is a finite, distributive interlaced bilattice, v and V are valuations of the language L in **B**, and φ is a sentence of L (hence without negations). If $v \leqslant_t V$ or $v \leqslant_k V$ then:

1. $(v \oplus V)(\varphi) = v(\varphi) \oplus V(\varphi)$,
2. $(v \otimes V)(\varphi) = v(\varphi) \otimes V(\varphi)$,
3. $(v \wedge V)(\varphi) = v(\varphi) \wedge V(\varphi)$,
4. $(v \vee V)(\varphi) = v(\varphi) \vee V(\varphi)$,

Proof. By induction on the complexity of φ. If φ is an atomic sentence, the result is immediate from the definition, and does not use the inequalities between v and V. We consider one of the induction cases, as a representative example. Suppose φ is $(\exists x)\psi(x)$, and 1) is known for formulas simpler than φ, in particular for $\psi(d)$ for each $d \in \mathbf{D}$. Then

$$\begin{aligned}
(v \oplus V)(\varphi) &= (v \oplus V)((\exists x)\psi(x)) \\
&= \bigvee_{d \in \mathbf{D}} (v \oplus V)(\varphi(d)) \\
&= \bigvee_{d \in \mathbf{D}} [v(\psi(d)) \oplus V(\psi(d))] \\
&\quad \text{(by induction hypothesis)} \\
&= \bigvee_{d \in \mathbf{D}} v(\psi(d)) \oplus \bigvee_{d \in \mathbf{D}} V(\psi(d)) \\
&\quad \text{(by Lemma 6.4)} \\
&= v((\exists x)\psi(x)) \oplus V((\exists x)\psi(x)) \\
&= v(\varphi) \oplus V(\varphi).
\end{aligned}$$

Finally we come to the proof of Theorem 6.2 itself.

Proof. Smallest and biggest fixed points are approached via a transfinite sequence of approximations. We use the following notation to represent this. $(v_t)_0$ is the smallest realistic valuation in the $\leqslant_t$ direction. For an arbitrary ordinal α, $(v_t)_{\alpha+1} = \Phi((v_t)_\alpha)$. For a limit ordinal λ, $(v_t)_\lambda = \bigvee_{\alpha<\lambda} (v_t)_\alpha$. As usual, the sequence $(v_t)_\alpha$ is monotone increasing in α ($\alpha < \beta$ implies $(v_t)_\alpha \leqslant_t (v_t)_\beta$). And for some ordinal ∞, $(v_t)_\infty = v_t$, the least fixed point of Φ in the $\leqslant_t$ ordering. In fact, from that stage on, things remain fixed; that is, if $\alpha \geqslant \infty$ then $(v_t)_\alpha = v_t$.

More notation. $(V_t)_0$ is the largest realistic valuation in the $\leqslant_t$ ordering. $(V_t)_{\alpha+1} = \Phi((V_t)_\alpha)$. And for limit λ, $(V_t)_\lambda = \bigwedge_{\alpha<\lambda} (V_t)_\alpha$. Then $(V_t)_\alpha$ decreases in the $\leqslant_t$ ordering, and for some ordinal ∞, $(V_t)_\infty = V_t$. Finally we use $(v_k)_\alpha$ and $(V_k)_\alpha$ analogously, but with $\leqslant_k$, Σ and Π playing the rules that $\leqslant_t$, $\bigvee$ and $\bigwedge$ played above.

Now to show item (1) of Theorem 6.2, $v_k = v_t \otimes V_t$, it is enough to show by transfinite induction that, for each ordinal α, $(v_k)_\alpha = (v_t)_\alpha \otimes (V_t)_\alpha$. Items (2), (3) and (4) are proved in exactly the same way.

Initial Case. Let A be an atomic sentence. The initial valuations are smallest and greatest in their respective orderings. If A is arithmetic, or in **R**, all realistic valuations will agree on A and so, trivially, $(v_k)_0 = (v_t)_0 \otimes (V_t)_0$. If A is neither arithmetic nor in **R**, then $(v_k)_0(A) = \bot$, $(v_t)_0(A) = \mathit{false}$ and $(V_t)_0(A) = \mathit{true}$, so $(v_k)_0(A) = (v_t)_0(A) \otimes (V_t)_0(A)$ by Proposition 4.1. Thus $(v_k)_0 = (v_t)_0 \otimes (V_t)_0$.

Induction Case. Suppose $(v_k)_\alpha = (v_t)_\alpha \otimes (V_t)_\alpha$. Let A be an atomic sentence; we show $(v_k)_{\alpha+1}(A) = (v_t)_{\alpha+1}(A) \otimes (V_t)_{\alpha+1}(A)$. The subcases where A is arithmetic or in **R** are trivial, and are treated as in the previous case. Now suppose A is $\mathbf{Tr}(\ulcorner X \urcorner)$ for some sentence X.

$$\begin{aligned}
(v_k)_{\alpha+1}(A) &= (v_k)_{\alpha+1}(\mathbf{Tr}(\ulcorner X \urcorner)) \\
&= \Phi((v_k)_\alpha)(\mathbf{Tr}(\ulcorner X \urcorner)) \\
&= (v_k)_\alpha(X) \\
&= [(v_t)_\alpha \otimes (V_t)_\alpha](X) \\
&\quad \text{(by induction hypothesis)} \\
&= (v_t)_\alpha(X) \otimes (V_t)_\alpha(X) \quad \text{(by Proposition 6.5)} \\
&= \Phi((v_t)_\alpha(\mathbf{Tr}(\ulcorner X \urcorner)) \otimes \Phi((V_t)_\alpha(\mathbf{Tr}(\ulcorner X \urcorner)) \\
&= (v_t)_{\alpha+1}(\mathbf{Tr}(\ulcorner X \urcorner)) \otimes (V_t)_{\alpha+1}(\mathbf{Tr}(\ulcorner X \urcorner)) \\
&= (v_t)_{\alpha+1}(A) \otimes (V_t)_{\alpha+1}(A).
\end{aligned}$$

Limit Case. Suppose $(v_k)_\alpha = (v_t)_\alpha \otimes (V_t)_\alpha$ for every $\alpha < \lambda$, where λ is a limit ordinal. Let A be an atomic sentence. Then $(v_k)_\lambda(A) = (\Sigma_{\alpha<\lambda} (v_k)_\alpha)(A) = \Sigma_{\alpha<\lambda} (v_k)_\alpha(A)$. Since the interlaced bilattice **B** is finite, and $(v_k)_\alpha$ is increasing with α in the $\leqslant_k$ ordering, there must be an $\alpha_0 < \lambda$ so that $\Sigma_{\alpha<\lambda}(v_k)_\alpha(A) = (v_k)_{\alpha_0}(A)$. Further, for any ordinal β with $\alpha_0 \leqslant \beta \leqslant \lambda$ we must have $(v_k)_{\alpha_0}(A) = (v_k)_\beta(A) = (v_k)_\lambda(A)$. Similarly, using the facts that $(v_t)_\alpha$ is increasing and $(V_t)_\alpha$ is decreasing in the $\leqslant_t$ ordering, there must be ordinals $\alpha_1, \alpha_2 < \lambda$ such that $\alpha_1 \leqslant \beta \leqslant \lambda \Rightarrow (v_t)_{\alpha_1}(A) = (v_t)_\beta(A) = (v_t)_\lambda(A) = \bigvee_{\alpha<\lambda}(v_t)_\alpha(A)$ and $\alpha_2 \leqslant \beta \leqslant \lambda \Rightarrow (V_t)_{\alpha_2}(A) = (V_t)_\beta(A) = (V_t)_\lambda(A) = \bigwedge_{\alpha<\lambda} (V_t)_\alpha(A)$. Now, let $\gamma = \max\{\alpha_0, \alpha_1, \alpha_2\}$. Then, since $\gamma < \lambda$ we can use the induction hypothesis, and so:

$$\begin{aligned}
(v_k)_\lambda(A) &= (v_k)_\gamma(A) \\
&= [(v_t)_\gamma \otimes (V_t)_\gamma](A)
\end{aligned}$$

$$= (v_t)_\gamma(A) \otimes (V_t)_\gamma(A)$$
$$= (v_t)_\lambda(A) \otimes (V_t)_\lambda(A).$$

Hence $(v_k)_\lambda = (v_t)_\lambda \otimes (V_t)_\lambda$.
This concludes the proof.

7. CONCLUSION

While Kripke's original paper on the theory of truth used a three-valued logic, we believe a four-valued version is more natural. Its use allows for possible inconsistencies in information about the world, yet contains Kripke's development within it. Moreover, using a four-valued logic makes it possible to work with complete *lattices* rather than complete *semi-lattices*, and thus the mathematics is somewhat simplified. But more strikingly, the four-valued version has a wide, natural generalization to the family of interlaced bilattices. Thus, with little more work, the theory is extended to a broad class of settings. Indeed, a result like Theorem 6.2 would not even be possible to state without the interlaced bilattice machinery. We hope the notion of interlaced bilattice will make apparent further such connections.

ACKNOWLEDGEMENT

Revised version of a paper entitled *Kripke's Theory of Truth, Generalized*, Research partly supported by NSF Grant CCR-8702307 and PSC-CUNY Grants 6-67295 and 6-68283.

REFERENCES

[1] J. Barwise, *Admissible Sets and Structures*, Springer-Verlag, Berlin (1975).
[2] N. D. Belnap, Jr. A useful four-valued logic, in *Modern Uses of Multiple-Valued Logic*, J. Michael Dunn and G. Epstein, eds, pp. 8–37, D. Reidel (1977).
[3] J. M. Dunn, Intuitive semantics for first-degree entailments and 'coupled trees', *Philosophical Studies*, vol. 29, pp. 149–168 (1976).
[4] M. C. Fitting, A Kripke/Kleene semantics for logic programs, *Journal of Logic Programming*, pp. 295–312 (1985).
[5] M. C. Fitting, Notes on the mathematical aspects of Kripke's Theory of Truth, *Notre Dame Journal of Formal Logic*, vol. 27, pp. 75–88 (1986).
[6] M. C. Fitting, Logic programming on a topological bilattice, *Fundamenta Informatica*, vol. 11, pp. 209–218 (1988).

[7] M. C. Fitting, Bilattices and the semantics of logic programming, *Journal of Logic Programming* (to appear).
[8] M. L. Ginsberg, Multi-valued logics, *Proc. AAAI-86, Fifth National Conference on Artificial Intelligence*, pp. 243–247, Morgan Kaufmann Publishers (1986).
[9] M. L. Ginsberg, Multivalued logics: A uniform approach to inference in artificial intelligence, *Computational Intelligence*, Vol. 4, no. 3.
[10] S. C. Kleene, *Introduction to Metamathematics*, Van Nostrand (1950).
[11] S. Kripke, Semantical analysis of modal logic I, normal propositional calculi, *Zeitschrift für mathematische Logik und Grundlagen der Mathematik*, Vol. 9, pp. 67–96 (1963).
[12] S. Kripke, Semantical considerations on modal logics, *Acta Philosophica Fennica, Modal and Many-Valued Logics*, pp. 83–94 (1963).
[13] S. Kripke, Outline of a theory of truth, *The Journal of Philosophy*, vol. 72, pp. 690–716 (1975); reprinted in *New Essays on Truth and The Liar Paradox*, R. L. Martin, ed., Oxford (1983).
[14] R. L. Martin and P. W. Woodruff, On representing "true-in-L" in L, *Philosophia (Israel)*, vol. 5, pp. 213–217 (1975).
[15] B. K. Matilal, *Logic, Language and Reality*, Motilal Banardsidass, Delhi, 1985.
[16] A. Tarski, A lattice-theoretical fixpoint theorem and its applications, *Pacific Journal of Mathematics*, vol. 5, pp. 285–309 (1955).
[17] A. Visser, Four valued semantics and the liar, *Journal of Philosophical Logic*, vol. 13, pp. 181–212 (1984).
[18] P. W. Woodruff, Paradox, truth and logic, part I: Paradox and truth, *Journal of Philosophical Logic*, vol. 13, pp. 213–232 (1984).

Department of Mathematics and Computer Science,
Lehman College (CUNY),
Bronx, NY 10468,
U.S.A.