TPLP **22** (2): 305–334, 2022.

doi:10.1017/S1471068421000168 First published online 1 September 2021

Alternating Fixpoint Operator for Hybrid MKNF Knowledge Bases as an Approximator of AFT

FANGFANG LIU
School of Computer Engineering and Science, Shanghai University, Shanghai, China
(*e-mail:* ffliu@shu.edu.cn)

JIA-HUAI YOU
Department of Computing Science, University of Alberta, Edmonton, Canada
(*e-mail:* jyou@ualberta.ca)

submitted 6 May 2020; revised 8 July 2021; accepted 8 July 2021

Abstract

Approximation fixpoint theory (AFT) provides an algebraic framework for the study of fixpoints of operators on bilattices and has found its applications in characterizing semantics for various classes of logic programs and nonmonotonic languages. In this paper, we show one more application of this kind: the alternating fixpoint operator by Knorr et al. for the study of the well-founded semantics for hybrid minimal knowledge and negation as failure (MKNF) knowledge bases is in fact an approximator of AFT in disguise, which, thanks to the abstraction power of AFT, characterizes not only the well-founded semantics but also two-valued as well as three-valued semantics for hybrid MKNF knowledge bases. Furthermore, we show an improved approximator for these knowledge bases, of which the least stable fixpoint is information richer than the one formulated from Knorr et al.'s construction. This leads to an improved computation for the well-founded semantics. This work is built on an extension of AFT that supports consistent as well as inconsistent pairs in the induced product bilattice, to deal with inconsistencies that arise in the context of hybrid MKNF knowledge bases. This part of the work can be considered generalizing the original AFT from symmetric approximators to arbitrary approximators.

KEYWORDS: approximation fixpoint theory, hybrid MKNF knowledge bases, logic programs, answer set semantics, description logics, inconsistencies

1 Introduction

Approximation fixpoint theory (AFT) is a framework for the study of semantics of nonmonotonic logics based on operators and their fixpoints (Denecker *et al*. 2004). Under this theory, the semantics of a logic theory is defined or characterized in terms of respective stable fixpoints constructed by employing an *approximator* on a (product) bilattice. The least stable fixpoint of such an approximator is called the well-founded fixpoint, which serves as the basis for a well-founded semantics, and the stable fixpoints that are total characterize a stable semantics, while partial stable fixpoints give rise to a partial stable semantics. The approach is highly general as it only depends on mild conditions on approximators, and highly abstract as well since the semantics is given in terms of an algebraic structure. As different approximators may represent different intuitions, AFT provides a powerful framework to treat semantics uniformly and allows to explore alternative semantics by different approximators.

Due to the underlying algebraic structure, a main feature of AFT is that we can understand some general properties of a semantics without referring to a concrete approximator. For example, the well-founded fixpoint approximates all other fixpoints, and mathematically, this property holds for all approximators. An implication of this property is that it provides the bases for building constraint propagators for solvers; for logic programs for example, it guarantees that the true and false atoms in the well-founded fixpoint remain to hold in all stable fixpoints, and as such, the computation for the well-founded fixpoint can be adopted as constraint propagation for the computation of stable fixpoints. For example, this lattice structure of stable fixpoints has provided key technical insights in building a DPLL-based solver for normal hybrid MKNF knowledge bases (Ji *et al*. 2017), while previously the only known computational method was based on guess-and-verify (Motik and Rosati 2010).

AFT has been applied to default logic as well as autoepistemic logic, and the study has shown how the fixpoint theory induces the main and sometimes new semantics and leads to new insights in these logics (Denecker *et al*. 2003), including the well-founded semantics for autoepistemic logic (Bogaerts *et al*. 2016). AFT has been adopted in the study of the semantics of logic programs with aggregates (Pelov *et al*. 2007) and disjunctive HEX programs (Antic *et al*. 2013). Vennekens *et al.* (2006) used AFT in a modularity study for a number of nonmonotonic logics, and by applying AFT, Strass (2013) showed that many semantics from Dung's argumentation frameworks and abstract dialectical frameworks (Dung 1995) can be obtained rather directly. More recently, AFT has been shown to play a key role in the study of semantics for database revision based on active integrity constraints (Bogaerts and Cruz-Filipe 2018) and in addressing semantics issues arising in weighted abstract dialectical frameworks, which are abstract argumentation frameworks that incorporate not only attacks but also support, joint attacks and joint support (Bogaerts 2019). AFT has also contributed to the study of induction (Bogaerts *et al*. 2018) and knowledge compilation (Bogaerts and den Broeck 2015).

In this paper, we add one more application to the above collection for hybrid MKNF (which stands for minimal knowledge and negation as failure). Hybrid MKNF was proposed by Motik and Rosati (2010) for integrating nonmonotonic rules with description logics (DLs). Since reasoning with DLs is based on classic, monotonic logic, there is no support of nonmonotonic features such as defeasible inheritance or default reasoning. On the other hand, rules under the stable model semantics (Gelfond and Lifschitz 1988) are formulated mainly to reason with ground knowledge, without supporting quantifiers or function symbols. It has been argued that such a combination draws strengths from both and the weaknesses of one are balanced by the strengths of the other. The formalism of hybrid MKNF knowledge bases provides a tight integration of rules with DLs.

A hybrid MKNF knowledge base $\mathscr{K}$ consists of two components, $\mathscr{K} = (\mathscr{O}, \mathscr{P})$, where $\mathscr{O}$ is a DL knowledge base, which is expressed by a decidable first-order theory, and $\mathscr{P}$ is a collection of MKNF rules based on the stable model semantics. MKNF structures in

https://doi.org/10.1017/S1471068421000168 Published online by Cambridge University Press

this case are two-valued, under which MKNF formulas are interpreted to be true or false. Knorr *et al.* (2011) formulated a three-valued extension of MKNF and defined three-valued MKNF models, where the least one is called the *well-founded MKNF model*. An alternating fixpoint operator was then formulated for the computation of the well-founded MKNF model for (non-disjunctive) hybrid MKNF knowledge bases. In this paper, our primary goal is to show that this alternating fixpoint operator is in fact an approximator of AFT. Due to the abstraction power of AFT, it turns out that Knorr et al.'s alternating fixpoint construction provides a uniform characterization of all semantics based on various kinds of three-valued MKNF models, including two-valued MKNF models of Motik and Rosati (2010).

As shown in previous research (Knorr *et al*. 2011; Liu and You 2017), not all hybrid MKNF knowledge bases possess a well-founded MKNF model, and in general, deciding the existence of a well-founded MKNF model is intractable even if the underlying DL knowledge base is polynomial (Liu and You 2017). On the other hand, we also know that alternating fixpoint construction provides a tractable means in terms of a linear number of iterations to compute the well-founded MKNF model for a subset of hybrid MKNF knowledge bases. A question then is whether this subset can be enlarged. In this paper, we answer this question positively by formulating an improved approximator, which is *more precise* than the one derived from Knorr et al.'s alternating fixpoint operator. As a result, the well-founded MKNF model can be computed iteratively for a strictly larger class of hybrid MKNF knowledge bases than what was known previously.

Hybrid MKNF combines two very different reasoning paradigms, namely closed world reasoning with nonmonotonic rules and open world reasoning with ontologies that are expressed in DLs. In this context, inconsistencies naturally arise. AFT was first developed for consistent approximations. In the seminal work (Denecker *et al*. 2004), the authors show that the theory of consistent approximations generalizes to a class of approximators beyond consistent pairs, which are called symmetric approximators. They also state that it is possible to develop a generalization of AFT without the symmetry assumption. These results and claims are given under the restriction that an approximator maps an exact pair on a product bilattice (which represents a two-valued interpretation) to an exact pair. Unfortunately, this assumption is too restrictive for hybrid MKNF since a two-valued interpretation for a hybrid MKNF knowledge base may well lead to an inconsistent state.

Approximations under symmetric approximators already provide a powerful framework for characterizing intended models of a logic theory. But we want to go beyond that. We do not only want to capture consistent approximations in the product bilattice, but also want to allow operators to map a consistent state to an inconsistent one, and even allow inconsistent stable fixpoints. This is motivated by the possible role that AFT may play in building constraint propagators for solvers of an underlying logic (e.g., Ji *et al*. 2017), where inconsistency not only guides the search via backtracking but also provides valuable information to prune the search space (e.g., by learned clauses in SAT/ASP solvers). One can also argue that inconsistent stable fixpoints may provide useful information for debugging purposes (a potential topic beyond the scope of this paper).

We show in this paper that all of the above requires only a mild generalization of AFT, which is defined for all pairs in the product bilattice without the assumption of symmetry.

We relax the condition for an approximator so that an approximator is required to map an exact pair to an exact pair only in the case of consistent approximation. Based on this revised definition of approximator, we present a definition of the stable revision operator, which is well defined, increasing, and monotone on the product bilattice of a complete lattice, that guarantees existence of fixpoints and a least fixpoint.

In summary, we extend AFT from consistent and symmetric approximators to arbitrary approximators for the entire product bilattice. The goal is to use stable fixpoints as candidates for intended models, or to provide useful information on stable states (in terms of fixpoints) that may contain consistent as well as inconsistent information. Such an extension is not without subtleties. We provide a detailed account of how such technical subtleties are addressed.

The paper is organized as follows. The next section introduces notations, basics of fixpoint theory, and the current state of AFT. In Section 3, we present an extended AFT. Section 4 gives a review of three-valued MKNF and hybrid MKNF knowledge bases along with the underlying semantics. Then, in Section 5 we show how Knorr *et al.*'s alternating fixpoint operator can be recast as an approximator and provide semantic characterizations, and in Section 6, we show an improved approximator. Section 7 is about related work, concluding remarks, and future directions.

This paper is revised and extended from a preliminary report of the work that appeared in Liu and You (2019). The current paper is reorganized by first presenting a detailed study of generalized AFT. Especially, we provide an elaborate account of the original AFT and contrast it with our generalization. In this extended version of the work, all claims are complete with a proof.

2 Preliminaries

In this section, we recall the basic definitions regarding lattices underlying our work based on the Knaster–Tarski fixpoint theory (Tarski 1955).

A *partially ordered set* $\langle L, \leq \rangle$ is a set L equipped with a partial order $\leq$, which is a reflexive, antisymmetric, and transitive relation. As usual, the strict order is expressed by $x < y$ as an abbreviation for $x \leq y$ and $x \neq y$. Given a subset $S \subseteq L$, an element $x \in L$ is an *upper bound* (resp. *a lower bound*) if $s \leq x$ (resp. $x \leq s$) for all $s \in S$. A *lattice* $\langle L, \leq \rangle$ is a *partially ordered set* (poset) in which every two elements have a *least upper bound* (lub) and a *greatest lower bound* (glb). A *complete lattice* is a lattice where every subset of L has a least upper bound and a greatest lower bound. A complete lattice has both a least element $\bot$ and a greatest element $\top$. A greatest lower bound of a subset $S \subseteq L$ is called a *meet* and a least upper bound of S is called a *join*, and we use the notations: $\bigwedge S = glb(S)$, $x \wedge y = glb(\{x, y\})$, $\bigvee S = lub(S)$, and $x \vee y = lub(\{x, y\})$. An operator O on L is *monotone* if for all $x, y \in L$, that $x \leq y$ implies $O(x) \leq O(y)$. An element $x \in L$ is a *pre-fixpoint* of O if $O(x) \leq x$; it is a *post-fixpoint* of O if $x \leq O(x)$. The Knaster–Tarski fixpoint theory (Tarski 1955) tells us the fact that a monotone operator O on a complete lattice has fixpoints and a least fixpoint, denoted $lfp(O)$, which coincides with its least pre-fixpoint. The following result of Knaster–Tarski fixpoint theory (Tarski 1955) serves as the basis of our work in this paper.

Theorem 1

Let $\langle L, \leq \rangle$ be a complete lattice and O a monotone operator on L. Then O has fixpoints, a least fixpoint, and a least pre-fixpoint. (i) The set of fixpoints of O is a complete lattice under order $\leq$. (ii) The least fixpoint and least pre-fixpoint of O coincide, that is, $lfp(O) = \wedge\{x \in L : O(x) \leq x\}$.

A *chain* in a poset $\langle L, \leq \rangle$ is a linearly ordered subset of L. A poset $\langle L, \leq \rangle$ is *chain-complete* if it contains a least element $\bot$ and every chain $C \subseteq L$ has a least upper bound in L. A complete lattice is chain-complete, but the converse does not hold in general. However, as pointed out by Denecker *et al.* (2004), the Knaster–Tarski fixpoint theory generalizes to chain-complete posets.

Theorem 2

(Markowsky 1976) Let $\langle L, \leq \rangle$ be a chain-complete poset and O a monotone operator on L. Then O has fixpoints, a least fixpoint, and a least pre-fixpoint. (i) The set of fixpoints of O is a chain-complete poset under order $\leq$. (ii) The least fixpoint and least pre-fixpoint of O coincide.

Given a complete lattice $\langle L, \leq \rangle$, AFT is built on the induced product bilattice $\langle L^2, \leq_p \rangle$, where $\leq_p$ is called the *precision order* and defined as: for all $x, y, x', y' \in L$, $(x, y) \leq_p (x', y')$ if $x \leq x'$ and $y' \leq y$. The $\leq_p$ ordering is a complete lattice ordering on L^2. Below, we often write a lattice $\langle L, \leq \rangle$ by L and its induced product bilattice by L^2.

We define two *projection functions* for pairs in L^2: $(x, y)_1 = x$ and $(x, y)_2 = y$. For simplicity, we write $A(x, y)_i$, where $i \in [1, 2]$, instead of more formal $(A(x, y))_i$ to refer to the corresponding projection of the value of the operator A on the pair (x, y). A pair $(x, y) \in L^2$ is *consistent* if $x \leq y$, *inconsistent* otherwise, and *exact* if $x = y$. A consistent pair (x, y) in L defines an *interval*, denoted $[x, y]$, which is identified by the set $\{z \mid x \leq z \leq y\}$. We therefore also use an interval to denote the corresponding set. A consistent pair (x, y) in L can be seen as an approximation of every $z \in L$ such that $z \in [x, y]$. In this sense, the precision order $\leq_p$ corresponds to the precision of approximation, while an exact pair approximates the only element in it. We denote by L^c the set of consistent pairs in L^2. Note that $\langle L^c, \leq_p \rangle$ is not a complete lattice in general.

On the other hand, an inconsistent pair (x, y) in L^2 can be viewed as a departure from some point $z \in L$, for which (z, z) is revised either by increasing the first component of the pair (w.r.t. the order $\leq$), or by decreasing its second component, or by performing both at the same time. Inconsistent pairs have a natural embedding of the notion of *the degree of inconsistency*. For two inconsistent pairs such that $(x_1, y_1) \leq_p (x_2, y_2)$, the latter is of higher degree of inconsistency than the former. Here, there is a natural notion of inconsistency being *partial* as in contrast with full inconsistency represented by the special pair $(\top, \bot)$. Intuitively, this means that an inconsistent pair in general may embody consistent as well as inconsistent information.

In logic programming for instance, L is typically the power set 2^Σ, where Σ is a set of (ground) atoms representing reasoning individuals. A consistent pair (T, P), where T and P are sets of atoms and $T \subseteq P$, is considered a three-valued interpretation, where T is the set of true atoms and P the set of possibly true atoms; thus the atoms in $\Sigma \setminus P$ are false. If $T \not\subseteq P$, the atoms that are in T but not in P are interpreted both true and false, resulting in inconsistency. This gives rise to the notion of inconsistency in various degrees.

https://doi.org/10.1017/S1471068421000168 Published online by Cambridge University Press

2.1 Approximation fixpoint theory: the previous development

At the center of AFT is the notion of approximator. We call an operator $A : L^2 \rightarrow L^2$ an *approximator* if A is $\leq_p$-monotone and maps exact pairs to exact pairs. To emphasize the role of an operator $O : L \rightarrow L$ whose fixpoints are approximated by an approximator, we say that A is an *approximator for* O if A is $\leq_p$-monotone and $A(x, x) = (O(x), O(x))$ for all $x \in L$.

In Denecker *et al.* (2004), AFT was first developed for consistent approximations, where an approximator is *consistent* if it maps consistent pairs to consistent pairs. We denote by $Appx(L^2)$ the set of all approximators on L^2 and by $Appx(L^c)$ the set of consistent approximators on L^c. Given an approximator $A \in Appx(L^2)$, we denote by A^c the restriction of A to L^c under the condition that A^c is an operator on L^c.[1]

For the study of semantics based on partial interpretations, we can focus on the fixpoints of approximators, independent of how they may approximate operators on L. First, since $\langle L^c, \leq_p\rangle$ is not a complete lattice, the Knaster–Tarski fixpoint theory does not apply. But L^c is a chain-complete poset (ordered by $\leq_p$), so according to Markowsky's theorem, an approximator $A \in Appx(L^c)$ has a least fixpoint, called *Kripke–Kleene fixpoint* of A, and other fixpoints. However, some of these fixpoints may not satisfy the minimality principle commonly adopted in knowledge representation.[2] To eliminate non-minimal fixpoints, we can focus on what are called the *stable fixpoints* of A, which are the fixpoints of a *stable revision operator* $St_A : L^c \rightarrow L^c$, which is defined as:

$$St_A(u, v) = (lfp(A(\cdot, v)_1), lfp(A(u, \cdot)_2)), \tag{1}$$

where $A(\cdot, v)_1$ denotes the operator $[\bot, v] \rightarrow [\bot, v] : z \mapsto A(z, v)_1$ and $A(u, \cdot)_2$ denotes the operator $[u, \top] \rightarrow [u, \top] : z \mapsto A(u, z)_2$.

Denecker *et al.* (2004) show that (1) is well defined for pairs in L^c under a desirable property. We call a pair $(u, v) \in L^c$ *A-reliable* if $(u, v) \leq_p A(u, v)$. Intuitively, if $A(u, v)$ is viewed as a revision of (u, v) for more accurate approximation, under A-reliability, $A(u, v)$ is at least as accurate as (u, v). Furthermore, Denecker *et al.* (2004) show that if a pair $(u, v) \in L^c$ is A-reliable, then $A(\cdot, v)_1$ is internal in $[\bot, v]$, hence we can safely define $A(\cdot, v)_1$ to be an operator on the lattice $[\bot, v]$; similarly, since $A(u, \cdot)_2$ is internal in $[u, \top]$, we can define it on lattice $[u, \top]$ (Proposition 3.3). Since the operators $A(\cdot, v)_1$ and $A(u, \cdot)_2$ are $\leq$-monotone on their respective domains, a least fixpoint for each exists; hence the stable revision operator St_A is well defined. Note that by definition, since a fixpoint of St_A is a fixpoint of A, a stable fixpoint of A is a fixpoint of A.

However, the notion of A-reliability is not strong enough to guarantee another desirable property: for any A-reliable pair (u, v), we want $(u, v) \leq_p St_A(u, v) (= (lfp(A(\cdot, v)_1), lfp(A(u, \cdot)_2)))$; that is, a stable fixpoint computed from a given pair should be at least as accurate. This property does not hold in general for A-reliable pairs. In addition, we also want $A(u, v) \leq_p St_A(u, v)$, so that there is a guarantee that the stable revision operator "revises even more", that is, stable revision is at least as accurate as revision by a single application of A. We therefore introduce a new property: an A-reliable

[1] Such A^c may not exist in general, but for symmetric approximators, it always does; cf. Proposition 14 of Denecker *et al.* (2000).

[2] The situation is analogue to the notion of Kripke–Kleene model of a logic program, which is a least fixpoint of a three-valued van Emden–Kowalski operator.

https://doi.org/10.1017/S1471068421000168 Published online by Cambridge University Press

pair $(u, v) \in L^c$ is called *A-prudent* if $u \leq lfp(A(\cdot, v)_1)$. We denote by L^{rp} the set of A-prudent pairs in L^c. Denecker *et al.* (2004) show that for all A-prudent pairs (u, v) in L^c, $(u, v) \leq_p St_A(u, v)$ and $A(u, v) \leq_p St_A(u, v)$ (Propositions 3.7 and 3.8).

Example 1
Consider a complete lattice $\langle L, \leq \rangle$ where $L = \{\bot, \top\}$ and $\leq$ is defined as usual. Define an operator A on L^c as: $A(\top, \top) = (\top, \top)$ and $A(\bot, \top) = A(\bot, \bot) = (\top, \top)$. It can be seen that A is $\leq_p$-monotone on L^c, the pairs $(\top, \top)$ and $(\bot, \top)$ are A-reliable, and $(\bot, \bot)$ is not. Both A-reliable pairs $(\top, \top)$ and $(\bot, \top)$ are A-prudent as well, thus $L^{rp} = \{(\top, \top), (\bot, \top)\}$.

Now let A' be the identify operator on L^c except $A'(\bot, \bot) = (\top, \top)$. The operator A' is $\leq_p$-monotone on L^c. The pairs $(\top, \top)$ and $(\bot, \top)$ are A'-reliable whereas $(\bot, \bot)$ is not. But the A'-reliable pair $(\top, \top)$ is not A'-prudent because $lfp(A'(\cdot, \top)_1) = \bot < \top$. Note that $(\top, \top)$ is a fixpoint of A' but not a stable fixpoint. Thus, for approximator A', $L^{rp} = \{(\bot, \top)\}$.

The above development has led to the following results of the properties of the stable revision operator.

Theorem 3 (*Theorem 3.11 of Denecker et al. 2004*)
Let L be a complete lattice, $A \in Appx(L^c)$. The set of A-prudent elements of L^c is a chain-complete poset under the precision order $\leq_p$, with least element $(\bot, \top)$. The stable revision operator is a well defined, increasing and monotone operator in this poset.

This theorem serves as the foundation for AFT as it guarantees that the stable revision operator has fixpoints and a least fixpoint, which we have called stable fixpoints of A.

The notion of approximator is then generalized to *symmetric approximators*, which are $\leq_p$-monotone operators A on L^2 such that $A(x, y)_1 = A(y, x)_2$, for all $x, y \in L$. As remarked in Denecker *et al.* (2004), this generalization is motivated by operators arising in knowledge representation that are symmetric.[3] A critical property of a symmetric approximator A is that $A(x, x)$ yields an exact pair, for all $x \in L$, that is, it maps an exact pair to an exact pair, which is consistent. This can be seen as follows: Since $A(x, x) = (A(x, x)_1, A(x, x)_2)$ for all $x \in L$, and by the symmetry of A, $A(x, x)_1 = A(x, x)_2$ and thus $A(x, x)$ is consistent.

3 Approximation fixpoint theory generalized

In this section, we generalize AFT as given in Denecker *et al.* (2004) from consistent and symmetric approximators to arbitrary approximators. This generalization is needed in order to define approximators for hybrid MKNF knowledge bases since an exact pair in this context is a two-valued interpretation which can be mapped to an inconsistent one. This is because a hybrid MKNF knowledge base allows predicates to appear both in the underlying DL knowledge base and in rules, inconsistencies may arise from the combination of classic negation in the former and derivations using nonmonotonic negation in the latter.

[3] For example, Fitting's immediate consequence operator for normal logic programs (Fitting 2002), placed in the context of bilattice $((2^{\Sigma})^2, \subseteq_p)$ where Σ is a set of ground atoms, induces a symmetric approximator.

https://doi.org/10.1017/S1471068421000168 Published online by Cambridge University Press

The current AFT is defined for consistent and symmetric approximators. As alluded earlier, a critical property of a symmetric approximator is that it maps an exact pair to an exact pair. However, a $\leq_p$-monotone operator on L^2 may not possess this property.

Example 2
(Bi *et al*. 2014) Consider a complete lattice where $L = \{\bot, \top\}$ and $\leq$ is defined as usual. Let O be the identity function on L. Then we have two fixpoints, $O(\bot) = \bot$ and $O(\top) = \top$. Let A be an identity function on L^2 everywhere except $A(\top, \top) = (\top, \bot)$. Thus, $A(\top, \top)$ is inconsistent. It is easy to check that A is $\leq_p$-monotone, especially, from $(\top, \top) \leq_p (\top, \bot)$ we have $A(\top, \top) \leq_p A(\top, \bot)$. There is exactly one exact pair $(\bot, \bot)$ for which $A(\bot, \bot)$ is consistent, and the condition $A(\bot, \bot) = (O(\bot), O(\bot))$ is satisfied. For the other exact pair $(\top, \top)$, $A(\top, \top)$ is inconsistent and $A(\top, \top) \neq (O(\top), O(\top))$, even though $O(\top) = \top$. The fixpoint $\top$ of O is not captured by the operator A because $A(\top, \top)$ is inconsistent.

Conclusion: Though the operator A above is $\leq_p$-monotone on L^2, it is not an approximator by the current definition because it fails to map an exact pair to an exact pair when inconsistency arises.[4]

In order to accommodate operators like A above, we present a generalization by relaxing the condition for an approximator.

Definition 1
We say that an operator $A : L^2 \to L^2$ is an approximator if A is $\leq_p$-monotone and for all $x \in L$, if $A(x, x)$ is consistent then A maps (x, x) to an exact pair. Let O be an operator on L. We say that $A : L^2 \to L^2$ is an *approximator for O* if A is an approximator and for all $x \in L$, if $A(x, x)$ is consistent then $A(x, x) = (O(x), O(x))$.

That is, we make the notion of approximation partial: $A(x, x)$ captures O only when $A(x, x)$ is consistent. Under this definition, the operator A in Example 2 is an approximator and it approximates, for example, the identify operator O on L.

Before we proceed to generalize the notion of stable revision operator, we need to extend the definition of A-reliability and A-prudence to pairs in L^2. Such a definition is already provided in the study of well-founded inductive definitions. Following Denecker and Vennekens (2007), given an operator A on L^2, we say that a pair $(u, v) \in L^2$ is *A-contracting* if $(u, v) \leq_p A(u, v)$.[5] The notion of A-prudence is generalized to L^2 as well. A pair $(u, v) \in L^2$ is *A-prudent* if $u \leq \mathit{lfp}(A(\cdot, v)_1)$ (when $\mathit{lfp}(A(\cdot, v)_1)$ exists). By an abuse of notation and without confusion, in the rest of this paper we will continue to use L^{rp} but this time to denote the set of A-contracting and A-prudent pairs in L^2.

Now, we relax the definition of the stable revision operator as follows: Given any pair $(u, v) \in L^2$, define

$$St_A(u, v) = (\mathit{lfp}(A(\cdot, v)_1), \mathit{lfp}(A(u, \cdot)_2)), \tag{2}$$

where $A(\cdot, v)_1$ denotes the operator $L \to L : z \mapsto A(z, v)_1$ and $A(u, \cdot)_2$ denotes the operator $L \to L : z \mapsto A(u, z)_2$. That is, both $A(\cdot, v)_1$ and $A(u, \cdot)_2$ are operators on L.

[4] This example specifies a system in which states are represented by a pair of factors – high and low. Here, all states are stable except the one in which both factors are high. This state may be transmitted to an "inconsistent state" with the first factor high and the second low. This state is the only inconsistent one, and it itself is stable.

[5] Earlier in this paper, A-contracting pairs were called A-reliable in the context of L^c.

Notation: Let $(u, v) \in L^2$ and $A \in Appx(L^2)$. We define

$$(C_1(v), C_2(u)) = (lfp(A(\cdot, v)_1), lfp(A(u, \cdot)_2)),$$

where $A(\cdot, v)_1$ and $A(u, \cdot)_2$ are the respective projection operators defined on L. We use the notation $(C_1(v), C_2(u))$ with the understanding that the underlying approximator is clear from the context.

Since A is $\leq_p$-monotone on L^2, the projection operators $A(\cdot, v)_1$ and $A(u, \cdot)_2$, for any pair $(u, v) \in L^2$, are both $\leq$-monotone on L, which guarantees the existence of a least fixpoint for each. Thus, the stable revision operator in equation (2) is well defined for all pairs in L^2. Note that in this case a stable fixpoint can be inconsistent. For example, consider lattice $L = \{\bot, \top\}$ and an operator A on L^2, which is identity on every pair except $A(\bot, \bot) = (\top, \bot)$. Clearly, A is $\leq_p$-monotone. The inconsistent pair $(\top, \bot)$ is a stable fixpoint of A since $St_A(\top, \bot) = (lfp(A(\cdot, \bot)_1), lfp(A(\top, \cdot)_2)) = (\top, \bot)$.

The definition of stable revision above has been proposed and adopted in the literature of AFT already,[6] for example, in Denecker *et al.* (2004) and more recently in Bogaerts and Cruz-Filipe (2018), Bogaerts *et al.* (2015), for consistent and symmetric approximators. It however differs from stable revision for consistent approximators with regard to the domains of the two projection operators. As mentioned earlier, in consistent AFT (where an approximator is from $Appx(L^c)$), we know from Denecker *et al.* (2004) that $A(\cdot, v)_1$ is internal in $[\bot, v]$ so we define $A(\cdot, v)_1$ to be an operator on the lattice $[\bot, v]$, and $A(u, \cdot)_2$ is internal in $[u, \top]$ so we define it on lattice $[u, \top]$. Now, let us generalize this to all approximators in $Appx(L^2)$ for consistent pairs in L^c.

Notation: Let $(u, v) \in L^c$ and $A \in Appx(L^2)$. We define

$$(D_1(v), D_2(u)) = (lfp(A(\cdot, v)_1), lfp(A(u, \cdot)_2)),$$

where $A(\cdot, v)_1$ is defined on $[\bot, v]$ and $A(u, \cdot)_2$ is defined on $[u, \top]$. In the sequel, the term *consistent stable fixpoints* refer to the fixpoints determined by this definition.

Since we consider the entire product bilattice L^2, we are interested in knowing which consistent pairs in it make the above projection operators well defined under our relaxed definition of approximators.

Proposition 1
Let $\langle L, \leq \rangle$ be a complete lattice and A an approximator on L^2. If a consistent pair $(u, v) \in L^2$ is A-contracting and $A(u, u)$ is consistent, then for every $x \in [u, \top]$, $A(u, x)_2 \in [u, \top]$, and for every $x \in [\bot, v]$, $A(u, v)_1 \in [\bot, v]$.

Proof
We can show that, for any $x \in [u, \top]$,

$$u \leq A(u, v)_1 \leq A(u, u)_1 = A(u, u)_2 \leq A(u, x)_2.$$

The first inequality is because $(u, v) \leq_p A(u, v)$ (i.e., (u, v) is A-contracting). The second is due to $A(u, v) \leq_p A(u, u)$, as (u, v) is consistent thus $(u, v) \leq_p (u, u)$ and A is $\leq_p$-monotone. The next equality is by the fact that since $A(u, u)$ is consistent, it maps a consistent pair to a consistent pair. The last inequality is due to $x \geq u$ and that A

[6] But notice a critical difference in our definition of an approximator discussed above.

https://doi.org/10.1017/S1471068421000168 Published online by Cambridge University Press

is $\leq_p$-monotone. For any $x \in [\bot, v]$, we can similarly show that $A(x,v)_1 \leq A(v,v)_1 = A(v,v)_2 \leq v$.[7] □

A question that arises is whether consistent stable fixpoints from consistent approximations are carried over to approximators on L^2. That is, assume $(u,v) \in L^c$ is a stable fixpoint as computed by $(D_1(v), D_2(u))$, and the question is whether (u,v) is also a stable fixpoint as computed by $(C_1(v), C_2(u))$. If $(D_1(v), D_2(u)) = (C_1(v), C_2(u))$, then the answer is yes for (u,v). In this way, a consistent stable fixpoint as computed by $(D_1(v), D_2(u))$ is preserved for the stable revision operator as defined by $(C_1(v), C_2(u))$.

The above question was answered positively by Denecker *et al.* (2004) (cf. Theorem 4.2) for symmetric approximators by restricting them to consistent pairs. The authors show that the theory of consistent approximations captures general AFT that treats consistent and symmetric approximators on the product bilattice, as long as we restrict our attention to consistent pairs. They show that for any symmetric approximator A, a consistent pair (u,v) is a stable fixpoint of A on L^2 (as defined in terms of $(C_1(v), C_2(u))$) if and only if it is a stable fixpoint of A^c (as defined in terms of $(D_1(v), D_2(u))$). They state that it is possible to develop a generalization of AFT for which these results hold without the assumption of symmetry. However, once we allow consistent pairs to be mapped to inconsistent ones and adopt the domain L for the projection operators, a discrepancy with consistent AFT emerges.

Example 3
Let $L = \{\bot, \top\}$ and A an identity function everywhere on L^2 except that $A(\bot, \top) = A(\bot, \bot) = (\top, \top)$. It is easy to verify that A is $\leq_p$-monotone. Clearly, $A^c \in Appx(L^c)$, that is, it maps consistent pairs to consistent pairs, it is $\leq_p$-monotone on L^c, and approximates, for example, the identify operator O on L. But A is not symmetric since $A(\bot, \top)_1 = \top$ and $A(\top, \bot)_2 = \bot$. Since $A^c \in Appx(L^c)$, $A^c(\top, \cdot)_2$ is an operator on $[\top, \top]$. Since $St_{A^c}(\top, \top) = (lfp(A^c(\cdot, \top)_1), lfp(A^c(\top, \cdot)_2) = (\top, \top)$, it follows that $(\top, \top)$ is a stable fixpoint of A^c. Now let us apply the definition of stable revision in equation (2) to approximator A, where both projection operators $A(\cdot, y)_1$ and $A(x, \cdot)_2$ are defined on L. In this case, since $St_A(\top, \top) = (lfp(A(\cdot, \top)_1), lfp(A(\top, \cdot)_2) = (\top, \bot)$, $(\top, \top)$ is not a stable fixpoint of A. This example is not a surprise since in general different domains may well lead to different least fixpoints.

Now consider another approximator $A' \in Appx(L^2)$ such that A' maps all pairs to $(\top, \top)$. It can be seen that A' is $\leq_p$-monotone and $(\top, \top)$ is a stable fixpoint of A' in both cases, where $A'(\cdot, T)_1$ is defined as an operator either on $[\bot, \top]$ or on L, and $A'(\top, \cdot)_2$ is defined as an operator either on $[\top, \top]$ or on L. That is, for each projection operator, the least fixpoints of it on two different domains coincide.

Conclusion: For an arbitrary approximator A on the product bilattice L^2, the stable revision operator $St_A(u,v)$ is well defined for all pairs $(u,v) \in L^2$, if we define both projection operators on L. However, consistent stable fixpoints under consistent AFT may not be preserved if we adopt the stable revision operator as defined in this paper (i.e., by equation (2) in terms of $(C_1(v), C_2(u))$).

[7] The proof is essentially the same as the poof of Proposition 3.3 in Denecker *et al.* (2004); but there is a subtle difference in the definition of approximator: in the case of Denecker *et al.* (2004), the claim is proved for $A^c \in Appx(L^c)$. But in our case, the claim is for arbitrary approximators in $Appx(L^2)$. This shows an argument in favor of our relaxed definition of approximators.

https://doi.org/10.1017/S1471068421000168 Published online by Cambridge University Press

Let us call the existence of a gap between the two pairs of least fixpoints, $(C_1(v), C_2(u))$ and $(D_1(v), D_2(u))$, discussed above an "anomaly". One can argue that a desirable approximator should not exhibit this anomaly so that accommodating inconsistent pairs does not have to sacrifice the preservation of consistent stable fixpoints.

Definition 2
Let $A \in Appx(L^2)$, and $(u, v) \in L^c$ such that $(u, v) = (lfp(A(\cdot, v)_1), lfp(A(u, \cdot)_2))$ where $A(\cdot, v)_1$ is an operator on $[\bot, v]$ and $A(u, \cdot)_2$ is an operator on $[u, \top]$. Approximator A is called *strong* for (u, v) if $(u, v) = (lfp(A(\cdot, v)_1), lfp(A(u, \cdot)_2))$ where both $A(\cdot, v)_1$ and $A(u, \cdot)_2$ are operators on L. Approximator A is called *strong* if it is strong for every $(u, v) \in L^c$ that satisfies the above condition.

In other words, a strong approximator preserves consistent stable fixpoints under the definition of stable revision adopted in this paper. For example, in Example 3, while the approximator A' is strong for $(\top, \top)$, the approximator A is not.

A question arises: are there natural approximators that are strong? For normal logic programs, it is known that Fitting's immediate consequence operator $\Theta_{\mathscr{P}}$ (Fitting 2002) induces a symmetric approximator. It can be shown that $\Theta_{\mathscr{P}}$ is also a strong approximator.[8] In addition, we show later in this paper that the approximators we formulate for hybrid MKNF knowledge bases are (essentially) strong approximators.[9] If we focus on strong approximators, the relaxed AFT as presented in this paper can be seen as a generalization of the original AFT.

Finally, as a generalization of the current AFT, we show that the properties of the stable revision operator as stated in Theorem 3 for consistent AFT can be generalized.

Theorem 4
Let $(L, \leq)$ be a complete lattice and A an approximator on L^2. Then, $\langle L^{rp}, \leq_p\rangle$ is a chain-complete poset under the precision order $\leq_p$, with least element $(\bot, \top)$. The stable revision operator as given in equation (2) is a well defined, increasing and monotone operator in this poset.

Proof
The least element $(\bot, \top)$ is naturally A-contracting and A-prudent. Let C be a chain in L^{rp}, and C_1, C_2 be the respective projections of C. First we show that the element $(lub(C_1), glb(C_2)) = (\bigvee C_1, \bigwedge C_2)$ is the least upper bound of C, which is also in L^{rp}. Since C is a chain in L^{rp} ordered by the relation $\leq_p$, it is easy to see that the least upper bound of C_1 exists, which is just the maximum element in C_1; similarly, the greatest lower bound of C_2 exists. Then, it is clear that the least upper bound of C is $lub(C) = (lub(C_1), glb(C_2))$.

To show $(lub(C_1), glb(C_2)) = (\bigvee C_1, \bigwedge C_2)$ is A-contracting and A-prudent, let $u_0 = \bigvee C_1$ and $v_0 = \bigwedge C_2$ and consider any $(a, b) \in C$. Since C is a chain in L^{rp} that contains

[8] We can apply Lemma 4.1 of Denecker *et al.* (2004), which says that for any symmetric approximator A and for any consistent pair (u, v), if (u, v) is A^c-prudent, then $(D_1(v), D_2(u)) = (C_1(v), C_2(u))$. Since a consistent, $\Theta_{\mathscr{P}}$-prudent stable fixpoint of $\Theta_{\mathscr{P}}$ is $\Theta^c_{\mathscr{P}}$-prudent, the conclusion follows.

[9] Technically, we need a mild condition: Given a consistent stable fixpoint (u, v), these approximators are strong for (u, v) if u is consistent with the given DL knowledge base. If the condition is not satisfied, the stable fixpoint (u, v) does not correspond to a three-valued MKNF model. Thus, the condition does not affect the preservation of consistent stable fixpoints that give three-valued MKNF models.

https://doi.org/10.1017/S1471068421000168 Published online by Cambridge University Press

(a,b), $a \in C_1$ and $b \in C_2$, we have $a \leq u_0$ and $v_0 \leq b$, from which we obtain $a \leq A(a,b)_1 \leq A(u_0,b)_1 \leq A(u_0,v_0)_1$, where the first inequality is because $(a,b) \in L^{rp}$ is A-contracting and the next two inequalities are due to the $\leq_p$-monotonicity of A. Since $a \in C_1$ is arbitrary, letting $a = u_0$, we then have $u_0 = \bigvee C_1 \leq A(u_0,v_0)_1$. Similarly, we can show $A(u_0,v_0)_2 \leq A(a,v_0)_2 \leq A(a,b)_2 \leq b$. Since $v_0 = \bigwedge C_2$, it follows that $A(u_0,v_0)_2 \leq v_0 = \bigwedge C_2$. Hence, $(\bigvee C_1, \bigwedge C_2)$ is A-contracting.

To show that $(\bigvee C_1, \bigwedge C_2)$ is A-prudent, let $u' = \mathit{lfp}(A(\cdot,v_0)_1)$. For any $(a,b) \in C$, we have $A(u',b)_1 \leq A(u',v_0)_1 = u'$, then u' is a pre-fixpoint of $A(\cdot,b)_1$ and thus $\mathit{lfp}(A(\cdot,b)_1) \leq u'$. Also since (a,b) is A-contracting, we have $a \leq \mathit{lfp}(A(\cdot,b)_1) \leq u'$. Since a is arbitrary from C_1, this applies to $u_0 = \bigvee C_1 \in C_1$ and thus $\bigvee C_1 \leq u'$. That is, $(\bigvee C_1, \bigwedge C_2)$ is A-prudent.

We therefore conclude that $(L^{rp}, \leq_p)$ is a chain-complete poset under order $\leq_p$.

Next, we show that the stable revision operator defined in equation (2) is a well defined, increasing and monotone operator in this poset. In the definition of the stable revision operator $St_A(u,v) = (\mathit{lfp}(A(\cdot,v)_1), \mathit{lfp}(A(u,\cdot)_2))$ in equation (2), we already argued that St_A is a well-defined mapping, due to the fact that both projection operators $A(\cdot,v)_1$ and $A(u,\cdot)_2$ are defined on L. We now show that

- St_A is well defined for L^{rp}, namely L^{rp} is closed under St_A, that is, for all $(a,b) \in L^{rp}$, $St_A(a,b) \in L^{rp}$,
- St_A is increasing, that is, $(a,b) \leq_p St_A(a,b)$ for all $(a,b) \in L^{rp}$, and
- St_A is $\leq_p$-monotone.

Let $(a,b),(c,d) \in L^{rp}$. For simplicity, let $u_1 = St_A(a,b)_1 = \mathit{lfp}(A(\cdot,b)_1)$ and $v_1 = St_A(a,b)_2 = \mathit{lfp}(A(a,\cdot)_2)$, $u_2 = St_A(c,d)_1 = \mathit{lfp}(A(\cdot,d)_1)$ and $v_2 = St_A(c,d)_2 = \mathit{lfp}(A(c,\cdot)_2)$.

For convenience, let us first show that St_A is increasing and $\leq_p$-monotone. By A-prudence of (a,b), $a \leq u_1$. Since (a,b) is A-contracting, $A(a,b)_2 \leq b$. Thus b is a pre-fixpoint of $A(a,\cdot)_2$, and since $v_1 = \mathit{lfp}(A(a,\cdot)_2)$, it follows $v_1 \leq b$. That is, $(a,b) \leq_p St_A(a,b)$.

For $\leq_p$-monotonicity, given $(a,b) \leq_p (c,d)$, we have $A(u_2,b)_1 \leq A(u_2,d)_1 = u_2$ by $d \leq b$, and thus u_2 is a pre-fixpoint of $A(\cdot,b)_1$ and $u_1 \leq u_2$. Similarly, $A(c,v_1)_2 \leq A(a,v_1)_2 = v_1$ by $a \leq c$, so v_1 is a pre-fixpoint of $A(c,\cdot)_2$ and thus $v_2 \leq v_1$. That is, $St_A(a,b) \leq_p St_A(c,d)$.

We now show that St_A maps a pair $(a,b) \in L^{rp}$ to a pair in L^{rp}. We observe that $u_1 = A(u_1,b)_1 \leq A(u_1,v_1)_1$, where the equality is because u_1 is a fixpoint of the operator $A(\cdot,b)_1$ and the inequality is because (a,b) is A-contracting. Similarly, $v_1 = A(a,v_1)_2 \leq A(u_1,v_1)_2$. Therefore, $(u_1,v_1) \leq_p A(u_1,v_1)$, that is, (u_1,v_1) is A-contracting. To prove A-prudence of (u_1,v_1), since (a,b) is A-contracting, $b \geq v_1$ and by $\leq_p$-mononononitcity of A, for any $x \in L$, $A(x,b) \leq_p A(x,v_1)$ and thus $A(x,b)_1 \leq A(x,v_1)_1$. Thus, every pre-fixpoint of $A(\cdot,v_1)_1$ is a pre-fixpoint of $A(\cdot,b)_1$, that is, for any $z \in L$, if $A(z,v_1)_1 \leq z$ then $A(z,b)_1 \leq A(z,v_1)_1 \leq z$. Since $A(\cdot,v_1)_1$ is a monotone operator on L, $\mathit{lfp}(A(\cdot,v_1)_1)$ exists and thus the set of pre-fixpoints of $A(\cdot,v_1)_1$ is non-empty. Therefore $u_1 \leq \mathit{lfp}(A(\cdot,v_1)_1)$ and (u_1,v_1) is A-prudent. □

https://doi.org/10.1017/S1471068421000168 Published online by Cambridge University Press

We now can apply Theorem 2 so that given an approximator A on the product bilattice L^2, the stable revision operator defined by equation (2) possesses fixpoints and a least fixpoint, the latter of which can be computed iteratively from the least element $(\bot, \top)$.

Note that by the Knaster–Tarski fixpoint theory, since $\langle L^2, \leq_p \rangle$ is a complete lattice and the stable revision operator St_A is $\leq_p$-monotone on L^2 (which can be shown by the same proof for the $\leq_p$-monotonicity on L^{rp} above), the operator St_A defined in equation (2) is already guaranteed to possess fixpoints and a least fixpoint. Nevertheless, Theorem 4 above is still relevant because it shows a generalization of the chain-completeness result from L^c to L^2, and in addition, it points to a smaller domain of pairs L^{rp} from which consistent as well as inconsistent stable fixpoints can be computed by the guess-and-verify method.

4 Hybrid MKNF knowledge bases

4.1 Minimal knowledge and negation as failure

The logic of MKNF (Lifschitz 1991) is based on a first-order language $\mathscr{L}$ (possibly with equality $\approx$) with two modal operators, **K**, for minimal knowledge, and **not**, for negation as failure. In MKNF, *first-order atoms* are defined as usual and *MKNF formulas* are first-order formulas with **K** and **not**. An MKNF formula φ is *ground* if it contains no variables, and $\varphi[t/x]$ denotes the formula obtained from φ by replacing all free occurrences of variable x with term t. Given a first-order formula ψ, $\mathbf{K}\psi$ is called a (modal) **K**-*atom* and $\mathbf{not}\,\psi$ called a (modal) **not**-*atom*. Both of these are also called *modal atoms*.

A *first-order interpretation* is understood as in first-order logic. The universe of a first-order interpretation I is denoted by $|I|$. A *first-order structure* is a non-empty set M of first-order interpretations with the universe $|I|$ for some fixed $I \in M$. An *MKNF structure* is a triple (I, M, N), where M and N are sets of first-order interpretations with the universe $|I|$. We extend the language $\mathscr{L}$ by adding object constants representing all elements of $|I|$, and call these constants *names*. The satisfaction relation $\models$ between an MKNF structure (I, M, N) and an MKNF formula φ is defined as follows:

$$
\begin{aligned}
&(I, M, N) \models \varphi \ (\varphi \text{ is a first-order atom) if } \varphi \text{ is true in } I,\\
&(I, M, N) \models \neg\varphi \text{ if } (I, M, N) \not\models \varphi,\\
&(I, M, N) \models \varphi_1 \wedge \varphi_2 \text{ if } (I, M, N) \models \varphi_1 \text{ and } (I, M, N) \models \varphi_2,\\
&(I, M, N) \models \exists x \varphi \text{ if } (I, M, N) \models \varphi[\alpha/x] \text{ for some name } \alpha,\\
&(I, M, N) \models \mathbf{K}\varphi \text{ if } (J, M, N) \models \varphi \text{ for all } J \in M,\\
&(I, M, N) \models \mathbf{not}\,\varphi \text{ if } (J, M, N) \not\models \varphi \text{ for some } J \in N.
\end{aligned}
$$

The symbols $\top$, $\bot$, $\vee$, $\forall$, and $\supset$ are interpreted as usual.

An *MKNF interpretation* M is a non-empty set of first-order interpretations over the universe $|I|$ for some $I \in M$. In MKNF, a notion called *standard name assumption* is imposed to avoid unintended behaviors (Motik and Rosati 2010). This requires an interpretation to be a Herbrand interpretation with a countably infinite number of additional constants, and the predicate $\approx$ to be a congruence relation.[10] Intuitively, given

[10] The requirement that the predicate $\approx$ be interpreted as a congruence relation overwrites the earlier assumption that $\approx$ is interpreted as equality.

https://doi.org/10.1017/S1471068421000168 Published online by Cambridge University Press

the assumption that each individual in the universe of an interpretation is denoted by a constant and the countability it implies, the standard name assumption becomes a convenient normalized representation of interpretations since each interpretation is isomorphic to the quotient (w.r.t. $\approx$) of a Herbrand interpretation and each quotient of a Herbrand interpretation is an interpretation. In the sequel, we assume the standard name assumption, and due to this assumption, in definitions we need not explicitly mention the universe associated with the underlying interpretations.

An MKNF interpretation M *satisfies* an MKNF formula φ, written $M \models_{MKNF} \varphi$, if $(I, M, M) \models \varphi$ for each $I \in M$. *Two-valued MKNF models* are defined as follows.

Definition 3
An MKNF interpretation M is an *MKNF model* of an MKNF formula φ if

(1) $M \models_{MKNF} \varphi$, and
(2) for all MKNF interpretations M' such that $M' \supset M$, $(I', M', M) \not\models \varphi$ for every $I' \in M'$.

For example, with the MKNF formula $\varphi = \textbf{not}\, b \supset \mathbf{K}a$, it is easy to verify that the MKNF interpretation $M = \{\{a\}, \{a, b\}\}$ is an MKNF model of φ.

Following Knorr *et al.* (2011), a *three-valued MKNF structure*, $(I, \mathscr{M}, \mathscr{N})$, consists of a first-order interpretation, I, and two pairs, $\mathscr{M} = \langle M, M_1 \rangle$ and $\mathscr{N} = \langle N, N_1 \rangle$, of sets of first-order interpretations, where $M_1 \subseteq M$ and $N_1 \subseteq N$. From the two component sets in $\mathscr{M} = \langle M, M_1 \rangle$, we can define three truth values for modal **K**-atoms in the following way: $\mathbf{K}\varphi$ is true w.r.t. $\mathscr{M} = \langle M, M_1 \rangle$ if φ is true in all interpretations in M; it is false if it is false in at least one interpretation in M_1; and it is undefined otherwise. For **not**-atoms, a symmetric treatment w.r.t. $\mathscr{N} = \langle N, N_1 \rangle$ is adopted. Let $\{\mathbf{t}, \mathbf{u}, \mathbf{f}\}$ be the set of truth values *true*, *undefined*, and *false* with the order $\mathbf{f} < \mathbf{u} < \mathbf{t}$, and let the operator *max* (resp. *min*) choose the greatest (resp. the least) element with respect to this ordering. Table 1 shows three-valued evaluation of MKNF formulas.

A *(three-valued) MKNF interpretation pair* (M, N) consists of two MKNF interpretations, M and N, with $\emptyset \subset N \subseteq M$. An MKNF interpretation pair satisfies an MKNF formula φ, denoted $(M, N) \models \varphi$, iff $(I, \langle M, N \rangle, \langle M, N \rangle)(\varphi) = \mathbf{t}$ for each $I \in M$. If $M = N$, the MKNF interpretation pair is called *total*.

Definition 4
An MKNF interpretation pair (M, N) is a *three-valued MKNF model* of an MKNF formula φ if

(a) $(M, N) \models \varphi$, and
(b) for all MKNF interpretation pairs (M', N') with $M \subseteq M'$ and $N \subseteq N'$, where at least one of the inclusions is proper and $M' = N'$ if $M = N$, $\exists I' \in M'$ such that $(I', \langle M', N' \rangle, \langle M, N \rangle)(\varphi) \neq \mathbf{t}$.

Condition (a) checks satisfiability while condition (b), with the evaluation of **not**-atoms fixed, constrains the evaluation of modal **K**-atoms to be minimal w.r.t the ordering $\mathbf{f} < \mathbf{u} < \mathbf{t}$ while maximizing falsity. That is, by enlarging M to M' we limit the derivation of **K**-atoms, and by enlarging N to N' we expand on falsity to reduce undefined. Thus, a three-valued MKNF model is one for which neither of these is possible under the assumption that **not**-atoms remain to be evaluated w.r.t. (M, N). If $M = N$, then

Table 1. *Evaluation in three-valued MKNF structure* $(I, \mathcal{M}, \mathcal{N})$

$$(I,\mathcal{M},\mathcal{N})(P(t_1,\ldots,t_n)) = \begin{cases} \mathbf{t} & \text{iff } P(t_1,\ldots,t_n) \text{ is true in } I \\ \mathbf{f} & \text{iff } P(t_1,\ldots,t_n) \text{ is false in } I \end{cases}$$

$$(I,\mathcal{M},\mathcal{N})(\neg\varphi) = \begin{cases} \mathbf{t} & \text{iff } (I,\mathcal{M},\mathcal{N})(\varphi) = \mathbf{f} \\ \mathbf{u} & \text{iff } (I,\mathcal{M},\mathcal{N})(\varphi) = \mathbf{u} \\ \mathbf{f} & \text{iff } (I,\mathcal{M},\mathcal{N})(\varphi) = \mathbf{t} \end{cases}$$

$$(I,\mathcal{M},\mathcal{N})(\varphi_1 \wedge \varphi_2) = \min\{(I,\mathcal{M},\mathcal{N})(\varphi_1), (I,\mathcal{M},\mathcal{N})(\varphi_2)\}$$

$$(I,\mathcal{M},\mathcal{N})(\varphi_1 \supset \varphi_2) = \begin{cases} \mathbf{t} & \text{iff } (I,\mathcal{M},\mathcal{N})(\varphi_2) \geq (I,\mathcal{M},\mathcal{N})(\varphi_1) \\ \mathbf{f} & \text{otherwise} \end{cases}$$

$$(I,\mathcal{M},\mathcal{N})(\exists x\colon \varphi) = \max\{(I,\mathcal{M},\mathcal{N})(\varphi[\alpha/x]) \mid \alpha \text{ is a name}\}$$

$$(I,\mathcal{M},\mathcal{N})(\mathbf{K}\varphi) = \begin{cases} \mathbf{t} & \text{iff } (J,\langle M, M_1\rangle,\mathcal{N})(\varphi) = \mathbf{t} \text{ for all } J \in M \\ \mathbf{f} & \text{iff } (J,\langle M, M_1\rangle,\mathcal{N})(\varphi) = \mathbf{f} \text{ for some } J \in M_1 \\ \mathbf{u} & \text{otherwise} \end{cases}$$

$$(I,\mathcal{M},\mathcal{N})(\mathbf{not}\,\varphi) = \begin{cases} \mathbf{t} & \text{iff } (J,\mathcal{M},\langle N, N_1\rangle)(\varphi) = \mathbf{f} \text{ for some } J \in N_1 \\ \mathbf{f} & \text{iff } (J,\mathcal{M},\langle N, N_1\rangle)(\varphi) = \mathbf{t} \text{ for all } J \in N \\ \mathbf{u} & \text{otherwise} \end{cases}$$

(M, M) is equivalent to a two-valued MKNF model. The requirement $M' = N'$ reduces the definition to one for two-valued MKNF models as given in Definition 3, which enables Knorr *et al.* (2011) to show that an MKNF interpretation pair (M, M) that is a three-valued MKNF model of φ corresponds to a two-valued MKNF model M of φ as defined in Motik and Rosati (2010).

Example 4
Consider the MKNF formula $\varphi = [(\mathbf{not}\, b \wedge \mathbf{not}\, a) \supset \mathbf{K}a] \wedge [\mathbf{K}a \supset \mathbf{K}d]$ and the MKNF interpretation pair (M, M) where $M = \{\{a, d\}, \{a, b, d\}\}$. We have $(M, M) \models \{\mathbf{not}\, b, \neg\mathbf{not}\, a, \mathbf{K}a, \mathbf{K}d\}$. Though $(M, M) \models \varphi$, it violates condition (b) of Definition 4, since the three-valued MKNF structure $(I, \langle M', M'\rangle, \langle M, M\rangle)$, where $M' = \{\emptyset, \{a, d\}, \{a, b, d\}\}$ and thus $M \subset M'$, evaluates $[\mathbf{not}\, b, \mathbf{not}\, a, \mathbf{K}a, \mathbf{K}d]$ to $[\mathbf{t}, \mathbf{f}, \mathbf{f}, \mathbf{f}]$, respectively, independent of I. It follows that $(I, \langle M', M'\rangle, \langle M, M\rangle)$ evaluates φ to $\mathbf{t}$, according to Table 1.

MKNF interpretation pairs can be compared by an *order of knowledge*. Let (M_1, N_1) and (M_2, N_2) be MKNF interpretation pairs. $(M_1, N_1) \succeq_k (M_2, N_2)$ iff $M_1 \subseteq M_2$ and $N_1 \supseteq N_2$. A three-valued MKNF model (M, N) of an MKNF formula φ is called a *well-founded MKNF model* of φ if $(M_1, N_1) \succeq_k (M, N)$ for all three-valued MKNF models (M_1, N_1) of φ.

https://doi.org/10.1017/S1471068421000168 Published online by Cambridge University Press

4.2 Hybrid MKNF knowledge bases

The critical issue of how to combine open and closed world reasoning is addressed in Motik and Rosati (2010) by seamlessly integrating rules with DLs. A hybrid MKNF knowledge base $\mathcal{K} = (\mathcal{O}, \mathcal{P})$ consists of a decidable DL knowledge base $\mathcal{O}$, translatable into first-order logic and a rule base $\mathcal{P}$, which is a finite set of rules with modal atoms. The original work on hybrid MKNF knowledge bases (Motik and Rosati 2007; 2010) defines a two-valued semantics for such knowledge bases with disjunctive rules. In this paper, following Knorr *et al.* (2011), our focus is on non-disjunctive rules as presented in Motik and Rosati (2007).

An MKNF rule (or simply a *rule*) r is of the form: $\mathbf{K}H \leftarrow \mathbf{K}A_1, \ldots, \mathbf{K}A_m$, $\mathbf{not}\,B_1, \ldots, \mathbf{not}\,B_n$, where H, A_i, and B_j are function-free first-order atoms. Given a rule r, we let $hd(r) = \mathbf{K}H$, $bd^+(r) = \{\mathbf{K}A_i \,|\, i = 1..m\}$, and $bd^-(r) = \{B_i \,|\, i = 1..n\}$. A rule is *positive* if it contains no **not**-atoms. When all rules in $\mathcal{P}$ are positive, $\mathcal{K} = (\mathcal{O}, \mathcal{P})$ is called *positive*.

For the interpretation of a hybrid MKNF knowledge base $\mathcal{K} = (\mathcal{O}, \mathcal{P})$ in the logic of MKNF, a transformation $\pi(\mathcal{K}) = \mathbf{K}\pi(\mathcal{O}) \wedge \pi(\mathcal{P})$ is performed to transform $\mathcal{O}$ into a first-order formula and rules $r \in \mathcal{P}$ into a conjunction of first-order implications to make each of them coincide syntactically with an MKNF formula. More precisely,

$$\pi(r) = \forall \vec{x}\colon (\mathbf{K}H \subset \mathbf{K}A_1 \wedge \ldots \wedge \mathbf{K}A_m \wedge \mathbf{not}\,B_1 \wedge \ldots \wedge \mathbf{not}\,B_n)$$
$$\pi(\mathcal{P}) = \textstyle\bigwedge_{r \in \mathcal{P}} \pi(r), \quad \pi(\mathcal{K}) = \mathbf{K}\pi(\mathcal{O}) \wedge \pi(\mathcal{P}),$$

where $\vec{x}$ is the vector of free variables in r.

Under the additional assumption of DL-safety a first-order rule base is semantically equivalent to a finite ground rule base, in terms of two-valued MKNF models (Motik and Rosati 2010) as well as in terms of three-valued MKNF models (Knorr *et al.* 2011); hence decidability is guaranteed. Given a hybrid MKNF knowledge base $\mathcal{K} = (\mathcal{O}, \mathcal{P})$, a rule r in $\mathcal{P}$ is said to be *DL-safe* if every variable in r occurs in at least one **K**-atom in the body of r whose predicate symbol does not appear in $\mathcal{O}$,[11] and $\mathcal{K}$ is DL-safe if all rules in $\mathcal{P}$ are DL-safe. In this paper, we assume that a given rule base is always DL-safe, and for convenience, when we write $\mathcal{P}$ we assume it is already grounded.

Given a hybrid MKNF knowledge base $\mathcal{K} = (\mathcal{O}, \mathcal{P})$, let $\mathsf{KA}(\mathcal{K})$ be the set of all (ground) **K**-atoms $\mathbf{K}\phi$ such that either $\mathbf{K}\phi$ occurs in $\mathcal{P}$ or $\mathbf{not}\,\phi$ occurs in $\mathcal{P}$. We generalize the notion of partition (Knorr *et al.* 2011) from consistent pairs to all pairs: A *partition* of $\mathsf{KA}(\mathcal{K})$ is a pair (T, P) such that $T, P \subseteq \mathsf{KA}(\mathcal{K})$; if $T \subseteq P$, then (T, P) is said to be consistent, otherwise it is inconsistent. A partition of the form (E, E) is said to be *exact*.

Intuitively, for a partition (T, P), T contains *true* modal **K**-atoms and P contains *possibly true* modal **K**-atoms. Thus, the complement of P is the set of *false* modal **K**-atoms and $P \backslash T$ the set of *undefined* modal **K**-atoms.

Partitions are closely related to MKNF interpretation pairs. It is shown in Knorr *et al.* (2011), Liu and You (2017) that an MKNF interpretation pair (M, N) induces a consistent partition (T, P) such that for any modal **K**-atom $\mathbf{K}\xi \in \mathsf{KA}(\mathcal{K})$,

[11] Such a modal **K**-atom is called a *non-DL-atom* in Knorr *et al.* (2011), Motik and Rosati (2010).

https://doi.org/10.1017/S1471068421000168 Published online by Cambridge University Press

1. $\mathbf{K}\xi \in T$ iff $\forall I \in M, (I, \langle M, N\rangle, \langle M, N\rangle)(\mathbf{K}\xi) = \mathbf{t}$,
2. $\mathbf{K}\xi \notin P$ iff $\forall I \in M, (I, \langle M, N\rangle, \langle M, N\rangle)(\mathbf{K}\xi) = \mathbf{f}$, and
3. $\mathbf{K}\xi \in P\backslash T$ iff $\forall I \in M, (I, \langle M, N\rangle, \langle M, N\rangle)(\mathbf{K}\xi) = \mathbf{u}$.

Given a set of first-order atoms S, we define the corresponding set of modal **K**-atoms as: $\mathbf{K}(S) = \{\mathbf{K}\phi \mid \phi \in S\}$.

Let S be a subset of $\mathsf{KA}(\mathscr{K})$. The *objective knowledge* of S relevant to $\mathscr{K}$ is the set of first-order formulas $\mathsf{OB}_{\mathscr{O},S} = \{\pi(\mathscr{O})\} \cup \{\xi \mid \mathbf{K}\xi \in S\}$.

Example 5
Consider a hybrid MKNF knowledge base $\mathscr{K} = (\mathscr{O}, \mathscr{P})$, where $\mathscr{O} = a \wedge (b \supset c) \wedge \neg f$ and $\mathscr{P}$ is

$$\mathbf{K}b \leftarrow \mathbf{K}a. \quad \mathbf{K}d \leftarrow \mathbf{K}c, \mathbf{not}\, e. \quad \mathbf{K}e \leftarrow \mathbf{not}\, d. \quad \mathbf{K}f \leftarrow \mathbf{not}\, b.$$

Reasoning with $\mathscr{K}$ can be understood as follows: since $\mathbf{K}\mathscr{O}$ implies $\mathbf{K}a$, by the first rule we derive $\mathbf{K}b$; then due to $b \supset c$ in $\mathscr{O}$ we derive $\mathbf{K}c$. Thus its occurrence in the body of the second rule is true and can be ignored. For the **K**-atoms $\mathbf{K}d$ and $\mathbf{K}e$ appearing in the two rules in the middle, without preferring one over the other, both can be undefined. Because both $\mathbf{not}\, b$ and $\mathbf{K}f$ are false (the latter is due to $\neg f$ in $\mathscr{O}$), the last rule is also satisfied. Now consider an MKNF interpretation pair $(M, N) = (\{I \mid I \models \mathscr{O} \wedge b\}, \{I \mid I \models \mathscr{O} \wedge b \wedge d \wedge e\})$, which corresponds to partition $(T, P) = (\{\mathbf{K}a, \mathbf{K}b, \mathbf{K}c\}, \{\mathbf{K}a, \mathbf{K}b, \mathbf{K}c, \mathbf{K}d, \mathbf{K}e\})$. For instance, we have that, for all $I \in M$, $(I, \langle M, N\rangle, \langle M, N\rangle)(\mathbf{K}a) = \mathbf{t}$ and $(I, \langle M, N\rangle, \langle M, N\rangle)(\mathbf{K}d) = \mathbf{u}$. The interpretation pair (M, N) is a three-valued MKNF model of $\mathscr{K}$; in fact, it is the well-founded MKNF model of $\mathscr{K}$.

It is known that in general the well-founded MKNF model may not exist.

Example 6
(Liu and You 2017) Let us consider $\mathscr{K} = (\mathscr{O}, \mathscr{P})$, where $\mathscr{O} = (a \supset h) \wedge (b \supset \neg h)$ and $\mathscr{P}$ consists of

$$\mathbf{K}a \leftarrow \mathbf{not}\, b. \quad \mathbf{K}b \leftarrow \mathbf{not}\, a.$$

Consider two partitions, $(\{\mathbf{K}a\}, \{\mathbf{K}a\})$ and $(\{\mathbf{K}b\}, \{\mathbf{K}b\})$. The corresponding MKNF interpretation pairs turn out to be two-valued MKNF models of $\mathscr{K}$. For example, for the former the interpretation pair is (M, M), where $M = \{\{a, h\}\}$. Since these two-valued MKNF models are not comparable w.r.t. undefinedness and there are no other three-valued MKNF models of $\mathscr{K}$, it follows that no well-founded MKNF model for $\mathscr{K}$ exists.

5 Approximators for hybrid MKNF knowledge bases

In this section, we first show that the alternating fixpoint operator defined by Knorr *et al.* (2011) can be recast as an approximator of AFT, and therefore can be applied to characterize all three-valued MKNF models automatically and naturally. We show that this approximator is a strong approximator. Since this approximator is not symmetric, we have discovered a strong approximator for an important application without the assumption of symmetry. Being strong guarantees that all consistent stable fixpoints are preserved. At the end, we show how stable fixpoints of this approximator serve as the candidates for three-valued MKNF models by a simple consistency test.

https://doi.org/10.1017/S1471068421000168 Published online by Cambridge University Press

Throughout this section, the underlying complete lattice is $(2^{\mathsf{KA}(\mathscr{K})}, \subseteq)$ and the induced product bilattice is $(2^{\mathsf{KA}(\mathscr{K})})^2$.

We define an operator on $2^{\mathsf{KA}(\mathscr{K})}$, which is to be approximated by our approximators introduced shortly.

Definition 5
Let $\mathscr{K} = (\mathscr{O}, \mathscr{P})$ be a hybrid MKNF knowledge base. We define an operator $\mathscr{T}_\mathscr{K}$ on $2^{\mathsf{KA}(\mathscr{K})}$ as follows: given $I \subseteq \mathsf{KA}(\mathscr{K})$,

$$\begin{aligned}\mathscr{T}_\mathscr{K}(I) = \{\mathbf{K}a \in \mathsf{KA}(\mathscr{K}) \mid \mathsf{OB}_{\mathscr{O},I} \models a\} \cup \\ \{hd(r) \mid r \in \mathscr{P} :\ bd^+(r) \subseteq I,\ \mathbf{K}(bd^-(r)) \cap I = \emptyset\}\end{aligned}$$

If $\mathscr{K}$ is a positive hybrid MKNF knowledge base, the operator $\mathscr{T}_\mathscr{K}$ is monotone and has a least fixpoint. If in addition $\mathscr{O}$ is an empty DL knowledge base, then $\mathscr{T}_\mathscr{K}$ is essentially the familiar *immediate consequence operator* of van Emden and Kowalski (1976).

Knorr *et al.* (2011) defined two kinds of transforms with consistent partitions. For the purpose of this paper, let us allow arbitrary partitions.

Definition 6
Let $\mathscr{K} = (\mathscr{O}, \mathscr{P})$ be a hybrid MKNF knowledge base and $S \in 2^{\mathsf{KA}(\mathscr{K})}$. Define two forms of reduct:

$\mathscr{K}/S = (\mathscr{O}, \mathscr{P}')$, where

$$\mathscr{P}' = \{\mathbf{K}a \leftarrow bd^+(r) \mid r \in \mathscr{P} : hd(r) = \mathbf{K}a, \mathbf{K}(bd^-(r)) \subseteq \mathsf{KA}(\mathscr{K}) \setminus S\}$$

$\mathscr{K}//S = (\mathscr{O}, \mathscr{P}'')$, where

$$\mathscr{P}'' = \{\mathbf{K}a \leftarrow bd^+(r) \mid r \in \mathscr{P} : hd(r) = \mathbf{K}a,\ \mathbf{K}(bd^-(r)) \subseteq \mathsf{KA}(\mathscr{K}) \setminus S, \mathsf{OB}_{\mathscr{O},S} \not\models \neg a\}$$

We call $\mathscr{K}/S$ *MKNF transform* and $\mathscr{K}//S$ *MKNF-coherent transform.*

Since in both cases of $\mathscr{K}/S$ and $\mathscr{K}//S$ the resulting rule base is positive, a least fixpoint in each case exists. Let us define $\Gamma_\mathscr{K}(S) = lfp(\mathscr{T}_{\mathscr{K}/S})$ and $\Gamma'_\mathscr{K}(S) = lfp(\mathscr{T}_{\mathscr{K}//S})$. Then, we can construct two sequences $\mathbf{P}_i$ and $\mathbf{N}_i$ as follows:

$$\begin{aligned}\mathbf{P}_0 = \emptyset, \ldots, \mathbf{P}_{n+1} = \Gamma_\mathscr{K}(\mathbf{N}_n), \ldots, \mathbf{P}_\Omega = \bigcup \mathbf{P}_i \\ \mathbf{N}_0 = \mathsf{KA}(\mathscr{K}), \ldots, \mathbf{N}_{n+1} = \Gamma'_\mathscr{K}(\mathbf{P}_n), \ldots, \mathbf{N}_\Omega = \bigcap \mathbf{N}_i\end{aligned}$$

Intuitively, starting from $\mathbf{P}_0$ where no modal $\mathbf{K}$-atoms are known to be true and $\mathbf{N}_0$ where all modal $\mathbf{K}$-atoms are possibly true, $\mathbf{P}_{i+1}$ computes the true modal $\mathbf{K}$-atoms given the set of possibly true modal $\mathbf{K}$-atoms in $\mathbf{N}_i$, and $\mathbf{N}_{i+1}$ computes the set of possibly true modal $\mathbf{K}$-atoms given the $\mathbf{K}$-atoms are known be true in $\mathbf{P}_i$. Now let us place this construction under AFT by formulating an approximator.

Definition 7
Let $\mathscr{K} = (\mathscr{O}, \mathscr{P})$ be a hybrid MKNF knowledge base. We define an operator $\Phi_\mathscr{K}$ on $(2^{\mathsf{KA}(\mathscr{K})})^2$ as follows: $\Phi_\mathscr{K}(T, P) = (\Phi_\mathscr{K}(T, P)_1, \Phi_\mathscr{K}(T, P)_2)$, where

$$\begin{aligned}\Phi_\mathscr{K}(T, P)_1 = \{\mathbf{K}a \in \mathsf{KA}(\mathscr{K}) \mid \mathsf{OB}_{\mathscr{O},T} \models a\} \cup \\ \{hd(r) \mid r \in \mathscr{P} :\ bd^+(r) \subseteq T,\ \mathbf{K}(bd^-(r)) \cap P = \emptyset\} \\ \Phi_\mathscr{K}(T, P)_2 = \{\mathbf{K}a \in \mathsf{KA}(\mathscr{K}) \mid \mathsf{OB}_{\mathscr{O},P} \models a\} \cup \\ \{hd(r) \mid r \in \mathscr{P} :\ hd(r) = \mathbf{K}a,\ \mathsf{OB}_{\mathscr{O},T} \not\models \neg a,\ bd^+(r) \subseteq P, \\ \mathbf{K}(bd^-(r)) \cap T = \emptyset\}\end{aligned}$$

https://doi.org/10.1017/S1471068421000168 Published online by Cambridge University Press

Intuitively, given a partition (T, P), the operator $\Phi_{\mathscr{K}}(\cdot, P)_1$, with P fixed, computes the set of true modal **K**-atoms w.r.t. (T, P) and operator $\Phi_{\mathscr{K}}(T, \cdot)_2$, with T fixed, computes the set of modal **K**-atoms that are possibly true w.r.t. (T, P).

Note that the least fixpoint of operator $\Phi_{\mathscr{K}}(\cdot, P)_1$ corresponds to an element in the sequence $\mathbf{P}_i$, that is, if P in $\Phi_{\mathscr{K}}(\cdot, P)_1$ is $\mathbf{N}_n$, then $lfp(\Phi_{\mathscr{K}}(\cdot, P)_1)$ is $\mathbf{P}_{n+1} = \Gamma_{\mathscr{K}}(\mathbf{N}_n)$. Similarly for operator $\Phi_{\mathscr{K}}(T, \cdot)_2$. In this way, the $\Phi_{\mathscr{K}}$ operator can be seen as a reformulation of the corresponding alternating fixpoint operator; namely, $\Phi_{\mathscr{K}}(\cdot, P)_1$ simulates operator $\mathscr{T}_{\mathscr{K}/P}$ and $\Phi_{\mathscr{K}}(T, \cdot)_2$ simulates operator $\mathscr{T}_{\mathscr{K}//T}$.

Proposition 2
$\Phi_{\mathscr{K}}$ is an approximator for $\mathscr{T}_{\mathscr{K}}$.

Proof
Let us check $\subseteq_p$-monotonicity of $\Phi_{\mathscr{K}}$. Let $(T_1, P_1) \subseteq_p (T_2, P_2)$. From $T_1 \subseteq T_2$ and $P_2 \subseteq P_1$, it is easy to verify that $\Phi_{\mathscr{K}}(T_1, P_1)_1 \subseteq \Phi_{\mathscr{K}}(T_2, P_2)_1$. For $\Phi_{\mathscr{K}}(T_2, P_2)_2 \subseteq \Phi_{\mathscr{K}}(T_1, P_1)_2$, note that $\Phi_{\mathscr{K}}(\cdot, \cdot)_2$ is defined in terms of two subsets. For the first subset, since $P_2 \subseteq P_1$, the set defined w.r.t. P_2 is a subset of the set defined w.r.t. P_1, that is, $\{\mathbf{K}a \in \mathsf{KA}(\mathscr{K}) \mid \mathsf{OB}_{\mathscr{O}, P_2} \models a\}$ is a subset of $\{\mathbf{K}a \in \mathsf{KA}(\mathscr{K}) \mid \mathsf{OB}_{\mathscr{O}, P_1} \models a\}$. For the second subset, along with $T_1 \subseteq T_2$, the set defined w.r.t. (T_2, P_2) is a subset of the set defined w.r.t. (T_1, P_1). Thus $\Phi_{\mathscr{K}}(T_1, P_1) \subseteq_p \Phi_{\mathscr{K}}(T_2, P_2)$. Furthermore, $\Phi_{\mathscr{K}}$ approximates $\mathscr{T}_{\mathscr{K}}$, since by definition $\Phi_{\mathscr{K}}(I, I)_1 \supseteq \Phi_{\mathscr{K}}(I, I)_2$, and it follows that whenever $\Phi_{\mathscr{K}}(I, I)$ is consistent, $\Phi_{\mathscr{K}}(I, I) = (\mathscr{T}_{\mathscr{K}}(I), \mathscr{T}_{\mathscr{K}}(I))$. □

Example 7
Consider a hybrid MKNF knowledge base $\mathscr{K} = (\mathscr{O}, \mathscr{P})$, where $\mathscr{O} = c \wedge (e \supset \neg r)$ and $\mathscr{P}$ consists of

$$\mathbf{K}r \leftarrow \mathbf{K}c, \mathbf{K}i, \mathbf{not}\, o, \mathbf{not}\, l. \qquad \mathbf{K}e \leftarrow . \qquad \mathbf{K}i \leftarrow .$$

One can derive that, for the exact pair $(T, T) = (\{\mathbf{K}c, \mathbf{K}i, \mathbf{K}e\}, \{\mathbf{K}c, \mathbf{K}i, \mathbf{K}e\})$, $\Phi_{\mathscr{K}}(T, T) = (\{\mathbf{K}c, \mathbf{K}i, \mathbf{K}e, \mathbf{K}r, \mathbf{K}o, \mathbf{K}l\}, \{\mathbf{K}c, \mathbf{K}i, \mathbf{K}e\})$. Operator $\Phi_{\mathscr{K}}$ maps the exact pair (T, T) to an inconsistent one, and it is therefore not a symmetric approximator. Note that the least stable fixpoint of $\Phi_{\mathscr{K}}$ is just the mapped inconsistent pair. It is interesting to see the information revealed in this stable fixpoint – while it is inconsistent, it provides consistent information on **K**-atoms, $\mathbf{K}c, \mathbf{K}i$, and $\mathbf{K}e$.

Our next goal is to show that the operator $\Phi_{\mathscr{K}}$ is a strong approximator under a mild condition. First, let us introduce some notations. Recall that we use the notation $(D_1(P), D_1(T)) = lfp(\Phi_{\mathscr{K}}(\cdot, P)_1, lfp(\Phi_{\mathscr{K}}(T, \cdot)_2)$, where $\Phi_{\mathscr{K}}(\cdot, P)_1$ is defined on $[\emptyset, P]$ and $\Phi_{\mathscr{K}}(T, \cdot)_2$ is defined on $[T, \mathsf{KA}(\mathscr{K})]$, and the notation $(C_1(P), C_2(T)) = lfp(\Phi_{\mathscr{K}}(\cdot, P)_1, lfp(\Phi_{\mathscr{K}}(T, \cdot)_2)$, where both $\Phi_{\mathscr{K}}(\cdot, P)_1$ and $\Phi_{\mathscr{K}}(T, \cdot)_2$ are operators on $\mathsf{KA}(\mathscr{K})$. We now give notations to refer to intermediate results in a least fixpoint construction (we define them here for $D_1(P)$ and $C_2(T)$; others are similar):

$$D_1^{\uparrow 0}(P) = \emptyset \qquad C_2^{\uparrow 0}(T) = \emptyset$$
$$D_1^{\uparrow k+1}(P) = \Phi_{\mathscr{K}}(D_1^{\uparrow k}(P), P)_1 \qquad C_2^{\uparrow k+1}(T) = \Phi_{\mathscr{K}}(T, C_2^{\uparrow k}(T))_2 \quad \text{for all } k \geq 0.$$

https://doi.org/10.1017/S1471068421000168 Published online by Cambridge University Press

Proposition 3

Let $\mathscr{K} = (\mathscr{O}, \mathscr{P})$ be a hybrid MKNF knowledge base and (T, P) be a consistent stable fixpoint of $\Phi_{\mathscr{K}}$ such that $\mathsf{OB}_{\mathscr{O},T}$ is satisfiable. Then $\Phi_{\mathscr{K}}$ is a strong approximator for (T, P).

Proof

Let (T, P) be a consistent stable fixpoint of $\Phi_{\mathscr{K}}$ such that $\mathsf{OB}_{\mathscr{O},T}$ is satisfiable. We show that $\Phi_{\mathscr{K}}$ is strong for (T, P). We need to show $(C_1(P), C_2(T)) = (D_1(P), D_2(T))$. That $C_1(P) = D_1(P)$ is immediate since the monotonicity of the projection operators implies that the construction of the least fixpoint in both cases starts with the same least element, $\emptyset$, and is carried out in tandem by the same mapping, and therefore terminates at the same fixpoint.

That $C_2(T) \subseteq D_2(T)$ is also easy to show by induction. The construction of the least fixpoint by $C_2(T)$ starts from $\emptyset$ and the one by $D_2(T)$ starts from T. So for the base case, $C_2^{\uparrow 0}(T) \subseteq D_2^{\uparrow 0}(T)$. Then, one can verify by definition that for any (fixed) $k \geq 0$, by the monotonicity of the projection operators on their respective domains, that $C_2^{\uparrow k}(T) \subseteq D_2^{\uparrow k}(T)$ implies $C_2^{\uparrow k+1}(T) \subseteq D_2^{\uparrow k+1}(T)$.

To show $D_2(T) \subseteq C_2(T)$, we first prove by induction that $D_1(P) \subseteq C_2(T)$. Since $D_1(P) = C_1(P) = T$, this is to show $T \subseteq C_2(T)$. The base case is immediate since both least fixpoint constructions start with the same least element $\emptyset$. Assume $D_1^{\uparrow k}(P) \subseteq C_2^{\uparrow k}(T)$ for any (fixed) $k \geq 0$, and we show it for $k+1$. By definition, a new $\mathbf{K}$-atom $\mathbf{K}\,a$ is added to $D_1^{k+1}(P)$ because (i) $\mathsf{OB}_{\mathscr{O},D_1^{\uparrow k}(P)} \models a$, or (ii) there is a rule $r \in \mathscr{P}$ with $hd(r) = \mathbf{K}\,a$ such that $bd^+(r) \subseteq D_1^{\uparrow k}(P)$ and $\mathbf{K}(bd^-(r)) \cap P = \emptyset$. If case (i) applies, by induction hypothesis (I.H.), it follows $\mathsf{OB}_{\mathscr{O},C_2^{\uparrow k}(T)} \models a$, and thus $\mathbf{K}\,a \in C_2^{\uparrow k+1}(T)$. Otherwise, $\mathbf{K}\,a$ is derived only by rules as in case (ii). Note that since $D_1^{\uparrow k}(P) \subseteq T$, case (ii) implies $\mathsf{OB}_{\mathscr{O},T} \models a$. If $\mathsf{OB}_{\mathscr{O},T} \models \neg a$, then $\mathsf{OB}_{\mathscr{O},T}$ is unsatisfiable, violating the assumption that $\mathsf{OB}_{\mathscr{O},T}$ is satisfiable. Thus, we must have $\mathsf{OB}_{\mathscr{O},T} \not\models \neg a$; then the same rule applied in case (ii) above applies in the construction of $C_2^{\uparrow k+1}$, since the condition ($bd^+(r) \subseteq D_1^{\uparrow k}(P)$ and $\mathbf{K}(bd^-(r)) \cap P = \emptyset$) becomes ($bd^+(r) \subseteq C_2^{\uparrow k}(T)$ and $\mathbf{K}(bd^-(r)) \cap T = \emptyset$)), which holds by I.H. and the fact that $T \subseteq P$. Thus, $D_1(P)$ $(= T) \subseteq C_2(T)$.

Once we obtain $T \subseteq C_2(T)$, we are ready to conclude that $lfp(\Phi_{\mathscr{K}}(T, \cdot)_2)$ with the operator $\Phi_{\mathscr{K}}(T, \cdot)_2$ defined on domain $[T, \mathsf{KA}(\mathscr{K})]$ is a subset of $lfp(\Phi_{\mathscr{K}}(T, \cdot)_2)$ with the operator $\Phi_{\mathscr{K}}(T, \cdot)_2$ defined on domain $\mathsf{KA}(\mathscr{K})$. This is because the construction of the former least fixpoint starts from the least element T of the domain $[T, \mathsf{KA}(\mathscr{K})]$, and the construction of the latter is guaranteed to reach a set $T' \supseteq T$, and by induction on both sequences in parallel, we have $D_2(T) \subseteq C_2(T)$. □

Stable fixpoints of the operator $\Phi_{\mathscr{K}}$ can be related to three-valued MKNF models of $\mathscr{K}$ in the following way.

Theorem 5

Let $\mathscr{K} = (\mathscr{O}, \mathscr{P})$ be a hybrid MKNF knowledge base and (T, P) be a partition. Let further $(M, N) = (\{I \mid I \models \mathsf{OB}_{\mathscr{O},T}\}, \{I \mid I \models \mathsf{OB}_{\mathscr{O},P}\})$. Then, (M, N) is a three-valued MKNF model of $\mathscr{K}$ iff (T, P) is a consistent stable fixpoint of $\Phi_{\mathscr{K}}$ and $\mathsf{OB}_{\mathscr{O},lfp(\Phi_{\mathscr{K}}(\cdot,T)_1)}$ is satisfiable.

https://doi.org/10.1017/S1471068421000168 Published online by Cambridge University Press

Note that in the formulation of approximator $\Phi_{\mathscr{K}}$, stable fixpoints are partitions that provide candidate interpretation pairs for three-valued MKNF models. The extra condition that $\mathsf{OB}_{\mathscr{O},lfp(\Phi_{\mathscr{K}}(\cdot,T)_1)}$ is satisfiable means that even if we make all **not**-atoms $\mathbf{not}\,\phi$ true when $\phi \notin T$, in the construction of $lfp(\Phi_{\mathscr{K}}(\cdot,T)_1)$, it still does not cause contradiction with the DL knowledge base $\mathscr{O}$. This provides a key insight in the semantics of hybrid MKNF knowledge bases.

Notice also that this theorem provides a naive method, based on guess-and-verify, to compute three-valued MKNF models of a given hybrid MKNF knowledge base $\mathscr{K}$ – guess a consistent partition (T, P) of $\mathsf{KA}(\mathscr{K})$ and check whether (T, P) is a stable fixpoint of $\Phi_{\mathscr{K}}$ and whether $\mathsf{OB}_{\mathscr{O},lfp(\Phi_{\mathscr{K}}(\cdot,T)_1)}$ is satisfiable. Observe that the complexity of checking for one guessed partition is polynomial if the underlying DL is polynomial.

Proof

($\Leftarrow$) Assume that (T, P) is a consistent stable fixpoint of $\Phi_{\mathscr{K}}$ and $\mathsf{OB}_{\mathscr{O},lfp(\Phi_{\mathscr{K}}(\cdot,T)_1)}$ is satisfiable. Let $\theta(x)$ denote $lfp(\Phi_{\mathscr{K}}(\cdot,x)_1)$, given $x \subseteq \mathsf{KA}(\mathscr{K})$. Thus, $\mathsf{OB}_{\mathscr{O},lfp(\Phi_{\mathscr{K}}(\cdot,T)_1)}$ is often written as $\mathsf{OB}_{\mathscr{O},\theta(T)}$. By the definition of the operator $\Phi_{\mathscr{K}}$ (cf. Definition 7), it can be seen that $\Phi_{\mathscr{K}}(\cdot,T)_1$ coincides with $\Phi_{\mathscr{K}}(T,\cdot)_2$ except for the extra condition $\mathsf{OB}_{\mathscr{O},T} \not\models \neg a$ (where $hd(r) = \mathbf{K}a$ for some $r \in \mathscr{P}$) in the definition of the latter. Note that the operator $\Phi_{\mathscr{K}}(T,\cdot)_2$ is defined on $\mathsf{KA}(\mathscr{K})$. It then follows $lfp(\Phi_{\mathscr{K}}(T,\cdot)_2) \subseteq \theta(T)$. Since (T, P) is a stable fixpoint of $\Phi_{\mathscr{K}}$, $lfp(\Phi_{\mathscr{K}}(T,\cdot)_2) = P$ and thus $P \subseteq \theta(T)$. Then, that $\mathsf{OB}_{\mathscr{O},\theta(T)}$ is satisfiable implies that $\mathsf{OB}_{\mathscr{O},P}$ is satisfiable, and because (T, P) is consistent and thus $T \subseteq P$, $\mathsf{OB}_{\mathscr{O},T}$ is satisafiable as well. It follows that the pair

$$(M, N) = (\{I \mid I \models \mathsf{OB}_{\mathscr{O},T}\}, \{I \mid I \models \mathsf{OB}_{\mathscr{O},P}\})$$

is an MKNF interpretation pair because $\emptyset \subset N \subseteq M$. As shown by Knorr *et al.* (2011), for any $\mathbf{K}\xi \in \mathsf{KA}(\mathscr{K})$, $\mathbf{K}\xi \in T$ iff $\mathbf{K}\xi$ evaluates to $\mathbf{t}$ (under (M, N)), $\mathbf{K}\xi \notin P$ iff $\mathbf{K}\xi$ evaluates to $\mathbf{f}$, and otherwise $\mathbf{K}\xi$ evaluates to $\mathbf{u}$ (also see the review of this property in Section 4.2, or Liu and You 2017 for more details).

We now show that (M, N) is a three-valued MKNF model of $\mathscr{K}$. First we show that (M, N) satisfies $\pi(\mathscr{K})$. Since $\mathsf{OB}_{\mathscr{O},T} = \{\pi(\mathscr{O})\} \cup \{\xi \mid \mathbf{K}\xi \in T\}$ and $\mathsf{OB}_{\mathscr{O},P} = \{\pi(\mathscr{O})\} \cup \{\xi \mid \mathbf{K}\xi \in P\}$, it follows $(M, N) \models \mathbf{K}\pi(\mathscr{O})$. Now consider any rule $r \in \mathscr{P}$. Let $hd(r) = \mathbf{K}a$. By the definition of $\Phi_{\mathscr{K}}(T,P)_1$, if $bd^+(r) \subseteq T$ and $\mathbf{K}(bd^-(r)) \cap P = \emptyset$, then $\mathbf{K}a \in T$; for $\Phi_{\mathscr{K}}(T,P)_2$, if $bd^+(r) \subseteq P$, $\mathbf{K}(bd^-(r)) \cap T = \emptyset$, and $\mathsf{OB}_{\mathscr{O},T} \not\models \neg a$, then $\mathbf{K}a \in P$. The case that $hd(r)$ evaluates to $\mathbf{t}$ (under (M, N)) is automatic. If $hd(r)$ evaluates to $\mathbf{u}$, that is, $\mathbf{K}a \in P$ and $\mathbf{K}a \notin T$, then $bd(r)$ evaluates to $\mathbf{u}$ or $\mathbf{f}$, since if $bd(r)$ evaluates to $\mathbf{t}$, $\mathbf{K}a \in lfp(\Phi_{\mathscr{K}}(\cdot,P)_1)(= T)$, resulting in a contradiction. If $hd(r)$ evaluates to $\mathbf{f}$, then $\mathsf{OB}_{\mathscr{O},T} \models \neg a$, in which case $bd(r)$ must evaluate to $\mathbf{f}$ as well, as otherwise $\mathbf{K}a \in \theta(T)$ $(= lfp(\Phi_{\mathscr{K}}(\cdot,T)_1))$ and thus $\mathsf{OB}_{\mathscr{O},\theta(T)}$ is unsatisfiable, leading to a contradiction. As this proof applies to all rules in $\mathscr{P}$, we have $(M, N) \models \pi(\mathscr{P})$, and with $(M, N) \models \mathbf{K}\pi(\mathscr{O})$, $(M, N) \models \pi(\mathscr{K})$.

Next, assume for the sake of contradiction (M, N) is not a three-valued MKNF model of $\mathscr{K}$. Then there exists a pair (M', N') with $M \subseteq M'$ and $N \subseteq N'$, where at least one of the inclusions is proper and $M' = N'$ if $M = N$, such that

$$(I, \langle M', N' \rangle, \langle M, N \rangle)(\pi(\mathscr{K})) = \mathbf{t} \tag{3}$$

https://doi.org/10.1017/S1471068421000168 Published online by Cambridge University Press

for some $I \in M'$. Let (T', P') be induced by (M', N'), that is,

$$(M', N') = (\{I \,|\, I \models \mathsf{OB}_{\mathcal{O},T'}\}, \{I \,|\, I \models \mathsf{OB}_{\mathcal{O},P'}\})$$

Clearly, $T' \subseteq T$ and $P' \subseteq P$, where at least one of the inclusions is proper and $T' = P'$ if $T = P$. We show that $(I, \langle M', N'\rangle, \langle M, N\rangle)(\pi(\mathcal{K})) \neq \mathbf{t}$ (independent of I), which leads to a contradiction.

Consider the case where $T' \subset T$. Let the sequence of intermediate sets of $\mathbf{K}$-atoms in the construction of $\mathit{lfp}(\Phi_{\mathcal{K}}(\cdot, P)_1)$ be $S_0, \ldots, S_n (= T)$. Assume step i $(0 \leq i \leq n-1)$ is the *first* iteration in which at least one $\mathbf{K}$-atom in $T \setminus T'$, say $\mathbf{K}a$, is derived. By the definition of $\Phi_{\mathcal{K}}(T, P)_1$, the derivation of $\mathbf{K}a$ is either by $\mathsf{OB}_{\mathcal{O},S_i} \models a$, or by a rule $r \in \mathcal{P}$ such that $\mathit{hd}(r) = \mathbf{K}a$, $\mathit{bd}^+(r) \subseteq S_i$, and $\mathbf{K}(\mathit{bd}^-(r)) \cap P = \emptyset$. For the latter case, by the assumption that i is the first iteration to derive any $\mathbf{K}$-atoms in $T \setminus T'$, $S_i \subseteq T'$. It follows that $\mathit{bd}(r)$ evaluates to $\mathbf{t}$ under $(I, \langle M', N'\rangle, \langle M, N\rangle)$ (independent of I), but its head $\mathit{hd}(r)$ evaluates to either $\mathbf{f}$ or $\mathbf{u}$; thus rule r is not satisfied, resulting in a contradiction to equation (3). If no $\mathbf{K}$-atom in $T \setminus T'$ is ever derived by a rule in iteration i, then it must be $\mathsf{OB}_{\mathcal{O},S_i} \models a$, and along with $S_i \subseteq T'$ and $\mathsf{OB}_{\mathcal{O},T'} \not\models a$, we derive a contradiction.

For the case where $P' \subset P$, the proof is similar. Consider the sequence of intermediate sets of $\mathbf{K}$-atoms $Q_0, \ldots, Q_n (= P)$ in the iterative construction of $\mathit{lfp}(\Phi_{\mathcal{K}}(T, \cdot)_2)$. Let j $(0 \leq j \leq n-1)$ be the first iteration in which at least one $\mathbf{K}$-atom in $P \setminus P'$ is derived (thus $Q_j \subseteq P'$). Let $\mathbf{K}a$ be such a $\mathbf{K}$-atom. Assume it is derived by a rule $r \in \mathcal{P}$ with $\mathit{hd}(r) = \mathbf{K}a$, such that $\mathsf{OB}_{\mathcal{O},T} \not\models \neg a$, $\mathit{bd}^+(r) \subseteq Q_j$, and $\mathbf{K}(\mathit{bd}^-(r)) \cap T = \emptyset$. Then, the body of rule r evaluates to $\mathbf{t}$ or $\mathbf{u}$ in $(I, \langle M', N'\rangle, \langle M, N\rangle)$. Since $\mathbf{K}a \notin P'$, $\mathit{hd}(r) = \mathbf{K}a$ evaluates to $\mathbf{f}$, and thus $(I, \langle M', N'\rangle, \langle M, N\rangle)(\pi(r)) \neq \mathbf{t}$. If in iteration j no fresh $\mathbf{K}$-atoms are derived by rules, then we must have $\mathsf{OB}_{\mathcal{O},Q_j} \models a$, and along with $Q_j \subseteq P'$ and $\mathsf{OB}_{\mathcal{O},P'} \not\models a$, we reach a contradiction. Note that the above proof is naturally applicable when $T' = P'$ because $T = P$, in which case evaluation reduces to two-valued. As both cases lead to a contradiction, (M, N) is therefore a three-valued MKNF model of $\mathcal{K}$.

($\Rightarrow$) Let $(M, N) = (\{I \,|\, I \models \mathsf{OB}_{\mathcal{O},T}\}, \{I \,|\, I \models \mathsf{OB}_{\mathcal{O},P}\})$ be a three-valued MKNF model of $\mathcal{K}$. Recall again that given an MKNF interpretation pair (M', N'), there exists a partition (X, Y) induced by (M', N'), in that $(M', N') = (\{I \,|\, I \models \mathsf{OB}_{\mathcal{O},X}\}, \{I \,|\, I \models \mathsf{OB}_{\mathcal{O},Y}\})$, such that for any $\mathbf{K}a \in \mathsf{KA}(\mathcal{K})$, $\mathbf{K}a \in X$ iff $\mathbf{K}a$ evaluates to $\mathbf{t}$ (under (M', N')), $\mathbf{K}a \notin Y$ iff $\mathbf{K}a$ evaluates to $\mathbf{f}$, and otherwise $\mathbf{K}a$ evaluates to $\mathbf{u}$. When (M, N) is a three-valued MKNF model of $\mathcal{K}$, the partition induced by (M, N) is just (T, P) such that $(M, N) = (\{I \,|\, I \models \mathsf{OB}_{\mathcal{O},T}\}, \{I \,|\, I \models \mathsf{OB}_{\mathcal{O},P}\})$. Since (M, N) is a three-valued MKNF model, (T, P) is consistent.

We now show that (T, P) is a stable fixpoint of $\Phi_{\mathcal{K}}$. First, we show that (T, P) is a fixpoint of $\Phi_{\mathcal{K}}$. By definition, $T \subseteq \Phi_{\mathcal{K}}(T, P)_1$. Assume $T \subset \Phi_{\mathcal{K}}(T, P)_1$ and let $\mathbf{K}a \notin T$ and $\mathbf{K}a \in \Phi_{\mathcal{K}}(T, P)_1$. Then there exists a rule r with $\mathit{hd}(r) = \mathbf{K}a$ such that $\mathbf{K}a$ can be derived due to satisfied body of rule r. It then follows $(M, N) \not\models \pi(r)$, contradicting to the three-valued MKNF model condition; thus $T = \Phi_{\mathcal{K}}(T, P)_1$. Similarly, we can show $P = \Phi_{\mathcal{K}}(T, P)_2$. If (T, P) is not a stable fixpoint, then either $T \neq \mathit{lfp}(\Phi_{\mathcal{K}}(\cdot, P)_1)$ or $P \neq \mathit{lfp}(\Phi_{\mathcal{K}}(T, \cdot)_2)$. For the former case, since T is a fixpoint of $\Phi_{\mathcal{K}}(\cdot, P)_1$, there exists $T' \subset T$ such that $T' = \mathit{lfp}(\Phi_{\mathcal{K}}(\cdot, P)_1)$. Consider the partition (T', P), for which we can construct an MKNF interpretation pair (M', N), where $M' = \{I \,|\, I \models \mathsf{OB}_{\mathcal{O},T'}\}$ and $M \subset M'$. It can be checked that $(I, \langle M', N\rangle, \langle M, N\rangle)(\pi(\mathcal{K})) = \mathbf{t}$, for any $I \in M'$. If $M = N$,

https://doi.org/10.1017/S1471068421000168 Published online by Cambridge University Press

one can verify that $(I, \langle M', M' \rangle, \langle M, M \rangle)(\pi(\mathcal{K})) = \mathbf{t}$ for any $I \in M'$. Thus, (M, N) is not a three-valued MKNF model of $\mathcal{K}$, a contradiction, and thus $T = \mathit{lfp}(\Phi_{\mathcal{K}}(\cdot, P)_1)$. Similarly, we can show $P = \mathit{lfp}(\Phi_{\mathcal{K}}(T, \cdot)_2)$. Therefore, (T, P) is a stable fixpoint of $\Phi_{\mathcal{K}}$, and a consistent one.

Finally, since (M, N) is a three-valued MKNF model of $\mathcal{K}$, in the construction of $\mathit{lfp}(\Phi_{\mathcal{K}}(T, \cdot)_2)$ the extra condition $\mathsf{OB}_{\mathcal{O},T} \not\models \neg a$ in the definition of $\Phi_{\mathcal{K}}(T, \cdot)_2$ always holds whenever the body of the relevant rule evaluates to $\mathbf{t}$. It follows $P = \mathit{lfp}(\Phi_{\mathcal{K}}(\cdot, T)_1) (= \theta(T))$, and since $\mathsf{OB}_{\mathcal{O},P}$ is satisfiable, $\mathsf{OB}_{\mathcal{O},\theta(T)}$ is satisfiable as well. □

Example 8
Consider a hybrid MKNF knowledge base $\mathcal{K} = (\{\neg a\}, \mathcal{P})$, where $\mathcal{P}$ consists of

$$\mathbf{K}a \leftarrow \mathbf{K}b. \qquad \mathbf{K}b \leftarrow \mathbf{not}\, b.$$

The least stable fixpoint of $\Phi_{\mathcal{K}}$ is $(T, P) = (\emptyset, \{\mathbf{K}b\})$, which is consistent but does not correspond to a three-valued MKNF model since $\mathsf{OB}_{\{\neg a\}, \mathit{lfp}(\Phi_{\mathcal{K}}(\cdot, T)_1)}$ is unsatisfiable.

6 A richer approximator for the well-founded semantics

A question arises whether richer approximators for MKNF knowledge bases exist. For any two approximators A and B on L^2, A is *richer than* B (or *more precise than* B, in the terminology of Denecker *et al.* 2004), denoted $B \leq_p A$, if for all $(x, y) \in L^2$, $B(x, y) \leq_p A(x, y)$.

There is a practical motivation for the question. Let (x, y) and (x', y') be the least stable fixpoints of B and A respectively. That A is richer than B means $(x, y) \leq_p (x', y')$. If A is strictly richer than B, and if (x', y') indeed corresponds to the well-founded MKNF model, then (x, y) cannot possibly correspond to the well-founded MKNF model. In this case, while (x', y') can be computed iteratively for A, it cannot be computed iteratively for B. Then, more complex reasoning method must be applied to compute the well-founded MKNF model. We now define such a richer approximator.

Definition 8
Let $\mathcal{K} = (\mathcal{O}, \mathcal{P})$ be a hybrid MKNF knowledge base. Define the operator $\Psi_{\mathcal{K}}$ on $(2^{\mathsf{KA}(\mathcal{K})})^2$ as follows: $\Psi_{\mathcal{K}}(T, P) = (\Psi_{\mathcal{K}}(T, P)_1, \Psi_{\mathcal{K}}(T, P)_2)$, where

$$\Psi_{\mathcal{K}}(T, P)_1 = \Phi_{\mathcal{K}}(T, P)_1$$
$$\Psi_{\mathcal{K}}(T, P)_2 = \{\mathbf{K}a \in \mathsf{KA}(\mathcal{K}) \mid \mathsf{OB}_{\mathcal{O},P} \models a\} \cup$$
$$\{\mathit{hd}(r) \mid r \in \mathcal{P} : \mathit{hd}(r) = \mathbf{K}a,\ \mathsf{OB}_{\mathcal{O},T} \not\models \neg a,\ \mathit{bd}^+(r) \subseteq P,\ \mathbf{K}(\mathit{bd}^-(r)) \cap T = \emptyset,$$
$$\nexists r' \in \mathcal{P} : \mathbf{K}b \leftarrow \mathit{bd}(r') \text{ is positive, where } \mathbf{K}a \in \mathit{bd}(r'), \text{ s.t. } \mathsf{OB}_{\mathcal{O},T} \models \neg b, \mathit{bd}(r') \setminus \{\mathbf{K}a\} \subseteq T\}$$

Operator $\Psi_{\mathcal{K}}$ differs from $\Phi_{\mathcal{K}}$ in the second projection operator, with an extra condition for deriving $\mathbf{K}a$ (the last two lines in the definition above), which says that if for some positive rule r' with $\mathbf{K}b$ as the head and $\mathbf{K}a$ in the body, the objective atom b is already false and the rule's body excluding $\mathbf{K}a$ is already true, then, since the rule must be satisfied, $\mathbf{K}a$ must be false and thus should not be derived as possibly true. Notice that this is like embedding the *unit propagation* rule in automated theorem proving into an approximator.

https://doi.org/10.1017/S1471068421000168 Published online by Cambridge University Press

Example 9
Let $\mathscr{K} = (\{\neg b\}, \mathscr{P})$, where $\mathscr{P}$ is

$$\mathbf{K}b \leftarrow \mathbf{K}a, \mathbf{K}e. \quad \mathbf{K}e \leftarrow \mathbf{not}\, p. \quad \mathbf{K}a \leftarrow \mathbf{not}\, c. \quad \mathbf{K}c \leftarrow \mathbf{not}\, a.$$

The least stable fixpoint of $\Phi_{\mathscr{K}}$ is $(T, P) = (\{\mathbf{K}e\}, \{\mathbf{K}e, \mathbf{K}a, \mathbf{K}c\})$, while the least stable fixpoint of operator $\Psi_{\mathscr{K}}$ is $(T', P') = (\{\mathbf{K}e, \mathbf{K}c\}, \{\mathbf{K}e, \mathbf{K}c\})$, which corresponds to the well-founded MKNF model of $\mathscr{K}$. Note that (T', P') is also a stable fixpoint of $\Phi_{\mathscr{K}}$; but because it is not the least, it cannot be computed by the standard iterative process.

Proposition 4
Operator $\Psi_{\mathscr{K}}$ is an approximator for $\mathscr{T}_{\mathscr{K}}$.

Proof
We can verify that $\Psi_{\mathscr{K}}$ is $\subseteq$-monotone on $(2^{\mathsf{KA}(\mathscr{K})})^2$. Let $(T_1, P_1) \subseteq_p (T_2, P_2)$. That $\Psi_{\mathscr{K}}(T_1, P_1)_1 \subseteq \Psi_{\mathscr{K}}(T_2, P_2)_1$ is immediate by definition. To show $\Psi_{\mathscr{K}}(T_2, P_2)_2 \subseteq \Psi_{\mathscr{K}}(T_1, P_1)_2$, we check all conditions in the definition of $\Psi_{\mathscr{K}}(T, P)_2$; in particular, let us consider the following conditions in the definition of $\Psi_{\mathscr{K}}(T, P)_2$:

$$\mathsf{OB}_{\mathscr{O},T} \not\models \neg a \tag{4}$$

$$\nexists r' \in \mathscr{P} : \mathbf{K}b \leftarrow bd(r') \text{ is positive, where } \mathbf{K}a \in bd(r'),$$
$$\text{s.t. } \mathsf{OB}_{\mathscr{O},T} \models \neg b \text{ and } bd(r') \setminus \{\mathbf{K}a\} \subseteq T \tag{5}$$

which may block a **K**-atom $\mathbf{K}a$ in the definition (cf. the second subset in the definition) to be included. There are three conditions in these expressions (one in (4) and two in (5)) that are determined by the following relationships under $(T_1, P_1) \subseteq_p (T_2, P_2)$: $\mathsf{OB}_{\mathscr{O},T_2} \models \mathsf{OB}_{\mathscr{O},T_1}$ and $T_2 \models T_1$. It is then easy to check that $\Psi_{\mathscr{K}}(T_2, P_2)_2 \subseteq \Psi_{\mathscr{K}}(T_1, P_1)_2$, and we therefore have $\Psi_{\mathscr{K}}(T_1, P_1) \subseteq_p \Psi_{\mathscr{K}}(T_2, P_2)$. Furthermore, $\Psi_{\mathscr{K}}$ approximates $\mathscr{T}_{\mathscr{K}}$, since by definition $\Psi_{\mathscr{K}}(I, I)_1 \supseteq \Psi_{\mathscr{K}}(I, I)_2$, and it follows that whenever $\Psi_{\mathscr{K}}(I, I)$ is consistent, $\Psi_{\mathscr{K}}(I, I) = (\mathscr{T}_{\mathscr{K}}(I), \mathscr{T}_{\mathscr{K}}(I))$. □

We show that $\Psi_{\mathscr{K}}$ is more precise than $\Phi_{\mathscr{K}}$.

Proposition 5
Given any hybrid MKNF knowledge base $\mathscr{K}$, $\Phi_{\mathscr{K}} \subseteq_p \Psi_{\mathscr{K}}$.

Proof
This is due to the extra condition in the definition of $\Psi_{\mathscr{K}}(T, \cdot)_2$, which is not present in the definition of $\Phi_{\mathscr{K}}(T, \cdot)_2$. A stronger condition produces a subset for the second component of the resulting pair. Thus, if $(T_1, P_1) \subseteq_p (T_2, P_2)$, from $\Psi_{\mathscr{K}}(T, x)_1 = \Phi_{\mathscr{K}}(T, x)_1$ for any $x \subseteq \mathsf{KA}(\mathscr{K})$, it follows that $\Phi_{\mathscr{K}}(T_1, P_1) \subseteq_p \Psi_{\mathscr{K}}(T_2, P_2)$. □

Operator $\Psi_{\mathscr{K}}$ is strong as well.

Proposition 6
Let $\mathscr{K} = (\mathscr{O}, \mathscr{P})$ be a hybrid MKNF knowledge base and (T, P) be a consistent stable fixpoint of $\Psi_{\mathscr{K}}$ such that $\mathsf{OB}_{\mathscr{O},T}$ is satisfiable. Then $\Psi_{\mathscr{K}}$ is a strong approximator for (T, P).

Proof

Let (T, P) be a consistent stable fixpoint of $\Psi_{\mathcal{K}}$ such that $\mathsf{OB}_{\mathcal{O},T}$ is satisfiable. We need to prove $(C_1(P), C_1(T)) = (D_1(P), D_2(T))$. The proof is identical to that of Proposition 3 except for the proof of $D_1(P) \subseteq C_2(T)$, which needs to be updated according to the definition of operator $\Psi_{\mathcal{K}}(T, \cdot)_2$. Recall that the goal is to show $T \subseteq C_2(T)$. The base case is again immediate. Assume $D_1^{\uparrow k}(P) \subseteq C_2^{\uparrow k}(T)$ for any (fixed) $k \geq 0$, and we show it for $k+1$. By definition, a $\mathbf{K}$-atom $\mathbf{K}\, a$ is added to $D_1^{k+1}(P)$ because (i) $\mathsf{OB}_{\mathcal{O},D_1^{\uparrow k}(P)} \models a$, or (ii) there is a rule $r \in \mathcal{P}$ with $hd(r) = \mathbf{K}\, a$ such that $bd^+(r) \subseteq D_1^{\uparrow k}(P)$ and $\mathbf{K}(bd^-(r)) \cap P = \emptyset$. Again, if case (i) applies, by induction hypothesis (I.H.), $\mathsf{OB}_{\mathcal{O},C_2^{\uparrow k}(T)} \models a$ and thus $\mathbf{K}\, a \in C_2^{\uparrow k+1}(T)$. Otherwise, $\mathbf{K}\, a$ is derived by a rule r, as in case (ii). Since $D_1^{\uparrow k}(P) \subseteq T$ and $\mathbf{K}(bd^-(r)) \cap P = \emptyset$, case (ii) implies $\mathsf{OB}_{\mathcal{O},T} \models a$. If $\mathsf{OB}_{\mathcal{O},T} \models \neg a$, then $\mathsf{OB}_{\mathcal{O},T}$ is unsatisfiable, which is a contradiction. Thus, we must have $\mathsf{OB}_{\mathcal{O},T} \not\models \neg a$. Now consider the condition

$$\nexists r' \in \mathcal{P} : \mathbf{K}\, b \leftarrow bd(r') \text{ is positive, where } \mathbf{K}\, a \in bd(r'), \text{ s.t. } \mathsf{OB}_{\mathcal{O},T} \models \neg b, bd(r') \setminus \{\mathbf{K}\, a\} \subseteq T$$

in the definition of $\Psi_{\mathcal{K}}(T, \cdot)_2$. By definition, $\mathbf{K} a$ is not derived by rule r in applying $\Psi_{\mathcal{K}}(T, \cdot)_2$ if such a rule r' exists. Since $\mathsf{OB}_{\mathcal{O},T} \models a$, from $bd(r') \setminus \{\mathbf{K} a\} \subseteq T$ we derive $bd(r') \subseteq T$, and therefore $\mathbf{K} b$ can be derived by rule r' resulting in $\mathbf{K} b \in T$, but at the same time we have $\mathsf{OB}_{\mathcal{O},T} \models \neg b$, and thus $\mathsf{OB}_{\mathcal{O},T}$ is unsatisfiable, a contradiction. Hence, such a rule r' does not exist. Therefore, the same rule applied in case (ii) for $D_1^{\uparrow k+1}$ above applies in the construction of $C_2^{\uparrow k+1}$, since the condition $(bd^+(r) \subseteq D_1^{\uparrow k}(P)$ and $\mathbf{K}(bd^-(r)) \cap P = \emptyset)$ becomes $(bd^+(r) \subseteq C_2^{\uparrow k}(T)$ and $\mathbf{K}(bd^-(r)) \cap T = \emptyset))$, which holds by I.H. and $T \subseteq P$. Thus, $D_1(P)\ (= T) \subseteq C_2(T)$. □

Finally, like Theorem 5, the stable fixpoints of the operator $\Psi_{\mathcal{K}}$ can be related to three-valued MKNF models of $\mathcal{K}$ as well.

Theorem 6

Let $\mathcal{K} = (\mathcal{O}, \mathcal{P})$ be a hybrid MKNF knowledge base and (T, P) be a partition. Let further $(M, N) = (\{I \mid I \models \mathsf{OB}_{\mathcal{O},T}\}, \{I \mid I \models \mathsf{OB}_{\mathcal{O},P}\})$. Then, (M, N) is a three-valued MKNF model of $\mathcal{K}$ iff (T, P) is a consistent stable fixpoint of $\Psi_{\mathcal{K}}$ and $\mathsf{OB}_{\mathcal{O},lfp(\Psi_{\mathcal{K}}(\cdot,T)_1)}$ is satisfiable.

Proof

The proof here follows the structure of the proof of Theorem 5, but with critical differences in dealing with approximator $\Psi_{\mathcal{K}}$. If a part of proof of Theorem 5 can be applied, we will make a reference to it, otherwise we will give a detailed proof even if parts of it repeat the same from the proof for Theorem 5.

($\Leftarrow$) Assume that (T, P) is a consistent stable fixpoint of $\Psi_{\mathcal{K}}$ and $\mathsf{OB}_{\mathcal{O},lfp(\Psi_{\mathcal{K}}(\cdot,T)_1)}$ is satisfiable. Let $\theta(x)$ denote $lfp(\Psi_{\mathcal{K}}(\cdot, x)_1)$, given $x \subseteq \mathsf{KA}(\mathcal{K})$. Let $P^* = lfp(\Phi_{\mathcal{K}}(T, \cdot)_2)$, then by the definition of operator $\Phi_{\mathcal{K}}$, $P^* \subseteq \theta(T)$, and by Proposition 5, $\Psi_{\mathcal{K}}(T, P^*)_2 \subseteq \Phi_{\mathcal{K}}(T, P^*)_2$, that is, $\Psi_{\mathcal{K}}(T, P^*)_2 \subseteq P^*$ and thus P^* is a pre-fixpoint of $\Psi_{\mathcal{K}}(T, \cdot)_2$; therefore $P \subseteq P^*$ and then $P \subseteq \theta(T)$. Since $\mathsf{OB}_{\mathcal{O},\theta(T)}$ is satisfiable, it follows that $\mathsf{OB}_{\mathcal{O},P}$ is satisfiable, and because $T \subseteq P$, $\mathsf{OB}_{\mathcal{O},T}$ is also satisfiable. It follows that the pair

$$(M, N) = (\{I \mid I \models \mathsf{OB}_{\mathcal{O},T}\}, \{I \mid I \models \mathsf{OB}_{\mathcal{O},P}\})$$

https://doi.org/10.1017/S1471068421000168 Published online by Cambridge University Press

is an MKNF interpretation pair because $\emptyset \subset N \subseteq M$. Recall the property by Knorr *et al.* (2011): given the above interpretation pair, for any $\mathbf{K}\xi \in \mathsf{KA}(\mathscr{K})$, $\mathbf{K}\xi \in T$ iff $\mathbf{K}\xi$ evaluates to $\mathbf{t}$ (under (M, N)), $\mathbf{K}\xi \notin P$ iff $\mathbf{K}\xi$ evaluates to $\mathbf{f}$, and otherwise $\mathbf{K}\xi$ evaluates to $\mathbf{u}$.

We show that (M, N) is a three-valued MKNF model of $\mathscr{K}$. First we show that (M, N) satisfies $\pi(\mathscr{K})$. The proof that $(M, N) \models \mathbf{K}\pi(\mathscr{O})$ is straightforward. For rules in $\mathscr{P}$, recall the following definition of $\Psi_{\mathscr{K}}(x, y)_2$:

$$\begin{aligned}\Psi_{\mathscr{K}}(x, y)_2 = \{&\mathbf{K}a \in \mathsf{KA}(\mathscr{K}) \mid \mathsf{OB}_{\mathscr{O},y} \models a\} \cup \\ &\{hd(r) \mid r \in \mathscr{P} :\ hd(r) = \mathbf{K}a,\ \mathsf{OB}_{\mathscr{O},x} \not\models \neg a,\ bd^+(r) \subseteq y,\ \mathbf{K}(bd^-(r)) \cap x = \emptyset, \\ &\nexists r' \in \mathscr{P} : \mathbf{K}b \leftarrow bd(r') \text{ is positive, where } \mathbf{K}a \in bd(r'),\ \text{s.t. } \mathsf{OB}_{\mathscr{O},x} \models \neg b, bd(r') \\ &\setminus \{\mathbf{K}a\} \subseteq x\}\end{aligned}$$

where the only difference from $\Phi_{\mathscr{K}}$ is the extra condition in the last two lines above. It can be checked that this extra condition does not effect the proof used for Theorem 5. Namely, for any rule $r \in \mathscr{P}$, it is satisfied if $hd(r)$ evaluates to $\mathbf{t}$; if $hd(r)$ evaluates to $\mathbf{u}$, which means $\mathbf{K}a \in P$ and $\mathbf{K}a \notin T$, then $bd(r)$ evaluates to $\mathbf{u}$ or $\mathbf{f}$, since if $bd(r)$ evaluates to $\mathbf{t}$, $\mathbf{K}a \in lfp(\Psi_{\mathscr{K}}(\cdot, P)_1)(= T)$, resulting in a contradiction; if $hd(r)$ evaluates to $\mathbf{f}$, then $\mathsf{OB}_{\mathscr{O},T} \models \neg a$, in which case $bd(r)$ evaluates to $\mathbf{f}$ as well, as otherwise $\mathbf{K}a \in \theta(T)$ $(= lfp(\Psi_{\mathscr{K}}(\cdot, T)_1))$ and thus $\mathsf{OB}_{\mathscr{O},\theta(T)}$ is unsatisfiable, again a contradiction. Hence, we conclude that $(M, N) \models \pi(\mathscr{P})$, and therefore $(M, N) \models \pi(\mathscr{K})$.

Now, assume for the sake of contradiction that (M, N) is not a three-valued MKNF model of $\mathscr{K}$. Then there exists an MKNF interpretation pair (M', N') with $M \subseteq M'$ and $N \subseteq N'$, where at least one of the inclusions is proper and $M' = N'$ if $M = N$, such that

$$(I, \langle M', N'\rangle, \langle M, N\rangle)(\pi(\mathscr{K})) = \mathbf{t} \tag{6}$$

for some $I \in M'$. Let (T', P') be induced by (M', N'), that is,

$$(M', N') = (\{I \,|\, I \models \mathsf{OB}_{\mathscr{O},T'}\}, \{I \,|\, I \models \mathsf{OB}_{\mathscr{O},P'}\})$$

Clearly, $T' \subseteq T$ and $P' \subseteq P$, where at least one of the inclusions is proper and $T' = P'$ if $T = P$. We show that $(I, \langle M', N'\rangle, \langle M, N\rangle)(\pi(\mathscr{K})) \neq \mathbf{t}$ (independent of I), which leads to contradiction.

Consider the case where $T' \subset T$. As in the proof of Theorem 5, this part of proof relies on the fixpoint construction of $lfp(\Psi_{\mathscr{K}}(\cdot, P)_1)$. Since by definition $\Psi_{\mathscr{K}}(T, P)_1 = \Phi_{\mathscr{K}}(T, P)_1$, the construction of $lfp(\Psi_{\mathscr{K}}(\cdot, P)_1)$ is identical to that of $lfp(\Phi_{\mathscr{K}}(\cdot, P)_1)$, the same proof of Theorem 5 for this part can be applied here, which shows $(I, \langle M', N'\rangle, \langle M, N\rangle)(\pi(\mathscr{K})) \neq \mathbf{t}$.

For the case of $P' \subset P$, consider the sequence of intermediate sets of $\mathbf{K}$-atoms $Q_0, \ldots, Q_n (= P)$ in the iterative construction of $lfp(\Psi_{\mathscr{K}}(T, \cdot)_2)$. Let j $(0 \leq j \leq n-1)$ be the first iteration in which at least one $\mathbf{K}$-atom, say $\mathbf{K}a \in P \setminus P'$, is derived (thus $Q_j \subseteq P'$). Assume further it is derived by a rule $r \in P$ with $hd(r) = \mathbf{K}a$, such that $\mathsf{OB}_{\mathscr{O},T} \not\models \neg a$, $bd^+(r) \subseteq Q_j$, $\mathbf{K}(bd^-(r)) \cap T = \emptyset$, and $\nexists r' \in \mathscr{P} : \mathbf{K}b \leftarrow bd^+(r')$ which is positive, where $\mathbf{K}a \in bd(r')$, s.t. $\mathsf{OB}_{\mathscr{O},T} \models \neg b$ and $bd^+(r') \setminus \{\mathbf{K}a\} \subseteq T$. Then, it can be seen that the body of rule r evaluates to $\mathbf{t}$ or $\mathbf{u}$ in $(I, \langle M', N'\rangle, \langle M, N\rangle)$. Since $\mathbf{K}a \notin P'$, $hd(r) = \mathbf{K}a$ evaluates to $\mathbf{f}$, and thus $(I, \langle M', N'\rangle, \langle M, N\rangle)(\pi(r)) \neq \mathbf{t}$. If in iteration j no fresh $\mathbf{K}$-atoms are derived by rules, then we must have $\mathsf{OB}_{\mathcal{O},Q_j} \models a$, and along with $Q_j \subseteq P'$ and $\mathsf{OB}_{\mathcal{O},P'} \not\models a$, we have a contradiction. Note that the above proof naturally applies when $T' = P'$ because of $T = P$, in which case evaluation reduces to two-valued. We therefore conclude that (M, N) is a three-valued MKNF model of $\mathcal{K}$.

($\Rightarrow$) Let $(M, N) = (\{I \mid I \models \mathsf{OB}_{\mathcal{O},T}\}, \{I \mid I \models \mathsf{OB}_{\mathcal{O},P}\})$ be a three-valued MKNF model of $\mathcal{K}$. As mentioned earlier, the following property holds: for any $\mathbf{K}\xi \in \mathsf{KA}(\mathcal{K})$, $\mathbf{K}\xi \in T$ iff $\mathbf{K}\xi$ evaluates to $\mathbf{t}$ (under (M, N)), $\mathbf{K}\xi \notin P$ iff $\mathbf{K}\xi$ evaluates to $\mathbf{f}$, and otherwise $\mathbf{K}\xi$ evaluates to $\mathbf{u}$.

By Theorem 5, (T, P) is a consistent stable fixpoint of $\Phi_{\mathcal{K}}$ and $\mathsf{OB}_{\mathcal{O},lfp(\Phi_{\mathcal{K}}(\cdot,T)_1)}$ is satisfiable. By definition, $\Psi_{\mathcal{K}}(\cdot, x)_1$ is the same operator as $\Phi_{\mathcal{K}}(\cdot, x)_1$ for all $x \subseteq \mathsf{KA}(\mathcal{K})$, and it follows $\mathsf{OB}_{\mathcal{O},lfp(\Psi_{\mathcal{K}}(\cdot,T)_1)}$ is also satisfiable and $T = lfp(\Psi_{\mathcal{K}}(\cdot, P)_1)$. Thus, for (T, P) to be a stable fixpoint of $\Psi_{\mathcal{K}}$, we only need to show $P = lfp(\Psi_{\mathcal{K}}(T, \cdot)_2)$. Let $P' = lfp(\Psi_{\mathcal{K}}(T, \cdot)_2)$. By definition (due to the extra condition in the definition of $\Psi_{\mathcal{K}}(T, \cdot)_2$), $P' \subseteq P$. For a contradiction, assume $P' \subset P$. Let $\mathbf{K}a \in P$ and $\mathbf{K}a \notin P'$. Then, by the definition of $\Psi_{\mathcal{K}}(T, \cdot)_2$, the reason for $\mathbf{K}a \notin P'$ is that, for any rule $r \in \mathcal{P}$ with $hd(r) = \mathbf{K}a$ such that $\mathsf{OB}_{\mathcal{O},T} \not\models \neg a$, $bd^+(r) \subseteq P$, and $\mathbf{K}(bd^-(r)) \cap T = \emptyset$, there exists a rule $r' \in \mathcal{P} : \mathbf{K}b \leftarrow bd^+(r')$ which is positive, where $\mathbf{K}a \in bd(r')$ s.t. $\mathsf{OB}_{\mathcal{O},T} \models \neg b$ and $bd^+(r') \setminus \{\mathbf{K}a\} \subseteq T$. If $\mathbf{K}a \in T$, then $\mathsf{OB}_{\mathcal{O},T} \models b$ and $\mathsf{OB}_{\mathcal{O},T}$ is thus unsatisfiable, contradicting to the fact that (M, N) is an MKNF interpretation pair. If $\mathbf{K}a \notin T$, since $\mathbf{K}a \in P$, the truth value of $\mathbf{K}a$ is undefined in (M, N), and thus r' is not satisfied by (M, N); again a contradiction. Thus, $P' = P$, and therefore (T, P) is a stable fixpoint of $\Psi_{\mathcal{K}}$. □

7 Summary, related work and remarks

The primary goal of this paper is to show that the alternating fixpoint operator formulated by Knorr *et al.* (2011) for hybrid MKNF knowledge bases is in fact an approximator of AFT, which can therefore be applied to characterize the well-founded semantics, two-valued semantics, as well as three-valued semantics for hybrid MKNF knowledge bases, and enables a better understanding of the relationships between these semantics in terms of a lattice structure.

Since this alternating fixpoint operator can map a consistent state to an inconsistent one, the desire to support operators like this motivated us to develop a mild generalization of AFT. As a result, all approximators defined on the entire product bilattice are well defined without the assumption of symmetry as required in the original AFT. In this paper, we studied the subtle issue whether consistent stable fixpoints can be preserved in the generalized AFT, and showed that for both approximators formulated in this paper for hybrid MKNF knowledge bases, consistent stable fixpoints are indeed carried over.

The alternating fixpoint construction by Knorr et al. aims at an iterative computation of the well-founded model. In Liu and You (2017), this construction is related to a notion called *stable partition* which exhibits properties corresponding to three-valued MKNF models. Based on the notion of stable partition, the relations between the alternating fixpoint construction and three-valued MKNF models are established. In this work, we do not use stable partition, instead we characterize three-valued MKNF models directly in terms of stable fixpoints of two appropriate approximators. In this way, we are able to show that the two approximators that are defined on the entire product bilattice capture

https://doi.org/10.1017/S1471068421000168 Published online by Cambridge University Press

the consistent stable fixpoints that lead to three-valued MKNF models, even though these approximators may have inconsistent stable fixpoints.

The only other work that treats inconsistency in AFT explicitly is Bi *et al.* (2014), where in case of inconsistency, instead of computing $(lfp(A(\cdot, v)_1), lfp(A(u, \cdot)_2))$ on the respective domains $[\bot, v]$ and $[u, \top]$, one computes $(lfp(A(\cdot, v)_1), A(u, v)_2)$ because $lfp(A(u, \cdot)_2)$ may no longer be defined on $[u, \top]$. By computing $A(u, v)_2$ for the second component of the resulting pair, non-minimal elements may be computed as sets of possibly true atoms when inconsistency arises.

The possibility of accommodating inconsistencies in AFT was first raised in Denecker *et al.* (2000). The precision order when applied to inconsistent pairs can be regarded as an order that measures the "degree of inconsistency", or "degree of doubt". If two inconsistent pairs satisfy $(x, y) \leq_p (x', y')$, the latter can be viewed at least as inconsistent as the former. In a more general context, researches have been trying to address questions like "where is the inconsistency", "how severer is it", and how to make changes to an inconsistent theory (see, e.g., Bona and Hunter 2017). A deeper understanding of inconsistencies in the context of AFT presents an interesting future direction.

In answer set programming, researchers have studied paraconsistent semantics. A noticeable example is the semi-stable semantics proposed by Sakama and Inoue (1995) for extended disjunctive logic programs, where a program transformation, called epistemic transformation, is introduced which embodies a notion of "believed to hold". The semantics is then characterized and enhanced by Amendola *et al.* 2016 using pairs of interpretations in the context of the logic of here-and-there (Pearce and Valverde 2008). For hybrid MKNF knowledge bases, Kaminski *et al.* (2015) propose a five-valued and a six-valued semantics for paraconsistent reasoning with different kinds of inconsistencies. An interpretation in this context is called a p-interpretation which evaluates a formula to true, false, or inconsistent. Since these semantics are formulated using the semantic structure consisting of pairs of interpretations, it is interesting to see whether appropriate approximators can be formulated to characterize intended models (of course, for the non-disjunctive case only since current AFT does not support disjunctive rules). In particular, since the alternating fixpoint constructions are already defined for Kaminski *et al.*'s five-valued and six-valued semantics, it may be possible to recast such an alternating fixpoint operator by an approximator. If successful, an interesting result would be that the underlying approximator defines not only the well-founded semantics but also five-valued and six-valued stable semantics. Furthermore, like Ji *et al.* (2017), due to the lattice structure of stable fixpoints, it may be possible to develop a DPLL-style solver for these knowledge bases based on a computation of unfounded atoms.

For disjunctive hybrid MKNF knowledge bases, the state-of-the-art reasoning method is still based on guess-and-verify as formulated by Motik and Rosati (2010). The lack of conflict-directed reasoning methods has prevented the theory from being tested in practice. Before any attempt to build a solver, one critical issue to study is the notion of unfounded sets for disjunctive hybrid MKNF knowledge bases, which has recently been investigated by Killen and You (2021), and another one is to develop a conflict-driven search engine for computing MKNF models.

References

Amendola, G., Eiter, T., Fink, M., Leone, N. and Moura, J. 2016. Semi-equilibrium models for paracoherent answer set programs. *Artificial Intellegence 234*, 219–271.

Antic, C., Eiter, T. and Fink, M. 2013. Hex semantics via approximation fixpoint theory. In *Proceedings of the 12th International Conference on Logic Programming and Nonmonotonic Reasoning, LPNMR-13*. Corunna, Spain, 102–115.

Bi, Y., You, J. and Feng, Z. 2014. A generalization of approximation fixpoint theory and application. In *Proceedings of the 8th International Conference on Web Reasoning and Rule Systems, RR-14*. Athens, Greece, 45–59.

Bogaerts, B. 2019. Weighted abstract dialectical frameworks through the lens of approximation fixpoint theory. In *Proceedings of the 33rd AAAI Conference on Artificial Intelligence, AAAI-19*. AAAI Press, Hawaii, USA, 2686–2693.

Bogaerts, B. and Cruz-Filipe, L. 2018. Fixpoint semantics for active integrity constraints. *Artificial Intellegence 255*, 43–70.

Bogaerts, B. and den Broeck, G. V. 2015. Knowledge compilation of logic programs using approximation fixpoint theory. *Theory and Practice of Logic Programming 15,* 4–5, 464–480.

Bogaerts, B., Jansen, J., Cat, B. D., Janssens, G., Bruynooghe, M.,and Denecker, M. 2016. Bootstrapping inference in the IDP knowledge base system. *New Generation Computing 34,* 3, 193–220.

Bogaerts, B., Vennekens, J. and Denecker, M. 2015. Grounded fixpoints and their applications in knowledge representation. *Artificial Intelligence 224*, 51–71.

Bogaerts, B., Vennekens, J. and Denecker, M. 2018. Safe inductions and their applications in knowledge representation. *Artificial Intelligence 259*, 167–185.

Bona, G. D. and Hunter, A. 2017. Localising iceberg inconsistencies. *Artificial Intelligence 246*, 118–151.

Denecker, M., Marek, V. and Truszczyński, M. 2000. Approximations, stable operators, well-founded fixpoints and applications in nonmonotonic reasoning. In *Logic-Based Artificial Intelligence*. Springer, 127–144.

Denecker, M., Marek, V. W. and Truszczynski, M. 2003. Uniform semantic treatment of default and autoepistemic logics. *Artificial Intelligence 143,* 1, 79–122.

Denecker, M., Marek, V. W. and Truszczynski, M. 2004. Ultimate approximation and its application in nonmonotonic knowledge representation systems. *Information and Computation 192,* 1, 84–121.

Denecker, M. and Vennekens, J. 2007. Well-founded semantics and the algebraic theory of non-monotone inductive definitions. In *Proceedings of the 9th International Conference on Logic Programming and Nonmonotonic Reasoning, LPNMR-07. Tempe, USA*. 84–96.

Dung, P. M. 1995. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificail Intelligence 77,* 2, 321–358.

Fitting, M. 2002. Fixpoint semantics for logic programming a survey. *Theoretical Computer Science 278,* 1–2, 25–51.

Gelfond, M. and Lifschitz, V. 1988. The stable model semantics for logic programming. In *Proceedings of the Fifth International Conference and Symposium on Logic Programming.* MIT Press, Seattle, Washington, 1070–1080.

Ji, J., Liu, F. and You, J. 2017. Well-founded operators for normal hybrid MKNF knowledge bases. *Theory and Practice of Logic Programming 17,* 5–6, 889–905.

Kaminski, T., Knorr, M. and Leite, J. 2015. Efficient paraconsistent reasoning with ontologies and rules. In *Proceedings of the 24th International Joint Conference on Artificial Intelligence, IJCAI 2015, Buenos Aires, Argentina, 25–31 July 2015*, Q. Yang and M. J. Wooldridge, Eds. AAAI Press, 3098–3105.

Killen, S. and You, J. 2021. Unfounded sets for disjunctive hybrid MKNF knowledge bases. In *Proceedings of the 18th International Conference on Principles of Knowledge Representation and Reasoning, KR-21 (to appear)*. IJCAI Inc.

Knorr, M., Alferes, J. J. and Hitzler, P. 2011. Local closed world reasoning with description logics under the well-founded semantics. *Artificial Intelligence 175,* 9–10, 1528–1554.

Lifschitz, V. 1991. Nonmonotonic databases and epistemic queries. In *Proceedings of the 12th Joint Conference on Artificial Intelligence, IJCAI-91*. Sydney, Australia, 381–386.

Liu, F. and You, J. 2017. Three-valued semantics for hybrid MKNF knowledge bases revisited. *Artificial Intelligence 252*, 123–138.

Liu, F. and You, J. 2019. Alternating fixpoint operator for hybrid MKNF knowledge bases as an approximator of AFT. In *Proceedings of the Third International Joint Conference on Rules and Reasoning, RuleML+RR 2019*, P. Fodor, M. Montali, D. Calvanese, and D. Roman, Eds. LNCS, vol. 11784. Springer, Bolzano, Italy, 113–127.

Markowsky, G. 1976. Chain-complete posets and directed sets with applications. *Algebra Universalis 6,* 1, 53–68.

Motik, B. and Rosati, R. 2007. A faithful integration of description logics with logic programming. In *Proceedings of the 19th International Joint Conference on Artificial Intelligence, IJCAI-07*. Hyderabad, India, 477–482.

Motik, B. and Rosati, R. 2010. Reconciling description logics and rules. *Journal of the ACM 57,* 5, 1–62.

Pearce, D. and Valverde, A. 2008. Quantified equilibrium logic and foundations for answer set programs. In *Proceedings of the 24th International Conference on Logic Programming, ICLP-08, Udine, Italy*, M. G. de la Banda and E. Pontelli, Eds. Lecture Notes in Computer Science, vol. 5366. Springer, 546–560.

Pelov, N., Denecker, M. and Bruynooghe, M. 2007. Well-founded and stable semantics of logic programs with aggregates. *Theory and Practice of Logic Programming 7,* 3, 301–353.

Sakama, C. and Inoue, K. 1995. Paraconsistent stable semantics for extended disjunctive programs. *Journal of Logic and Computation 5,* 3, 265–285.

Strass, H. 2013. Approximating operators and semantics for abstract dialectical frameworks. *Artificial Intelligence 205*, 39–70.

Tarski, A. 1955. A lattice-theoretical fixpoint theorem and its applications. *Pacific Journal of Mathematics 5*, 2, 285–309.

van Emden, M. H. and Kowalski, R. A. 1976. The semantics of predicate logic as a programming language. *The Journal of the ACM 23,* 4, 733–742.

Vennekens, J., Gilis, D. and Denecker, M. 2006. Splitting an operator: Algebraic modularity results for logics with fixpoint semantics. *ACM Transactions on Computational Logic 7,* 4, 765–797.