Approximations for Explanations of Inconsistency in Partially Known Multi-Context Systems*

Thomas Eiter, Michael Fink, and Peter Schüller

Institute of Information Systems
Vienna University of Technology
Favoritenstrasse 11, A-1040 Vienna, Austria
{eiter,fink,schueller}@kr.tuwien.ac.at

Abstract. Multi-context systems are a formalism to interlink decentralized and heterogeneous knowledge based systems (contexts), which interact via possibly nonmonotonic bridge rules. Inconsistency is a major problem, as it renders such systems useless. In applications involving confidentiality or trust, it is likely that complete knowledge about all system parts is unavailable. To address inconsistency in such scenarios, we extend existing notions for characterizing inconsistency in multi-context systems: we propose a representation of partial knowledge, and introduce a formalism for approximating reasons of inconsistency. We also discuss query selection strategies for improving approximations in situations where a limited number of queries can be posed to a partially known context.

1 Introduction

In recent years, there has been an increasing interest in interlinking knowledge bases, in order to enhance the capabilities of systems. Based on McCarthy's idea of contextual reasoning [11], the Trento School around Giunchiglia and Serafini has developed multi-context systems in many works, in which the components (called contexts) can be interlinked via so called bridge rules for information exchange, cf. [9, 5]. Generalizing this work, Brewka and Eiter [4] presented nonmonotonic multi-context systems (MCSs) as a generic framework for interlinking possibly heterogeneous and nonmonotonic knowledge bases.

Typically, an MCS is not built from scratch, but assembled from components which were not specifically designed to be part of a more complex system. Unintended interactions between contexts thus may easily arise and cause inconsistency, which renders an MCS useless. Making bridge rules defeasible [2] avoids inconsistency and cures faults in silent service. However, underlying reasons for inconsistency may remain unnoticed and cause unpleasant side-effects that are difficult to track.

Therefore, to help the user analyze, understand and eventually repair inconsistencies, suitable notions of consistency-based diagnosis and entailment-based explanation for inconsistency were introduced in [8]. Intuitively, diagnoses represent possible system repairs, while explanations characterize sources of inconsistency. An omniscient view of the system was assumed, where the user has full information about all contexts

* Supported by the Vienna Science and Technology Fund (WWTF) under grant ICT08-020.

including their knowledge bases and semantics. However, in many real world scenarios full information is not available [3], and some contexts are black boxes with internal knowledge bases or semantics that are not disclosed due to intellectual property or privacy issues (e.g., banks will not disclose their full databases to credit card companies). Partial behavior of such contexts may be known, however querying might be limited, e.g., by contracts or costs. In such scenarios, inconsistencies can only be explained given the knowledge of the system one has, and since this is partial, the explanations obtained just approximate the actual situation, i.e., those explanations one would obtain if one would have full insight.

In other words, this calls for explaining inconsistency in an MCS with *partial knowledge* about contexts, which raises the following technical challenges:

- how to represent partial knowledge about the system, and
- how to obtain reasonable *approximations* forexplanations of inconsistency in the actual system (under full knowledge), ideally in an efficient way.

The first issue depends on the nature of this knowledge, and a range of possibilities exists. The second issue requires an assessment methodto determine such approximations. We tackle both issues and make the following contributions.

• We develop a representation of partially known contexts, which is based on context abstraction with Boolean functions. Partially defined Boolean functions [15, 7] are used to capture partially known behavior of a context.
• We exploit these representations to define *over-* and *underapproximations* of diagnoses and explanations for inconsistency according to [8], in the presence of partially known contexts. The approximations target either the whole set of diagnoses, or one diagnosis at a time; analogously for explanations.
• For scenarios where partially known contexts can be asked a limited number of queries, we consider query selection strategies.
• Finally, we discuss computational complexity of recognizing approximate explanations. In contrast to semantic approximations for efficient evaluation [12], our approximations handle incompleteness, which usually increases complexity. Fortunately, our approach does not incur higher computational cost than the case of full information.

Our results extend methods for inconsistency handling in MCSs to more realistic settings, e.g., in health-care where privacy issues need to be respected, without increasing computational cost. In practical applications our approximations reduce the set of system parts relevant for restoring consistency, and allow for better focussing of repair efforts.

2 Preliminaries

A heterogeneous nonmonotonic MCS [4] consists of *contexts*, each composed of a knowledge base with an underlying *logic*, and a set of *bridge rules*

A logic $L = (\mathbf{KB}_L, \mathbf{BS}_L, \mathbf{ACC}_L)$ is an abstraction, which allows to capture many monotonic and nonmonotonic logics, e.g., classical logic, description logics, default logics, etc. It consists of the following components:

- $\mathbf{KB}_L$ is the set of well-formed knowledge bases of L. We assume each element of $\mathbf{KB}_L$ is a set of "formulas".
- $\mathbf{BS}_L$ is the set of possible belief sets, where a belief set is a set of "beliefs".
- $\mathbf{ACC}_L : \mathbf{KB}_L \rightarrow 2^{\mathbf{BS}_L}$ is a function describing the semantics of the logic by assigning to each knowledge base a set of acceptable belief sets.

Each context has its own logic, which allows to model heterogeneous systems.

A *bridge rule* models information flow between contexts: it can add information to a context, depending on the belief sets accepted at other contexts. Let $L = (L_1, \ldots, L_n)$ be a tuple of logics. An L_k-bridge rule r over L is of the form

$$(k : s) \leftarrow (c_1 : p_1), \ldots, (c_j : p_j), \mathbf{not}\ (c_{j+1} : p_{j+1}), \ldots, \mathbf{not}\ (c_m : p_m). \quad (1)$$

where $1 \leq c_i \leq n$, p_i is an element of some belief set of L_{c_i}, and k refers to the context receiving formula s. We denote by $hd\,(r)$ the formula s in the head of r.

Definition 1. *A multi-context system $M = (C_1, \ldots, C_n)$ is a collection of contexts $C_i = (L_i, kb_i, br_i)$, $1 \leq i \leq n$, where $L_i = (\mathbf{KB}_i, \mathbf{BS}_i, \mathbf{ACC}_i)$ is a logic, $kb_i \in \mathbf{KB}_i$ a knowledge base, and br_i is a set of L_i-bridge rules over $(L_1, \ldots, L_n)$. By $IN_i = \{hd\,(r) \mid r \in br_i\}$ we denote the set of possible* inputs *of context C_i added by bridge rules, and by $br_M = \bigcup_{i=1}^{n} br_i$ the set of all bridge rules of M.*

In addition, for each $H \subseteq IN_i$ we must have $kb_i \cup H \in \mathbf{KB}_{L_i}$.

The following running example involves policies and trust information which are often non-public and distributed [3], and thus demonstrates the necessity of reasoning under incomplete information. For more examples of MCSs see [4, 8].

Example 1. Consider an MCS M consisting of a permission database $C_1 = C_{\mathbf{perm}}$ and a credit card clearing context $C_2 = C_{\mathbf{cc}}$, and the following bridge rules:

$$\begin{aligned}
r_1 &: (\mathbf{perm} : \mathit{person}(\mathit{Person})) && \leftarrow \top. \\
r_2 &: (\mathbf{cc} : \mathit{card}(\mathit{CreditCard})) && \leftarrow (\mathbf{perm} : \mathit{person}(\mathit{Person})), \\
& && \quad\ \mathbf{not}\ (\mathbf{perm} : \mathit{grant}(\mathit{Person})), \\
& && \quad\ (\mathbf{perm} : \mathit{ccard}(\mathit{Person}, \mathit{CreditCard})). \\
r_3 &: (\mathbf{perm} : \mathit{ccValid}(\mathit{CreditCard})) && \leftarrow (\mathbf{cc} : \mathit{valid}(\mathit{CreditCard})).
\end{aligned}$$

Here r_1 defines a set of persons which is relevant for permission evaluation in $C_{\mathbf{perm}}$; r_2 specifies that, if some person is not granted access, credit cards of that person have to be checked; and r_3 translates validation results to $C_{\mathbf{perm}}$.

The MCS formalism is defined on ground bridge rules, which are in the following denoted by $r_{i,<constants>}$, e.g., $r_{2,moe,cnr2}$ denotes r_2 with $\mathit{Person} \mapsto \mathit{moe}$ and $\mathit{CreditCard} \mapsto \mathit{cnr2}$. Unless stated otherwise, we assume that bridge rules are grounded with $\mathit{Person} \in \{\mathit{nina}, \mathit{moe}\}$ and $\mathit{CreditCard} \in \{\mathit{cnr1}, \mathit{cnr2}\}$.

We next describe the context internals: $C_{\mathbf{perm}}$ is a datalog program with the following logic: $\mathbf{KB}_{\mathbf{perm}}$ contains all syntactically correct datalog programs, $\mathbf{BS}_{\mathbf{perm}}$ contains all possible answer sets, and $\mathbf{ACC}_{\mathbf{perm}}$ returns for each datalog program the corresponding answer sets. The knowledge base $kb_{\mathbf{perm}}$ is as follows:

$group(nina, vip)$. $\quad ccard(nina, cnr1)$. $\quad ccard(moe, cnr2)$.

$$\begin{aligned} igrant(Person) &\leftarrow person(Person), group(Person, vip). \\ grant(Person) &\leftarrow igrant(Person). \\ grant(Person) &\leftarrow ccValid(CreditCard), ccard(Person, CreditCard). \end{aligned}$$

Context $C_{\mathbf{cc}}$ is a credit card clearing facility, which typically is neither fully disclosed to the operator, nor can it be queried without significant cost. Hence, one obviously has to deal with partial knowledge: $C_{\mathbf{cc}}$ accepts $valid(X)$ iff card X is valid and validation is requested by $card(X)$. Without full insight or a history of past requests, weonly know the behavior of $C_{\mathbf{cc}}$ when no bridge rules are applicable: $\mathbf{ACC_{cc}}(kb_{\mathbf{cc}} \cup \emptyset) = \{\emptyset\}$. $\square$

Equilibrium semantics selects certain belief states of an MCS $M = (C_1, \ldots, C_n)$ as acceptable. A *belief state* is a sequence $S = (S_1, \ldots, S_n)$, s.t. $S_i \in \mathbf{BS}_i$. A bridge rule (1) is *applicable* in S iff for $1 \leq i \leq j$: $p_i \in S_{c_i}$ and for $j < l \leq m$: $p_l \notin S_{c_l}$. Let $app(R, S)$ denote the set of bridge rules in R that are applicable in belief state S.

Intuitively, an equilibrium is a belief state S, where each context C_i takes into account the heads of all bridge rules that are applicable in S, and accepts S_i.

Definition 2. *A belief state $S = (S_1, \ldots, S_n)$ of M is an equilibrium iff, for $1 \leq i \leq n$, the following condition holds: $S_i \in \mathbf{ACC}_i(kb_i \cup \{hd(r) \mid r \in app(br_i, S)\})$. By $\mathrm{EQ}(M)$ we denote the set of equilibria of M.*

Example 2 (ctd). Assume that M_1 is the MCS M with just $person(nina)$ present at $C_{\mathbf{perm}}$. As $nina$ is in the vip group there is no need to verify a credit card, and M_1 has the following equilibrium (we omit facts, that are present in $kb_{\mathbf{perm}}$): $(\{person(nina), igrant(nina), grant(nina)\}, \emptyset)$. $\square$

Inconsistency in an MCS is the lack of an equilibrium. No information can be obtained from an inconsistent MCS, i.e., reasoning tasks on equilibria become trivial. Therefore we analyze inconsistency in order to explain and eventually repair it.

Explanation of Inconsistency. We use the notions of consistency-based *diagnosis* and entailment-based *inconsistency explanation* in MCSs [8], which aim at describing inconsistency by sets of involved bridge rules.

Given an MCS M and a set R of bridge rules, by $M[R]$ we denote the MCS obtained from M by replacing its set of bridge rules br_M with R (in particular, $M[br_M] = M$ and $M[\emptyset]$ is M with no bridge rules). By $M \models \bot$ we denote that M is inconsistent, i.e., $\mathrm{EQ}(M) = \emptyset$, and by $M \not\models \bot$ the opposite. For any set of bridge rules A, $heads(A) = \{\alpha \leftarrow \mid \alpha \leftarrow \beta \in A\}$ are the rules in A in unconditional form. For pairs $A = (A_1, A_2)$ and $B = (B_1, B_2)$ of sets, the pointwise subset relation $A \subseteq B$ holds iff $A_1 \subseteq B_1$ and $A_2 \subseteq B_2$. We denote by $S|_A$ the projection of all sets X in set S to set A, formally $S|_A = \{X \cap A \mid X \in S\}$.

Definition 3. *Given an MCS M, a* diagnosis *of M is a pair (D_1, D_2), $D_1, D_2 \subseteq br_M$, s.t. $M[br_M \setminus D_1 \cup heads(D_2)] \not\models \bot$. $D^{\pm}(M)$ is the set of all such diagnoses. $D^{\pm}_m(M)$ is the set of all pointwise subset-minimal diagnoses of an MCS M.*

A diagnosis points out bridge rules which need to be modified to restore consistency; each rule can either be deactivated, or added unconditionally. Diagnoses represent concrete system repairs by these two basic actions, but thus also characterize more sophisticated ways of repair [8]. Moreover, weassume that context knowledge bases are consistent if no bridge rule heads are added (i.e., $\forall C_i : \mathbf{ACC}_i(kb_i) \neq \emptyset$) so restoring consistency is always possible (by removing all bridge rules). For more background and discussion of this notion, we refer to [8]; for inconsistency explanations cf. also Definition 9. We next give an example of an inconsistent MCS and its diagnoses.

Example 3 (ctd). Let M_2 be the MCS M with just $person(moe)$ present at $C_{\mathbf{perm}}$, and assume the following full knowledge about $C_{\mathbf{cc}}$: all credit cards are valid.

M_2 is inconsistent: moe is not in the vip group, card verification is required by $r_{2,moe,cnr2}$, and $C_{\mathbf{cc}}$ accepts $valid(cnr2)$. This allows $C_{\mathbf{perm}}$ to derive $grant(moe)$, which blocks applicability of $r_{2,moe,cnr2}$. Therefore, M_2 contains an unstable cycle and is inconsistent. Two $\subseteq$-minimal diagnoses of M_2 are then as follows: $(\{r_{2,moe,cnr2}\}, \emptyset)$ (do not validate $cnr2$), and $(\emptyset, \{r_{2,moe,cnr2}\})$ (always validate $cnr2$).[1] This points out r_2 as a likely culprit of inconsistency. Indeed, r_2 should intuitively contain $igrant(Person)$ in its body instead of $grant(Person)$. □

In this work we develop an approach which is able to point out a problem in r_2, without requiring complete knowledge.

3 Information Hiding

In this section, we introduce an abstraction of contexts which allows us to calculate diagnoses and explanations. We generalize this abstraction to represent partial knowledge, i.e., contexts C_i where either kb_i, or $\mathbf{ACC}_i$ is only partially known.

Context Abstraction. We abstract from a context's knowledge base kb_i and logic L_i by a Boolean function over the context's inputs IN_i (see Definition 1) and over the context's *output beliefs* OUT_i, which are those beliefs p in $\mathbf{BS}_i$ that occur in some bridge rule body in br_M as "$(i{:}p)$" or as "$\mathbf{not}\ (i{:}p)$" (see also [8]).

Recall that a Boolean function (BF) is a map $f : \mathbb{B}^k \to \mathbb{B}$ where $k \in \mathbb{N}$ and $\mathbb{B} = \{0, 1\}$. Such a BF can also be characterized either by its true points $T(f) = \{\vec{x} \mid f(\vec{x}) = 1\}$, or by its false points $F(f) = \{\vec{x} \mid f(\vec{x}) = 0\}$.

Given a set $X \subseteq U = \{u_1, \ldots, u_k\}$, we denote by $\vec{x}_U$ the characteristic vector of X wrt. some universe U (i.e. $\vec{x}_U = (b_1, \ldots, b_k)$, where $b_i = 1$ if $u_i \in X$, 0 otherwise). If understood, we omit U. Using this notation, we characterize sets of bridge rule heads $I \subseteq IN_i$ and sets of output beliefs $O \subseteq OUT_i$ by vectors $\vec{\imath}_{IN_i}$ and $\vec{o}_{OUT_i}$, respectively. For example, given $O = \{a, c\}$, and $OUT_i = \{a, b, c\}$, we have $\vec{o} = (1, 0, 1)$.

Example 4 (ctd). We use the following (ordered) sets for inputs and output beliefs: $IN_{\mathbf{cc}} = \{card(cnr1), card(cnr2)\}$, and $OUT_{\mathbf{cc}} = \{valid(cnr1), valid(cnr2)\}$. □

Definition 4. *The unique BF $f^{C_i} : \mathbb{B}^{|IN_i|+|OUT_i|} \to \mathbb{B}$ corresponds to the semantics of context C_i in an MCS M as follows:*

$$\forall I \subseteq IN_i, O \subseteq OUT_i : f^{C_i}(\vec{\imath}, \vec{o}) = 1 \textit{ iff } O \in \mathbf{ACC}_i(kb_i \cup I)\big|_{OUT_i}.$$

[1] Other $\subseteq$-minimal diagnoses of M_2 are $(\{r_{1,moe}\}, \emptyset)$, $(\{r_{3,cnr2}\}, \emptyset)$, and $(\emptyset, \{r_{3,cnr2}\})$.

Example 5 (ctd). With full knowledge (see Example 3), $C_{\mathbf{cc}}$ has as corresponding BF the function $f^{C_{\mathbf{cc}}}(X, Y, X, Y) = 1$ for all $X, Y \in \mathbb{B}$, 0 otherwise. □

If a context accepts a belief set O' for a given input I, we obtain the true point $(\vec{\imath}, \vec{o})$ of f with $O = O' \cap OUT_i$. Similarly, each non-accepted belief set yields a false point of f. Due to projection, different accepted belief sets can characterize the same true point.

Consistency Checking. Context abstraction provides sufficient information to calculate *output-projected equilibria* of the given MCS. Hence, it also allows for checking consistency and calculating diagnoses and explanations.

Given a belief state $S = (S_1, \ldots, S_n)$ in MCS M, the *output-projected belief state* $S' = (S'_1, \ldots, S'_n)$, $S'_i = S_i \cap OUT_i$, $1 \leq i \leq n$, is the projection of S to the output beliefs of M. In the following, we implicitly use the prime "$'$" to denote output-projection.

Definition 5 (see also [8]). *An output-projected belief state $S' = (S'_1, \ldots, S'_n)$ of an MCS M is an* output-projected equilibrium *iff, for $1 \leq i \leq n$, it holds that $S'_i \in \mathbf{ACC}_i(kb_i \cup \{hd(r) \mid r \in app(br_i, S')\})\big|_{OUT_i}$.*
By $\mathrm{EQ}'(M)$ we denote the set of output-projected equilibria of M.

Since $app(br_i, S) = app(br_i, S')$, a simple consequence is:

Lemma 1 ([8]). *For each equilibrium S of an MCS M, S' is an output-projected equilibrium; conversely, for each output-projected equilibrium S' of M there exists at least one equilibrium T of M such that $T' = S'$.*

This means, that output-projected equilibria provide precise characterizations of equilibria on beliefs which are relevant for bridge rule applicability, i.e., on output beliefs, but are indifferent on all other beliefs. The representation of a context by a BF provides an input/output oracle, projected to output beliefs. Therefore, the BF is sufficient for consistency checking as well.

Thus, towards a representation of an MCS with partial knowledge of certain contexts, we next provide a notation for an MCS M where the knowledge of a context C_i is given by BF f, rather than kb_i.

Definition 6. *Given MCS $M = (C_1, \ldots, C_n)$, BF f and index $1 \leq i \leq n$. We denote by $M[i/f]$ the MCS M where context C_i is replaced by a context $C(f)$ which contains the set br_i of bridge rules, a logic with a signature that contains $IN_i \cup OUT_i$, and $kb_{C(f)}$ and $\mathbf{ACC}_{C(f)}$, such that $f^{C(f)} = f^{C_i}$.*

For instance, $C(f)$ could be based on classical logic or logic programming, with $kb_{C(f)}$ over $IN \cup OUT$ as atoms encoding f by clauses (rules) that realize the correspondence.

We now show that a BF representation of a context is sufficient for calculating output-projected equilibria. We denote by $M[i_1, \ldots, i_k/f_1, \ldots, f_k]$ the substitution of pairwise distinct contexts $C_{i_1}, \ldots, C_{i_k}$ by $C(f_1), \ldots, C(f_k)$, respectively.

Theorem 1. *Let $M = (C_1, \ldots, C_n)$ be an MCS, and let $f_{i_1}, \ldots, f_{i_k}$ be BFs that correspond to $C_{i_1}, \ldots, C_{i_k}$. Then, $\mathrm{EQ}'(M) = \mathrm{EQ}'(M[i_1, \ldots, i_k/f_{i_1}, \ldots, f_{i_k}])$.*

Partially Known Contexts. As the BF representation concerns only output beliefs, it already hides part of the context, while we are still able to analyze inconsistency. Now we generalize the BF representation to *partially defined Boolean functions* (pdBFs) (cf. [15, 7]), to represent contexts where we have only partial knowledge about their output-projected behavior.

In applications, existence of such partial knowledge is realistic: for some bridge rule firings one may know an accepted belief set of a context, but not whether other accepted belief sets exist. Similarly one may know that a context is inconsistent for some input combination, but not whether it accepts some belief set for other input combinations.

Formally, a pdBF pf is a function from $\mathbb{B}^k$ to $\mathbb{B} \cup \{\star\}$, where $\star$ stands for undefined (cf. [15]). It is equivalently characterized by two sets [7]: its true points $T(pf) = \{\vec{\mathrm{x}} \mid pf(\vec{\mathrm{x}}) = 1\}$ and its false points $F(pf) = \{\vec{\mathrm{x}} \mid pf(\vec{\mathrm{x}}) = 0\}$. We denote by $U(pf) = \{\vec{\mathrm{x}} \mid pf(\vec{\mathrm{x}}) = \star\}$ the *unknown points* of pf. A BF f is an *extension* of a pdBF pf, formally $pf \leq f$, iff $T(pf) \subseteq T(f)$ and $F(pf) \subseteq F(f)$.

We connect partial knowledge of context semantics and pdBFs as follows.

Definition 7. *A pdBF $pf : \mathbb{B}^k \to \mathbb{B} \cup \{\star\}$ is compatible with a context C_i in an MCS M iff $pf \leq f^{C_i}$ (where f^{C_i} is as in Definition 4).*

Therefore, if a pdBF is compatible with a context, one extension of this pdBF is exactly f^{C_i}, which corresponds to the context's exact semantics.

Example 6 (ctd). Partial knowledge as given in Example 1 can be expressed by the pdBF $pf_{\mathbf{cc}}$ with $T(pf_{\mathbf{cc}}) = \{(0, 0, 0, 0)\}$ and $F(pf_{\mathbf{cc}}) = \{(0, 0, A, B) \mid A, B \in \mathbb{B}, (A, B) \neq (0, 0)\}$. (See Example 4 for the variable ordering.) ☐

In the following, a *partially known MCS* (M, i, pf) consists of an MCS M, where context C_i is partially known, given by pdBF pf which is compatible with C_i.

4 Approximations

In this section, we develop a method for calculating under- and overapproximations of diagnoses and explanations, using the pdBF representation for a partially known context C_i. For simplicity, we only consider the case that a single context in the system is partially known (the generalization is straightforward).

Diagnoses. Each diagnosis is defined in terms of consistency, which is witnessed by an output-projected equilibrium. Such an equilibrium requires a certain set of output beliefs O to be accepted by the context C_i, in the presence of certain bridge rule heads I. This means that f_{C_i} has true point $(\vec{\mathrm{I}}, \vec{\mathrm{o}})$. For existence of an equilibrium where C_i gets I as input and accepts O, no more information is required from f_{C_i} than this single true point. We thus can approximate the set of diagnoses of M as follows:

- Completing pf with false points, we obtain the extension $\underline{pf}$ with $T(\underline{pf}) = T(pf)$. The set of diagnoses witnessed by $T(pf)$ contains a *subset* of the diagnoses which actually occur in M, therefore we obtain an *under*approximation.
- Completing pf with true points, we obtain the extension $\overline{pf}$ as the extension of pf with the largest set of true points. The set of diagnoses witnessed by $\overline{pf}$ contains a *superset* of the diagnoses which actually occur in M, providing an *over*approximation. Formally,

Theorem 2. *Given a partially known MCS* (M, i, pf)*, the following holds:*

$$D^{\pm}(M[i/\underline{pf}]) \subseteq D^{\pm}(M) \subseteq D^{\pm}(M[i/\overline{pf}]).$$

Example 7 (ctd). The extensions $\overline{pf}_{\mathbf{cc}}$ and $\underline{pf}_{\mathbf{cc}}$ are as follows:

$$T(\overline{pf}_{\mathbf{cc}}) = \mathbb{B}^4 \setminus F(pf_{\mathbf{cc}}), \qquad F(\overline{pf}_{\mathbf{cc}}) = F(pf_{\mathbf{cc}}),$$
$$T(\underline{pf}_{\mathbf{cc}}) = T(pf_{\mathbf{cc}}), \text{ and } \qquad F(\underline{pf}_{\mathbf{cc}}) = \mathbb{B}^4 \setminus T(pf_{\mathbf{cc}}).$$

The underapproximation $D^{\pm}(M_2[\mathbf{cc}/\underline{pf}_{\mathbf{cc}}])$ yields several diagnoses, for instance, $D_\alpha = (\{r_{1,moe}\}, \emptyset)$, $D_\beta = (\{r_{2,moe,cnr2}\}, \emptyset)$, and $D_\gamma = (\emptyset, \{r_{3,cnr2}\})$.

The overapproximation $D^{\pm}(M_2[\mathbf{cc}/\overline{pf}_{\mathbf{cc}}])$ contains the empty diagnosis, i.e., $D_\delta = (\emptyset, \emptyset)$, because $M_2[\mathbf{cc}/\overline{pf}_{\mathbf{cc}}]$ is consistent; the latter has the following two equilibria: $(\{person(moe)\}, \emptyset)$ and $(\{person(moe)\}, \{valid(cnr1)\})$. □

Subset-minimality. If we approximate ⊆-minimal diagnoses, the situation is different. Obtaining additional diagnoses may cause an approximated diagnosis to be subset-minimal which is no diagnosis under full knowledge. However, at least one minimal diagnosis under full knowledge is a superset of the former.Vice versa, missing certain diagnoses can yield an approximated subset-minimal diagnoses which is a superset of (at least one) minimal diagnosis. However, if a diagnoses is subset-minimal under both, over- and underapproximation, then it is also a minimal diagnosis under full knowledge.

Theorem 3. *Given a partially known MCS* (M, i, pf)*, the following hold:*

$$\forall D \in D^{\pm}_m(M[i/\underline{pf}])\ \exists D' \in D^{\pm}_m(M) : D' \subseteq D \tag{2}$$
$$\forall D \in D^{\pm}_m(M)\ \exists D' \in D^{\pm}_m(M[i/\overline{pf}]) : D' \subseteq D \tag{3}$$
$$D^{\pm}_m(M[i/\underline{pf}]) \cap D^{\pm}_m(M[i/\overline{pf}]) \subseteq D^{\pm}_m(M) \tag{4}$$

Example 8 (ctd). Note that the diagnoses in Example 7 are in fact the ⊆-minimal diagnoses of the under- and overapproximation, and they are actual ⊆-minimal diagnoses. Under complete knowledge (Example 3), additional ⊆-diagnoses exist which are not obtained by underapproximation. Overapproximation, on the other hand, yields consistency and therefore an empty ⊆-minimal diagnosis D_δ. In Section 5 we develop a strategy for improving this approximation if limited querying of the context is possible. □

We can use the overapproximation to reason about the necessity of bridge rules in actual diagnoses: a necessary bridge rule is present in all diagnoses.[2]

Definition 8. *For a set of diagnoses* $\mathcal{D}$*, the set of* necessary bridge rules *is* $nec(\mathcal{D}) = \{r \mid \forall (D_1, D_2) \in \mathcal{D} : r \in D_1 \cup D_2\}$.

Proposition 1. *Given a partially known MCS* (M, i, pf)*, the set of necessary bridge rules for the overapproximation is necessary in the actual set of diagnoses. This is true for both arbitrary and ⊆-minimal diagnoses:*

$$nec(D^{\pm}(M[i/\overline{pf}])) \subseteq nec(D^{\pm}(M)), \text{ and } nec(D^{\pm}_m(M[i/\overline{pf}])) \subseteq nec(D^{\pm}_m(M)).$$

[2] Note that we do not consider the dual notion of relevance, as it is trivial in our definition of diagnosis: all bridge rules are relevant in any $D^{\pm}(M)$.

While simple, this property is useful in practice: in a repair of an MCS according to a diagnosis, necessary bridge rules need to be fixed in any case.

Inconsistency explanations. So far we have only described approximations for *diagnoses*. We now extend our notions to *inconsistency explanations* (in short 'explanations'), which are dual characterizations to diagnoses [8]. Intuitively, they point out bridge rules such that in the presence of bridge rules E_1 and the absence of bridge rules E_2 the MCS necessarily is inconsistent. Thus explanations allow to separate independent sources of inconsistency, while diagnoses characterize repairs. We first recall their definition.

Definition 9. *Given an MCS M, an* inconsistency explanation *of M is a pair (E_1, E_2) s.t. for all (R_1, R_2) where $E_1 \subseteq R_1 \subseteq br_M$ and $R_2 \subseteq br_M \setminus E_2$, it holds that $M[R_1 \cup heads(R_2)] \models \bot$. By $E^{\pm}(M)$ we denote the set of all inconsistency explanations of M, and by $E^{\pm}_m(M)$ the set of all pointwise subset-minimal ones.*

Example 9. With complete knowledge as in Example 3, there is one $\subseteq$-minimal explanation: $(\{r_{1,moe}, r_{2,moe,cnr2}, r_{3,cnr2}\}, \{r_{2,moe,cnr2}, r_{3,cnr2}\})$. □

Explanations are defined in terms of non-existing equilibria, therefore we can use witnessing equilibria as counterexamples. From the definitions we get:

Proposition 2. *For a given MCS M and a pair $(D_1, D_2) \subseteq br_M \times br_M$ of sets of bridge rules, the following statements are equivalent:*

(i) *(D_1, D_2) is a diagnosis, i.e., $(D_1, D_2) \in D^{\pm}(M)$,*
(ii) *$M[br_M \setminus D_1 \cup heads(D_2)]$ has an equilibrium, and*
(iii) *$(R_1, R_2) = (br_M \setminus D_1, D_2)$ is a counterexample for all explanation candidates $(E_1, E_2) \subseteq (br_M \setminus D_1, br_M \setminus D_2)$.*

Furthermore, such pairs (D_1, D_2) characterize all counterexamples *that can exist for explanation candidates.*

As a consequence, it is possible to characterize explanations in terms of diagnoses.

Lemma 2. *Given an MCS M, a pair (E_1, E_2) with $E_1, E_2 \subseteq br_M$ is an inconsistency explanation of M iff there exists no diagnosis $(D_1, D_2) \in D^{\pm}(M)$ such that $(D_1, D_2) \subseteq (br_M \setminus E_1, br_M \setminus E_2)$.*

In fact we can sharpen the above by replacing $D^{\pm}$ with $D^{\pm}_m$.

Using this characterization, we can infer the following: a subset of the actual set of diagnoses characterizes a superset of the actual set of explanations. This is true since a subset of diagnoses will rule out a subset of explanations, allowing more candidates to become explanations. Conversely, a superset of diagnoses characterizes a subset of the explanations. Applying Theorem 2, we obtain:

Theorem 4. *Given a partially known MCS (M, i, pf), the following hold:*

$$E^{\pm}(M[i/\overline{pf}]) \subseteq E^{\pm}(M) \subseteq E^{\pm}(M[i/\underline{pf}])$$
$$\forall E \in E^{\pm}_m(M[i/\overline{pf}])\ \exists E' \in E^{\pm}_m(M) : E' \subseteq E$$
$$\forall E \in E^{\pm}_m(M)\ \exists E' \in E^{\pm}_m(M[i/\underline{pf}]) : E' \subseteq E$$

Therefore, the extensions $\overline{pf}$ and $\underline{pf}$ allow to underapproximate and overapproximate diagnoses as well as inconsistency explanations.

Example 10 (ctd). From $\underline{pf}_{\mathbf{cc}}$ as in Example 7 we obtain one $\subseteq$-minimal explanation: $E_\mu = (\{r_{1,moe}, r_{2,moe,cnr2}\}, \{r_{3,cnr2}\})$. This explanation is a subset of the actual minimal explanation in Example 9. $\square$

5 Limited Querying

Up to now we used existing partial knowledge to approximate diagnoses, assuming that more information is simply not available. However, in practical scenarios like our running example, one can imagine that a (small) limited number of queries to a partially known context can be issued. Therefore we next aim at identifying queries to contexts, such that incorporating their answers into the pdBF will yield the best guarantee of improvement in approximation accuracy.

Given a partially known MCS (M, i, pf), let $D^{\pm}_{\Delta}(M, i, pf) = D^{\pm}(M[i/\overline{pf}]) \setminus D^{\pm}(M[i/\underline{pf}])$ (in short: $D^{\pm}_{\Delta}(pf)$ or $D^{\pm}_{\Delta}$) be the set of *potential diagnoses*, which are possible from the overapproximation but unconfirmed by the underapproximation. A large set of potential diagnoses provides less information than a smaller set. Hence, we aim at identifying unknown points of pf which remove from $D^{\pm}_{\Delta}$ as many potential diagnoses as possible. To this end we introduce the concept of a *witness* as an unknown point and a potential diagnosis that is supported by this point if it is a true point.

Definition 10. *Given a partially known MCS (M, i, pf), a* witness *is a pair $(\vec{x}, D)$ s.t. $\vec{x} \in U(pf)$ and $D \in D^{\pm}(M[i/f_{\vec{x}}]) \cap D^{\pm}_{\Delta}$, where $f_{\vec{x}}$ is the BF with the single true point $T(f_{\vec{x}}) = \{\vec{x}\}$. We denote by $W_{(M,i,pf)}$ the set of all witnesses wrt. (M, i, pf). If clear from the context, we omit subscript (M, i, pf).*

Based on W we define the set $wnd(\vec{x}) = \{D \mid (\vec{x}, D) \in W\}$ of potential diagnoses witnessed by unknown point $\vec{x}$, and the set $ewnd(\vec{x}) = \{D \in wnd(\vec{x}) \mid \nexists \vec{x}' \neq \vec{x} : (\vec{x}', D) \in W\}$ of potential diagnoses exclusively witnessed by $\vec{x}$. These sets are used to investigate how much the set of potential diagnoses is reduced when adding information about the value of an unknown point $\vec{x}$ to pf.

Lemma 3. *Given a partially known MCS (M, i, pf), and $\vec{x} \in U(pf)$, let $pf_{\vec{x}:0}$ ($pf_{\vec{x}:1}$) the pdBF that results from pf by making $\vec{x}$ a false (true) point. Then $D^{\pm}_{\Delta}(pf_{\vec{x}:1}) = D^{\pm}_{\Delta}(pf) \setminus wnd(\vec{x})$, and $D^{\pm}_{\Delta}(pf_{\vec{x}:0}) = D^{\pm}_{\Delta}(pf) \setminus ewnd(\vec{x})$.*

Note that $ewnd(\vec{x}) \subseteq wnd(\vec{x}) \subseteq D^{\pm}_{\Delta}$. If $\vec{x}$ is a true point, $|wnd(\vec{x})|$ many potential diagnoses become part of the underapproximation; otherwise $|ewnd(\vec{x})|$ many potential diagnoses are no longer part of the overapproximation. Knowing the value of $\vec{x}$ therefore guarantees a reduction of $D^{\pm}_{\Delta}$ by $|ewnd(\vec{x})|$ diagnoses.

Proposition 3. *Given a partially known MCS (M, i, pf), for all $\vec{x} \in U(pf)$ such that the cardinality of $ewnd(\vec{x})$ is maximal, the following holds:*

$$\max_{u \in \mathbb{B}} \left| D^{\pm}_{\Delta}\left(pf_{\vec{x}:u}\right) \right| \leq \min_{\vec{y} \in U(pf)} \max_{v \in \mathbb{B}} \left| D^{\pm}_{\Delta}\left(pf_{\vec{y}:v}\right) \right| . \quad (5)$$

Proposition 3 suggests to query unknown points $\vec{x}$ where $|ewnd(\vec{x})|$ is maximum. If there are more false points than true points (e.g., for contexts that accept only one belief set for each input), using $ewnd$ instead of wnd is even more suggestive. If the primary interest are necessary bridge rules (cf. previous section), we can base query selection on the number of bridge rules which become necessary if a certain unknown point is a false point. Let $nwnd(\vec{x}) = nec(\overline{D^{\pm}} \setminus ewnd(\vec{x})) \setminus nec(\overline{D^{\pm}})$, where $\overline{D^{\pm}} = D^{\pm}(M[i/\overline{pf}])$, then $|nwnd(\vec{x})|$ many bridge rules become necessary if $\vec{x}$ is identified as a false point.

Another possible criterion for selecting queries can be based on the likelihood of errors, similar to the idea of *leading diagnoses* [10]. Although a different notion of diagnosis is used there, the basic idea is applicable to our setting: if multiple problematic bridge rules are less likely than single ones, or if we have confidence values for bridge rules (e.g., some were designed by an expert, others by a less experienced administrator), then we can focus confirming or discarding diagnoses that have a high probability. If we have equal confidence in all bridge rules, this amounts to using *cardinality-minimal* potential diagnoses for determining witnesses and guiding the selection of queries.

Example 11 (ctd). In our example, the set of potential diagnoses is large, but the cardinality-minimal diagnosis is the empty diagnosis, which has the following property: bridge rule input at $C_{\mathbf{cc}}$ is $\{card(cnr2)\}$, and $C_{\mathbf{cc}}$ either accepts $\emptyset$ or $\{valid(cnr1)\}$ (the unrelated credit card). Therefore, points $(0, 1, 0, 0)$ and $(0, 1, 1, 0)$ are the only witnesses for D_{δ}, and querying these two unknown points is sufficient for verifying or falsifying D_{δ}. (Note that $pf_{\mathbf{cc}}$ has 12 unknown points, the four known points (one true and three false points) are $(0, 0, X, Y)$ s.t. $X, Y \in \mathbb{B}$.) After updating pf with these points (false points, if all credit cards are valid), the overapproximation yields the $\subseteq$-minimal diagnoses; this result is optimal. □

Instead of membership queries which check whether $O \in \mathbf{ACC}(kb \cup I)$ for given $(\vec{\imath}, \vec{o})$, one could use stronger queries that provide the *value* of $\mathbf{ACC}(kb \cup I)$ for a given $\vec{\imath}$. On the one hand this allows for a better query selection, roughly speaking because combinations of unknown points witness more diagnoses exclusively than they do individually. On the other hand, considering such combinations increases computational cost. Another extension of limited querying is the usage of meta-information, e.g., monotonicity, or consistency properties, of a partially known context.

6 Discussion

Approximation Quality. In the previous section, we related unknown points to potential diagnoses. This correspondence allows to obtain an estimate for the quality of an approximation, simply by calculating the ratio between known and potential true (resp., false) points: a high value of $\frac{|T(pf)|}{|T(pf)|+|U(pf)|}$ indicates a high underapproximation quality, while a low value indicates an underapproximation distant from the actual system. This is analogous for overapproximation, exchanging $T(pf)$ with $F(pf)$. These estimates can be calculated efficiently and prior to calculating an approximation; a decision between under- and overapproximation could be based on this heuristic. Concerning quality note also that even if nothing is known about the behavior of some context C, the overapproximation accurately characterizes inconsistencies that do not involve C.

Complexity and Computation. Since our approximation methods deal with incomplete knowledge, it is important how their computational complexity compares to the full knowledge case. For the latter setting, the following results were established in [8], depending on the complexity of output checking for contexts C_i, which is deciding for C_i, $I \subseteq IN_i$ and $O \subseteq OUT_i$ whether $O \in \mathbf{ACC}(kb_i \cup I)|_{OUT_i}$. With output checking in $\mathbf{P}$ (resp., $\mathbf{NP}$, $\mathbf{\Sigma_k^P}$), recognizing correct diagnoses is in $\mathbf{NP}$ (resp., $\mathbf{NP}$, $\mathbf{\Sigma_k^P}$) while recognizing minimal diagnoses and minimal explanations is in $\mathbf{D^P}$ (resp., $\mathbf{D^P}$, $\mathbf{D_k^P}$); completeness holds in all cases.

Let us first consider the case where some contexts C_i are given by their corresponding BF f_i (such that $f_i(\vec{\imath}, \vec{o})$ can be evaluated efficiently). As we know that context C_i accepts only input/output combinations which are true points of f, we simply guess all possible output beliefs O_i of all contexts and evaluate bridge rules to obtain I_i; if for some C_i as above, $f_i(\vec{\imath}_i, \vec{o}_i)=0$ we reject, otherwise we continue checking context acceptance for other contexts. Overall, this leads to the same complexity as if all contexts were total. Thus, detecting explanations of inconsistency for an MCS M, where some contexts are given as BFs, has the same complexity as if M were given regularly.

Approximations are done on an MCS where a pdBF pf is given instead of a BF f, in a representation such that the value of $pf(\vec{\imath}, \vec{o})$ can be computed efficiently. This implies that the extensions $\underline{pf}$ and $\overline{pf}$ can be computed efficiently as well. Hence, approximations of diagnoses and explanations have the same complexity as the exact concepts. Dealing with incomplete information usually increases complexity, as customary for many nonmonotonic reasoning methods. Our approach, however, exhibits no such increase in complexity, even though it provides faithful under- and overapproximations.

Learning. To learn a BF seems suggestive for our setting of incomplete information. However, explaining inconsistency requires correct information, therefore pac-learning methods [15] are not applicable. On the other hand, exact methods [1] require properties of the contexts which are beneficial to learning and might not be present.[3] Furthermore, contexts may only allow membership queries, which are insufficient for efficient learning of many concept domains [1]. Furthermore, partially known contexts may not allow many, even less a polynomial number of queries (which is the target for learnability). Most likely it will thus not be possible to learn the complete function. Hence learning cannot replace our approach, but it can be useful as a preprocessing step to increase the amount of partial information.

7 Related Work and Conclusion

To the best of our knowledge, explaining inconsistency in multi-context systems with partial specification has not been addressed before. Weakly related to our work is [14], who aimed at approximating abductive diagnoses of a single knowledge base. They replaced classical entailment with approximate entailment of [12], motivated by computational efficiency. However, there is no lack of information about the knowledge base or semantics as in our case.

[3] Note that, even if a context's logic is monotonic (resp., positive) this does not imply that the BF corresponding to the context is monotonic (resp., positive).

Our over- and underapproximations of $D^{\pm}$ and $E^{\pm}$ are reminiscent of lower and upper bounds of classical theories (viewed as sets of models [13]), known as cores and envelopes. The latter also were used for (fast) sound, resp. complete, reasoning from classical theories.

The limited querying approach is related to optimal probing strategies [6]. However, we do not require probing to localize faults in the system, but to obtain information about the behavior of system parts, which have a much more fine grained inner structure and more intricate dependencies than the systems in [6]. (Those system parts have as possible states 'up', and 'down', while in MCSs each partially known context possibly accepts certain belief sets for certain inputs.)

Ongoing further work includes an implementation of the approach given in this paper, and the usage of metainformation about context properties to improve approximation accuracy. The incorporation of probabilistic information into the pdBF representation is another interesting topic for future research.

References

1. Angluin, D.: Queries and concept learning. Machine Learning 2, 319–342 (1988)
2. Bikakis, A., Antoniou, G.: Distributed defeasible contextual reasoning in ambient computing. In: Ambient Intelligence. pp. 308–325 (2008)
3. Bonatti, P.A., Olmedilla, D.: Rule-based policy representation and reasoning for the semantic web. In: Reasoning Web. pp. 240–268 (2007)
4. Brewka, G., Eiter, T.: Equilibria in heterogeneous nonmonotonic multi-context systems. In: AAAI. pp. 385–390 (2007)
5. Brewka, G., Roelofsen, F., Serafini, L.: Contextual default reasoning. In: IJCAI. pp. 268–273 (2007)
6. Brodie, M., Rish, I., Ma, S., Odintsova, N.: Active probing strategies for problem diagnosis in distributed systems. In: IJCAI. pp. 1337–1338 (2003)
7. Crama, Y., Hammer, P.L., Ibaraki, T.: Cause-effect relationships and partially defined boolean functions. Annals of Operations Research 16, 299–326 (1988)
8. Eiter, T., Fink, M., Schüller, P., Weinzierl, A.: Finding explanations of inconsistency in nonmonotonic multi-context systems. In: KR. pp. 329–339 (2010)
9. Giunchiglia, F., Serafini, L.: Multilanguage hierarchical logics, or: How we can do without modal logics. Artificial Intelligence 65(1), 29–70 (1994)
10. de Kleer, J.: Focusing on probable diagnoses. In: AAAI. pp. 842–848 (1991)
11. McCarthy, J.: Notes on formalizing context. In: IJCAI. pp. 555–562 (1993)
12. Schaerf, M., Cadoli, M.: Tractable reasoning via approximation. Artificial Intelligence 74(2), 249–310 (1995)
13. Selman, B., Kautz, H.: Knowledge Compilation and Theory Approximation. J. ACM 43(2), 193–224 (1996)
14. ten Teije, A., van Harmelen, F.: Computing approximate diagnoses by using approximate entailment. In: KR. pp. 256–265 (1996)
15. Valiant, L.G.: A theory of the learnable. Commun. ACM 27, 1134–1142 (1984)