Optimizing Computation of Repairs from Active Integrity Constraints

Luís Cruz-Filipe*

Dept. of Mathematics and Computer Science, University of Southern Denmark, Denmark
LabMag, Lisboa, Portugal

Abstract. Active integrity constraints (AICs) are a form of integrity constraints for databases that not only identify inconsistencies, but also suggest how these can be overcome. The semantics for AICs defines different types of repairs, but deciding whether an inconsistent database can be repaired is a NP- or Σ_p^2-complete problem, depending on the type of repairs one has in mind. In this paper, we introduce two different relations on AICs: an equivalence relation of *independence*, allowing the search to be parallelized among the equivalence classes, and a *precedence* relation, inducing a stratification that allows repairs to be built progressively. Although these relations have no impact on the worst-case scenario, they can make significant difference in the practical computation of repairs for inconsistent databases.

1 Introduction

Maintaining and guaranteeing database consistency is one of the major problems in knowledge management. Database dependencies have been since long a main tool in the fields of relational and deductive databases [2,3], used to express integrity constraints on databases. They formalize relationships between data in the database that need to be satisfied so that the database conforms to its intended meaning.

Whenever an integrity constraint is violated, the database must be repaired in order to regain consistency. Typically there are several sets of update actions that achieve this goal, leading to different revised consistent databases. Restricting the set of database repairs to those considered most adequate is therefore an important task. Minimality of change is commonly accepted as an essential characteristic of a repair [6,8,18], but it is not enough to narrow down the set of possible repairs sufficiently.

The most common approach to processing integrity constraints in database management systems is to use active rules (a kind of event-condition-action rules, or ECAs [17]), for which rule processing algorithms have been proposed and a procedural semantics has been defined. However, their lack of declarative

* Supported by the Danish Council for Independent Research, Natural Sciences and by Fundação para a Ciência e Tecnologia under contract PEst-OE/EEI/UI0434/2011.

C. Beierle and C. Meghini (Eds.): FoIKS 2014, LNCS 8367, pp. 361–380, 2014.

semantics makes it difficult to understand the behaviour of multiple ECAs acting together and to evaluate rule-processing algorithms in a principled way.

Active integrity constraints (AICs) [9] are special forms of production rules that encode both an integrity constraint and preferred update actions to be performed whenever the former is violated. The declarative semantics for AICs [4,5] is based on the concept of founded and justified repairs. Informally, justified repairs are the repairs that are the most strongly grounded in the given database and the given set of AICs, that is, those resulting strictly from combinations of the preferences expressed by the database designer for each of the integrity constraints and from the principle of minimal change. The operational semantics for AICs [7] allows direct computation of justified repairs by means of intuitive tree algorithms. Interaction between different AICs in a set that must be collectively satisfied makes the problem of repairing a database highly non-trivial, however, and in the worst case deciding whether a database can be repaired is NP-complete or Σ_P^2-complete on the number of AICs [5], depending on the criteria used to choose possible repairs. For this reason, it is important to be able to control the number of AICs being considered simultaneously.

In this paper we first present parallelization results that allow a set of AICs to be split in smaller, independent sets such that repairs for each smaller set can be computed independently and the results straightforwardly combined into a repair for the original set. Afterwards, we introduce a hierarchization mechanism on AICs that allows repairs to be computed progressively, starting with a small set of AICs and extending this set while simultaneously extending the computed repair. With these techniques, it is possible to speed up the problem of finding repairs significantly; and, although they do not help in the worst-case scenario, the typical structure of real-life databases indicates that parallelization and hierarchization should be widely applicable.

1.1 Related Work

When a database needs to be changed, it is necessary to find a way to make the relevant modifications while maintaining the consistency of the data. This problem, which has been the focus of intensive research for over thirty years, was extensively discussed in [1], where three main change operations were identified: insertion of new facts, deletion of existing facts, and modification of information, and the concept of "good" update was characterized.

There are two distinct scenarios where database change is required, leading to the distinction between *update* and *revision* [8,11]. An update occurs whenever the world changes and the knowledge bases needs to be changed to reflect this fact; a revision happens when new knowledge is obtained about a world that did not change. This distinction is especially relevant in deductive databases and open-world knowledge bases, where the known information is not assumed to be complete.

In spite of their differences, there are obvious similarities between updates and revisions, and in both cases one has to consider the problems that arise when the intended semantics of the database is taken into account. Typically, the changes that have to be made conflict with the integrity constraints associated with the database, and the database must be repaired in order to regain consistency. The ways in which this can be done are many, and several proposals have been around for years. One possibility is to read integrity constraints as rules that suggest possible actions to repair inconsistencies [1]; another is to express database dependencies through logic programming, namely in the setting of deductive databases [12,14,15]. A more algorithmic approach uses event-condition-action rules [16,17], where actions are triggered by specific events, and for which rule processing algorithms have been proposed and a procedural semantics has been defined.

Several algorithms for computing repairs of inconsistent databases have been proposed and studied throughout the years, focusing on the different ways integrity constraints are specified and on the different types of databases under consideration [10,12,14,15]. This multitude of approaches is not an accident: deciding whether an inconsistent database can be repaired is typically a Π_p^2- or co-Σ_p^2- complete problem, and it has been observed [8] that there is no reason to believe in the existence of general-purpose algorithms for this problem, but one should rather focus on developing more specific algorithms for particular interesting cases.

Regardless of the approach taken, when an inconsistent database can be repaired there are typically several sets of update actions that achieve this goal, leading to different revised consistent databases. Restricting the set of database repairs to those considered most adequate is therefore an important task. Among the criteria that have been proposed to obtain this restriction are minimality of change [6,8,18] – one should change as little as possible – and the common sense law of inertia [15] – one should only change something if there is a reason for it –, but these are not enough to narrow down the set of possible repairs sufficiently. Ultimately, it is usually assumed that some human interaction will be required to choose the "best" possible repair [16].

Because of the intrinsic complexity involved in the computation of repairs, techniques to split a problem in several smaller problems are of particular interest. As far as we know, this problem has received little consideration over the years. There is a reference to semantic independency in [14] that is not explored further, and syntactic precedence is used in that same paper in order to compute models – but within a scenario that is far more powerful than that of active integrity constraints. More recently, syntactic precedence between constraints was also discussed with the explicit goal of making the search for repairs more efficient [13], but the authors did not allow for cyclic dependencies. The results we prove are therefore a significant extension of previous work, and we believe they can be easily extended to different formalisms of integrity constraints.

2 Background

Active integrity constraints were originally introduced in [9] as a special type of integrity constraints, specifying not only the consistency requirements imposed upon a database, but also actions that can be taken to correct the database when such requirements are not met.

Within this framework, a database is a subset of a finite set of propositional atoms $\mathcal{A}t$. An *active integrity constraint* (AIC) is a rule of the form

$$L_1, \ldots, L_n \supset \alpha_1 \mid \ldots \mid \alpha_m$$

where $L_1, \ldots, L_n$ are literals in the language generated by $\mathcal{A}t$; $\alpha_1, \ldots, \alpha_m$ are *update actions* of the form $+a$ or $-a$, where a is an atom in the same language; and every update action must contradict some literal, i.e. if $+a$ (resp. $-a$) occurs among the α_i, then $\mathsf{not}\ a$ (resp. a) must occur among the L_i. The set $\{L_1, \ldots, L_n\}$ is the *body* of the rule and $\{\alpha_1, \ldots, \alpha_m\}$ is its *head*.

The close connection between literals and actions is made precise by means of two operators. The atom *underlying* an action α is $\mathsf{lit}(\alpha)$, defined by $\mathsf{lit}(+a) = a$ and $\mathsf{lit}(-a) = \mathsf{not}\ a$, whereas the update action *corresponding* to L is $\mathsf{ua}(L)$, defined by $\mathsf{ua}(a) = +a$ and $\mathsf{ua}(\mathsf{not}\ a) = -a$. The *dual* of a literal L, L^D, is defined as usual by $a^D = \mathsf{not}\ a$ and $(\mathsf{not}\ a)^D = a$. Using this notation, the requirement that valid AICs must satisfy can be stated as $\{\mathsf{lit}(\alpha_1), \ldots, \mathsf{lit}(\alpha_m)\}^D \subseteq \{L_1, \ldots, L_n\}$.

Being a set of propositional atoms, any database $\mathcal{I}$ induces a propositional interpretation of literals. We say that $\mathcal{I}$ *entails* literal L, $I \models L$, if L is a and $a \in \mathcal{I}$, or if L is $\mathsf{not}\ a$ and $a \notin \mathcal{I}$. Given an AIC r of the form $L_1, \ldots, L_n \supset \alpha_1, \ldots, \alpha_m$, we say that $\mathcal{I} \models r$ if $\mathcal{I} \not\models L_i$ for some i; otherwise, r is said to be *applicable* in $\mathcal{I}$. Finally, if η is a set of AICs, then $\mathcal{I} \models \eta$ iff $\mathcal{I} \models r$ for every $r \in \eta$.

The operational nature of rules is given by the notion of updating a database by a set of update actions, which captures the intuive idea conveyed above. The result of updating $\mathcal{I}$ with a set of update actions $\mathcal{U}$ is $\mathcal{I} \circ \mathcal{U}$, defined as

$$\mathcal{I} \circ \mathcal{U} = (\mathcal{I} \cup \{a \mid +a \in \mathcal{U}\}) \setminus \{a \mid -a \in \mathcal{U}\}\,.$$

In order for this definition to make sense, $\mathcal{U}$ must not contain $+a$ and $-a$ for the same atom a. A set of update actions satisfying this requirement is said to be *consistent*.

Given a set of AICs η and a database $\mathcal{I}$, a set of update actions $\mathcal{U}$ such that $\mathcal{I} \circ \mathcal{U} \models \eta$ achieves the task of making $\mathcal{I}$ consistent w.r.t. η. In general, for any given database that is inconsistent w.r.t. η there will be either none or several such $\mathcal{U}$. In order to compare different ways of repairing $\mathcal{I}$, Caroprese and Truszczyński [5] studied different semantics for AICs.

Minimality of change is commonly accepted as a desirable property [6,18]. This motivates the following notion: given a database $\mathcal{I}$ and a set of AICs η, a consistent set of update actions $\mathcal{U}$ such that (i) every action in $\mathcal{U}$ changes $\mathcal{I}$ and (ii) $\mathcal{I} \circ \mathcal{U} \models \eta$ is called a *weak repair* for $\langle \mathcal{I}, \eta \rangle$; a *repair* for $\langle \mathcal{I}, \eta \rangle$ is a weak repair for $\langle \mathcal{I}, \eta \rangle$ that is minimal w.r.t. inclusion (so it contains no proper subset that is also a weak repair). Condition (i) states that weak repairs only include actions that change the database, and may be formally stated as $(\{+a \mid a \in \mathcal{I}\} \cup \{-a \mid a \in \mathcal{A}t \setminus \mathcal{I}\}) \cap \mathcal{U} = \emptyset$, or equivalently as $\mathcal{I} \circ \alpha \neq \mathcal{I}$ for every $\alpha \in \mathcal{U}$. Condition (ii) simply states that weak repairs make the database consistent w.r.t. η.

None of these conditions takes into account the operational nature of AICs, however, since they ignore the actions in the heads of the rules in η. For this purpose, one needs to consider the more sophisticated notion of founded (weak) repairs [5]. The intuition behind these is that they should contain only actions that are motivated (founded) by the application of some rule. An update action α is *founded*[1] w.r.t. $\langle \mathcal{I}, \eta \rangle$ and a set of update actions $\mathcal{U}$ if there is a rule $r \in \eta$ such that $\alpha \in \mathsf{head}(r)$ and $\mathcal{I} \circ \mathcal{U} \models L$ for every literal $L \in \mathsf{body}(r) \setminus \{\mathsf{lit}(\alpha)^D\}$. Quoting [5], "if $\mathcal{U}$ is to enforce r, then it must contain α" – if α is removed from $\mathcal{U}$, then all literals in the body of r are true and the rule is violated. A set of update actions $\mathcal{U}$ is *founded* w.r.t $\langle \mathcal{I}, \eta \rangle$ if every action in $\mathcal{U}$ is founded w.r.t. $\langle \mathcal{I}, \eta \rangle$ and $\mathcal{U}$. A set of update actions $\mathcal{U}$ is a founded (weak) repair for $\langle \mathcal{I}, \eta \rangle$ if (i) $\mathcal{U}$ is a weak repair for $\langle \mathcal{I}, \eta \rangle$ and (ii) $\mathcal{U}$ is founded w.r.t. $\langle \mathcal{I}, \eta \rangle$.

It is important to stress that founded repairs are minimal weak repairs that are founded. Indeed, there are founded weak repairs that do not contain any founded repair as subset (see [5] for an example). Also, being founded does not imply being a weak repair, so these two tests must be performed independently.

Founded repairs, however, sometimes exhibit unexpected properties, such as *circularity of support* [5] – e.g. they contain two actions α and β such that α is founded by means of a rule r whose body only holds because of β, and β is founded by means of a rule r' whose body only holds because of α –, and it is therefore interesting to consider a more complex type of repairs: justified repairs. In order to define these, we need some auxiliary notions. The set of *non-updatable literals* of a rule r is defined as $\mathsf{nup}(r) = \mathsf{body}(r) \setminus (\mathsf{lit}(\mathsf{head}(r)))^D$, were lit is extended to sets in the obvious way. A set of update actions $\mathcal{U}$ is closed for rule r if $\mathsf{nup}(r) \subseteq \mathsf{lit}(\mathcal{U})$ implies $\mathsf{head}(r) \cap \mathcal{U} \neq \emptyset$, and $\mathcal{U}$ is closed for η if $\mathcal{U}$ is closed for every rule in η.

An update action $+a$ (resp. $(-a)$) is a *no-effect action* w.r.t. $\mathcal{I}$ and $\mathcal{J}$ if $a \in \mathcal{I} \cap \mathcal{J}$ (resp. $a \notin (\mathcal{I} \cup \mathcal{J})$) – in other words, both $\mathcal{I}$ and $\mathcal{J}$ are unaffected by the action. The set of all no-effect actions w.r.t. $\mathcal{I}$ and $\mathcal{J}$ is denoted by $\mathsf{ne}(\mathcal{I}, \mathcal{J})$. Given a database $\mathcal{I}$ and a set of AICs η, a consistent set of update actions $\mathcal{U}$ is a *justified action set* for $\langle \mathcal{I}, \eta \rangle$ if $\mathcal{U}$ is a minimal set of update actions containing $\mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U})$ and closed for η. In that case, the set $\mathcal{U} \setminus \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U})$ is a *justified weak repair* for $\langle \mathcal{I}, \eta \rangle$. Being closed for η implies being a weak repair for $\langle \mathcal{I}, \eta \rangle$, so this terminology is consistent with the previous usage of the latter term.

In spite of the minimality requirement in the definition of justified weak repair, there *are* justified weak repairs that contain a justified repair as a proper subset; this is because the minimality involved in this definition is within a different universe. All justified weak repairs are founded, but not conversely: indeed, these repairs successfully avoid circularity of support.

We will use the following alternative characterization of justified weak repair: a weak repair $\mathcal{U}$ for $\langle \mathcal{I}, \eta \rangle$ is justified if (i) $\mathcal{U} \cap \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U}) = \emptyset$ and (ii) $\mathcal{U} \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U})$ is a justified action set. Indeed, taking $\mathcal{W} = \mathcal{U} \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U})$, it can easily be checked that $\mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{W}) = \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U})$: the only differences between

[1] This equivalent characterization of founded action, which can be found in [7], is slightly different from that in [5], and simpler to use in practice.

$\mathcal{I}$ and $\mathcal{I} \circ \mathcal{W}$ must originate from $\mathcal{U}$ by definition of $\mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U})$. Therefore $\mathcal{W} \setminus \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{W}) = \mathcal{U}$.

The major problem with computing repairs for inconsistent databases lies in the complexity of deciding whether such repairs exist. Given $\mathcal{I}$ and η, the problem of deciding whether there exists a weak repair, a repair or a founded weak repair for $\langle \mathcal{I}, \eta \rangle$ is NP-complete (on the size of η), whereas deciding whether there is a founded repair, a justified weak repair or a justified repair for η is Σ_P^2-complete (again on the size of η).[2] In the special case where all AICs are *normalized* – they have only one action in their head – the last two problems also become NP-complete. Due to these ultimately bad complexity bounds, techniques to lower the size of the problem can be extremely useful in practice. The goal of this paper is to discuss how a set of AICs η can be divided into smaller sets such that the computation of (simple, founded, justified) repairs can be computed for each of those sets and the results combined in polynomial time.

3 Independent AICs

In this section, we introduce a notion of independence between active integrity constraints. The goal is the following: given a set of AICs η, to partition it in distinct independent sets $\eta_1, \ldots, \eta_n$ such that the search for repairs for a database $\mathcal{I}$ and η can be parallelized among the η_i. We define independent sets of AICs in such a way that (simple, founded, justified) repairs for the different sets can be combined into a (simple, founded, justified) repair for $\langle \mathcal{I}, \eta \rangle$.

The basic concept is that of independent AICs. Two AICs are independent if they do not share any atoms between their literals, so that applicability of one does not affect applicability of the other.

Definition 1.

1. *The atom* underlying *a literal L is $|L|$, defined as $|a| = |\mathsf{not}\ a| = a$.*
2. *Let r_1 and r_2 be two AICs, where r_1 is $L_1, \ldots, L_n \supset \alpha_1, \ldots, \alpha_p$ and r_2 is $M_1, \ldots, M_m \supset \beta_1, \ldots, \beta_q$. Then r_1 and r_2 are* independent*, $r_1 \perp\!\!\!\perp r_2$, if $\{|L_1|, \ldots, |L_n|\} \cap \{|M_1|, \ldots, |M_m|\} = \emptyset$.*
3. *Let η_1 and η_2 be sets of AICs. Then η_1 and η_2 are* independent*, $\eta_1 \perp\!\!\!\perp \eta_2$, if $r \perp\!\!\!\perp s$ whenever $r \in \eta_1$ and $s \in \eta_2$.*

Two comments are in place regarding this definition. First, the notion of independence does not take into account the actions in the rules (the "active" part of the AICs); this aspect will be dealt with in Section 5. Second, this concept only depends on the active integrity constraints themselves, and not on the underlying database. This issue has positive practical implications, as we will see later.

This notion of independence captures the spirit of parallelization, as the next lemmas state. Throughout the remainder of this section, let $\mathcal{I}$ be a database, η_1, η_2 be independent sets of AICs and $\eta = \eta_1 \cup \eta_2$.

[2] The size of $\mathcal{I}$ does not affect the complexity bounds for these problems [5].

Lemma 1. *Let* $\mathcal{U}_1$ *and* $\mathcal{U}_2$ *be sets of update actions such that every action in* $\mathcal{U}_i$ *corresponds to a literal (or its dual) in a rule in* η_i*,*[3] *and take* $\mathcal{U} = \mathcal{U}_1 \cup \mathcal{U}_2$*. For every literal* L *such that* $L \in \mathsf{body}(r)$ *with* $r \in \eta_i$*,* $\mathcal{I} \circ \mathcal{U} \models L$ *iff* $\mathcal{I} \circ \mathcal{U}_i \models L$*. In particular, for every* $r \in \eta_i$*,* $I \circ \mathcal{U} \models r$ *iff* $\mathcal{I} \circ \mathcal{U}_i \models r$*.*

Proof. Let $L \in \mathsf{body}(r)$ for some $r \in \eta_1$. If $\alpha \in \mathcal{U}_2$, then $|\mathsf{lit}(\alpha)| \neq |L|$ because $\eta_1 \perp\!\!\!\perp \eta_2$, whence $\mathcal{I} \circ \mathcal{U} \models L$ iff $\mathcal{I} \circ \mathcal{U}_1 \models L$ (note that $\mathcal{I} \circ \mathcal{U} = (\mathcal{I} \circ \mathcal{U}_1) \circ \mathcal{U}_2$). The result for rules is a straightforward consequence. The argument for $\mathcal{U}_2$ is similar. □

Lemma 2. *Let* $\mathcal{U}_1$ *and* $\mathcal{U}_2$ *be weak repairs for* $\langle \mathcal{I}, \eta_1 \rangle$ *and* $\langle \mathcal{I}, \eta_2 \rangle$*, respectively, such that the actions in* $\mathcal{U}_i$ *are all duals of literals in the body of some rule in* η_i*. Then* $\mathcal{U} = \mathcal{U}_1 \cup \mathcal{U}_2$ *is a weak repair for* $\langle \mathcal{I}, \eta \rangle$*.*

Proof. We first show that $\mathcal{U}$ is a consistent set of actions containing only essential actions. For consistency, note that the set of atoms underlying the actions in $\mathcal{U}_1$ is disjoint from that of the atoms underlying the atoms in $\mathcal{U}_2$, from the hypothesis and the fact that $\eta_1 \perp\!\!\!\perp \eta_2$; hence, if $+\alpha$ and $-\alpha$ were both in $\mathcal{U} = \mathcal{U}_1 \cup \mathcal{U}_2$ for some a, this would mean that $+\alpha, -\alpha \in \mathcal{U}_i$ for some i, whence $\mathcal{U}_i$ would be inconsistent. Furthermore, if $\alpha \in \mathcal{U}_i$ then α must change the state of $\mathcal{I}$ (since $\mathcal{U}_i$ is a weak repair for $\langle \mathcal{I}, \eta_i \rangle$), so $\mathcal{U}$ consists only of essential update actions.

Finally, we show that $\mathcal{U}$ is a weak repair. Without loss of generality, let $r \in \eta_1$. Then $\mathcal{I} \circ \mathcal{U}_1 \models r$, since $\mathcal{U}_1$ is a weak repair for $\langle \mathcal{I}, \eta_1 \rangle$, and by Lemma 1 $\mathcal{I} \circ \mathcal{U} \models r$. □

The hypothesis that the actions in each $\mathcal{U}_i$ are all duals of literals in the body of some rule in η_i is essential: if it were not required, then $\mathcal{U}_1$ could "break" satisfaction of some rule in η_2 or reciprocally, or there might be inconsistencies from joining $\mathcal{U}_1$ and $\mathcal{U}_2$. Although this hypothesis could be weakened, it is actually a (very) reasonable assumption: no reasonable algorithm for computing weak repairs should include actions that do not affect the semantics of the integrity constraints that should hold, since this verification can be done very efficiently.

If $\mathcal{U}_1$ and $\mathcal{U}_2$ are repairs, we get the following stronger result.

Lemma 3. *If* $\mathcal{U}_1$ *and* $\mathcal{U}_2$ *are repairs for* $\langle \mathcal{I}, \eta_1 \rangle$ *and* $\langle \mathcal{I}, \eta_2 \rangle$*, respectively, then* $\mathcal{U} = \mathcal{U}_1 \cup \mathcal{U}_2$ *is a repair for* $\langle \mathcal{I}, \eta \rangle$*.*

Proof. By Lemma 2, $\mathcal{U}$ is a weak repair for $\langle \mathcal{I}, \eta \rangle$. For any $\mathcal{U}' \subsetneq \mathcal{U}$, define $\mathcal{U}'_i = \mathcal{U}' \cap \mathcal{U}_i$ for $i = 1, 2$. Note that one of the inclusions $\mathcal{U}'_i \subseteq \mathcal{U}_i$ must be strict; without loss of generality, assume that $\mathcal{U}'_1 \subsetneq \mathcal{U}_1$. Since $\mathcal{U}_1$ is a repair, this means that $\mathcal{U}'_1$ cannot be a weak repair, hence there is a rule $r \in \eta_1$ such that $\mathcal{U}'_1 \not\models r$. By Lemma 1 $\mathcal{U}'_1 \cup \mathcal{U}'_2 \not\models r$, hence $\mathcal{U}' = \mathcal{U}'_1 \cup \mathcal{U}'_2$ cannot be a weak repair for $\langle \mathcal{I}, \eta \rangle$. Therefore $\mathcal{U}$ is a repair for $\langle \mathcal{I}, \eta \rangle$. □

The converse result also holds: if we split the actions in a weak repair $\mathcal{U}$ according to whether they affect rules in η_1 or η_2, we get weak repairs for those sets of AICs.

[3] Formally, $\{|\mathsf{lit}(\alpha)| \mid \alpha \in \mathcal{U}_i\} \subseteq \{|L| \mid \exists r \in \eta_i . L \in \mathsf{body}(r)\}$.

Lemma 4. *Let* $\mathcal{U}$ *be a weak repair for* $\langle \mathcal{I}, \eta \rangle$. *Then*

$$\mathcal{U}_i = \{\alpha \in \mathcal{U} \mid \exists r \in \eta_i . \mathsf{lit}(\alpha)^D \in \mathsf{body}\,(r)\}$$

are weak repairs for $\langle \mathcal{I}, \eta_i \rangle$. *Furthermore,* $\mathcal{U} = \mathcal{U}_1 \cup \mathcal{U}_2$ *if the actions in* $\mathcal{U}$ *are all duals of literals in the body of some rule in* η.

Proof. Assume that $\mathcal{U}$ is a weak repair for $\langle \mathcal{I}, \eta \rangle$ and let $\mathcal{U}_i$ be as stated. Since $\mathcal{U}$ is a weak repair for $\langle \mathcal{I}, \eta \rangle$, $\mathcal{I} \circ \mathcal{U} \models r$ for every rule $r \in \eta_i$. By Lemma 1, $\mathcal{I} \circ \mathcal{U}_i \models r$. Therefore $\mathcal{U}_i$ is a weak repair for $\langle \mathcal{I}, \eta_i \rangle$.

If the actions in $\mathcal{U}$ are all duals of literals in the body of some rule in η, then they occur in the body of a rule in η_1 or η_2, so they will all occur either in $\mathcal{U}_1$ or $\mathcal{U}_2$, whence $\mathcal{U} = \mathcal{U}_1 \cup \mathcal{U}_2$. □

The stated equality can be made to hold in the general case by adding the actions that do not affect any rule to either $\mathcal{U}_1$ or $\mathcal{U}_2$; however, this is not an interesting situation, and we will not consider it any further.

If $\mathcal{U}$ is minimal, then the same result can be made stronger.

Lemma 5. *If* $\mathcal{U}$ *is a repair for* $\langle \mathcal{I}, \eta \rangle$, *then* $\mathcal{U}_1$ *and* $\mathcal{U}_2$ *as defined above are repairs for* $\langle \mathcal{I}, \eta_1 \rangle$ *and* $\langle \mathcal{I}, \eta_2 \rangle$, *respectively, and furthermore* $\mathcal{U} = \mathcal{U}_1 \cup \mathcal{U}_2$.

Proof. By Lemma 4, each $\mathcal{U}_i$ is a weak repair for $\langle \mathcal{I}, \eta_i \rangle$. Suppose that $\mathcal{U}_1' \subsetneq \mathcal{U}_1$ is also a weak repair for $\langle \mathcal{I}, \eta_1 \rangle$. By Lemma 1, $\mathcal{U}' = \mathcal{U}_1' \cup \mathcal{U}_2$ is a weak repair for $\langle \mathcal{I}, \eta \rangle$ with $\mathcal{U}' \subsetneq \mathcal{U}$, which is absurd. Therefore $\mathcal{U}_1'$ is not a weak repair, hence $\mathcal{U}_1$ is a repair. The case for $\mathcal{U}_2$ is similar. Finally, $\mathcal{U}$ cannot contain actions that are not duals of literals in the body of rules in η, since these can always be removed without affecting the property of being a weak repair; therefore $\mathcal{U}_1 \cup \mathcal{U}_2 = \mathcal{U}$. □

These results also hold if we consider founded or justified (weak) repairs.

Lemma 6. *Let* $\mathcal{U}_1$ *and* $\mathcal{U}_2$ *be founded w.r.t.* $\langle \mathcal{I}, \eta_1 \rangle$ *and* $\langle \mathcal{I}, \eta_2 \rangle$, *respectively. Then* $\mathcal{U} = \mathcal{U}_1 \cup \mathcal{U}_2$ *is founded w.r.t.* $\langle \mathcal{I}, \eta \rangle$.

Proof. In order for $\mathcal{U}$ to be founded w.r.t. $\langle \mathcal{I}, \eta \rangle$, every action in $\mathcal{U}$ must be founded w.r.t. $\langle \mathcal{I}, \eta \rangle$ and $\mathcal{U}$. Let $\alpha \in \mathcal{U}$ and assume that $\alpha \in \mathcal{U}_1$ (the case when $\alpha \in \mathcal{U}_2$ is similar).

Since $\mathcal{U}_1$ is founded w.r.t. $\langle \mathcal{I}, \eta_1 \rangle$, there is a rule $r \in \eta_1$ such that $\alpha \in \mathsf{head}\,(r)$ and $\mathcal{I} \circ \mathcal{U}_1 \models L$ for every $L \in \mathsf{body}\,(r) \setminus \{\mathsf{lit}(\alpha)^D\}$. By Lemma 1, $I \circ \mathcal{U} \models L$ for every such L. Since $\eta_1 \subseteq \eta$, this means that α is founded w.r.t. $\langle \mathcal{I}, \eta \rangle$ and $\mathcal{U}$. □

Corollary 1. *If* $\mathcal{U}_1$ *and* $\mathcal{U}_2$ *are founded (weak) repairs, then* $\mathcal{U}$ *is also a founded (weak) repair.*

Proof. Consequence of Lemmas 2, 3 and 6. □

Lemma 7. *Let* $\mathcal{U}$ *be founded w.r.t. for* $\langle \mathcal{I}, \eta \rangle$. *Then* $\mathcal{U}_1$ *and* $\mathcal{U}_2$ *as defined in Lemma 4 are such that* $\mathcal{U} = \mathcal{U}_1 \cup \mathcal{U}_2$ *and each* $\mathcal{U}_i$ *is founded w.r.t.* $\langle \mathcal{I}, \eta_i \rangle$.

Proof. Let $\alpha \in \mathcal{U}_1$; since $\mathcal{U}$ is founded w.r.t. $\langle \mathcal{I}, \eta \rangle$, there is a rule $r \in \eta$ such that $\alpha \in \mathsf{head}(r)$ and $\mathcal{I} \circ \mathcal{U} \models L$ for every $L \in \mathsf{body}(r) \setminus \{\mathsf{lit}(\alpha)^D\}$. But if $\alpha \in \mathsf{head}(r)$, then necessarily $r \in \eta_1$; and in that case $\mathcal{I} \circ \mathcal{U}_1 \models L$ for every $L \in \mathsf{body}(r) \setminus \{\mathsf{lit}(\alpha)^D\}$ by Lemma 1. Therefore α is founded w.r.t. $\langle \mathcal{I}, \eta_1 \rangle$ and $\mathcal{U}_1$, whence $\mathcal{U}_1$ is founded w.r.t. $\langle \mathcal{I}, \eta_1 \rangle$. The case when $\alpha \in \mathcal{U}_2$ is similar.

By definition of founded set, all actions in $\mathcal{U}$ must necessarily be in either $\mathcal{U}_1$ or $\mathcal{U}_2$, so $\mathcal{U} = \mathcal{U}_1 \cup \mathcal{U}_2$. □

Corollary 2. *If $\mathcal{U}$ is a (weak) founded repair, then $\mathcal{U}_1$ and $\mathcal{U}_2$ are also (weak) founded repairs.*

Proof. Consequence of Lemmas 4, 5 and 7. □

Lemma 8. *Let $\mathcal{U}_1$ and $\mathcal{U}_2$ be justified (weak) repairs for $\langle \mathcal{I}, \eta_1 \rangle$ and $\langle \mathcal{I}, \eta_2 \rangle$, respectively. Then $\mathcal{U} = \mathcal{U}_1 \cup \mathcal{U}_2$ is a justified (weak) repair for $\langle \mathcal{I}, \eta \rangle$.*

Proof. We begin by making some observations that will be used recurrently throughout the proof.

(a) For $i = 1, 2$, $\mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U}) \subseteq \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U}_i)$, since $\mathcal{U}_i \subseteq \mathcal{U}$.
(b) For $i = 1, 2$, $\mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U}_i) \subseteq (\mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U}) \cup \mathcal{U}_{3-i})$: $\mathcal{U}$ can only change literals that changed either by $\mathcal{U}_1$ or by $\mathcal{U}_2$. In particular, since $\eta_1 \perp\!\!\!\perp \eta_2$, if $\mathsf{nup}(r) \subseteq \mathsf{lit}(\mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U}_i))$ for some $r \in \eta_i$, then $L \in \mathsf{lit}(\mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U}))$; and if $\alpha \in \mathsf{head}(r)$ for some $r \in \eta_i$ and $\alpha \in \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U}_i)$, then $\alpha \in \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U})$.
(c) For $i = 1, 2$, if $L \in \mathsf{body}(r)$ with $r \in \eta_i$ and $L \in \mathsf{lit}(\mathcal{U})$, then $L \in \mathsf{lit}(\mathcal{U}_i)$: since every justified weak repair is founded [5], $\mathcal{U}_i$ only contains actions in the heads of rules of η_i, and the thesis follows from $\eta_1 \perp\!\!\!\perp \eta_2$.

We first show that $\mathcal{U} \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U})$ is closed for η. Let $r \in \eta_1$; the case when $r \in \eta_2$ is similar. Suppose $\mathsf{nup}(r) \subseteq \mathsf{lit}(\mathcal{U} \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U}))$, and let $L \in \mathsf{nup}(r)$. If $L \in \mathsf{lit}(\mathcal{U})$, then $L \in \mathsf{lit}(\mathcal{U}_1)$ by (c), hence $\mathsf{nup}(r) \subseteq \mathsf{lit}(\mathcal{U}_1 \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U}))$, whence $\mathsf{nup}(r) \subseteq \mathsf{lit}(\mathcal{U}_1 \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U}_1))$ by (a). But $\mathcal{U}_1 \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U}_1)$ is closed for η_1, so $\mathsf{head}(r) \cap (\mathcal{U}_1 \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U}_1)) \neq \emptyset$. By $\mathcal{U}_1 \subseteq \mathcal{U}$ and (b), also $\mathsf{head}(r) \cap (\mathcal{U} \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U})) \neq \emptyset$.

To check minimality, suppose that $\mathcal{U}' \subsetneq \mathcal{U}$ is such that $\mathcal{U}' \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U})$ is closed for η and take $\mathcal{U}'_i = \mathcal{U}' \cap \mathcal{U}_i$ for $i = 1, 2$. Note that one of the inclusions $\mathcal{U}'_i \subseteq \mathcal{U}_i$ must be strict. Without loss of generality, assume this is the case when $i = 1$, and take $r \in \eta_1$. If $\mathsf{nup}(r) \subseteq \mathsf{lit}(\mathcal{U}'_1 \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U}_1))$, then $\mathsf{nup}(r) \subseteq \mathsf{lit}(\mathcal{U}' \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U}))$, consequence of $\mathcal{U}'_1 \subseteq \mathcal{U}'$ and (b). Since $\mathcal{U}' \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U})$ is closed for η and $\eta_1 \subseteq \eta$, it follows that $\mathsf{head}(r) \cap (\mathcal{U}' \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U})) \neq \emptyset$. By definition of $\mathcal{U}_1$ and (a), it follows that $\mathsf{head}(r) \cap (\mathcal{U}'_1 \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U}_1)) \neq \emptyset$. Then $\mathcal{U}'_1 \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U}_1)$ is closed for η_1, contradicting minimality of $\mathcal{U}_1$.

Hence $\mathcal{U}$ is a justified weak repair for $\langle \mathcal{I}, \eta \rangle$. By Lemma 3, if $\mathcal{U}_1$ and $\mathcal{U}_2$ are both justified repairs for $\langle \mathcal{I}, \eta_1 \rangle$ and $\langle \mathcal{I}, \eta_2 \rangle$, respectively, then $\mathcal{U}$ is also a justified repair for $\langle \mathcal{I}, \eta \rangle$. □

Lemma 9. *Let $\mathcal{U}$ be a justified (weak) repair for $\langle \mathcal{I}, \eta \rangle$. Then $\mathcal{U}_1$ and $\mathcal{U}_2$ as defined in Lemma 4 are such that $\mathcal{U} = \mathcal{U}_1 \cup \mathcal{U}_2$ and each $\mathcal{U}_i$ is a justified (weak) repair for $\langle \mathcal{I}, \eta_i \rangle$.*

Proof. Again note that properties (a), (b) and (c) from the previous proof hold. We begin by showing that $\mathcal{U}_1 \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U}_1)$ is closed under η_1. Take $r \in \eta_1$ and suppose that $\mathsf{nup}(r) \subseteq \mathsf{lit}(\mathcal{U}_1 \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U}_1))$. Then $\mathsf{nup}(r) \subseteq \mathsf{lit}(\mathcal{U} \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U}))$ by $\mathcal{U}_1 \subseteq \mathcal{U}$ and (b). Since $\mathcal{U} \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U})$ is closed for η, it follows that $\mathsf{head}(r) \cap (\mathcal{U} \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U})) \neq \emptyset$. By construction of $\mathcal{U}_1$ and (a), we conclude that $\mathsf{head}(r) \cap (\mathcal{U}_1 \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U}_1)) \neq \emptyset$. The case for $\mathcal{U}_2$ is similar.

To check minimality, suppose that $\mathcal{U}_1' \subsetneq \mathcal{U}_1$ is such that $\mathcal{U}_1' \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U}_1)$ is closed for η_1 and take $\mathcal{U}' = \mathcal{U}_1' \cup \mathcal{U}_2$. Let $r \in \eta$ and assume $\mathsf{nup}(r) \subseteq \mathsf{lit}(\mathcal{U}' \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U}))$; there are two possible cases.

- Suppose $r \in \eta_1$ and let $L \in \mathsf{nup}(r)$. Note that $L \in \mathsf{lit}(\mathcal{U}_2)$ is impossible, since $\eta_1 \perp\!\!\!\perp \eta_2$. Therefore $\mathsf{nup}(r) \subseteq \mathsf{lit}(\mathcal{U}_1' \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U}))$, whence by (a) $\mathsf{nup}(r) \subseteq \mathsf{lit}(\mathcal{U}_1' \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U}_1))$, and therefore $\mathsf{head}(r) \cap (\mathcal{U}_1' \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U}_1)) \neq \emptyset$. From $\mathcal{U}_1' \subseteq \mathcal{U}'$ and (b), also $\mathsf{head}(r) \cap (\mathcal{U}' \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U})) \neq \emptyset$.
- Suppose $r \in \eta_2$ and let $L \in \mathsf{nup}(r)$. Since $L \in \mathsf{lit}(\mathcal{U}_1')$ is impossible, it follows that $L \in \mathcal{U}_2 \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U})$, and since $\mathcal{U}_2 \subseteq \mathcal{U}$ we conclude that $\mathsf{nup}(r) \subseteq \mathsf{lit}(\mathcal{U} \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U}))$. Since $\mathcal{U} \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U})$ is closed for η (which contains η_2), it follows that $\mathsf{head}(r) \cap (\mathcal{U} \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U})) \neq \emptyset$, and since $\mathsf{head}(r)$ does not contain actions in $\mathcal{U}_1$ necessarily $\mathsf{head}(r) \cap (\mathcal{U}_2 \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U})) \neq \emptyset$, whence $\mathsf{head}(r) \cap (\mathcal{U}' \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U})) \neq \emptyset$.

In either case, from $\mathsf{nup}(r) \subseteq (\mathcal{U}' \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U}))$ one concludes that $\mathsf{head}(r) \cap (\mathcal{U}' \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U})) \neq \emptyset$, whence $\mathcal{U}' \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U})$ is closed for η, contradicting minimality of $\mathcal{U}$. This is absurd, so $\mathcal{U}_1$ is a justified weak repair. Again the case for $\mathcal{U}_2$ is similar.

Since justified weak repairs are founded, Lemma 7 guarantees that $\mathcal{U} = \mathcal{U}_1 \cup \mathcal{U}_2$. Furthermore, if $\mathcal{U}$ is a justified repair for $\langle \mathcal{I}, \eta \rangle$, then each $\mathcal{U}_i$ is a justified repair for $\langle \mathcal{I}, \eta_i \rangle$ by Lemma 5. □

The practical significance of the results in this section is a parallelization algorithm: if $\eta = \eta_1 \cup \eta_2$ with $\eta_1 \perp\!\!\!\perp \eta_2$, then all (simple, founded, justified) repairs for $\langle \mathcal{I}, \eta \rangle$ can be expressed as unions of (simple, founded, justified) repairs for $\langle \mathcal{I}, \eta_1 \rangle$ and $\langle \mathcal{I}, \eta_2 \rangle$ by Lemmas 5, 7 and 9, so one can search for these repairs instead and combine them in at the end; Lemmas 3, 6 and 8 guarantee that no spurious results are obtained. The next section expands on these ideas, and discusses how η can be adequately split.

4 Finding Independent Sets of AICs

The results in the previous section show that splitting a set of AICs η into two independent sets η_1 and η_2 allows one to parallelize the search for repairs of a database $\mathcal{I}$, by searching independently for repairs for $\langle \mathcal{I}, \eta_1 \rangle$ and $\langle \mathcal{I}, \eta_2 \rangle$. In this section we address a complementary issue: how can one find these sets? We begin by formulating the results in the previous section in a more general way.

Definition 2. *A* partition *of a set of AICs η is a set $\boldsymbol{\eta} = \{\eta_1, \ldots, \eta_n\}$ such that $\eta = \cup_{i=1}^n \eta_i$ and $\eta_i \perp\!\!\!\perp \eta_j$ for $i \neq j$.*

Theorem 1. *Let $\boldsymbol{\eta}$ be a partition of η.*

1. *If $\mathcal{U}$ is a simple/founded/justified (weak) repair for $\langle\mathcal{I},\eta\rangle$, then there exist sets $\mathcal{U}_1,\ldots,\mathcal{U}_n$ with $\mathcal{U}=\cup_{i=1}^{n}\mathcal{U}_i$ such that $\mathcal{U}_i$ is a simple/founded/justified (weak) repair for $\langle\mathcal{I},\eta_i\rangle$.*
2. *If $\mathcal{U}_i$ is a simple/founded/justified (weak) repair for $\langle\mathcal{I},\eta_i\rangle$ for $i=1,\ldots,n$ and $\mathcal{U}=\cup_{i=1}^{n}\mathcal{U}_i$, then $\mathcal{U}$ is a simple/founded/justified (weak) repair for $\langle\mathcal{I},\eta\rangle$.*

Proof. By induction on n. For $n=1$, the results are trivial. Assume that the result is true for n; applying the induction hypothesis to $\eta_1,\ldots,\eta_n$, on the one hand, and the adequate lemma from Section 3 to $\eta'=\bigcup_{i=1}^{n}$ and η_{n+1}, yields the result for $\eta_1,\ldots,\eta_{n+1}$, since $\eta' \perp\!\!\!\perp \eta_{n+1}$. □

To find a partition of η (actually, the best partition of η), we will define an auxiliary relation on AICs. Two AICs r_1 and r_2 are *dependent*, $r_1 \not\perp\!\!\!\perp r_2$, if there exist literals $L_1\in\mathsf{body}\,(r_1)$ and $L_2\in\mathsf{body}\,(r_2)$ such that $|L_1|=|L_2|$.

Lemma 10. *Let $\boldsymbol{\eta}$ be a partition of η. Then η_i is closed under $\not\perp\!\!\!\perp$ for every i, i.e. for every rule $r,r'\in\eta$, if $r\in\eta_i$ and $r \not\perp\!\!\!\perp r'$, then $r'\in\eta_i$.*

Proof. Let r be a rule in η_i and let $r'\in\eta$ be such that $r \not\perp\!\!\!\perp r'$. Since $\boldsymbol{\eta}$ is a partition of η, $r'\in\eta_k$ for some k. But $i\neq k$ would contradict $\eta_i \perp\!\!\!\perp \eta_k$ (since r and r' are not independent), hence $i=k$. Therefore η_i is closed under $\not\perp\!\!\!\perp$. □

This relation is reflexive and symmetric, so its transitive closure $\not\perp\!\!\!\perp^{+}$ is an equivalence relation. This equivalence relation defines the best partition of η.

Theorem 2. *The quotient set $^{\eta}/_{\not\perp\!\!\!\perp^{+}}$ is a partition of η. Furthermore, for any other partition $\boldsymbol{\eta}'$ of η, if $\eta_i'\in\boldsymbol{\eta}'$, there exists $\eta_j\in {}^{\eta}/_{\not\perp\!\!\!\perp^{+}}$ such that $\eta_j\subseteq\eta_i'$.*

Proof. Let $^{\eta}/_{\not\perp\!\!\!\perp^{+}}=\{\eta_1,\ldots,\eta_n\}$. By definition of quotient set, $\bigcup_{i=1}^{n}\eta_i=\eta$. By definition of $\not\perp\!\!\!\perp$, $\eta_i \perp\!\!\!\perp \eta_j$. Given η_i' as in the statement of the theorem and choosing $r\in\eta_i'$ and observing that η_i' is closed under $\not\perp\!\!\!\perp$ (Lemma 10) and $[r]$ is the minimal set containing r and closed under $\not\perp\!\!\!\perp$, it follows that $[r]\subseteq\eta_i'$. □

Furthermore, $^{\eta}/_{\not\perp\!\!\!\perp^{+}}$ can be computed efficiently and in an incremental way.

Theorem 3. *Let η be a set of AICs such that every rule in η contains at most k literals in its body. Then $^{\eta}/_{\not\perp\!\!\!\perp^{+}}$ can be computed in $\mathcal{O}\,(k\times|\eta|)$.*

Proof. Consider the undirected graph whose nodes are both the rules in η and the atoms occurring in those rules, and where there is an edge between an atom and a rule if that atom occurs in that rule. This graph has at most $k\times|\eta|$ nodes and can be constructed in $\mathcal{O}\,(k\times|\eta|)$ time; it is a well-known fact that its connected components can again be computed in $\mathcal{O}\,(k\times|\eta|)$ time, and the rules in each component coincide precisely with the equivalence classes in $^{\eta}/_{\not\perp\!\!\!\perp^{+}}$. □

Three important remarks are due. First, k typically does not grow with η and is usually small, so essentially this algorithm is linear in the number of AICs. Also, the algorithm is independent of the underlying database, which is useful since the database typically changes more often than η. Finally, if one wishes to add new rules to η one can reuse the existing partition for η as a starting point, which makes the algorithm incremental.

5 Stratified Active Integrity Constraints

In this section, we show how to define a finer relation among active integrity constraints that will allow an incremental construction of these repairs that can again substantially reduce the time required to find them.

Throughout this section we assume a fixed set of AICs η, so all definitions are within the universe of this set.

Definition 3. *Let r_1 and r_2 be active integrity constraints. Then $r_1 \prec r_2$ (r_1* precedes *r_2) if* $\{|\mathsf{lit}(\alpha)| \mid \alpha \in \mathsf{head}(r_1)\} \cap \{|L| \mid L \in \mathsf{body}(r_2)\} \neq \emptyset$.

Intuitively, r_1 precedes r_2 if ensuring r_1 may affect applicability of r_2. In particular, $r_1 \prec r_2$ implies $r_1 \not\perp r_2$.

By definition of AIC, $\prec$ is a reflexive relation. Let $\preceq$ be its transitive closure (within η) and $\approx$ be the equivalence relation induced by $\preceq$, i.e. $r_1 \approx r_2$ iff $r_1 \preceq r_2$ and $r_2 \preceq r_1$. It is a well-known result that $\langle {}^{\eta}/_{\approx}, \preceq\rangle$ is a partial order, where $[r_1] \preceq [r_2]$ iff $r_1 \preceq r_2$.

Definition 4. *Let $\eta_1, \eta_2 \subseteq \eta$ be closed under $\approx$. Then $\eta_1 \prec \eta_2$ (η_1* precedes *η_2) if (i) some rule in η_1 precedes some other rule in η_2, but (ii) no rule in η_2 precedes a rule in η_1.*[4]

In particular, if $\eta_1 \prec \eta_2$ then η_1 and η_2 must be disjoint. Note that, if η_1 and η_2 are distinct minimal sets closed under $\approx$ (i.e. elements of $\eta/\approx$), then $\eta_1 \preceq \eta_2$ iff $\eta_1 \prec \eta_2$.

This stratification allows us to search for weak repairs as follows: if $\eta_1 \prec \eta_2$, then we can look for weak repairs for $\eta_1 \cup \eta_2$ by first looking for weak repairs for η_1 and then extending these to $\eta_1 \cup \eta_2$.

Lemma 11. *Let $\eta_1, \eta_2 \subseteq \eta$ with $\eta_1 \prec \eta_2$, $\mathcal{I}$ be a database and $\mathcal{U}$ be a set of update actions such that all actions in $\mathcal{U}$ occur in the head of some rule in $\eta_1 \cup \eta_2$. Let $\mathcal{U}_i$ be the restriction of $\mathcal{U}$ to the actions in the heads of rules in η_i. If $\mathcal{U}$ is a weak repair for $\langle \mathcal{I}, \eta_1 \cup \eta_2\rangle$, then $\mathcal{U}_1$ and $\mathcal{U}_2$ are weak repairs for $\langle \mathcal{I}, \eta_1\rangle$ and $\langle \mathcal{I} \circ \mathcal{U}_1, \eta_2\rangle$, respectively.*

Proof. Since $\eta_1 \prec \eta_2$, (a) actions in the head of a rule in η_2 cannot change literals in the body of rules in η_1 and in particular (b) $\mathcal{U}_1$ and $\mathcal{U}_2$ are disjoint.

By (a), $\mathcal{I} \circ \mathcal{U}_1 \models r$ iff $\mathcal{I} \circ \mathcal{U} \models r$ for every $r \in \eta_1$, so $\mathcal{U}_1$ is a weak repair for $\langle \mathcal{I}, \eta_1\rangle$. By (b), $\mathcal{I} \circ \mathcal{U} = \mathcal{I} \circ (\mathcal{U}_1 \cup \mathcal{U}_2) = (\mathcal{I} \circ \mathcal{U}_1) \circ \mathcal{U}_2$, hence $\mathcal{U}_2$ is a weak repair for $\langle \mathcal{I} \circ \mathcal{U}_1, \eta_2\rangle$. □

[4] Formally: $\eta_1 \prec \eta_2$ if (i) $r_1 \prec r_2$ for some $r_1 \in \eta_1$ and $r_2 \in \eta_2$, but (ii) $r_2 \not\prec r_1$ for every $r_1 \in \eta_1$ and $r_2 \in \eta_2$.

Lemma 12. *In the conditions of Lemma 11, if $\mathcal{U}$ is founded w.r.t. $\langle\mathcal{I},\eta_1\cup\eta_2\rangle$, then $\mathcal{U}_1$ and $\mathcal{U}_2$ are founded w.r.t. $\langle\mathcal{I},\eta_1\rangle$ and $\langle\mathcal{I}\circ\mathcal{U}_1,\eta_2\rangle$, respectively.*

Proof. (i) Let $\alpha\in\mathcal{U}_1$. Since $\mathcal{U}$ is founded w.r.t. $\langle\mathcal{I},\eta_1\cup\eta_2\rangle$, there is a rule $r\in\eta_1\cup\eta_2$ such that $\alpha\in\mathsf{head}(r)$ and $\mathcal{I}\circ\mathcal{U}\models L$ for every $L\in\mathsf{body}(r)\setminus\{\mathsf{lit}(\alpha)^D\}$. Since $\eta_1\prec\eta_2$, necessarily $r\in\eta_1$. By (b) from the previous proof, $\mathcal{I}\circ\mathcal{U}_1\models L$ for every $L\in\mathsf{body}(r)\setminus\{\mathsf{lit}(\alpha)^D\}$, whence α is founded w.r.t. $\langle\mathcal{I},\eta_1\rangle$ and $\mathcal{U}_1$. Thus $\mathcal{U}_1$ is founded w.r.t. $\langle\mathcal{I},\eta_1\rangle$.

(ii) Let $\alpha\in\mathcal{U}_2$. Again there must be a rule $r\in\eta$ such that $\alpha\in\mathsf{head}(r)$ and $\mathcal{I}\circ\mathcal{U}\models L$ for every $L\in\mathsf{body}(r)\setminus\{\mathsf{lit}(\alpha)^D\}$, and as before necessarily $r\in\eta_2$. Since $I\circ\mathcal{U}=(\mathcal{I}\circ\mathcal{U}_1)\circ\mathcal{U}_2$, it follows that α is founded w.r.t. $\langle\mathcal{I}\circ\mathcal{U}_1,\eta_2\rangle$ and $\mathcal{U}_2$, hence $\mathcal{U}_2$ is founded w.r.t. $\langle\mathcal{I}\circ\mathcal{U}_1,\eta_2\rangle$. □

Corollary 3. *If $\mathcal{U}$ is a founded weak repair for $\langle\mathcal{I},\eta\rangle$, then $\mathcal{U}_1$ and $\mathcal{U}_2$ are founded weak repairs for $\langle\mathcal{I},\eta_1\rangle$ and $\langle\mathcal{I}\circ\mathcal{U}_1,\eta_2\rangle$, respectively.*

Proof. Immediate consequence of Lemmas 11 and 12. □

Lemma 13. *In the conditions of Lemma 11, if $\mathcal{U}$ is a justified weak repair for $\langle\mathcal{I},\eta_1\cup\eta_2\rangle$, then $\mathcal{U}_1$ and $\mathcal{U}_2$ are justified weak repairs for $\langle\mathcal{I},\eta_1\rangle$ and $\langle\mathcal{I}\circ\mathcal{U}_1,\eta_2\rangle$, respectively.*

Proof. We first make some remarks that will be relevant throughout the proof.

(a) $\mathcal{I}\circ\mathcal{U}_1\models L$ iff $\mathcal{I}\circ\mathcal{U}\models L$ for every literal $L\in\mathsf{body}(r)$ with $r\in\eta_1$, as argued in the proof of Lemma 11.
(b) $\mathsf{ne}(\mathcal{I},\mathcal{I}\circ\mathcal{U})\subseteq\mathsf{ne}(\mathcal{I},\mathcal{I}\circ\mathcal{U}_1)$, as in the proof of Lemma 8.
(c) $\mathsf{ne}(\mathcal{I},\mathcal{I}\circ\mathcal{U}_1)\subseteq\mathsf{ne}(\mathcal{I},\mathcal{I}\circ\mathcal{U})\cup\mathcal{U}_2$, since actions in $\mathcal{U}_2$ may not affect literals in the body of rules in η_1 (this would contradict $\eta_1\prec\eta_2$). In particular, if $\mathsf{nup}(r)\subseteq\mathsf{lit}(\mathsf{ne}(\mathcal{I},\mathcal{I}\circ\mathcal{U}_1))$ for some $r\in\eta_1$, then $L\in\mathsf{lit}(\mathsf{ne}(\mathcal{I},\mathcal{I}\circ\mathcal{U}))$; and if $\alpha\in\mathsf{head}(r)$ for some $r\in\eta_1$ and $\alpha\in\mathsf{ne}(\mathcal{I},\mathcal{I}\circ\mathcal{U}_1)$, then $\alpha\in\mathsf{ne}(\mathcal{I},\mathcal{I}\circ\mathcal{U})$.

(i) Let $r\in\eta_1$ be such that $\mathsf{nup}(r)\subseteq\mathsf{lit}(\mathcal{U}_1\cup\mathsf{ne}(\mathcal{I},\mathcal{I}\circ\mathcal{U}_1))$. From $\mathcal{U}_1\subseteq\mathcal{U}$ and (c), one gets $\mathsf{nup}(r)\subseteq\mathsf{lit}(\mathcal{U}\cup\mathsf{ne}(\mathcal{I},\mathcal{I}\circ\mathcal{U}))$; since $\mathcal{U}$ is closed under η, $\mathsf{head}(r)\cap(\mathcal{U}\cup\mathsf{ne}(\mathcal{I},\mathcal{I}\circ\mathcal{U}))\neq\emptyset$. By definition of $\mathcal{U}_1$ and (c), also $\mathsf{head}(r)\cap(\mathcal{U}_1\cup\mathsf{ne}(\mathcal{I},\mathcal{I}\circ\mathcal{U}_1))\neq\emptyset$, whence $\mathcal{U}_1\cup\mathsf{ne}(\mathcal{I},\mathcal{I}\circ\mathcal{U}_1)$ is closed under η_1.

For minimality, suppose that $\mathcal{U}_1'\subsetneq\mathcal{U}_1$ is such that $\mathcal{U}_1'\cup\mathsf{ne}(\mathcal{I},\mathcal{I}\circ\mathcal{U}_1)$ is closed under η_1 and take $\mathcal{U}'=\mathcal{U}_1'\cup\mathcal{U}_2$. We show that $\mathcal{U}'\cup\mathsf{ne}(\mathcal{I},\mathcal{I}\circ\mathcal{U})$ is closed under η. Assume $\mathsf{nup}(r)\subseteq\mathsf{lit}(\mathcal{U}'\cup\mathsf{ne}(\mathcal{I},\mathcal{I}\circ\mathcal{U}))$; there are two possible cases.

- $r\in\eta_1$: since $\eta_1\prec\eta_2$, no literal in $\mathsf{nup}(r)$ can occur in $\mathsf{lit}(\mathcal{U}_2)$; therefore, from (b) it follows that $\mathsf{nup}(r)\subseteq\mathsf{lit}(\mathcal{U}_1'\cup\mathsf{ne}(\mathcal{I},\mathcal{I}\circ\mathcal{U}_1))$, and thus $\mathsf{head}(r)\cap(\mathcal{U}_1'\cup\mathsf{ne}(\mathcal{I},\mathcal{I}\circ\mathcal{U}_1))\neq\emptyset$. From $\mathcal{U}_1'\subseteq\mathcal{U}'$ and (c), also $\mathsf{head}(r)\cap(\mathcal{U}'\cup\mathsf{ne}(\mathcal{I},\mathcal{I}\circ\mathcal{U}))\neq\emptyset$.
- $r\in\eta_2$: from $(\mathcal{U}_1'\cup\mathcal{U}_2)\subseteq\mathcal{U}$, we conclude that $\mathsf{nup}(r)\subseteq\mathsf{lit}(\mathcal{U}\cup\mathsf{ne}(\mathcal{I},\mathcal{I}\circ\mathcal{U}))$, hence $\mathsf{head}(r)\cap(\mathcal{U}\cup\mathsf{ne}(\mathcal{I},\mathcal{I}\circ\mathcal{U}))\neq\emptyset$ because $\eta_2\subseteq\eta$ and $\mathcal{U}$ is closed for η. But $\mathsf{head}(r)$ cannot contain actions in $\mathcal{U}_1$, so $\mathsf{head}(r)\cap(\mathcal{U}_2\cup\mathsf{ne}(\mathcal{I},\mathcal{I}\circ\mathcal{U}))\neq\emptyset$, whence $\mathsf{head}(r)\cap(\mathcal{U}'\cup\mathsf{ne}(\mathcal{I},\mathcal{I}\circ\mathcal{U}))\neq\emptyset$.

In either case, from $\mathsf{nup}(r) \subseteq (\mathcal{U}' \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U}))$ one concludes that $\mathsf{head}(r) \cap (\mathcal{U}' \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U})) \neq \emptyset$, whence $\mathcal{U}' \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U})$ is closed for η, which contradicts $\mathcal{U}$ being a justified weak repair for $\langle \mathcal{I}, \eta \rangle$. This is absurd, therefore $\mathcal{U}_1$ is a justified weak repair for $\langle \mathcal{I}, \eta_1 \rangle$.

(ii) Denote by $\mathcal{N}$ the set $\mathsf{ne}(\mathcal{I} \circ \mathcal{U}_1, \mathcal{I} \circ \mathcal{U}_1 \circ \mathcal{U}_2)$. To show that $\mathcal{U}_2$ is a justified weak repair for $\langle \mathcal{I} \circ \mathcal{U}_1, \eta_2 \rangle$, we need to show that $\mathcal{U}_2 \cup \mathcal{N}$ is closed for η_2 and that it is the minimal such set containing $\mathcal{N}$. Note that (d) $\mathcal{N} = \mathcal{U}_1 \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U})$, since $I \circ \mathcal{U}_1$ is "between" $\mathcal{I}$ and $\mathcal{I} \circ \mathcal{U}$ (as $\mathcal{U}_1 \subseteq \mathcal{U}$).

First we show that $\mathcal{U}_2 \cup \mathcal{N}$ is closed for η_2. Let $r \in \eta_2$ and assume that $\mathsf{nup}(r) \subseteq \mathsf{lit}(\mathcal{U}_2 \cup \mathcal{N})$. By (d), $\mathsf{nup}(r) \subseteq \mathsf{lit}(\mathcal{U}_2 \cup \mathcal{U}_1 \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U})) = (\mathcal{U} \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U}))$, whence $\mathsf{head}(r) \cap (\mathcal{U} \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U})) \neq \emptyset$ because $\mathcal{U} \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U})$ is closed under η. By construction of $\mathcal{U}_2$ and (d), also $\mathsf{head}(r) \cap (\mathcal{U}_2 \cup \mathcal{N})$, whence $\mathcal{U}_2 \cup \mathcal{N}$ is closed under η_2.

Now let $\mathcal{U}_2' \subsetneq \mathcal{U}_2$ be such that $\mathcal{U}_2' \cup \mathcal{N}$ is closed for η_2 and take $\mathcal{U}' = \mathcal{U}_1 \cup \mathcal{U}_2'$. We show that $\mathcal{U}' \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U})$ is closed under η. Let $r \in \eta$ be such that $\mathsf{nup}(r) \subseteq \mathsf{lit}(\mathcal{U}' \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U}))$. Yet again, there are two cases to consider.

- $r \in \eta_1$: since $\mathcal{U}' \subseteq \mathcal{U}$, also $\mathsf{nup}(r) \subseteq \mathsf{lit}(\mathcal{U} \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U}))$, whence $\mathsf{head}(r) \cap (\mathcal{U} \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U})) \neq \emptyset$ because $\mathcal{U} \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U})$ is closed for η. But actions in $\mathsf{head}(r)$ may not occur in $\mathcal{U}_2$, hence $\mathsf{head}(r) \cap (\mathcal{U}' \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U})) \neq \emptyset$ since $(\mathcal{U} \setminus \mathcal{U}') \subseteq \mathcal{U}_2$.
- $r \in \eta_2$: by (d), $\mathcal{U}' \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U}) = \mathcal{U}_1 \cup \mathcal{U}_2' \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U}) = \mathcal{U}_2' \cup \mathcal{N}$, whence $\mathsf{head}(r) \cap (\mathcal{U}_2' \cup \mathcal{N}) \neq \emptyset$ because $\mathcal{U}_2' \cup \mathcal{N}$ is closed for η_2, which amounts to saying that that $\mathsf{head}(r) \cap (\mathcal{U}' \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U})) \neq \emptyset$.

In either case, $\mathsf{head}(r) \cap (\mathcal{U}' \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U})) \neq \emptyset$, so $\mathcal{U}' \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U})$ is closed under η, contradicting the fact that $\mathcal{U}$ is a justified weak repair for $\langle \mathcal{I}, \eta \rangle$. Therefore $\mathcal{U}_2$ is a justified weak repair for $\langle \mathcal{I} \circ \mathcal{U}_1, \eta_2 \rangle$. □

Lemmas 11, 12 and 13 are analogue to Lemmas 4, 7 and 9, respectively. Interestingly, the analogue of Lemma 5 does not hold in this setting: it may happen that $\mathcal{U}$ is a repair, but $\mathcal{U}_1$ is a weak repair. The reason is that there may be a repair for $\langle \mathcal{I}, \eta_1 \rangle$ such that there is no (weak) repair for $\langle \mathcal{I} \circ \mathcal{U}_1, \eta_2 \rangle$.

Example 1. Let $\mathcal{I} = \emptyset$ and consider the following active integrity constraints.

$$r_1 : \mathsf{not}\ a \supset +a \qquad r_4 : a, \mathsf{not}\ b, \mathsf{not}\ c, d \supset -d$$
$$r_2 : \mathsf{not}\ b, c \supset +b \qquad r_5 : a, \mathsf{not}\ b, \mathsf{not}\ c, \mathsf{not}\ d \supset +d$$
$$r_3 : b, \mathsf{not}\ c \supset +c$$

Taking $\eta_1 = \{r_1, r_2, r_3\}$ and $\eta_2 = \{r_4, r_5\}$, one has $\eta_1 \prec \eta_2$. Furthermore, $\{+a\}$ and $\{+a, +b, +c\}$ are weak repairs for $\langle \mathcal{I}, \eta_1 \rangle$, the first of which is a repair. However, the only repair for $\langle \mathcal{I}, \eta_1 \cup \eta_2 \rangle$ is $\{+a, +b, +c\}$, which is not the union of $\{+a\}$ with a repair for $\langle \mathcal{I} \circ \{+a\}, \eta_2 \rangle$.

However, if both steps succeed then we can combine their results as before.

Lemma 14. *Let $\eta_1, \eta_2 \subseteq \eta$ with $\eta_1 \prec \eta_2$, $\mathcal{I}$ be a database, and $\mathcal{U}_1$ and $\mathcal{U}_2$ be sets of update actions such that all actions in $\mathcal{U}_i$ occur the head of some rule in η_i. If $\mathcal{U}_1$ is a weak repair for $\langle \mathcal{I}, \eta_1 \rangle$ and $\mathcal{U}_2$ is a weak repair for $\langle \mathcal{I} \circ \mathcal{U}_1, \eta_2 \rangle$, then $\mathcal{U} = \mathcal{U}_1 \cup \mathcal{U}_2$ is a weak repair for $\langle \mathcal{I}, \eta_1 \cup \eta_2 \rangle$.*

Proof. Since $\eta_1 \prec \eta_2$, the hypothesis over $\mathcal{U}_2$ imply that (a) actions in $\mathcal{U}_2$ cannot change literals in the body of rules in η_1 and in particular (b) $\mathcal{U}_1$ and $\mathcal{U}_2$ are disjoint.

Take $r \in \eta_1$. Then $\mathcal{I} \circ \mathcal{U}_1 \models r$, whence $\mathcal{I} \circ \mathcal{U} \models r$ by (a).

Take $r \in \eta_2$. Then $(\mathcal{I} \circ \mathcal{U}_1) \circ \mathcal{U}_2 \models r$, and by (b) $(\mathcal{I} \circ \mathcal{U}_1) \circ \mathcal{U}_2 = \mathcal{I} \circ \mathcal{U}$.

Therefore $\mathcal{U}_1 \circ \mathcal{U}_2$ is a weak repair for $\langle \mathcal{I}, \eta_1 \cup \eta_2 \rangle$. □

Lemma 15. *In the conditions of Lemma 14, if $\mathcal{U}_1$ is a repair for $\langle \mathcal{I}, \eta_1 \rangle$ and $\mathcal{U}_2$ is a repair for $\langle \mathcal{I} \circ \mathcal{U}_1, \eta_2 \rangle$, then $\mathcal{U}$ is a repair for $\langle \mathcal{I}, \eta_1 \cup \eta_2 \rangle$.*

Proof. By Lemma 14, $\mathcal{U}$ is a weak repair for $\langle \mathcal{I}, \eta_1 \cup \eta_2 \rangle$. Suppose $\mathcal{U}$ is not a repair; then there is $\mathcal{U}' \subsetneq \mathcal{U}$ such that $\mathcal{U}'$ is also a weak repair for $\langle \mathcal{I}, \eta_1 \cup \eta_2 \rangle$.

Take $\mathcal{U}'_1 = \mathcal{U}' \cap \mathcal{U}_1$ and $\mathcal{U}'_2 = \mathcal{U}' \cap \mathcal{U}_2$; by Lemma 11, $\mathcal{U}'_1$ is a weak repair for $\langle \mathcal{I}, \eta_1 \rangle$ and $\mathcal{U}'_2$ is a weak repair for $\langle \mathcal{I} \circ \mathcal{U}_1, \eta_2 \rangle$. But at least one of the inclusions $\mathcal{U}'_1 \subseteq \mathcal{U}_1$ and $\mathcal{U}'_2 \subseteq \mathcal{U}_2$ must be strict, contradicting the hypothesis that $\mathcal{U}_1$ and $\mathcal{U}_2$ are both repairs. Therefore $\mathcal{U}$ is a repair for $\langle \mathcal{I}, \eta_1 \cup \eta_2 \rangle$. □

In this setting, the condition that $\mathcal{U}_1$ and $\mathcal{U}_2$ be repairs is sufficient but not necessary, as illustrated by the example above – unlike in Lemma 3 earlier.

Lemma 16. *In the conditions of Lemma 14, if $\mathcal{U}_1$ is founded w.r.t. $\langle \mathcal{I}, \eta_1 \rangle$ and $\mathcal{U}_2$ is founded w.r.t. $\langle \mathcal{I} \circ \mathcal{U}_1, \eta_2 \rangle$, then $\mathcal{U}$ is founded w.r.t. $\langle \mathcal{I}, \eta_1 \cup \eta_2 \rangle$.*

Proof. Take $\alpha \in \mathcal{U}_1$. Since $\mathcal{U}_1$ is founded w.r.t. $\langle \mathcal{I}, \eta_1 \rangle$, there is a rule $r \in \eta_1$ such that $\alpha \in \mathsf{head}\,(r)$ and $\mathcal{I} \circ \mathcal{U}_1 \models L$ for every $L \in \mathsf{body}\,(r) \setminus \{\mathsf{lit}(\alpha)^D\}$. By (b) from the proof of Lemma 14, also $\mathcal{I} \circ \mathcal{U} \models L$ for every $L \in \mathsf{body}\,(r) \setminus \{\mathsf{lit}(\alpha)^D\}$, whence α is founded w.r.t. $\langle \mathcal{I}, \eta_1 \cup \eta_2 \rangle$ and $\mathcal{U}$.

Take $\alpha \in \mathcal{U}_2$. Since $\mathcal{U}_2$ is founded w.r.t. $\langle \mathcal{I} \circ \mathcal{U}_1, \eta_2 \rangle$, there is a rule $r \in \eta_2$ such that $(\mathcal{I} \circ \mathcal{U}_1) \circ \mathcal{U}_2 \models L$ for every $L \in \mathsf{body}\,(r) \setminus \{\mathsf{lit}(\alpha)^D\}$, and since $(\mathcal{I} \circ \mathcal{U}_1) \circ \mathcal{U}_2 = \mathcal{I} \circ \mathcal{U}$ this implies that α is founded w.r.t. $\langle \mathcal{I}, \eta_1 \cup \eta_2 \rangle$ and $\mathcal{U}$.

Therefore $\mathcal{U}$ is founded w.r.t. $\langle \mathcal{I}, \eta_1 \cup \eta_2 \rangle$. □

As before, Lemmas 14, 15 and 16 can be combined in the following corollary.

Corollary 4. *In the conditions of Lemma 14, if $\mathcal{U}_1$ is a founded (weak) repair for $\langle \mathcal{I}, \eta_1 \rangle$ and $\mathcal{U}_2$ is a founded (weak) repair for $\langle \mathcal{I} \circ \mathcal{U}_1, \eta_2 \rangle$, then $\mathcal{U}$ is a founded (weak) repair for $\langle \mathcal{I}, \eta_1 \cup \eta_2 \rangle$.*

Lemma 17. *In the conditions of Lemma 14, if $\mathcal{U}_1$ is a justified weak repair for $\langle \mathcal{I}, \eta_1 \rangle$ and $\mathcal{U}_2$ is a justified weak repair for $\langle \mathcal{I} \circ \mathcal{U}_1, \eta_2 \rangle$, then $\mathcal{U}$ is a justified weak repair for $\langle \mathcal{I}, \eta_1 \cup \eta_2 \rangle$.*

Proof. First observe that properties (a–d) of the proof of Lemma 13 all hold in this context. Define $\mathcal{N} = \mathsf{ne}\,(\mathcal{I} \circ \mathcal{U}_1, \mathcal{I} \circ \mathcal{U}_1 \circ \mathcal{U}_2)$ as in that proof.

To see that $\mathcal{U} \cup \mathsf{ne}\,(\mathcal{I}, \mathcal{I} \circ \mathcal{U})$ is closed for $\langle \mathcal{I}, \eta \rangle$, let $r \in \eta_1 \cup \eta_2$ be such that $\mathsf{nup}(r) \subseteq \mathsf{lit}(\mathcal{U} \cup \mathsf{ne}\,(\mathcal{I}, \mathcal{I} \circ \mathcal{U}))$. We need to consider two cases.

– If $r \in \eta_1$, then $\mathsf{nup}(r) \subseteq \mathsf{lit}(\mathcal{U}_1 \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U}_1))$ by (a) and (b), and since $\mathcal{U}_1 \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U}_1)$ closed for η_1 this implies that $\mathsf{head}(r) \cap (\mathcal{U}_1 \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U}_1)) \neq \emptyset$, whence also $\mathsf{head}(r) \cap (\mathcal{U} \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U})) \neq \emptyset$ by $\mathcal{U}_1 \subseteq \mathcal{U}$ and (c).
– If $r \in \eta_2$, then by equality (d) we have $\mathcal{U} \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U}) = \mathcal{U}_2 \cup \mathcal{U}_1 \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U}) = \mathcal{U}_2 \cup \mathcal{N}$; then $\mathsf{nup}(r) \subseteq \mathsf{lit}(\mathcal{U}_2 \cup \mathcal{N})$, whence $\mathsf{head}(r) \cap (\mathcal{U}_2 \cup \mathcal{N}) \neq \emptyset$ because $\mathcal{U}_2 \cup \mathcal{N}$ is closed for η_2, and the latter condition is precisely $\mathsf{head}(r) \cap (\mathcal{U} \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U})) \neq \emptyset$.

In either case $\mathcal{U} \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U})$ is closed for r, whence $\mathcal{U} \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U})$ is closed for $\eta_1 \cup \eta_2$.

For minimality, let $\mathcal{U}' \subseteq \mathcal{U}$ be such that $\mathcal{U}' \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U})$ is closed for $\eta_1 \cup \eta_2$ and take $\mathcal{U}'_i = \mathcal{U}' \cap \mathcal{U}_i$ for $i = 1, 2$. We show that $\mathcal{U}'_1 = \mathcal{U}_1$ and $\mathcal{U}'_2 = \mathcal{U}_2$.

– Let $r \in \eta_1$ be such that $\mathsf{nup}(r) \subseteq \mathsf{lit}(\mathcal{U}'_1 \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U}_1))$. Since $\mathcal{U}'_1 \subseteq \mathcal{U}'$, from (c) and the fact that $\mathsf{nup}(r) \cap \mathsf{lit}(\mathcal{U}_2) = \emptyset$ (because $\eta_1 \prec \eta_2$) we conclude that $\mathsf{nup}(r) \subseteq \mathsf{lit}(\mathcal{U}' \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U})$, whence $\mathsf{head}(r) \cap (\mathcal{U}' \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U})) \neq \emptyset$. By (b) and the fact that $\mathsf{head}(r) \cap \mathcal{U}_2 = \emptyset$, also $\mathsf{head}(r) \cap (\mathcal{U}'_1 \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U}_1)) \neq \emptyset$. Therefore $\mathcal{U}'_1 \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U}_1)$ contains $\mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U}_1)$ and is closed for η_1; since $\mathcal{U}_1 \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U}_1)$ is the minimal set with this property and $\mathcal{U}_1 \cap \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U}_1) = \emptyset$, it follows that $\mathcal{U}'_1 = \mathcal{U}_1$.
– Let $r \in \eta_2$ be such that $\mathsf{nup}(r) \subseteq \mathsf{lit}(\mathcal{U}'_2 \cup \mathcal{N})$. From (d) and the equality $\mathcal{U}'_1 = \mathcal{U}_1$ established above, $\mathsf{nup}(r) \subseteq \mathsf{lit}(\mathcal{U}' \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U}))$, whence $\mathsf{head}(r) \cap (\mathcal{U}' \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U})) \neq \emptyset$. Again by (d) and $\mathcal{U}'_1 = \mathcal{U}_1$ this amounts to saying that $\mathsf{head}(r) \cap (\mathcal{U}'_2 \cup \mathcal{N}) \neq \emptyset$. Therefore $\mathcal{U}'_2 \cup \mathcal{N}$ contains $\mathcal{N}$ and is closed for η_2, whence as before necessarily $\mathcal{U}'_2 = \mathcal{U}_2$.

Therefore $\mathcal{U}' = \mathcal{U}$, hence the set $\mathcal{U} \cup \mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U})$ is the minimal set containing $\mathsf{ne}(\mathcal{I}, \mathcal{I} \circ \mathcal{U})$ and closed for $\eta_1 \cup \eta_2$. Therefore $\mathcal{U}$ is a justified weak repair for $\langle \mathcal{I}, \eta_1 \cup \eta_2 \rangle$. □

Lemmas 11, 12 and 13 allow us to split the search for (weak) repairs into smaller steps, while Lemmas 14, 15, 16 and 17 allow us to combine the results. However, $\langle {}^{\eta}/_{\approx}, \preceq \rangle$ is in general not a total order. Therefore, to obtain (weak) repairs for η, we need to be able to combine weak repairs of sets η_1 and η_2 that are not related via $\prec$ (see example below).

Let η_1, η_2 be two such sets, and consider a weak repair $\mathcal{U}$ for $\langle \mathcal{I}, \eta_1 \cup \eta_2 \rangle$. By Lemma 11, restricting $\mathcal{U}$ to the actions in $\eta_1 \cap \eta_2$ yields a weak repair $\mathcal{U}'$ for $\langle \mathcal{I}, \eta_1 \cap \eta_2 \rangle$; furthermore, restricting $\mathcal{U}$ to the actions in $(\eta_1 \cup \eta_2) \setminus (\eta_1 \cap \eta_2)$ yields a weak repair for $\langle \mathcal{I} \circ (\eta_1 \cap \eta_2), (\eta_1 \cup \eta_2) \setminus (\eta_1 \cap \eta_2) \rangle$. This allows us to restrict ourselves, without loss of generality, to the analysis of the situation where $\eta_1 \cap \eta_2 = \emptyset$. Since in this case the application of rules in η_1 does not affect the semantics of rules in η_2 and vice-versa, the proofs of Lemmas 2, 3, 6 and 8 can be straightforwardly adapted[5] to prove the following result.

[5] Although these lemmas assume that $\eta_1 \perp\!\!\!\perp \eta_2$, the key argument is that applying rules in $\mathcal{U}_1$ does not affect the semantics of rules in η_2 and conversely, which still remains true if η_1 and η_2 are closed under $\approx$ and neither $\eta_1 \prec \eta_2$ nor $\eta_2 \prec \eta_1$.

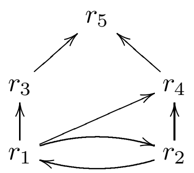
Lemma 18. *Let $\eta_1, \eta_2 \subseteq \eta$ be closed under $\approx$ and such that $\eta_1 \not\preceq \eta_2$ and $\eta_2 \not\preceq \eta_1$. Let $\mathcal{U}$ be a weak repair for $\langle \mathcal{I}, \eta_1 \cup \eta_2 \rangle$ such that $\mathcal{U}$ only consists of actions in the heads of rules in $\eta_1 \cup \eta_2$. Define $\mathcal{U}_i$ to be the restriction of $\mathcal{U}$ to the actions in the heads of rules in η_i. Then:*

1. *each $\mathcal{U}_i$ is a weak repair for $\langle \mathcal{I}, \eta_i \rangle$;*
2. *if $\mathcal{U}$ is a repair for $\langle \mathcal{I}, \eta_1 \cup \eta_2 \rangle$, then $\mathcal{U}_i$ is a repair for $\langle \mathcal{I}, \eta_i \rangle$;*
3. *if $\mathcal{U}$ is founded w.r.t. $\langle \mathcal{I}, \eta_1 \cup \eta_2 \rangle$, then $\mathcal{U}_i$ is founded w.r.t. $\langle \mathcal{I}, \eta_i \rangle$;*
4. *if $\mathcal{U}$ is a justified (weak) repair for $\langle \mathcal{I}, \eta_1 \cup \eta_2 \rangle$, then $\mathcal{U}_i$ is a justified (weak) repair for $\langle \mathcal{I}, \eta_i \rangle$.*

Example 2. To understand how these results can be applied, consider the following set of AICs η.

$$r_1 : a, b \supset -a \mid -b \qquad r_4 : a, \mathsf{not}\ b, \mathsf{not}\ e \supset +e$$
$$r_2 : \mathsf{not}\ a, c \supset +a \qquad r_5 : d, e, \mathsf{not}\ f \supset +f$$
$$r_3 : b, c, d \supset -d$$

The precedence relation between these rules, omitting the reflexive edges, can be summarized in the following diagram.



The equivalence classes are $\eta_1 = \{r_1, r_2\}$, $\eta_2 = \{r_3\}$, $\eta_3 = \{r_4\}$ and $\eta_4 = \{r_5\}$, with (direct) precedence relation $\eta_1 \preceq \eta_2 \preceq \eta_4$ and $\eta_1 \preceq \eta_3 \preceq \eta_4$. In order to find e.g. a founded weak repair for $\langle \mathcal{I}, \eta \rangle$, we would:

1. find all founded weak repairs for $\langle \mathcal{I}, \{r_1, r_2\} \rangle$;
2. extend each such $\mathcal{U}$ to founded weak repairs for $\langle \mathcal{I} \circ \mathcal{U}, \{r_3\} \rangle$ and $\langle \mathcal{I} \circ \mathcal{U}, \{r_4\} \rangle$, using Lemma 16;
3. for each pair of weak repairs $\mathcal{U}_2$ for $\langle \mathcal{I}, \{r_1, r_2, r_3\} \rangle$ and $\mathcal{U}_3$ for $\langle \mathcal{I}, \{r_1, r_2, r_4\} \rangle$ such that $\mathcal{U}_2$ and $\mathcal{U}_3$ coincide on the actions from heads of rules in $\{r_1, r_2\}$ (i.e. $-a$, $+a$ and $-b$), find weak repairs for $\langle \mathcal{I} \circ (\mathcal{U}_2 \cup \mathcal{U}_3), \{r_5\} \rangle$, using Lemma 18.

In the last step, we are using the fact that any weak repair $\mathcal{U}$ for $\langle \mathcal{I}, \eta \rangle$ must contain a weak repair $\mathcal{U}'$ for $\langle \mathcal{I}, \{r_1, r_2, r_3, r_4\} \rangle$; in turn, this can be split into a weak repair $\mathcal{U}_1$ for $\langle \mathcal{I}, \{r_1, r_2\} \rangle$ and weak repairs $\mathcal{U}_2'$ for $\langle \mathcal{I} \circ \mathcal{U}_1, \{r_3\} \rangle$ and $\mathcal{U}_3'$ for $\langle \mathcal{I} \circ \mathcal{U}_1, \{r_4\} \rangle$; defining $\mathcal{U}_2 = \mathcal{U}_2' \cup \mathcal{U}_1$ and $\mathcal{U}_3 = \mathcal{U}_3' \cup \mathcal{U}_1$, we must have $\mathcal{U}' = \mathcal{U}_1 \cup \mathcal{U}_2' \cup \mathcal{U}_3' = \mathcal{U}_2 \cup \mathcal{U}_3$. Lemma 12 guarantees that this algorithm finds all founded weak repairs for $\langle \mathcal{I}, \eta \rangle$.

6 Conclusions

We introduced independence and precedence relations among active integrity constraints that allow parallelization and sequentialization of the computation of repairs for inconsistent databases. These two processes allow us to speed up the process of finding these repairs: the advantages of parallelization are well-known, whereas the sequentialization herein presented allows a complex problem to be split in several small (and simpler) problems. Since size is a key issue in the search for repairs of a database – this being an NP- or Σ_P^2-comlete problem – it is in general much more efficient to solve several small problems than a single one as big as all of those taken together. Furthermore, the relations proposed are well-behaved w.r.t. the different kinds of repairs considered in the denotational semantics for AICs [5], so these results apply to all of them.

Using all the results presented in this paper, the strategy for computing repairs for a set η of AICs can be summarized as follows.

1. Compute ${}^{\eta}/_{\not\perp\!\!\!\perp^+}$
2. For each $\eta_i \in {}^{\eta}/_{\not\perp\!\!\!\perp^+}$
   (a) Compute $\eta_i / \approx$
   (b) Find (founded/justified) weak repairs for the minimal elements of $\eta_i / \approx$
   (c) For each non-minimal element η_j, find its (founded/justified) weak repairs by (i) combining the weak repairs for its predecessors, (ii) applying each result to $\mathcal{I}$, with result $\mathcal{I}'$, and (iii) computing (founded/justified) weak repairs for $\langle \mathcal{I}', \eta_j \rangle$ (as in the example at the end of the last section).

   This yields all (founded/justified) (weak) repairs for each element of ${}^{\eta}/_{\not\perp\!\!\!\perp^+}$.
3. Combine these (weak) repairs into a single (founded/justified) (weak) repair for η.

The only catch regards the situation depicted in Example 1: if one is interested in computing repairs, then one may restrict the search in the outer cycle to repairs. However, in step 2, whenever a repair cannot be extended when moving upwards in ${}^{\eta_i}/_{\approx}$, one must also consider weak repairs including that repair, since the end result may be a repair for the larger set. Also, if one does not want founded or justified repairs, the precedence relation cannot be used. The applicability of these techniques is summarized in Table 1.

In the worst case scenario, the set ${}^{\eta}/_{\not\perp\!\!\!\perp^+}$ will be a singleton (so there will be no parallelization) and likewise for ${}^{\eta}/_{\approx}$ (so there will be no sequentialization). However, in practical settings these are extremely unlikely situations: in typical databases concepts are built from more primitive ones, suggesting that the structure of these sets will be quite rich. Since finding repairs is an NP-complete or Σ_p^2-complete problem, this division can play a key role in making this search process much faster.

Work is in progress to implement these optimizations in order to obtain a more precise understanding of their benefits.

Table 1. Applicability of parallelization and stratification techniques to the different kinds of repairs

Type	Parallelization	Stratification
weak repairs	yes	no
repairs	yes	no
founded weak repairs	yes	yes
founded repairs	yes	yes[†]
justified weak repairs	yes	yes
justified repairs	yes	yes[†]

[†] may require computation of weak repairs

Acknowledgements. The author wishes to thank Patrícia Engrácia, Graça Gaspar and Isabel Nunes for their input and fruitful discussions on the topic of active integrity constraints. A special word of thanks goes to the anonymous referees, who provided me with very useful pointers that contributed to a much more comprehensive section on related work; and also for suggesting a simplification of the original proof of Theorem 3.

References

1. Abiteboul, S.: Updates, a new frontier. In: Gyssens, M., Paredaens, J., Van Gucht, D. (eds.) ICDT 1988. LNCS, vol. 326, pp. 1–18. Springer, Heidelberg (1988)
2. Abiteboul, S., Hull, R., Vianu, V.: Foundations of Databases. Addison-Wesley (1995)
3. Beeri, C., Vardi, M.Y.: The implication problem for data dependencies. In: Even, S., Kariv, O. (eds.) ICALP 1981. LNCS, vol. 115, pp. 73–85. Springer, Heidelberg (1981)
4. Caroprese, L., Greco, S., Sirangelo, C., Zumpano, E.: Declarative semantics of production rules for integrity maintenance. In: Etalle, S., Truszczyński, M. (eds.) ICLP 2006. LNCS, vol. 4079, pp. 26–40. Springer, Heidelberg (2006)
5. Caroprese, L., Truszczyński, M.: Active integrity constraints and revision programming. Theory Pract. Log. Program. 11(6), 905–952 (2011)
6. Chomicki, J.: Consistent query answering: Five easy pieces. In: Schwentick, T., Suciu, D. (eds.) ICDT 2007. LNCS, vol. 4353, pp. 1–17. Springer, Heidelberg (2006)
7. Cruz-Filipe, L., Engrácia, P., Gaspar, G., Nunes, I.: Computing repairs from active integrity constraints. In: Wang, H., Banach, R. (eds.) TASE 2013, pp. 183–190. IEEE (2013)
8. Eiter, T., Gottlob, G.: On the complexity of propositional knowledge base revision, updates, and counterfactuals. Artif. Intell. 57(2-3), 227–270 (1992)
9. Flesca, S., Greco, S., Zumpano, E.: Active integrity constraints. In: Moggi, E., Scott Warren, D. (eds.) PPDP, pp. 98–107. ACM (2004)
10. Kakas, A.C., Mancarella, P.: Database updates through abduction. In: McLeod, D., Sacks-Davis, R., Schek, H.-J. (eds.) VLDB 1990, pp. 650–661. Morgan Kaufmann (1990)
11. Katsuno, H., Mendelzon, A.O.: On the difference between updating a knowledge base and revising it. In: Allen, J.F., Fikes, R., Sandewall, E. (eds.) KR 1991, pp. 387–394. Morgan Kaufmann (1991)

12. Marek, V.W., Truszczynski, M.: Revision programming, database updates and integrity constraints. In: Gottlob, G., Vardi, M.Y. (eds.) ICDT 1995. LNCS, vol. 893, pp. 368–382. Springer, Heidelberg (1995)
13. Mayol, E., Teniente, E.: Addressing efficiency issues during the process of integrity maintenance. In: Bench-Capon, T.J.M., Soda, G., Tjoa, A.M. (eds.) DEXA 1999. LNCS, vol. 1677, pp. 270–281. Springer, Heidelberg (1999)
14. Naqvi, S.A., Krishnamurthy, R.: Database updates in logic programming. In: Edmondson-Yurkanan, C., Yannakakis, M. (eds.) PODS 1988, pp. 251–262. ACM (1988)
15. Przymusinski, T.C., Turner, H.: Update by means of inference rules. J. Log. Program. 30(2), 125–143 (1997)
16. Teniente, E., Olivé, A.: Updating knowledge bases while maintaining their consistency. VLDB J. 4(2), 193–241 (1995)
17. Widom, J., Ceri, S. (eds.): Active Database Systems: Triggers and Rules For Advanced Database Processing. Morgan Kaufmann (1996)
18. Winslett, M.: Updating Logical Databases. Cambridge Tracts in Theoretical Computer Science. Cambridge University Press (1990)