An overview of the ABC Repair System for Datalog-like Theories

Xue Li*, Alan Bundy*

School of Informatics, University of Edinburgh, UK

Abstract

Humans are smart in revising their knowledge and concepts based on observations when they find conflicts. This ability to repair representations is also important for AI agents so that they can represent their environment correctly. This paper gives an overview of the domain-independent ABC system for repairing faulty logical theories by combining three existing techniques: abduction, belief revision and conceptual change. (A) Given an observation, represented as an assertion, and a current theory, abduction adds axioms, or deletes preconditions, which explain that observation by making the corresponding assertion derivable from the expanded theory. (B) Belief revision incorporates a new piece of information which conflicts with the input theory by either deleting old axioms or adding new preconditions to them. (C) Conceptual change uses the reformation algorithm for blocking unwanted proofs or unblocking wanted proofs. The former two techniques change an axiom as a whole, while reformation changes the language in which the theory is written. These three techniques are complementary so they are combined into one system: the ABC repair system, which is capable of repairing logical theories with better result than each individual technique alone and has been applied to applications in multiple domains. Datalog is used as ABC's underlying logic of theories, but the proposed system has the potential to be adapted to theories in other logics.

Keywords

Automated theory repair, Abduction, Belief revision, Conceptual change, Reformation, Knowledge representation, Automated reasoning

1. Introduction

Automated agents use a representation of their environment (i) to interpret incoming sensory data, (ii) to infer new knowledge from old, (iii) to make plans to achieve their goals, and (iv) to predict the consequences of their actions and those of other agents. These environmental representations, which can be formalised as logical theories, are not static. They must change (i) when the environment changes, (ii) when the agent must deal with new kinds of goals, or (iii) when the agent detects that they are erroneous. Automated theory repair systems are proposed to address representation changes based on the proof of falsehoods [1, 2, 3]. Some repair systems are complementary in terms of the faults they tackled and the repair operations they generated. Thus, the ABC repair system is developed to combine abduction, belief revision and conceptual change for

International Workshop on Human-Like Computing (HLC 2022), 28 -30 Sept. 2022, United Kingdom
*The corresponding authors.
xue.shirley.li@ed.ac.uk (X. Li); A.Bundy@ed.ac.uk (A. Bundy)

generating more diverse repairs in a domain independent manner. No human interaction is involved in ABC's procedure but it may be necessary in particular applications, e.g., ABC is applied twice in a row in the root-cause analysis (RCA) [4], where domain experts need to select the best repair from ABC's output in the first step and then use it as ABC's input in the second step of RCA.

There are abduction techniques based on machine learning approaches, e.g., [5], but the ABC repair system is a proof driven approach so its results are explainable. On the other hand, abduction based explanation generators, such as [6], do not consider negative examples so cannot restrict their learnt rules while ABC abduces explanations for true examples under the restriction of negative examples. In addition, none of other the diagnosis and repair systems including [7] is able to conduct conceptual change, which modifies the language in which the theory is formalised. Some examples of ABC's conceptual change are given in appendix.

In ABC, a Datalog-like theory [8] is diagnosed to be faulty w.r.t. a given benchmark if it proves any negative examples or fails to prove any positive examples, which are defined as incompatibility and insufficiency, respectively [9]. Then the original theory is repaired by breaking proofs of negative examples ($\mathcal{F}(\$)$) and/or building proofs for positive examples ($\mathcal{T}(\$)$). Note that by 'negative', we mean the examples in $\mathcal{F}(\$)$ are not observed, i.e., false assertions, rather than being negations. As a result, repaired theories will contain neither insufficiencies nor incompatibilities. The diverse repairs allow ABC a wide range of applications to modelling human behaviours[1]:

Example 1. Virtual Bargaining Game Theory.

$$mark(X,\ Y) \implies select(X,Y) \qquad \text{(A1)}$$
$$\implies <(g1,hp,hm) \qquad \text{(A2)}$$
$$\implies <(g2,hm,hp) \qquad \text{(A3)}$$
$$\implies mark(g1,b1) \qquad \text{(A4)}$$
$$\implies mark(g2,b1) \qquad \text{(A5)}$$
$$\implies b1 \neq b2 \qquad \implies b1 \neq b3 \qquad \implies b2 \neq b3 \qquad \text{(A6-8)}$$

$$\mathcal{T}(\$) = \{select(g1,b1),\ select(g2,b2),\ select(g2,b3)\}$$
$$\mathcal{F}(\$) = \{select(g1,b2), select(g1,b3),\ select(g2,b1)\}$$

- Modelling virtual bargaining game theory. In this game, human players need to guess or adjust the winning strategy based only on others' game moves. Given this limited bandwidth, each player has to imagine what the other is thinking and plan their play too, so called virtual bargaining. To model this process, the game setups and a candidate of original winning strategies are given as ABC's input theory and

[1] We only give the example of the virtual bargaining in this paper. Other examples can be found in the cited papers.

the desired game movement as the benchmark of $\mathcal{T}(\$)$ and $\mathcal{F}(\$)$, as in Example 1. ABC will repair the candidate into correct winning strategies, given by Example 2. It is particularly useful when the game setup is changed so the previous winning strategy (A1 in Example 1) is outdated [10] so is evolved by ABC into a stronger strategy represented by (A1') and (TR) in Example 2. It can be seen that all examples in $\mathcal{T}(\$)$ and none in $\mathcal{F}(\$)$ are theorems of the repaired theory.

- Discovering the cause of a mathematical mistake. By giving the incorrect mathematical results $\mathbb{I}$ as the benchmark and the correct mathematical calculation rules $\mathbb{R}$ as the original theory, ABC will repair $\mathbb{R}$ into a model of the student's incorrect mathematical calculation rules $\mathbb{R}'$ which results in $\mathbb{I}$ [11].
- Root cause analysis (RCA) based on system logs of network systems, where single causes can trigger multiple failures. Taking the input theory containing the information from system logs and domain rules, ABC system can detect missing information that is essential to cause failures and then suggest repairs to fix root causes [4].
- Model physical theories: Equations describing a new domain (say electro-static force) can be based on an analogy with equations for an old domain (say, gravity). Then reformation, one of the repair approaches combined by ABC, can be used to correct discrepancies between the new equations and observations of the environment, leading to a correct theory of the new domain. This work is described in another paper in this volume [12].

Example 2. Repaired Virtual Bargaining Game Theory.

$$
\begin{array}{rrl}
<(X,\ hp,\ hm) \wedge mark(X,\ Y) \Longrightarrow & select(X,\ Y) & \text{(A1')} \\
<(Z,\ hm,\ hp) \wedge \neq (X,\ Y) \wedge mark(Z,\ X) \Longrightarrow & select(Z,\ Y) & \text{(TR)} \\
\Longrightarrow & <(g1, hp, hm) & \text{(A2)} \\
\Longrightarrow & <(g2, hm, hp) & \text{(A3)} \\
\Longrightarrow & mark(g1, b1) & \text{(A4)} \\
\Longrightarrow & mark(g2, b1) & \text{(A5)} \\
\Longrightarrow b1 \neq b2 \qquad \Longrightarrow b1 \neq b3 \qquad \Longrightarrow & b2 \neq b3 & \text{(A6-8)}
\end{array}
$$

$\mathcal{T}(\$) = \{select(g1, b1),\ select(g1, b3),\ select(g2, b2)\}$
$\mathcal{F}(\$) = \{select(g1, b2),\ select(g2, b1),\ select(g2, b3)\}$

2. The ABC Repair System

ABC repairs incompatibility and insufficiency by combining different repair approaches. Thus, ABC has richer repair operations that are not only adding or deleting axioms; it can also rewrite the language of the theory, e.g., rename a constant/predicate or adapt the arity of a predicate. Particularly, ABC can add/delete preconditions from existing rules or create new rules. More details about ABC's repair operations w.r.t. a (possibly failed) proof can be found in [9] and [13] §5.

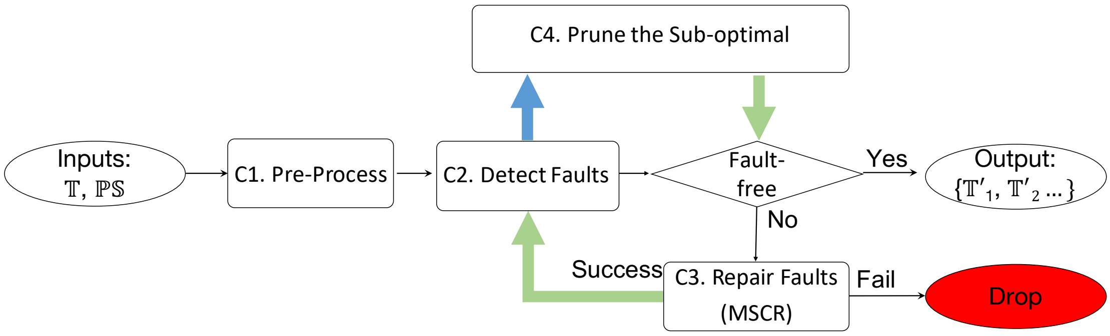

Figure 1: Pipeline of the ABC repair system: $\mathbb{T}$ is the input theory and $\mathbb{PS}$ is the benchmark; green arrows deliver a set of theories one by one to the next process; the blue arrow collects and delivers theories as a set; a faulty-theory will be dropped if it is not repairable.

The pipeline of ABC is given by Figure 1. Datalog is a declarative logic programming language in first-order logic. Currently, ABC is restricted to repair Datalog-like theories, i.e., a piece of Datalog program is seen as a logical theory based on Datalog logic. Datalog is chosen because its deduction is decidable [14] but it is sufficiently expressive to allow a wide range of practical applications [13] and §1.

C1 is a pre-process which checks whether the $\mathbb{S}$ is self-contradictory and calculates minimal sets of axioms by pruning redundancy. It reduces the search space of both fault detection and repair generation. Given a minimal axiom set, faults are detected by C2. If there is no fault, then the theory is collected as an output. Otherwise, the information about the fault, which could be either proofs or failed proofs, is provided to C3 to generate repairs. If no repairs can be found or the resource threshold[2] is reached, the process will be terminated with a failure to find any repaired theories.

Otherwise, C3 generates all possible repairs for all detected faults in parallel. When repairs which individually aim at different faults change different parts of the theory, they can be applied at the same time, called commuting repairs. C3 computes maximal sets of commuting repairs (MSCRs) and then applies each MSCR respectively to reduce the search space[3]. Then C2 will check the remaining faults of these semi-repaired theories one by one. Based on the number of the remaining faults and the applied repair operations, sub-optimal theories will be pruned by C4 [15]. This process of C2, C3 and C4 will repeat until no repairable faulty theories are left in the process, or some threshold is exceeded. Entrenchment scores, that represents how valuable a piece of information is, are estimated [16, 17] and used to rank repaired theories. In addition, a set of optional heuristics are implemented for users to choose in order to restrict repairs[4].

---

[2]The depth limit of search branches and the maximum number of repair operations given by the user.
[3]More details can be found in §6.3 in [13].
[4]More details can be found in §6.5 in [13].

3. Conclusion

An overview of the ABC repair system, which combines abduction, belief revision and conceptual change, is given in this paper. Detecting faults based on a given benchmark of positive examples and negative examples allows ABC wide-ranging applications. Meanwhile, ABC's richer repairs make it powerful in terms of adapting knowledge representation. ABC's successful applications of modelling virtual bargaining game theory, mathematics mistake causes, root-cause analysis in software system maintenance and adapting physics theories to new domain, show its power in modelling human behaviours by representation changes.

Acknowledgments

This research has been supported by the Huawei funded TREAT project CIENG4721/LSC. The first author gratefully acknowledge the support of ELIAI (The Edinburgh Laboratory for Integrated Artificial Intelligence) EPSRC (grant no EP/W002876/1) and the second author of UKRI grant EP/V026607/1. Meanwhile, great thanks to the HCL 2022 reviewers for their useful feedback, which improved the paper.

Appendix: Examples of Conceptual Change

Two examples of conceptual change are given in this appendix, together with one of their best repaired theories generated by the ABC system. Changes in repaired theories are highlighted in red.

- The *Families* Theory is repaired by enriching the constant *birth* into a variable in $A1$, which illustrates a conceptual change.

Example 3. Families Theory.

$$parent(X, Y, birth) \implies families(X, Y) \tag{A1}$$
$$\implies parent(a, b, birth) \tag{A2}$$
$$\implies parent(a, c, step)] \tag{A3}$$

$\mathscr{T}(\$) = \{families(a, b),\ \ families(a, c)\}$
$\mathscr{F}(\$) = \emptyset$

Example 4. Repaired Families Theory.

$$parent(X,Y,Z) \implies families(X,Y) \quad (A1)$$
$$\implies parent(a,b,birth) \quad (A2)$$
$$\implies parent(a,c,step)] \quad (A3)$$

$\mathcal{T}(\$) = \{families(a,b),\ families(a,c)\}$
$\mathcal{F}(\$) = \emptyset$

Example 5. Tweety Theory.

$$penguin(X) \implies bird(X) \quad (A1)$$
$$bird(X) \implies feathered(X) \quad (A2)$$
$$bird(X) \implies fly(X) \quad (A3)$$
$$\implies bird(polly) \quad (A4)$$
$$\implies penguin(tweety) \quad (A5)$$

$\mathcal{T}(\$) = \{feathered(tweety),\ feathered(polly), fly(polly)\}$
$\mathcal{F}(\$) = \{fly(tweety)\}$

- The *Tweety* theory is repaired by increasing the arity of *bird*. The repaired theory says that all birds have feathers and only normal types of birds fly, while penguin is a special type.

Example 6. Repaired Tweety Theory.

$$penguin(X) \implies bird(X, dummy2) \quad (A1)$$
$$bird(X,Y) \implies feathered(X) \quad (A2)$$
$$bird(X, dummy1) \implies fly(X) \quad (A3)$$
$$\implies bird(polly, dummy1) \quad (A4)$$
$$\implies penguin(tweety) \quad (A5)$$

$\mathcal{T}(\$) = \{feathered(tweety),\ feathered(polly), fly(polly)\}$
$\mathcal{F}(\$) = \{fly(tweety)\}$

References

[1] P. Gärdenfors, Belief revision, volume 29, Cambridge University Press, 2003.

[2] T. P.T.Cox, General Diagnosis by Inductive Inference, Technical Report, ????

[3] A. Bundy, B. Mitrovic, Reformation: A Domain-Independent Algorithm for Theory Repair, Technical Report, University of Edinburgh, 2016.

[4] X. Li, A. Bundy, ABC repair system in root cause analysis by adding missing information, in: The 8th International Online & Onsite Conference on Machine Learning, Optimization, and Data Science, special session of AI for Network/Cloud Management, 2022.

[5] S. H. Muggleton, Learning efficient logical robot strategies involving composable objects, in: Proceedings of the 24th International Joint Conference Artificial Intelligence (IJCAI 2015), 2015, pp. 3423–3429.

[6] R. Mooney, S. Bennett, A domain independent explanation-based generalizer, in: Proceedings of AAAI-86, Morgan Kaufmann, 1986, pp. 551–555.

[7] R. Reiter, A theory of diagnosis from first principles, Artificial intelligence 32 (1987) 57–95.

[8] S. Ceri, G. Gottlob, L. Tanca, Logic Programming and Databases, Surveys in Computer Science, Springer-Verlag, Berlin, 1990.

[9] X. Li, A. Bundy, A. Smaill, ABC repair system for Datalog-like theories, in: KEOD, 2018, pp. 333–340.

[10] A. Bundy, E. Philalithis, X. Li, Modelling virtual bargaining using logical representation change, in: Machine Intelligence 21 workshop, 2020, pp. 1135–1149.

[11] J. W. Q. Tang, Arithmetic Errors Revisited: Diagnosis and Remediation of Erroneous Arithmetic Performance as Repair of Faulty Representations, Msc thesis, School of Informatics, University of Edinburgh, 2016.

[12] C.-H. Cai, A. Bundy, Repairing numerical equations in analogically blended theories using reformation, in: A. Bundy, D. Mareschal (Eds.), Proceedings of HLC 2022, CEUR, 2022.

[13] X. Li, Automating the Repair of Faulty Logical Theories, Ph.D. thesis, School of Informatics, University of Edinburgh, 2021.

[14] F. Pfenning, Datalog, Lecture 26, 15-819K: Logic Programming, 2006. URL: https://www.cs.cmu.edu/~fp/courses/lp/lectures/26-datalog.pdf.

[15] M. Urbonas, A. Bundy, J. Casanova, X. Li, The use of max-sat for optimal choice of automated theory repairs, in: M. Bramer, R. Ellis (Eds.), Artificial Intelligence XXXVII, Springer International Publishing, Cham, 2020, pp. 49–63.

[16] X. Li, A. Bundy, Measuring axiom and precondition entrenchment for Datalog theory repair, in: Technical report, 2022.

[17] X. Li, A. Bundy, E. Philalithis, Signature entrenchment and conceptual changes in automated theory repair, in: The Ninth Annual Conference on Advances in Cognitive Systems, 2021.