Actual Causality in Contextual Abduction*

Degoldie Sonny[1], Ari Saptawijaya[2], and Raja Oktovin Parhasian Damanik[2]

[1] Shopee, Singapore
degoldiesonny98@gmail.com
[2] Faculty of Computer Science, Universitas Indonesia
{saptawijaya, rajaoktovin}@cs.ui.ac.id

1 Introduction

In daily life, causality is often employed to understand of why particular things happen. Different from general causality that captures the general laws that describe the cause-and-effect relationships, *actual causality* aims at explaining why a certain event happened using observations of what have happened prior to the event.

One of the well-known approaches to deal with actual causality has been developed by Halpern and Pearl (HP) [7]. HP causal model is based on structural models [14], where structural equations are introduced to capture the causal influence of some variables to another. The variables are distinguished according to how their values are determined: those whose values are determined by factors outside of the model are called *exogenous variables*, whereas *endogenous variables* have their values determined by other variables within the model. A signature $\mathcal{S}$ is a triple $(\mathcal{U}, \mathcal{V}, \mathcal{R})$, where $\mathcal{U}$ is the set of exogenous variabels, $\mathcal{V}$ is the set of endogenous variables, and $\mathcal{R}$ is a function such that for each $X \in \mathcal{U} \cup \mathcal{V}$, $\mathcal{R}(X)$ is the set of possible values that X can take. In this extended abstract, we assume that the variable is *binary* or *boolean*. A *causal model* is then a tuple $M = (\mathcal{S}, \mathcal{F})$ where $\mathcal{S} = (\mathcal{U}, \mathcal{V}, \mathcal{R})$ is a signature and $\mathcal{F} = \{F_X : X \in \mathcal{V}\}$ is a set of causal functions F_X, called *structural equations*, one for each endogenous variables X.

Bochman and Lifschitz (BL) shows that structural equations can be represented in causal calculus [4]. In boolean causal model, function F_X corresponds to *boolean structural equations* of the form $X = F$, where X is an endogenous variable and F is a propositional formula in $\mathcal{U} \cup (\mathcal{V} \setminus \{X\})$. An interpretation of propositions satisfying biconditional $A \Leftrightarrow F$ for every boolean structural equation $A = F$ in a boolean causal model M gives a *causal world* for M. Given a boolean causal model M, a *(propositional) causal theory* Δ_M can be defined as the set of the rules $F \Rightarrow A$ and $\neg F \Rightarrow \neg A$ for each structural equation $A = F$ in M and all rules $A \Rightarrow A$ and $\neg A \Rightarrow \neg A$ for every exogenous atoms A in M.

As actual causality amounts to explaining a specific event, *abduction* seems to be a natural approach in finding such explanations. In abduction, one chooses from available hypotheses those that would best explain the observed evidence

[11]. Abduction has already been well studied in logic programming with various applications [5]. *Abductive logic programming* is realized by extending logic programs with hypotheses, called *abducibles*. An *abducible* is an atom Ab or its negation not_Ab (syntactically an atom, but denoting $not\ Ab$) whose truth value are initially not assumed. Typically, abduction may be restricted by integrity constraints, but they are not needed for the purpose of this extended abstract and therefore are omitted.

The goal of this ongoing research is therefore to further examine the notion of actual causation in abductive logic programming. In particular, we look into *contextual abduction* with its implementation TABDUAL [17, 16] to model actual causality. Contextual abductions allows an observation to be explained with respect to a given context. This context can be viewed as a constraint in finding plausible explanations. In the first contribution, we propose *a causal abductive framework* to represent causal model of Halpern and Pearl [7] by translating their structural equations into abductive logic programs. In the second contribution, we provide a practical procedure, on top of TABDUAL, to enact actual causality directly from an abductive logic program by contextual abduction. Therein, the given actual causal world is treated as the context of abduction and actual causes are then computed by finding consistent abductive solutions with respect to that given context.

2 Technical Means

The rules in BL causal theory that correspond to structural equations in HP boolean causal model cannot immediately be represented as a logic program. This is because rules with a negative head, $\neg F \Rightarrow \neg A$, do not qualify as logic program's rules by definition. Nevertheless, abduction in logic programming lends itself to capturing boolean structural equations. Recall that exogenous variables differ from endogenous ones, in that the values of the former are determined by factors outside of the model. That is, in contrast to endogenous variables, they are not causally explained by variables in the model. In abductive logic programming, this notion of exogenous variables can suitably be mapped into abducibles, thus BL's rule $A \Rightarrow A$ for an exogenous variable A needs no corresponding logic program's rule.

Definition 1 (Abducibles wrt. Boolean Causal Model). *Let M be a boolean causal model. The set of abducibles with respect to model M, denoted by $\mathcal{A}_M$, is the set of exogenous atoms U in M.*

Definition 2 (Logic Program wrt. Boolean Causal Model). *Let M be a boolean causal model. The logic program with respect to model M, denoted by P_M, is the set of rules, $A \leftarrow F$, obtained from the structural equation $A = F$ in model M for each endogenous atom A.*

Note that the logic program P_M corresponds only to a subset of rules in the causal theory Δ_M for the reason explained in the beginning of this section.

Next, we state the correspondence between an abductive framework and HP boolean causal model.

Definition 3 (Causal Abductive Framework). *Let M be a boolean causal model. A causal abductive framework $F_M = \langle P_M, \mathcal{A}_M \rangle$ is an abductive framework that corresponds to M, where P_M and $\mathcal{A}_M$ are the logic program and the set of abducibles wrt. model M, respectively.*

We adopt NESS (Necessary Element of a Sufficient Set) test [19] for defining actual causality as post factum attribution of causal responsibility for actual outcome: "a particular condition c was a cause of event e". That is, in addition to a causal abductive framework F, the definition of actual causality also assumes that an *actual world* (an exact causal world) α of F is observed.

NESS test has also been used recently by Bochman [3] to define actual causality albeit using BL causal theory as its logical setting. We now refer to our causal abductive framework and define actual cause by formalizing the following NESS test: "a particular condition was a cause of (condition contributing to) a specific consequence if and only if it was a necessary element of a set of antecedent actual conditions that was sufficient for the occurrence of the consequence".

Definition 4 (Actual Cause via NESS Test). *Let F be a causal abductive framework and α be an actual world wrt. F. A literal $L \in \alpha$ is an actual cause of a literal $G \in \alpha$ wrt. F if S is a minimal abductive solution to query G wrt. F, S is consistent wrt. α, and $L \in S$.*

The consistency requirement for a minimal abductive solution wrt. an actual world can be achieved in contextual abduction by imposing the given actual world as the abductive context of the query.

Definition 5 (Abductive solution wrt. actual world). *Let F_M be a causal abductive framework corresponds to boolean causal model M. Given a query G wrt. F_M, the abductive solution S_α wrt. an actual world α is obtained by having the abductive context I for G, where I is formed from literals in α whose atoms are exogenous variables in M.*

We now set up a procedure for determining, in a particular actual world α, actual causes of an event $l \in \alpha$ wrt. a causal abductive framework F. The procedure relies on TABDUAL to carry out contextual abduction and to automatically compute the set of dual rules $dual(P)$ of the abductive logic program P [17]. In TABDUAL, a set of dual rules are introduced, by means of the *dual program transformation* [1], to deal with abduction under negative goals. The idea of the dual transformation is to define, for each atom A and its set of rules R in a logic program P, a set of dual rules whose head not_A is true if and only if A is false by R in the employed semantics of P. Note that, instead of having a negative goal $not\ A$ as the rules' head, its corresponding 'positive' literal not_A is used, thus conforming the syntax of rules in a logic program.

In the following procedure, $compl_{\mathcal{A}}(A)$ denotes the negation complement of an abducible A, where the complement of a positive abducible Ab and its negation not_Ab is defined as $compl_{\mathcal{A}}(Ab) = not\_Ab$ and $compl_{\mathcal{A}}(not\_Ab) = Ab$, respectively.

Algorithm 1: Finding actual causes via contextual abduction in TABDUAL

Input : $F = \langle P, \mathcal{A} \rangle$, actual world α, literal $l \in \alpha$
Output: actual cause of l wrt. F and causal world α

1. Compute $dual(P)$. Let $P^+ = P \cup dual(P)$.
2. Compute minimal abductive solution S of query l with empty context under P^+. Let W be the set of all such minimal abductive solutions.
3. Compute the abductive solution S_α of query l w.r.t. the given actual world α under P^+.
4. Construct set W_α that consists of only abductive solution $S \in W$ satisfying for every $t \in S_\alpha$, $\mathrm{compl}_{\mathcal{A}}(t) \notin S$.
5. Every literal $c \in S$ where $S \in W_\alpha$ is an actual cause of l in α.

The experiment in modeling examples from the literature (including Loader [10], Window [3], Backup [9], Bottle [3], Bogus Prevention [8], Push [13], Inevitable Shock [13], Purple Flame [6]) shows that the returned actual causes mostly agree with those delivered by Bochman's approach [3], albeit restricted to abducibles only.

3 Conclusions

In this extended abstract, we have tried to examine the notion of actual causation in a causal abductive framework, i.e., by translating the structural equations in the causal model of Halpern and Pearl [7] into abductive logic programs. Employing this framework, a procedure has been proposed that makes use of the tabled abduction system TABDUAL, to enact actual causality directly from an abductive logic program by contextual abduction.

Our approach is similar to the causal theory of Bochman and Lifschitz (BL) [4] in that situations are modeled using a set of rules. Different from Bochman's approach in computing actual causes [3], we focus on abduction in defining and computing actual causality. Moreover, our framework consists of normal logic programs, whereas BL causal theory allows rules with a negative head. But thanks to the dual program transformation of TABDUAL, such rules with negative head (dual rules) are automatically computed.

This research is an ongoing work. An important next step will be to evaluate further the obtained results and conduct a comparative analysis with other approaches that define actual causality in logic programming, e.g., [12, 18, 2]. It is also part of future work to explore the application of the present approach to machine ethics, extending our previous counterfactual approach [15]—also built from abduction in logic programming—for distinguishing causes and side effects in justifying morally permissible actions.

References

1. Alferes, J.J., Pereira, L.M., Swift, T.: Abduction in well-founded semantics and generalized stable models via tabled dual programs. Theory and Practice of Logic Programming **4**(4), 383–428 (2004)
2. Beckers, S., Vennekens, J.: A general framework for defining and extending actual causation using CP-logic. International Journal of Approximate Reasoning **77**, 105–126 (2016)
3. Bochman, A.: Actual causality in a logical setting. In: IJCAI. pp. 1730–1736 (2018)
4. Bochman, A., Lifschitz, V.: Pearl's causality in a logical setting. In: AAAI (2015)
5. Denecker, M., Kakas, A.C.: Abduction in logic programming. In: Computational Logic: Logic Programming and Beyond, Essays in Honour of Robert A. Kowalski, Part I. p. 402–436. Springer-Verlag, Berlin, Heidelberg (2002)
6. Ehring, D.: Causal relata. Synthese pp. 319–328 (1987)
7. Halpern, J.Y., Pearl, J.: Causes and explanations: A structural-model approach. Part I: Causes. British Journal for the Philosophy of Science **56**(4), 843–887 (2005)
8. Hiddleston, E.: Causal powers. British Journal for the Philosophy of Science **56**(1), 27–59 (2005)
9. Hitchcock, C.: Prevention, preemption, and the principle of sufficient reason. The Philosophical Review **116**(4), 495–532 (2007)
10. Hopkins, M., Pearl, J.: Clarifying the usage of structural models for commonsense causal reasoning. In: Proc. AAAI Spring Symposium on Logical Formalizations of Commonsense Reasoning. pp. 83–89 (2003)
11. Josephson, J.R., Josephson, S.G.: Abductive Inference: Computation, Philosophy, Technology. Cambridge U. P. (1995)
12. LeBlanc, E., Balduccini, M., Vennekens, J.: Explaining actual causation via reasoning about actions and change. In: JELIA. Springer (2019)
13. McDermott, M.: Redundant causation. British Journal for the Philosophy of Science pp. 523–544 (1995)
14. Pearl, J.: Causality: Models, Reasoning and Inference. Cambridge University Press (2000)
15. Pereira, L.M., Saptawijaya, A.: Counterfactuals, logic programming and agent morality. In: Urbaniak, R., Payette, G. (eds.) Applications of Formal Philosophy, vol. 14. Springer (2017)
16. Perkasa, S., Saptawijaya, A., Pereira, L.: Tabling in contextual abduction with answer subsumption. In: ICACSIS. pp. 459–464. IEEE (2017)
17. Saptawijaya, A., Pereira, L.M.: TABDUAL: a tabled abduction system for logic programs. IfCoLog Journal of Logics and their Applications **2**(1), 69–124 (2015)
18. Vennekens, J.: Actual causation in CP-logic. Theory and Practice of Logic Programming **11**(4-5), 647–662 (2011)
19. Wright, R.: Causation in tort law. California Law Review **73**, 1735 (1985)