Operational Causality – Necessarily Sufficient and Sufficiently Necessary*

Christel Baier, Clemens Dubslaff, Florian Funke, Simon Jantsch, Jakob Piribauer, and Robin Ziemek

Technische Universität Dresden, Dresden, Germany
{christel.baier, clemens.dubslaff, florian.funke, simon.jantsch, jakob.piribauer, robin.ziemek}@tu-dresden.de

Abstract. Necessity and sufficiency are well-established notions in logic and causality analysis, but have barely received attention in the formal methods community. In this paper, we present temporal logic characterizations of necessary and sufficient causes in terms of state sets in operational system models. We introduce degrees of necessity and sufficiency as quality measures for sufficient and necessary causes, respectively, along with a versatile weight-based approach to find "good causes". The resulting optimization problems of finding optimal causes are shown to be solvable in polynomial time.

1 Introduction

The classical model-checking task is to verify whether a given formal system satisfies a property usually expressed in some temporal logic [19, 66]. Much effort has been devoted to enriching classical yes/no answers of model checkers with useful diagnostic information. If the system does not meet the prescribed condition, many model checkers produce *counterexample traces* [21] that can further be investigated in order to localize precisely where the error lies or how far the trace is from satisfying the formula [8, 73, 37, 69, 35, 36]. However, realistic system models can usually produce errors for a variety of reasons so that more diverse analysis techniques are required. In the case of a positive model-checking result, *coverage estimation* aims at determining which parts of the system are essential to ensure satisfaction [45, 16, 18, 17], and *vacuity detection* analyzes whether it is due to some unintended, trivial behavior [12, 54, 65].

In this paper we tackle the explication of the behavior of transition systems through novel notions of cause–effect relationships. Both cause and effect are represented as subsets of the state space of the transition system, and formulas in linear temporal logic (LTL, [64]) are used to express the principles of necessity

* The authors are supported by the DFG through the Collaborative Research Center TRR 248 (CPEC, project ID 389792660, https://perspicuous-computing.science), the Cluster of Excellence EXC 2050/1 (CeTI, project ID 390696704, as part of Germany's Excellence Strategy) and the Research Training Groups QuantLA (GRK 1763) and RoSI (GRK 1907).

and sufficiency in causal reasoning. A *necessary cause* is a state set that necessarily needs to be passed before reaching the effect set. A *sufficient cause* is a state set where every extension of a path reaching that set eventually sees an effect state. Therefore, necessary and sufficient causes provide orthogonal views on causality in operational systems. To estimate the explanatory power of such causes and determine "good causes", we exploit counterbalances on these orthogonal views: We determine necessary causes with maximal *degree of sufficiency* and sufficient causes with maximal *degree of necessity*. In order to admit use-case specific quality criteria for necessary causes, a rather general weight-based approach is finally presented. Weight-minimal necessary causes in this framework can be computed in polynomial time via a reduction to a min-cut problem in weighted graphs.

Despite being loosely inspired by philosophical theories of causation, the theory put forth in this paper concentrates on formal operational system models and does not transcend the borders of computer science. There have been philosophical attempts to unterstand causality in terms of necessity and sufficiency [59, 60, 71, 31]. Perhaps most elaborate in this direction is the INUS condition ("insufficient but necessary part of a condition which is itself unnecessary but sufficient") [59] that is closely related to the NESS test (necessary element of sufficient subset) from jurisprudence [42, 72]. Our contributions are in some sense also orthogonal to Halpern and Pearl's *actual causality*, the perhaps most influential instance of causality in the computer science community [40, 41, 39]. Halpern and Pearl express causal dependencies in structural equation models [63, 29, 30, 38] and employ the counterfactuality principle that has a rich history in philosophical theories of causal reasoning [46, 47, 58]. Counterfactuality proclaims to consider alternative worlds in which the cause has not occurred and then check whether the effect still happened. To what extend necessity, sufficiency, counterfactuality, and conditionality etc. relate to each other and emerge to meaningful notions of causality is a matter of ongoing debate.

Related Work. Notions of causality inspired by Halpern and Pearl's actual causes have been employed in the verification landscape to analyze counterexample traces for temporal logic specifications in transition systems [11], LTL model checking [57, 10, 52, 13], concurrent interacting programs [23], and timed systems [53]. To deal with the limited expressive power of propositional structural equation models, Hopkins and Pearl [44] introduced a notion of actual causality defined in the framework of the *situation calculus* [68]. This line of work has recently been picked up again [9, 48]. Causal reasoning in component-based systems [32, 33, 34] and causality-based notions on responsibility [14] have also been considered in the model-checking community [15, 61, 26].

Rather than defining cause–effect relationships *within* a system, there are also approaches to use causal reasoning as a basis for verification algorithms *on* transition systems [55, 56] and two-player reachability games [2]. From a conceptual viewpoint, the latter article defines necessary and sufficient subgoals in the same spirit as our formalization of necessary and sufficient causes (which, nevertheless, serve a different purpose there).

Recently, notions of causality have been considered in the realm of stochastic operational systems. Based on the probability-raising principle [67], Kleinberg and Mishra [50, 51, 49] presented an approach towards causal inference in time series modeled as Markov chains. This has recently sparked novel probabilistic causality notions [4, 5], including notions of precision and recall that are closely connected to our notion of degrees of sufficiency and necessity [5]. Probabilistic causation has also been expressed in terms of hyperproperties [1, 25].

Finally, the survey article [3] exhibits how the notion of causality entered and influenced the verification landscape over the course of the past two decades.

2 Preliminaries

In the sequel, we briefly present our notation regarding transition systems, Markov chains, and linear temporal logic (LTL). For more details, see standard textbooks on systems modeling and verification [6].

A *transition system* $\mathcal{T}$ is a tuple (S, R, I) comprising a finite set of states S, a transition relation $R \subseteq S \times S$, and a set of initial states $I \subseteq S$. A state that does not have any outgoing transition is called terminal. A path π in $\mathcal{T}$ is a sequence of states $s_0 s_1 \dots$ such that $s_0 \in I$ and $(s_i, s_{i+1}) \in R$ for all appropriate i and where π is either infinite or ends in a terminal state. A state $s \in S$ is called reachable if there is a path that contains s. We assume that all states in a transition system are reachable.

A *Markov chain* $\mathcal{M}$ is a tuple $(S, \mathbf{P}, \iota)$ comprising a finite set of states S, a transition probability function $\mathbf{P}\colon S \times S \to [0,1]$ where we require $\sum_{s' \in S} \mathbf{P}(s, s') \in \{0, 1\}$ for all $s \in S$, and an initial state distribution $\iota\colon S \to [0,1]$ satisfying $\sum_{s \in S} \iota(s) = 1$. We say that a state s is terminal if $\sum_{s' \in S} \mathbf{P}(s, s') = 0$. A path π in $\mathcal{M}$ is a state sequence $s_0 s_1 \dots$ such that $\iota(s_0) > 0$ and $\mathbf{P}(s_i, s_{i+1}) > 0$ for all appropriate i, and π is either infinite or ends in a terminal state. The σ-algebra of the probability space over sets of paths of $\mathcal{M}$ is generated by cylinder sets $\mathrm{Cyl}(\hat{\pi})$ comprising all path extensions of path prefixes $\hat{\pi}$. The probability measure $\Pr_{\mathcal{M}}$ on paths of $\mathcal{M}$ is induced by $\Pr_{\mathcal{M}}\big(\mathrm{Cyl}(s_0 \dots s_n)\big) = \iota(s_0) \cdot \mathbf{P}(s_0, s_1) \cdot \ldots \cdot \mathbf{P}(s_{n-1}, s_n)$ [6, Chapter 10]. We write $\Pr_s$ for the probability measure that arises for $\mathcal{M}$ with $\iota(s) = 1$.

A formula in *linear temporal logic* (LTL) over a set AP of atomic propositions is formed according to the following grammar

$$\varphi ::= \mathit{true} \mid a \mid \varphi \wedge \varphi \mid \neg\varphi \mid \bigcirc\varphi \mid \varphi \,\mathsf{U}\, \varphi$$

where $a \in \mathrm{AP}$. In this paper, we consider LTL over sets of states as atomic propositions with the intended meaning that the atomic proposition A holds in a state s iff $s \in A$. We use the standard syntactic derivations $\varphi \vee \psi \equiv \neg(\neg\varphi \wedge \neg\psi)$, $\varphi \to \psi \equiv \neg\varphi \vee \psi$, $\Diamond\varphi \equiv \mathit{true} \,\mathsf{U}\, \varphi$ ("eventually"), $\Box\varphi \equiv \neg\Diamond\neg\varphi$ ("always"), $\varphi \,\mathsf{W}\, \psi \equiv \Box\varphi \vee \varphi \,\mathsf{U}\, \psi$ ("weak until") and $\varphi \,\mathsf{R}\, \psi \equiv \neg(\neg\varphi \,\mathsf{U}\, \neg\psi)$ ("release").

The semantics of LTL over sequences of atomic proposition sets is defined the standard way (see, e.g., [6]). For example, a path $\pi = s_0 s_1 \ldots$ satisfies $\varphi \mathsf{R} \psi$, denoted by $\pi \models \varphi \mathsf{R} \psi$, iff there is a position $k \in \mathbb{N}$ such that $s_i s_{i+1} \ldots$ satisfies ψ for $i \leq k$ and $s_k s_{k+1} \ldots$ satisfies φ. A transition system $\mathcal{T}$ is said to satisfy an LTL formula φ, denoted by $\mathcal{T} \models \varphi$, if all paths π of $\mathcal{T}$ satisfy φ. We write $\mathcal{T}, s \models \varphi$ in case $\mathcal{T}$ satisfies φ under the assumption that $I = \{s\}$ is the only initial state of $\mathcal{T}$. The set of states satisfying a formula φ in $\mathcal{T}$ is denoted by $\mathrm{Sat}_{\mathcal{T}}(\varphi) = \{s \in S \mid \mathcal{T}, s \models \varphi\}$, or simply $\mathrm{Sat}(\varphi)$ if $\mathcal{T}$ is clear from the context.

3 Necessary and Sufficient Causes

In this section, we define two notions of causes in transition systems, namely necessary and sufficient causes. Both notions lead to a binary relation on events, stating that an event is a cause for an effect event. Here, we focus on reachability events as causes and effects, such that they can be represented by sets of states. Our focus is motivated by the fact that numerous properties can be expressed by reachability properties on transition systems obtained by well-known automata-theoretic transformations [43, 70, 22, 24].

3.1 Necessary Causes

Informally spoken, an event C is considered to be a *necessary cause* of an event E whenever the presence of E necessarily implies the prior occurrence of C. The presence of C, on the other hand, does not necessarily imply that E will occur. This idea can be expressed formally using LTL formulas over state sets:

Definition 1 (Necessary cause). *Let $\mathcal{T} = (S, R, I)$ be a transition system and let $C, E \subseteq S$ be sets of states. We say that C is a* necessary cause *for E, denoted by $C \prec_{\mathsf{nec}} E$, if E is non-empty and*

$$\mathcal{T} \models C \,\mathsf{R}\, \neg E \qquad (\equiv \ \Box \neg E \vee \big(\neg E \,\mathsf{U}\, (\neg E \wedge C)\big)).$$

The formula $C \,\mathsf{R}\, \neg E$ is fulfilled whenever E is not reached *before* reaching C. In particular, there needs to be at least one transition between reaching C and E.

Note that if the set E consists only of terminal states, i.e., states without any outgoing transitions, and C and E are disjoint, then C is a necessary cause of E iff $\mathcal{T} \models \Diamond E \rightarrow \Diamond C$. The set I of initial states is a trivial necessary cause for any effect $E \subseteq S$ if its intersection with the effect states E is empty. For any effect E not containing an initial state, it is thus clear that necessary causes always exist. Saying that the set of initial states is a necessary cause, however, does of course not carry much explanatory information.

Example 1. Consider the transition system $\mathcal{T}$ depicted in Figure 1. We are interested in necessary causes for the effect $E = \{e\}$. Any set containing the initial state s_0 is trivially a necessary cause. More interesting are necessary causes that do not contain s_0. There are two such causes containing two states: $C_1 = \{a_1, b\}$

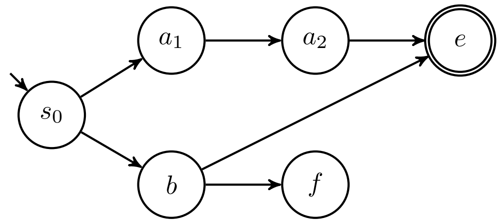

Fig. 1. The transition system $\mathcal{T}$ for Example 1.

and $C_2 = \{a_2, b\}$. C_1 occurs at least as early as C_2 on all paths. Nevertheless, $C_1 \prec_{\mathsf{nec}} C_2$ does not hold since we required causes to occur strictly before their effects: when entering state b the events to reach C_1 and C_2 occur simultaneously.

In order to compare simultaneously occurring causes for the same effect, we introduce a second type of 'necessary cause'-relation between sets of states that we call *necessary quasi-cause* in the following definition. For a quasi-cause, we do not require that to occur *strictly* before its effect.

Definition 2 (Necessary quasi-cause). *Let $\mathcal{T} = (S, R, I)$ be a transition system and let $C, E \subseteq S$ be sets of states. We say that C is a* necessary quasi-cause *for E, denoted by $C \preceq^q_{\mathsf{nec}} E$, if E is non-empty and*

$$\mathcal{T} \models \neg E \,\mathsf{W}\, C \qquad (\equiv\ \Box\neg E \vee (\neg E \,\mathsf{U}\, C)).$$

For a quasi-cause, we only require non-strict temporal priority. Hence, on any path reaching an effect E, it is sufficient if the quasi-cause C is reached simultaneously with E. Returning to Example 1, we therefore have $C_1 \preceq^q_{\mathsf{nec}} C_2$ even though $C_1 \prec_{\mathsf{nec}} C_2$ does not hold. As indicated by the name quasi-cause, we do not claim that this notion itself constitutes a meaningful cause–effect relationship. For example, any effect set provides a quasi-cause for itself. The notion is useful, however, when comparing different causes for the same effect.

We now establish first fundamental properties of the relations $\prec_{\mathsf{nec}}$ and $\preceq^q_{\mathsf{nec}}$:

Lemma 1. *Let $\mathcal{T} = (S, R, I)$ be a transition system. Then:*

(1) The relation $\prec_{\mathsf{nec}}$ is a strict partial order (irreflexive, asymmetric, and transitive) on the powerset of the state space S of $\mathcal{T}$.
(2) The relation $\preceq^q_{\mathsf{nec}}$ is a preorder (reflexive and transitive) on the powerset of the state space S of $\mathcal{T}$.
(3) For all $C, E \subseteq S$, we have that $C \prec_{\mathsf{nec}} E$ implies $C \preceq^q_{\mathsf{nec}} E$.
(4) For all $C_1, C_2, E \subseteq S$, if $C_1 \preceq^q_{\mathsf{nec}} C_2$ and $C_2 \prec_{\mathsf{nec}} E$, then $C_1 \prec_{\mathsf{nec}} E$.
(5) For all $C, E_1, E_2 \subseteq S$, if $C \prec_{\mathsf{nec}} E_1$ and $E_1 \preceq^q_{\mathsf{nec}} E_2$, then $C \prec_{\mathsf{nec}} E_2$.

Proof. Ad (1): $\prec_{\mathsf{nec}}$ is irreflexive since $C \,\mathsf{R}\, \neg C$ does not hold on paths that reach C (recall that $\neg C$ still has to hold when C releases the requirement of $\neg C$ to hold) and every state of $\mathcal{T}$ is assumed to be reachable in $\mathcal{T}$. Similarly asymmetry of $\prec_{\mathsf{nec}}$ is clear as any path π with $\pi \models C \,\mathsf{R}\, \neg E$ cannot satisfy $E \,\mathsf{R}\, \neg C$. For transitivity, assume that $A \prec_{\mathsf{nec}} B$ and $B \prec_{\mathsf{nec}} C$ for three sets A, B, and C of states of $\mathcal{T}$. To show $\mathcal{T} \models A \,\mathsf{R} \neg C$, let $\pi = s_0 s_1 s_2 \ldots$ be a path in $\mathcal{T}$. If $\pi \models \Box\neg C$, we have $\pi \models A \,\mathsf{R} \neg C$. So, suppose that $\pi \models \Diamond C$. Let s_i be the first state in π that is in C. As $\mathcal{T} \models B \,\mathsf{R}\, \neg C$, there is a position $j < i$ with $s_j \in B$. Analogously, there is $k < j$ such that $s_k \in A$. So, $\pi \models A \,\mathsf{R} \neg C$. We conclude that $\prec_{\mathsf{nec}}$ is transitive.

Ad (2): As for any set of states A the formula $\neg A \mathsf{W} A$ is a tautology, reflexivity of $\preceq^q_{\mathsf{nec}}$ is clear. Transitivity is shown analogously to the proof of transitivity above, where the strict inequalities on the positions i, j, and k are replaced by non-strict ones.

Ad (3): This is a direct consequence of $C \,\mathsf{R}\, \neg E \equiv \Box\neg E \vee \big(\neg E \,\mathsf{U}\, (\neg E \wedge C)\big)$ entailing $\neg E \,\mathsf{W}\, C \equiv \Box\neg E \vee (\neg E \,\mathsf{U}\, C)$.

Ad (4) and (5): The proofs are again analogous to the proof of transitivity above, where this time one of the strict inequalities between positions is replaced by a non-strict one. $\square$

These definitions and basic properties of the two relations will help us to find "good causes" later on. There is no gold standard what precisely constitutes a good necessary cause. One common approach also within other notions of causality is to only consider minimal representatives as causes [39, 26], i.e., events where removing some part leads to loosing the property of being a cause. In our setting, necessary causes may contain redundant states that do not affect the causal relationships to potential effect sets and could be removed towards more concise causes. To provide an intuition, consider again the transition system $\mathcal{T}$ depicted in Figure 1. The necessary cause $C_3 = \{a_1, a_2, b\}$ contains the redundant state a_2. This state can only be reached if the set C_3 is visited already before in state a_1. As only the first visit to the set is relevant in the relations $\prec_{\mathsf{nec}}$ and $\preceq^q_{\mathsf{nec}}$, the fact that a_2 belongs to C_3 does not play a role at all for causal relationships. To remove such redundant states, we define the following pruning of sets of states.

Definition 3 (Pruning of state sets). *Let $\mathcal{T} = (S, R, I)$ be a transition system and let $A \subseteq S$ be a set of states. We define the* pruning $\lfloor A \rfloor$ *of A by*

$$\lfloor A \rfloor = \{\ a \in A \ \mid\ \textit{there is a path } \pi \textit{ in } \mathcal{T} \textit{ with } \pi \models \neg A \,\mathsf{U}\, a\ \}.$$

Recall that paths always start in an initial state of the transition system. The pruning $\lfloor A \rfloor$ includes precisely those states in A that are reachable without previously seeing A. It satisfies the following properties related to the necessary (quasi-)cause relations defined above.

Lemma 2. *Let $\mathcal{T} = (S, R, I)$ be a transition system.*

(1) For all $A \subseteq S$, we have $A \preceq^q_{\mathsf{nec}} \lfloor A \rfloor$ and $\lfloor A \rfloor \preceq^q_{\mathsf{nec}} A$.
(2) For $A, B \subseteq S$, we have that $A \preceq^q_{\mathsf{nec}} B$ and $B \preceq^q_{\mathsf{nec}} A$ implies $\lfloor A \rfloor = \lfloor B \rfloor$.
(3) For all $C, E \subseteq S$ with $C \prec_{\mathsf{nec}} E$, we have $C \prec_{\mathsf{nec}} \lfloor E \rfloor$.

Proof. Ad (1): First, we will show $\mathcal{T} \models (\neg A) \mathsf{W} \lfloor A \rfloor$ which is equivalent to $\mathcal{T} \models \Diamond A \to (\neg A) \mathsf{U} \lfloor A \rfloor$. Let $\pi = s_0 s_1 \ldots$ be a path in $\mathcal{T}$ that satisfies $\Diamond A$ and let $i \in \mathbb{N}$ be the first position such that $s_i \in A$. Then, by definition of $\lfloor A \rfloor$, we have $s_i \in \lfloor A \rfloor$. This shows $\pi \models (\neg A) \mathsf{U} \lfloor A \rfloor$. In the other direction, $\mathcal{T} \models (\neg \lfloor A \rfloor) \mathsf{W} A$ holds because $\lfloor A \rfloor \subseteq A$.

Ad (2): Assume $A \preceq^q_{\mathsf{nec}} B$ and $B \preceq^q_{\mathsf{nec}} A$ and suppose towards a contradiction that $\lfloor A \rfloor \neq \lfloor B \rfloor$. Assume w.l.o.g. that there is an $a \in \lfloor A \rfloor \setminus \lfloor B \rfloor$. By the definition of $\lfloor A \rfloor$, there is a path $\pi = s_0\, s_1 \ldots s_n \ldots$ with $s_0 \in I$ and $s_n = a$ such that $s_i \notin A$ for all $i < n$. As $A \preceq^q_{\mathsf{nec}} B$, it follows that also $s_i \notin B$ for all $i < n$. Since $a \notin \lfloor B \rfloor$, also $a \notin B$ since otherwise the path π would witness that a also belongs to $\lfloor B \rfloor$. Thus, $\pi \not\models (\neg A) \mathsf{W} B$ and hence $B \not\preceq^q_{\mathsf{nec}} A$, which yields a contradiction.

Ad (3): The claim follows from $E \preceq^q_{\mathsf{nec}} \lfloor E \rfloor$ by (1) and Lemma 1(5). □

The preorder $\preceq^q_{\mathsf{nec}}$ induces an equivalence relation defined by

$$A \sim B \qquad \text{iff} \qquad A \preceq^q_{\mathsf{nec}} B \;\text{ and }\; B \preceq^q_{\mathsf{nec}} A.$$

Statements (1) and (2) of Lemma 2 tell us that in each of these equivalence classes, there is exactly one pruned set. Choosing the respective pruned set as representative for each equivalence class, we obtain that $\preceq^q_{\mathsf{nec}}$ is a partial order (reflexive, transitive, and anti-symmetric) on the set of pruned subsets of S. In the light of Lemma 1 and Lemma 2(3), we can conclude that for sets $C_1, C_2, E_1, E_2 \subseteq S$ with $C_1 \sim C_2$ and $E_1 \sim E_2$, we have

$$C_1 \prec_{\mathsf{nec}} E_1 \qquad \text{iff} \qquad C_2 \prec_{\mathsf{nec}} E_2.$$

In words, $\prec_{\mathsf{nec}}$ is well-defined on the equivalence classes induced by $\preceq^q_{\mathsf{nec}}$ and it is therefore reasonable to restrict ourselves to the canonical representatives for necessary causes in terms of pruned sets.

3.2 Sufficient Causes

Intuitively, a sufficient cause C for an event E means that the presence of C necessarily implies the subsequent occurrence of E. This intuition can be formalized using LTL formulas over state sets:

Definition 4 (Sufficient cause). *Let $\mathcal{T} = (S, R, I)$ be a transition system. A non-empty set $C \subseteq S$ is a* sufficient cause *for $E \subseteq S$ if*

$$\mathcal{T} \models \Box\big(C \to \bigcirc\Diamond E\big).$$

The formula basically states that whenever we see a state $c \in C$ we will also see E at some point in the future. Note that if E comprises terminal states only and C and E are disjoint, the above characterization of sufficient causes is equivalent to $\Diamond C \to \Diamond E$.

Example 2. Consider the transition system depicted in Figure 2, modeling a coffee machine that has a defect and sometimes only produces hot water instead

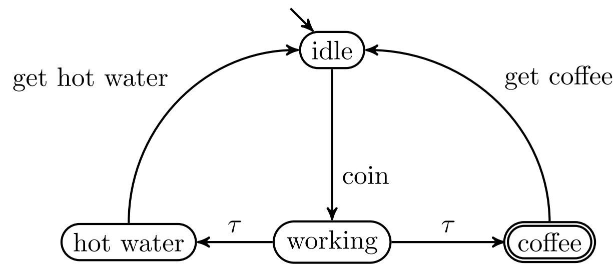

Fig. 2. A defect coffee machine that sometimes produces hot water

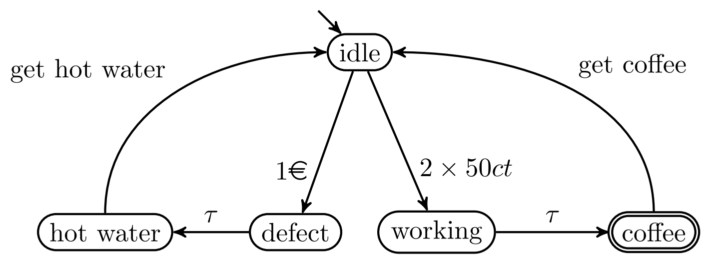

Fig. 3. A refined transition system for the defect coffee machine

of delicious coffee. We consider the effect $E = \{\text{coffee}\}$. Within this model, there are no sufficient causes for E since it is unclear how the non-deterministic choice in the working state is resolved. However, both $C_1 = \{\text{working}\}$ and $C_2 = \{\text{idle}\}$ are necessary causes for E.

Suppose now that we have additional knowledge about the defect and can refine the transition system model towards the one in Figure 3. Here, we assume that a person with strong desire of getting more insights about the defect, let us call him Frits, figured out a trick: when using two coins instead of one, the defect does not occur and the machine always delivers coffee. For the effect $E = \{\text{coffee}\}$, we then still have $C_1 = \{\text{working}\}$ and $C_2 = \{\text{idle}\}$ as necessary causes. In this model, C_1 is additionally a sufficient cause since all paths that visit C_1 will visit E afterwards eventually.

In analogy to necessary causes, we can observe sufficient causes are transitive:

Lemma 3 (Transitivity of sufficient causes.). *Let $\mathcal{T} = (S, R, I)$ be a transition system and $C, W, E \subseteq S$. If C is a sufficient cause for W and W is a sufficient cause for E, then C is a sufficient cause for E.*

Proof. Assume that C is a sufficient cause for W, which in turn is a sufficient cause for E. Let $\pi = s_0\, s_1 \ldots$ be a path in $\mathcal{T}$. Then, since $\mathcal{T} \models \Box(C \to \bigcirc\Diamond W)$ we have for each $i \in \mathbb{N}$ with $s_i \in C$ that there is $j > i$ with $s_j \in W$. Likewise, by $\mathcal{T} \models \Box(W \to \bigcirc\Diamond E)$ for each $j \in \mathbb{N}$ with $s_j \in W$ there is $k > j$ with $s_k \in E$. We conclude that $\pi \models \Box(C \to \bigcirc\Diamond E)$ for all paths π in $\mathcal{T}$. Hence, C is a sufficient cause for E. $\square$

Since all states in a transition system $\mathcal{T}$ are assumed to be reachable, a set C is a sufficient cause for E in $\mathcal{T}$ iff the sufficiency condition holds for all states included in C. That is, for all $a \in C$ the formula $\Box(a \to \bigcirc\Diamond E)$ holds in $\mathcal{T}$. Equivalently, a set $C \subseteq S$ is a sufficient cause for E iff $\varnothing \neq C \subseteq \mathrm{Sat}_{\mathcal{T}}\big(\bigcirc\Diamond E\big)$. Therefore, existence of a sufficient cause can be checked in polynomial time with standard model-checking algorithms [20, 6].

The satisfaction set of $\bigcirc\Diamond E$ is consequently the inclusion-maximal sufficient cause for the effect E in $\mathcal{T}$. This set, however, might be very large and does not necessarily point to "good causes" for the effect. To this end, we define the *canonical sufficient* cause as $C_c^E = \lfloor \mathrm{Sat}_{\mathcal{T}}\big(\bigcirc\Diamond E\big) \rfloor$, i.e., the set of all states in $\mathrm{Sat}_{\mathcal{T}}\big(\bigcirc\Diamond E\big)$ that are either initial or are reachable from some initial state by visiting only non-sufficient states. The name "canonical" sufficient cause is justified by the following observation:

Proposition 1. *Let $\mathcal{T}$ and E be as above. The canonical sufficient cause C_c^E is the unique pruned and $\preceq_{\mathsf{nec}}^q$-least sufficient cause for E.*

Proof. By Lemma 1(1), we know that $C_c^E \preceq_{\mathsf{nec}}^q \mathrm{Sat}_{\mathcal{T}}\big(\bigcirc\Diamond E\big)$, and therefore also $C_c^E \preceq_{\mathsf{nec}}^q C$ for all $C \subseteq \mathrm{Sat}_{\mathcal{T}}\big(\bigcirc\Diamond E\big)$ by the definition of $\preceq_{\mathsf{nec}}^q$. But the sufficient causes of E are exactly the non-empty subsets of $\mathrm{Sat}_{\mathcal{T}}\big(\bigcirc\Diamond E\big)$. Thus, the canonical sufficient cause C_c^E is a $\preceq_{\mathsf{nec}}^q$-least sufficient cause. By Lemma 2, we know that there is only one pruned cause in the equivalence class (induced by $\preceq_{\mathsf{nec}}^q$) of $\preceq_{\mathsf{nec}}^q$-least sufficient causes, rendering C_c^E unique. $\square$

While Proposition 1 already shows that C_c^E is a distinguished sufficient cause, we will see later on that it is also optimal with respect to other criteria, namely the degree of necessity introduced in the next section.

4 Finding Good Causes

We have seen that causes may differ in their information they provide and their ability to concisely explain reasons for the effect. For example, the set of initial states is a necessary cause for any effect E that does not contain an initial state. In this section, we introduce different ways to quantify the quality of a cause and show how to find optimal causes with respect to the introduced quality measures.

4.1 Degrees of Sufficiency and Necessity

The notions of sufficiency and necessity defined in the previous section are qualitative: either a set satisfies the corresponding criterion, or it does not. However, one can think of situations where a set C is *almost* sufficient or necessary, e.g., that a very large part of the executions which see the effect set E are preceded by C in the case of necessity. To quantify how close a set is a necessary cause for an effect (resp. a sufficient cause), we define degrees of necessity (resp. degree of sufficiency). Here, we rely on the probability measure on paths that we obtain by equipping the outgoing transitions from each state with a uniform probability distribution. In particular, we are interested in the trade-off between sufficiency and necessity and aim toward sufficient causes with a high degree of necessity, and vice versa.

For the remainder of this section, let us fix a transition system $\mathcal{T} = (S, R, I)$ and a non-empty effect set $E \subseteq S$. Then we can construct the Markov chain $\mathcal{M}_{\mathcal{T},E} = (S, \mathbf{P}, \iota)$ as follows. For each transition $(s, s') \in R$ with $s \notin E$, we have $\mathbf{P}(s, s') = 1/|\mathrm{Post}(s)|$, where $\mathrm{Post}(s)$ denotes the set of direct successors of s. For all $s, s' \in S$ where $(s, s') \notin R$ or $s \in E$, we set $\mathbf{P}(s, s') = 0$, i.e., all effect states are terminal in $\mathcal{M}_{\mathcal{T},E}$. Further, we set $\iota(s) = 1/|I|$ for all $s \in I$. In the following, we denote by Pr the probability measure $\Pr_{\mathcal{M}_{\mathcal{T},E}}$ on measurable sets of paths of $\mathcal{M}_{\mathcal{T},E}$.

The degree of sufficiency of a non-empty candidate cause $C \subseteq S \setminus E$ intuitively provides a measure how many of the paths that see C will also see E. It is defined as a conditional probability in the following way:

$$\text{suff-deg}(C, E) \;=\; \Pr(\lozenge E \mid \lozenge C) \;=\; \frac{\Pr(\lozenge E \wedge \lozenge C)}{\Pr(\lozenge C)}$$

With a similar reasoning, the degree of necessity of C is defined as:

$$\text{nec-deg}(C, E) \;=\; \Pr(\lozenge C \mid \lozenge E) \;=\; \frac{\Pr(\lozenge E \wedge \lozenge C)}{\Pr(\lozenge E)}$$

Note that these degrees can be computed in polynomial time by standard techniques for computing conditional probabilities on Markov chains [7].

If C is a sufficient cause as defined above, then its degree of sufficiency clearly is 1. The analogous statement holds for necessary causes, but the reverse directions do not hold in general. Since multiple sufficient causes may exist, it makes sense to look for those with maximal degree of necessity. In case C is a sufficient cause, the above expression for nec-deg(C, E) simplifies to

$$\text{nec-deg}(C, E) \;=\; \frac{\Pr(\lozenge C)}{\Pr(\lozenge E)} \tag{$*$}$$

as the formula $\lozenge C \to \lozenge E$ holds in $\mathcal{M}_{\mathcal{T},E}$ with E comprising terminal states only by construction. Analogously, if C is a necessary cause we have

$$\text{suff-deg}(C, E) \;=\; \frac{\Pr(\lozenge E)}{\Pr(\lozenge C)} \tag{$**$}$$

The above definitions raise the question of how to find a sufficient cause with maximal degree of necessity, or, a necessary cause with maximal degree of sufficiency. Observe that causes that are sufficient *and* necessary may not exist in general. The following lemma connects the degrees of necessity and sufficiency to the necessary quasi-cause relation $\preceq^{q}_{\mathsf{nec}}$.

Lemma 4. *Let $C_1, C_2 \subseteq S$ be two necessary causes for E, i.e., $C_1 \prec_{\mathsf{nec}} E$ and $C_2 \prec_{\mathsf{nec}} E$. Then, $C_1 \preceq^q_{\mathsf{nec}} C_2$ implies that* $\text{suff-deg}(C_1, E) \leq \text{suff-deg}(C_2, E)$.

Let $D_1, D_2 \subseteq S$ be two sufficient causes for E. Then, $D_1 \preceq^q_{\mathsf{nec}} D_2$ implies that $\text{nec-deg}(D_1, E) \geq \text{nec-deg}(D_2, E)$.

Proof. For any sets $A_1, A_2 \subseteq S$ with $A_1 \preceq^q_{\mathsf{nec}} A_2$, we have that $\mathcal{T} \models \Diamond A_2 \to \Diamond A_1$. So, $\Pr(\Diamond A_1) \geq \Pr(\Diamond A_2)$. Applied to the necessary causes C_1 and C_2, we conclude the claim due to equation $(**)$. For the sufficient causes D_1 and D_2, the claim follows analogously using equation $(*)$. $\square$

Sufficient causes with maximal degree of necessity. The story of how to find a sufficient cause with maximal degree of necessity is quickly told: By Lemma 4, we know that $\preceq^q_{\mathsf{nec}}$-least sufficient causes have maximal degree of necessity. In Proposition 1, we have seen that the canonical sufficient cause $C^E_c = \lfloor \text{Sat}_{\mathcal{T}}(\bigcirc \Diamond E) \rfloor$, is a $\preceq^q_{\mathsf{nec}}$-least sufficient cause. We conclude:

Proposition 2. *Let $\mathcal{T} = (S, R, I)$ a transition system and $E \subseteq S$. The canonical sufficient cause C^E_c has maximal degree of necessity among all sufficient causes for E.*

Necessary causes with maximal degree of sufficiency. While sufficient causes are always non-empty subsets of an LTL satisfaction set, this is not the case for necessary causes. Indeed, the set of all states S is always a necessary cause for any effect that is disjoint from the initial states but not all state sets have to be a necessary cause. Following the definition of a canonical sufficient cause suggests considering the pruned maximal necessary cause as a candidate. However, in the case above, $I = \lfloor S \rfloor$, which does not attain the maximal degree of sufficiency among all necessary causes (on the contrary, it achieves the minimal degree of sufficiency). A necessary cause with maximal degree of sufficiency is the *direct-predecessor cause*: It is denoted by $C^E_{\text{dp}} = \{s \in S \mid \text{there is } e \in E \text{ such that } (s, e) \in R\}$ and comprises all those states that have at least one transition to E.

Proposition 3. *Let $\mathcal{T} = (S, R, I)$ a transition system and $E \subseteq S \setminus I$. The direct-predecessor cause C^E_{dp} is a necessary cause that achieves the maximal degree of sufficiency among all necessary causes for E.*

Proof. Clearly, C^E_{dp} is a necessary cause by definition, since for all paths π in $\mathcal{T}$ that visit E we clearly have $\pi \models \neg E \mathbin{\mathsf{U}} (\neg E \wedge C^E_{\text{dp}})$ (recall that $E \cap I = \varnothing$). We show that $\Pr(\Diamond C^E_{\text{dp}}) \leq \Pr(\Diamond C)$ for every necessary cause $C \subseteq S$ by proving

$$\{\pi \mid \pi \models \Diamond C^E_{\text{dp}}\} \quad \subseteq \quad \{\pi \mid \pi \models \Diamond C\}.$$

Let $\pi = s_0 s_1 \ldots$ be a path in $\mathcal{T}$ with $\pi \models \Diamond C^E_{\text{dp}}$ and let $i \in \mathbb{N}$ be the smallest position such that $s_i \in C^E_{\text{dp}}$. Then clearly $s_j \notin E$ for all $j \leq i$ and there is a path $\pi' = s_0 s_1 \ldots s_i s'_{i+1} \ldots$ where $s'_{i+1} \in E$ and thus, $\pi' \models \neg E \mathbin{\mathsf{U}} (\neg E \wedge C)$ as C is a necessary cause. But then there is $k \leq i$ with $s_k \in C$ and thus $\pi \models \Diamond C$. $\square$

The motivation for pruned necessary causes is also applicable to direct-predecessor causes, asking for the "earliest" necessary cause C that has the same degree of sufficiency as C^E_{dp}. To this end, we consider the set of states $C^E_{\Diamond\,\mathrm{dp}} = \{s \in S \mid \Pr_s(\Diamond C^E_{\mathrm{dp}}) = 1\}$, which is a necessary cause due to $C^E_{\mathrm{dp}} \subseteq C^E_{\Diamond\,\mathrm{dp}}$. By Lemma 1, its pruned set is also a necessary cause, i.e. $\lfloor C^E_{\Diamond\,\mathrm{dp}} \rfloor \prec_{\mathsf{nec}} E$. In $\mathcal{M}_{\mathcal{T},E}$ we have $\Pr(\Diamond \lfloor C^E_{\Diamond\,\mathrm{dp}} \rfloor) \leq \Pr(\Diamond C^E_{\Diamond\,\mathrm{dp}}) = \Pr(\Diamond C^E_{\mathrm{dp}})$. On the other hand, $\lfloor C^E_{\Diamond\,\mathrm{dp}} \rfloor \preceq^q_{\mathsf{nec}} C^E_{\Diamond\,\mathrm{dp}}$ (again by Lemma 1) implies $\Pr(\Diamond C^E_{\mathrm{dp}}) \leq \Pr(\Diamond \lfloor C^E_{\Diamond\,\mathrm{dp}} \rfloor)$. Therefore, we have $\Pr(\Diamond C^E_{\mathrm{dp}}) = \Pr(\Diamond \lfloor C^E_{\Diamond\,\mathrm{dp}} \rfloor)$ and hence the degrees of sufficiency of C^E_{dp} and $\lfloor C^E_{\Diamond\,\mathrm{dp}} \rfloor$ are the same.

Moreover, $\lfloor C^E_{\Diamond\,\mathrm{dp}} \rfloor$ is a necessary quasi-cause for all necessary causes of E that achieve the same degree of sufficiency:

Proposition 4. *Let $\mathcal{T} = (S, R, I)$ a transition system and $E \subseteq S \setminus I$. For all necessary causes C of E that satisfy* $\text{suff-deg}(C, E) = \text{suff-deg}(C^E_{\mathrm{dp}}, E)$ *we have $\lfloor C^E_{\Diamond\,\mathrm{dp}} \rfloor \preceq^q_{\mathsf{nec}} C$.*

Proof. It suffices to show $C \subseteq C^E_{\Diamond\,\mathrm{dp}}$. Since C is a necessary cause for E, we can apply the same argumentation as in the proof of Proposition 3, showing that

$$\{\pi \mid \pi \models \Diamond(C \wedge \Diamond C^E_{\mathrm{dp}})\} \quad = \quad \{\pi \mid \pi \models \Diamond C^E_{\mathrm{dp}}\} \quad \subseteq \quad \{\pi \mid \pi \models \Diamond C\}.$$

Due to $\text{suff-deg}(C, E) = \text{suff-deg}(C^E_{\mathrm{dp}}, E)$, we have $\Pr(\Diamond C) = \Pr(\Diamond C^E_{\mathrm{dp}})$ and thus, $\{\pi \mid \pi \models \Diamond(C \wedge \Diamond C^E_{\mathrm{dp}})\} = \{\pi \mid \pi \models \Diamond C\}$. Now fix some arbitrary $s \in C$. Then, every path that visits s has to visit C^E_{dp} eventually afterwards. Thus, $\Pr_s(\Diamond C^E_{\mathrm{dp}}) = 1$, which is equivalent to $s \in C^E_{\Diamond\,\mathrm{dp}}$, leading to $C \subseteq C^E_{\Diamond\,\mathrm{dp}}$. □

This leads to a $\preceq^q_{\mathsf{nec}}$-least necessary cause $\lfloor C^E_{\Diamond\,\mathrm{dp}} \rfloor$ with maximal degree of sufficiency that can be computed in polynomial time by standard methods.

4.2 Weight-minimal Necessary Causes

The previous section showed how to determine necessary causes with maximal degree of sufficiency and $\preceq^q_{\mathsf{nec}}$-least ones among them. We now describe a different technique to find optimal necessary causes with respect to a generic optimization criterion, employing a natural connection to minimal cuts from flow networks.

Let $\mathcal{T} = (S, R, I)$ be a transition system and $A, B \subseteq S$ be two sets of states. We call $X \subseteq S \setminus B$ an *AB-separator* if every finite path through $\mathcal{T}$ that starts in A and ends in B sees a vertex in X. The following observation follows directly from the definition of necessary causes.

Proposition 5. *The necessary causes for E which do not intersect E in $\mathcal{T}$ are exactly the IE-separators of $\mathcal{T}$.*

Let us augment $\mathcal{T}$ by a weight function $w\colon S \to \mathbb{Q}_{\geq 0}$. The weight of a set $X \subseteq S$ is defined to be $w(X) = \sum_{v \in X} w(v)$. In the presence of such a weight function, it makes sense to ask for weight-minimal AB-separators in $\mathcal{T}$, for some given

$A, B \subseteq S$. Via a polynomial reduction to the problem of computing minimal cuts, we get the following result.[1]

Proposition 6. *Weight-minimal AB-separators can be computed in polynomial time.*

Proof. We reduce the problem of computing minimal AB-separators to the problem of computing a weight-minimal s-t-cut. An s-t-cut of $\mathcal{T} = (S, R, I)$ is a partition S_1, S_2 of S such that $s \in S_1$, $t \in S_2$. Let $w\colon R \to \mathbb{Q}$ be a weight function on the edges of $\mathcal{T}$. The *bridging edges* of an s-t-cut S_1, S_2 are defined to be $\mathrm{br}(S_1, S_2) = R \cap (S_1 \times S_2)$, and its weight is $\sum_{(u,v)\in\mathrm{br}(S_1,S_2)} w(u, v)$. Weight-minimal cuts can be computed in polynomial time [62].

We show how to reduce the problem of computing weight-minimal AB-separators to the problem of computing weight-minimal cuts. Let $\mathcal{T} = (S, R, I)$, w, $A, B \subseteq S$ be an instance of the weight-minimal AB-separator problem. We may assume that $A \cap B = \varnothing$ and that B is a singleton set $\{b\}$. If $A \cap B \neq \varnothing$, then there are no AB-separators by definition. If B is not singleton, we can first collapse all states in B into a single state b, and let $B = \{b\}$. This transformation preserves AB-separators and their weights.

Now we transform the transition system $\mathcal{T}$ as follows. Define $\mathcal{T}' = (S \cup S' \cup \{a\}, R', I)$, where $S' = \{s' \mid s \in S\}$, and with edges

$$v \to v' \qquad \text{for all } v \in S \tag{1}$$

$$v' \to u \qquad \text{for all } (v, u) \in R \tag{2}$$

$$a \to v \qquad \text{for all } v \in A \tag{3}$$

Consider the weight function $w'\colon R' \to \mathbb{Q}_{\geq 0}$ defined by $w'(v, v') = w(v)$ for all $v \in S \setminus B$ and $w'(x, y) = w(S) + 1$ for all other edges (x, y) of $\mathcal{T}'$. Note that these transformations are all possible in polynomial time.

Each AB-separator X in $\mathcal{T}$ induces an a-b-cut in $\mathcal{T}'$ as follows. Take S_1 to be the union of X and the states of $\mathcal{T}'$ reachable from a without seeing X. As X is an AB-separator in $\mathcal{T}$, the partition $(S_1, (S \cup S') \setminus S_1)$ forms an a-b-cut in $\mathcal{T}'$. Furthermore, as the outgoing edges of S_1 are exactly $\{(u, u') \mid u \in X\}$, the weight of this cut is $w(X)$.

Conversely, every a-b-cut (S_1, S_2) in $\mathcal{T}'$ satisfying $\mathrm{br}(S_1, S_2) \subseteq \{(u, u') \mid u \in S \setminus B\}$ induces the AB-separator $X = \{u \in S \mid (u, u') \in \mathrm{br}(S_1, S_2)\}$ with the same weight. Finally, any a-b-cut (S_1, S_2) in $\mathcal{T}'$ which does not satisfy $\mathrm{br}(S_1, S_2) \subseteq \{(u, u') \mid u \in S \setminus B\}$ cannot be weight-minimal, as it has larger weight than any cut with this property. The a-b-cut induced by the set $A \cup \{a\}$ has this property, and hence such an a-b-cut exists (this uses our assumption $A \cap B = \varnothing$). Hence, a weight-minimal a-b-cut in $\mathcal{T}'$ induces a weight-minimal AB-separator in $\mathcal{T}$. ◻

[1] The problem of finding *balanced vertex separators*, as studied by Feige et al. [28, 27], is NP-complete and differs from the one we study in that it requires that the vertex separator partitions the graph into approximately equally sized components.

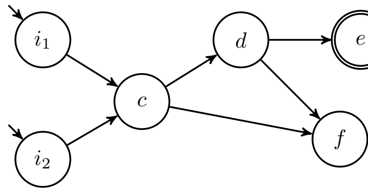

Fig. 4. Transition system $\mathcal{T}_1$ from Example 3

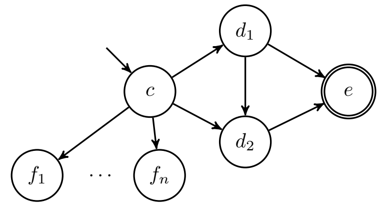

Fig. 5. Transition system $\mathcal{T}_2$ from Example 4

This gives us a tool to compute weight-optimal necessary causes in polynomial time. In the following, we consider two natural choices for weight functions which lead to different notions of optimality for necessary causes.

State-minimal necessary causes. Let $\mathcal{T} = (S, R, I)$ be a transition system as above and $E \subseteq S$ a set of states. Consider the weight function $w \colon S \to \mathbb{Q}_{\geq 0}$ where $w(s) = 1$ for all $s \in S$. Then, a weight-minimal necessary cause with respect to w is a necessary cause C such that $|C|$ is minimal among all necessary causes. By the above observations, such a cause can be computed in polynomial time.

If $|I| = 1$, then I itself is always a state-minimal necessary cause, which renders the optimization problem trivial. However, I has the worst possible degree of sufficiency among all necessary causes due to $\Pr(\lozenge I) = 1$ in the corresponding Markov chain $\mathcal{M}_{\mathcal{T},E}$. The following paragraph considers a weight function that aims to achieve a trade-off between the size of a necessary cause and its degree of sufficiency.

A trade-off between size and degree of sufficiency. Consider the weight function w defined by $w(v) = \Pr(\lozenge v)$, again with respect the probability measure in the Markov chain $\mathcal{M}_{\mathcal{T},E}$. A weight-minimal necessary cause with respect to this weight function is a necessary cause X minimizing

$$w(X) = \sum_{v \in X} \Pr(\lozenge v). \tag{†}$$

Recall that the degree of sufficiency of X is given by $\text{suff-deg}(X, E) = \frac{\Pr(\lozenge E)}{\Pr(\lozenge X)}$ if X is a necessary cause. We have $w(X) \geq \Pr(\lozenge X)$, and therefore minimizing the weight encourages necessary causes with high degree of sufficiency. At the same time, the number of states corresponds to the number of summands in $w(X)$, and hence few states are also encouraged.

Example 3. Consider the transition system $\mathcal{T}_1$ from Figure 4, where $I = \{i_1, i_2\}$ are the initial states and consider $E = \{e\}$ as the effect. Then both $C = \{c\}$ and $D = \{d\}$ are necessary causes of minimal size $|C| = |D| = 1$, as removing these states would separate I and E. This means that under the state-counting weight function, both C and D are weight-minimal necessary causes. However, D has a higher degree of sufficiency and would thus be optimal for both size and degree of sufficiency. Specifically, we have $\Pr(\Diamond C) = 1$ and $\Pr(\Diamond D) = 1/2$. Hence, $\text{suff-deg}(C, E) = \Pr(\Diamond E) = 1/4$, and $\text{suff-deg}(D, E) = 1/2$. Under the weight function defined in (†), only D is optimal, since $w(C) = \Pr(\Diamond C) = 1$ and $w(D) = \Pr(\Diamond D) = 1/2$.

Example 4. Consider the transition system $\mathcal{T}_2$ from Figure 5, with initial states $I = \{c\}$ and the effect $E = \{e\}$. Then I and $D = \{d_1, d_2\}$ are the only inclusion-minimal necessary causes. According to the weight function defined in (†) we have $w(I) = 1$ and $w(D) = \frac{5}{2(n+2)}$. For $n = 0$ we have $w(D) = 5/4 > 1 = w(I)$ and thus, I would be trade-off optimal. On the other hand, for $n > 0$ we have $w(D) < 1 = w(I)$, which turns D into the trade-off optimal necessary cause. Intuitively, increasing n makes I less sufficient, as it increases the set of paths that start in c but never reach E.

5 Conclusion

We have formalized well-known notions of necessity and sufficiency in the context of transition systems using temporal logic formulas over state sets that stand for causes and effects. Based on these formalizations, we addressed several trade-offs between necessity and sufficiency and presented three optimality criteria that differ in their properties with respect to conciseness and explainability: the degree of necessity, the degree of sufficiency, and through state weights. Causes that maximize the former two were explicitly characterized, and a polynomial-time algorithm for the computation of weight-optimal causes was described relying on known algorithms to determine minimal cuts in flow networks. Which notion of causality is appropriate to identify the reason for an effect, e.g., such that the imaginary person Frits from Example 2 can fix the broken coffee machine, highly depends on the considered system and it might be required to consider all our notions of causality to draw a conclusion.

In practice, also state sets with high degree of necessity *and* sufficiency might be interesting to consider also when they are neither sufficient nor necessary causes. In this direction it is promising to investigate trade-off values between the two degrees such as the f-score from statistics, as done for probability-raising causes in MDPs [5]. In future work we also plan to examine relaxations of the cause conditions studied here, following the more articulate INUS condition [59] or the NESS test [42, 72].

References

1. E. Ábrahám and B. Bonakdarpour. HyperPCTL: A temporal logic for probabilistic hyperproperties. In *Proc. of the 15th Intern. Conf. on Quantitative Evaluation of Systems (QEST)*, pages 20–35. Springer, 2018.

2. C. Baier, N. Coenen, B. Finkbeiner, F. Funke, S. Jantsch, and J. Siber. Causality-based game solving. In A. Silva and K. R. M. Leino, editors, *Computer Aided Verification*, pages 894–917, Cham, 2021. Springer International Publishing.
3. C. Baier, C. Dubslaff, F. Funke, S. Jantsch, R. Majumdar, J. Piribauer, and R. Ziemek. From Verification to Causality-Based Explications. In N. Bansal, E. Merelli, and J. Worrell, editors, *48th International Colloquium on Automata, Languages, and Programming (ICALP 2021)*, volume 198 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 1:1–1:20, Dagstuhl, Germany, 2021. Schloss Dagstuhl – Leibniz-Zentrum für Informatik.
4. C. Baier, F. Funke, S. Jantsch, J. Piribauer, and R. Ziemek. Probabilistic Causes in Markov Chains. In Z. Hou and V. Ganesh, editors, *Automated Technology for Verification and Analysis*, pages 205–221, Cham, 2021. Springer International Publishing.
5. C. Baier, F. Funke, J. Piribauer, and R. Ziemek. On probability-raising causality in Markov decision processes. In P. Bouyer and L. Schröder, editors, *Foundations of Software Science and Computation Structures*, pages 40–60, Cham, 2022. Springer International Publishing.
6. C. Baier and J.-P. Katoen. *Principles of Model Checking*. MIT Press, 2008.
7. C. Baier, J. Klein, S. Klüppelholz, and S. Märcker. Computing conditional probabilities in Markovian models efficiently. In *Proceedings of the 20th Conference on Tools and Algorithms for the Construction and Analysis of Systems (TACAS)*, volume LNCS:8413, pages 515–530, 2014.
8. T. Ball, M. Naik, and S. K. Rajamani. From symptom to cause: Localizing errors in counterexample traces. *SIGPLAN Not.*, 38(1):97–105, 2003.
9. V. Batusov and M. Soutchanski. Situation Calculus Semantics for Actual Causality. *Proceedings of the AAAI Conference on Artificial Intelligence*, 32(1), Apr. 2018.
10. A. Beer, S. Heidinger, U. Kühne, F. Leitner-Fischer, and S. Leue. Symbolic Causality Checking Using Bounded Model Checking. In *Model Checking Software*, pages 203–221. Springer, 2015.
11. I. Beer, S. Ben-David, H. Chockler, A. Orni, and R. J. Trefler. Explaining counterexamples using causality. *Formal Methods in System Design*, 40(1):20–40, 2012.
12. I. Beer, S. Ben-David, C. Eisner, and Y. Rodeh. Efficient Detection of Vacuity in ACTL Formulas. In *Proceedings of the 9th International Conference on Computer Aided Verification (CAV)*, pages 279–290, 1997.
13. G. Caltais, S. L. Guetlein, and S. Leue. Causality for General LTL-definable Properties. In *Proc. of the 3rd Workshop on formal reasoning about Causation, Responsibility, and Explanations in Science and Technology (CREST)*, pages 1–15, 2018.
14. H. Chockler and J. Y. Halpern. Responsibility and Blame: A Structural-Model Approach. *J. Artif. Int. Res.*, 22(1):93–115, Oct. 2004.
15. H. Chockler, J. Y. Halpern, and O. Kupferman. What causes a system to satisfy a specification? *ACM Transactions on Computational Logic*, 9(3), June 2008.
16. H. Chockler, O. Kupferman, R. P. Kurshan, and M. Y. Vardi. A Practical Approach to Coverage in Model Checking. In *Computer Aided Verification*, pages 66–78, 2001.
17. H. Chockler, O. Kupferman, and M. Vardi. Coverage Metrics for Formal Verification. *Intern. Journal on Software Tools for Technology Transfer (STTT)*, 8(4–5):373–386, 2006.
18. H. Chockler, O. Kupferman, and M. Y. Vardi. Coverage metrics for temporal logic model checking. In *Tools and Algorithms for the Construction and Analysis of Systems (TACAS)*, pages 528–542, 2001.

19. E. M. Clarke and E. A. Emerson. Design and synthesis of synchronization skeletons using branching time temporal logic. In D. Kozen, editor, *Logics of Programs*, pages 52–71, Berlin, Heidelberg, 1982. Springer Berlin Heidelberg.
20. E. M. Clarke, E. A. Emerson, and A. P. Sistla. Automatic verification of finite-state concurrent systems using temporal logic specifications. *ACM Trans. Program. Lang. Syst.*, 8:244–263, 1986.
21. E. M. Clarke, O. Grumberg, K. L. McMillan, and X. Zhao. Efficient Generation of Counterexamples and Witnesses in Symbolic Model Checking. In *Proc. of the 32nd Annual ACM/IEEE Design Automation Conf. (DAC)*, pages 427–432, New York, NY, USA, 1995. ACM.
22. C. Courcoubetis and M. Yannakakis. The Complexity of Probabilistic Verification. *Journal of the ACM*, 42(4):857–907, 1995.
23. A. Datta, D. Garg, D. Kaynar, D. Sharma, and A. Sinha. Program Actions as Actual Causes: A Building Block for Accountability. In *Proc. of the 28th IEEE Computer Security Foundations Symp. (CSF)*, pages 261–275, 2015.
24. L. de Alfaro. Temporal logics for the specification of performance and reliability. In *STACS 97*, pages 165–176, Berlin, Heidelberg, 1997. Springer.
25. R. Dimitrova, B. Finkbeiner, and H. Torfah. Probabilistic Hyperproperties of Markov Decision Processes. In *Proc. of the 18th Intern. Symp. on Automated Technology for Verification and Analysis (ATVA)*, volume 12302 of *LNCS*, pages 484–500. Springer, 2020.
26. C. Dubslaff, K. Weis, C. Baier, and S. Apel. Causality in configurable software systems. In *Proceedings of the 44th International Conference on Software Engineering (ICSE)*, 2022.
27. U. Feige, M. Hajiaghayi, and J. R. Lee. Improved Approximation Algorithms for Minimum Weight Vertex Separators. *SIAM J. Comput.*, 38(2):629–657, 2008.
28. U. Feige and M. Mahdian. Finding small balanced separators. In J. M. Kleinberg, editor, *Proceedings of the 38th Annual ACM Symposium on Theory of Computing, Seattle, WA, USA, May 21-23, 2006*, pages 375–384. ACM, 2006.
29. D. Galles and J. Pearl. Axioms of Causal Relevance. *Artif. Intell.*, 97(1–2):9–43, Dec. 1997.
30. D. Galles and J. Pearl. An Axiomatic Characterization of Causal Counterfactuals. *Foundations of Science*, 3:151–182, 1998.
31. G. Gomes. Are necessary and sufficient conditions converse relations? *Australasian Journal of Philosophy*, 87(3):375–387, 2009.
32. G. Gössler and D. Le Métayer. A General Trace-Based Framework of Logical Causality. In *Formal Aspects of Component Software*, pages 157–173. Springer, 2014.
33. G. Gössler and D. Le Métayer. A General Framework for Blaming in Component-Based Systems. *Science of Computer Programming*, 113:223–235, Dec. 2015.
34. G. Gössler and J.-B. Stefani. Causality Analysis and Fault Ascription in Component-Based Systems. *Theoretical Computer Science*, 837:158–180, Oct. 2020.
35. A. Groce. Error Explanation with Distance Metrics. In *Tools and Algorithms for the Construction and Analysis of Systems*, pages 108–122, Berlin, Heidelberg, 2004. Springer Berlin Heidelberg.
36. A. Groce, S. Chaki, D. Kroening, and O. Strichman. Error explanation with distance metrics. *International Journal on Software Tools for Technology Transfer*, 8(3):229–247, 2006.

37. A. Groce and W. Visser. What Went Wrong: Explaining Counterexamples. In T. Ball and S. K. Rajamani, editors, *Model Checking Software*, pages 121–136, Berlin, Heidelberg, 2003. Springer Berlin Heidelberg.
38. J. Y. Halpern. Axiomatizing Causal Reasoning. *Journal of Artif. Intell. Research*, 12(1):317–337, May 2000.
39. J. Y. Halpern. A Modification of the Halpern-Pearl Definition of Causality. In *Proc. of the 24th Intern. Joint Conf. on AI (IJCAI)*, pages 3022–3033. AAAI Press, 2015.
40. J. Y. Halpern and J. Pearl. Causes and Explanations: A Structural-Model Approach. Part I: Causes. *The British Journal for the Philosophy of Science*, 56(4):843–887, 2005.
41. J. Y. Halpern and J. Pearl. Causes and Explanations: A Structural-Model Approach. Part II: Explanations. *The British Journal for the Philosophy of Science*, 56(4):889–911, 2005.
42. H. L. A. Hart and A. M. Honoré. *Causation in the law.* Oxford University Press, 1959.
43. S. Hart, M. Sharir, and A. Pnueli. Termination of probabilistic concurrent program. *ACM Transactions on Programming Languages and Systems*, 5(3):356–380, 1983.
44. M. Hopkins and J. Pearl. Causality and Counterfactuals in the Situation Calculus. *Journal of Logic and Computation*, 17(5):939–953, Oct. 2007.
45. Y. Hoskote, T. Kam, P.-H. Ho, and X. Zhao. Coverage estimation for symbolic model checking. In *Proceedings of the 36th Annual ACM/IEEE Design Automation Conference (DAC)*, pages 300–305, 1999.
46. D. Hume. *A Treatise of Human Nature.* John Noon, 1739.
47. D. Hume. *An Enquiry Concerning Human Understanding.* London, 1748.
48. S. M. Khan and M. Soutchanski. Necessary and Sufficient Conditions for Actual Root Causes. In *ECAI 2020 - 24th European Conference on Artificial Intelligence, Including 10th Conference on Prestigious Applications of Artificial Intelligence (PAIS 2020)*, pages 800–808, 2020.
49. S. Kleinberg. A Logic for Causal Inference in Time Series with Discrete and Continuous Variables. In *Proc. of the 22nd Intern. Joint Conf. on AI (IJCAI)*, pages 943–950, 2011.
50. S. Kleinberg and B. Mishra. The Temporal Logic of Causal Structures. In *Proc. of the 25th Conf. on Uncertainty in AI (UAI)*, pages 303–312, 2009.
51. S. Kleinberg and B. Mishra. The Temporal Logic of Token Causes. In *Proc. of the 12th Intern. Conf. on Principles of Knowledge Representation and Reasoning (KR)*, 2010.
52. M. Kölbl and S. Leue. An Efficient Algorithm for Computing Causal Trace Sets in Causality Checking. In *Proc. of the 17th Intern. Symp. on Automated Technology for Verification and Analysis (ATVA)*, pages 171–186. Springer, 2019.
53. M. Kölbl, S. Leue, and R. Schmid. Dynamic Causes for the Violation of Timed Reachability Properties. In *Proc. of the 18th Intern. Conf. on Formal Modeling and Analysis of Timed Systems*, pages 127–143. Springer, 2020.
54. O. Kupferman and M. Y. Vardi. Vacuity Detection in Temporal Model Checking. In *Proc. of the 10th IFIP Working Conf. on Correct Hardware Design and Verification Methods (CHARME)*, pages 82–96, 1999.
55. A. Kupriyanov and B. Finkbeiner. Causality-Based Verification of Multi-threaded Programs. In *Proc. of the 24th Intern. Conf. on Concurrency Theory (CONCUR)*, LNCS, pages 257–272. Springer, 2013.

56. A. Kupriyanov and B. Finkbeiner. Causal Termination of Multi-threaded Programs. In *Proc. of the 26th Intern. Conf. on Computer Aided Verification (CAV)*, LNCS, pages 814–830. Springer, 2014.
57. F. Leitner-Fischer and S. Leue. Causality Checking for Complex System Models. In *Proc. of the 14th Intern. Conf. on Verification, Model Checking, and Abstract Interpretation (VMCAI)*, pages 248–267, 2013.
58. D. Lewis. Causation. *Journal of Philosophy*, 70(17):556–567, 1973.
59. J. L. Mackie. Causes and conditions. *American Philosophical Quarterly*, 2(4):245–264, 1965.
60. J. L. Mackie. *The Cement of the Universe: A Study of Causation.* Oxford, England: Oxford, Clarendon Press, 1974.
61. C. Mascle, C. Baier, F. Funke, S. Jantsch, and S. Kiefer. Responsibility and verification: Importance value in temporal logics. In *2021 36th Annual ACM/IEEE Symposium on Logic in Computer Science (LICS)*, pages 1–14, 2021.
62. C. H. Papadimitriou and K. Steiglitz. *Combinatorial Optimization: Algorithms and Complexity.* Prentice-Hall, 1982.
63. J. Pearl. Causal diagrams for empirical research. *Biometrika*, 82(4):669–688, 1995.
64. A. Pnueli. The temporal logic of programs. In *Proc. of the 18th Annual Symp. on Foundations of Computer Science (FOCS)*, pages 46–57, 1977.
65. M. Purandare and F. Somenzi. Vacuum Cleaning CTL Formulae. In *Proc. of the 14th Intern. Conf. on Computer Aided Verification (CAV)*, pages 485–499, 2002.
66. J. P. Queille and J. Sifakis. Specification and verification of concurrent systems in CESAR. In M. Dezani-Ciancaglini and U. Montanari, editors, *International Symposium on Programming*, pages 337–351, Berlin, Heidelberg, 1982. Springer Berlin Heidelberg.
67. H. Reichenbach. *The Direction of Time.* Berkeley and Los Angeles: University of California Press, 1956.
68. R. Reiter. *Knowledge in action: logical foundations for specifying and implementing dynamical systems.* MIT press, 2001.
69. M. Renieres and S. P. Reiss. Fault localization with nearest neighbor queries. In *Proc. of the 18th IEEE Intern. Conf. on Automated Software Engineering (ASE)*, pages 30–39, 2003.
70. M. Y. Vardi. Automatic Verification of Probabilistic Concurrent Finite State Programs. In *Proceedings of the 26th Annual Symposium on Foundations of Computer Science*, SFCS '85, pages 327–338. IEEE Computer Society, 1985.
71. R. Wertheimer. Conditions. *The Journal of Philosophy*, 65(12):355–364, 1968.
72. R. W. Wright. Causation in tort law. *California Law Review*, 73(6):1735–1828, 1985.
73. A. Zeller. Isolating Cause-Effect Chains from Computer Programs. In *Proc. of the 10th ACM SIGSOFT Symp. on Foundations of Software Engineering (FSE)*, pages 1–10, New York, NY, USA, 2002. ACM.