Formalising Overdetermination in a Labelled Transition System

Camilo Sarmiento
Lip6, Sorbonne Université
Paris, France
camilo.sarmiento@lip6.fr

Gauvain Bourgne
Lip6, Sorbonne Université
Paris, France
gauvain.bourgne@lip6.fr

Jean-Gabriel Ganascia
Lip6, Sorbonne Université
Paris, France
jean-gabriel.ganascia@lip6.fr

ABSTRACT

In the Knowledge Representation and Reasoning field, causality has mainly been used to determine the effects of actions. However, in legal or ethical reasoning, causality is used to determine the causal origin of a consequence. Such a posteriori reasoning is the focus of the Actual Causality field which has been extensively studied by lawyers, philosophers, mathematicians, and computer scientists. While most of the situations are easy to solve, there are several overdetermination cases that are far from trivial and that are still a source of disagreements in the field. Recent works have undertaken to adapt causality results into languages for Reasoning about Action and Change (RAC). In this paper, we aim to provide means to effectively address overdetermination issues to researchers seeking to incorporate actual causality into their RAC methods. To do so, we propose a definition of overdetermination in a Labelled Transition System and we formalise and enrich the existing typology of classical cases of overdetermination within this formalisation. The formal typology obtained enables the description of axiomatic properties of RAC methods that deal with overdetermination. To illustrate this, we describe the properties of a recent RAC formalisation of causality in light of this typology. We think that this way of doing can be generalised to all causality RAC formalisations.

CCS CONCEPTS

• **Computing methodologies → Causal reasoning and diagnostics**; *Planning for deterministic actions.*

KEYWORDS

Actual causality, Reasoning about Action and Change

ACM Reference Format:
Camilo Sarmiento, Gauvain Bourgne, and Jean-Gabriel Ganascia. 2025. Formalising Overdetermination in a Labelled Transition System. In *Proc. of the 24th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2025), Detroit, Michigan, USA, May 19 – 23, 2025*, IFAAMAS, 9 pages.

1 INTRODUCTION

Overdetermination cases are essentially examples in which more than one cause might yield the outcome by itself [3, 5, 16, 21, 23, 27, 34]. A canonical example is the one of Suzy and Billy throwing each a rock at the same bottle. Suzy's rock hits first and shatters the bottle, but Billy's rock would have shattered it otherwise, so one cannot infer that Suzy's throw is the cause simply by looking at what would have happened if she did not throw [18]. Cases of overdetermination are commonly found in the complexity of reality—causes of pollution, causes of suicide, causes of economic loss, etc—they give rise to numerous questions in law. They have therefore been extensively studied by lawyers [1, 29, 34], philosophers [3, 16, 21, 27], and computer scientists [4, 6, 9, 11, 18, 22] in the fields of causation. Despite the existing typology described in natural language, there is still disagreement regarding the expected causality in overdetermination cases [1, 11, 35].

Describing changes caused by the execution of actions is one of the earliest problems that artificial intelligence researchers attempted to tackle [25]. The division of the world into states and actions is a common thread among most approaches developed in this field, including Event [30] and Situation Calculus [31], PDDL [20], STRIPS [12], Action Description Languages [13–15], Markov Decision Processes and others. This division is not only accepted in the field of AI, it is also acknowledged by philosophers. For instance, in moral theory, it is commonly accepted that evaluating a state of the world involves assessing if it is good or bad, while evaluating an action involves determining if the action is obligatory, optional, or forbidden [33]. There is a clear distinction made here. In actual causality, most works do not make such a distinction and describe everything as events [3, 18, 27]. Concentrating solely on events was perhaps necessary to isolate the causal problem from other challenges that have occupied the Knowledge Representation and Reasoning (KRR) community for many years [30]. Nevertheless, we believe it would be a missed opportunity not to explore how causality research can enrich Reasoning about Actions and Change (RAC) proposals that make such distinction. Recent approaches have shown that it is feasible to do so [4, 9, 22, 32].

The aim of this work is to utilise the benefits of KRR to provide means to effectively address overdetermination issues to researchers seeking to incorporate actual causality into their RAC methods. To do so, we propose a definition of overdetermination in a Labelled Transition System (LTS) and we formalise and enrich the existing typology of classical overdetermination cases within this formalisation. This formal typology helps to clarify existing ambiguities about this notion. The choice to use a LTS is motivated by our intention to make this clarification work as broadly useful as possible. We do not view a LTS as a language for RAC, but rather as a semantics. In this regard, we share the perspective of numerous works in which action description languages and STRIPS [14], as well as Situation and Event Calculus [28], can be seen as describing a LTS. We thus believe that formalising the typology using a LTS is the most relevant approach to benefit the KRR field. This ensures that our results can be used by a wide audience.

This paper is structured as follows. Section 2 describes the general LTS and the basic causal relations that are needed. Section 3

Proc. of the 24th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2025), Y. Vorobeychik, S. Das, A. Nowé (eds.), May 19 – 23, 2025, Detroit, Michigan, USA. © 2025 International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org).

proposes the main contributions of this paper: a definition of overdetermination in the general LTS and the formalisation and enrichment of the overdetermination cases existing typology. In Section 4 we discuss some examples showing how important is the formalisation of the problem and present some benefits of using the typology proposed. We conclude and give some perspectives in Section 5.

2 LABELLED TRANSITION SYSTEM AND CAUSALITY

Example 1 (modern firing squad). *A firing squad typically consists of multiple soldiers who are given instructions to fire their weapons together. This procedure prevents both a member from disrupting the process and the identification of the person who fired the fatal shot. Consider Figure 1 electric circuit corresponding to a firing squad execution—paradigmatic situation where overdetermination is used to dilute responsibility—adapted to electric chair. This circuit is made up of a voltage source, an individual strapped and connected to electrodes, and two switches connected in parallel. We assume that the situation involves four agents: the one strapped and three others.*

Formalising all cases of overdetermination from Example 1 is possible by examining various combinations of how the agents behave. This paper will explore the aforementioned problem and also provide some reference examples of actual causality. To deal with these examples, it appears necessary to formalise the problem. We achieve this with a LTS. In this section, we introduce the formal aspects of a LTS and the general causal concepts needed to formalise the typology.

2.1 Labelled Transition System

In RAC approaches for classical planning, the state of the world is classically defined at any given time as a collection of fluents, which are variables describing the properties of the world. The evolution of the world can then be described as a series of states transitioning from one to another as a result of a set of events. Such systems can be explicitly represented by LTS, where states are sets of fluents and transitions between these states are labeled by set of events. Typically, such LTS are built from a more concise representation such as action languages [14]. We denote by $\mathbb{F}$ the set of variables describing the state of the world, more precisely *ground fluents* representing time-varying properties, and by $\mathbb{E}$ the set of variables describing transitions, more precisely *ground events* that modify fluents.

A *state* $S \subseteq \mathbb{F}$ is defined as a set of fluents. We define *state formulas* $\mathcal{F}$ as formulas of fluents using classical logical operators. For $\psi \in \mathcal{F}$, $S \vDash \psi$ is classically defined.

Definition 1 (Labelled Transition System (LTS)). *A* Labelled Transition System *is a directed graph* $\langle \mathbb{S}, \tau \rangle$ *where* $\mathbb{S} \subseteq 2^{\mathbb{F}}$ *is a set of states and* τ *is a set of* labelled transition relations $\tau \subseteq \mathbb{S} \times 2^{\mathbb{E}} \times \mathbb{S}$.

The set $\mathbb{S}$ represents all the possible states of the world and for each of these states, τ indicates, for a each state, the different sets of concurrently occurring events that could happen and their associated possible outcomes. Here, we restrict ourselves to *deterministic* LTS, assuming that for a given state S and a given set of events E there is at most one state S' such that $(S, E, S') \in \tau$.

An event $e \in \mathbb{E}$ is an atomic formula. When considering causality and overdetermination, it is important to understand why an event e can occur. Thus, it is useful to make explicit the preconditions of events, i.e. the conditions that must be satisfied by a state S for the event to be able to occur. The function that associates preconditions with each event is defined as: $pre : \mathbb{E} \to \mathcal{F}$.

We assume that the LTS is *correct* wrt to these precondition, meaning that for any $(S, E, S') \in \tau$, we have $S \vDash pre(e)$ for all $e \in E$. This is easily achieved when building the LTS from a classical RAC representation such as action languages, situation calculus or event calculus as preconditions are well defined in all these formalisms.

Disjunctive preconditions and concurrency of events are two essential elements that a LTS needs to allow when formalising cases of overdetermination. The first is satisfied by the form of $\mathcal{F}$, and the second by the fact that transitions are labeled by sets of events..

Since we consider actual causality, we are interested in a particular sequence of events and states to which a time $t \in \mathbb{T}$ can be associated, where $\mathbb{T} = \{0, \ldots, N\}$. Such sequence can be assimilated to a path in the LTS. S_0 is the *initial state*. Moreover, since we specifically address overdetermination, we must engage in counterfactual reasoning. Therefore, to precisely identify a path while enabling counterfactual reasoning, we will define a policy for timed states.

Definition 2 (policy π). *A* policy π *for an LTS* $\langle \mathbb{S}, \tau \rangle$ *is a function that associates timed states with the set of events that are supposed to occur in each of these states. Formally,* $\pi : \mathbb{S} \times \mathbb{T} \to 2^{\mathbb{E}}$.

Such a policy is correct *iff for all* $(S, t) \in \mathbb{S} \times \mathbb{T}$, *there exists* S' *such that* $(S, \pi(S, t), S') \in \tau$.

Definition 3 (causal setting χ). *The* causal setting χ *is the couple* (π, κ), *with* π *a policy and* κ *a context being the quintuple* $(\mathbb{E}, \mathbb{F}, pre, S_0, \mathbb{T})$.

A path in a LTS corresponds to a trace where $S^{\chi}(t)$ are the states we cross and $E^{\chi}(t)$ the transitions that take us there. As we limit our LTS to deterministic cases, to each χ corresponds a unique trace satisfying: (i) $S^{\chi}(0) = S_0$, (ii) $\forall t \in \mathbb{T}$, $E^{\chi}(t) = \pi(S^{\chi}(t), t)$, and (iii) $\forall t \in \mathbb{T}$, $(S^{\chi}(t), E^{\chi}(t), S^{\chi}(t+1)) \in \tau$.

Example 1 (continued). *The fluents* $f_1, f_2, f_3 \in \mathbb{F}^3$ *represent the closed state of each switch and the voltage source respectively.* $\psi = (f_1 \wedge f_3) \vee (f_2 \wedge f_3)$ *where* $\psi \in \mathcal{F}$ *represents the preconditions for the strapped individual being electrocuted* e_{ψ} *which results in his death* $d \in \mathbb{F}$. $e_1, e_2 \in \mathbb{E}^2$ *are the events which result is to close each switch respectively, and* $e_3 \in \mathbb{E}$ *is an event which result is to turn on the voltage source. The above information is formalised as* $pre(e_{\psi}) = \psi$ *and* $\forall i \in \{1, 2, 3\}$, $pre(e_i) = \top$. *Below is an example of a trace corresponding to an overdetermination case:*

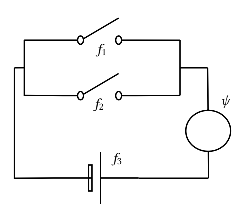

Figure 1: Electrical circuit consisting of a voltage source, two switches, and an individual connected to electrodes.

$S^\chi(0) = \{\neg f_1, \neg f_2, \neg f_3, \neg d\}$, $E^\chi(0) = \{e_3\}$;
$S^\chi(1) = \{\neg f_1, \neg f_2, f_3, \neg d\}$, $E^\chi(1) = \{e_1, e_2\}$;
$S^\chi(2) = \{f_1, f_2, f_3, \neg d\}$, $E^\chi(2) = \{e_\psi\}$;
$S^\chi(3) = \{f_1, f_2, f_3, d\}$.

To build a policy, one must consider which events should occur at each time step, depending on the current state. The policy thus defines both the actual scenario and counterfactual ones. Along the states of the actual path that happened, the policy indicates for each step what events took place. For states that are not part of the actual path, the policy indicates which actions should occur at a given time step in a counterfactual scenario where we reached this state at this time step. We propose here a method for building a simple policy from a plan associating a set of actions that must be done to each step, defined as a function $P : \mathbb{T} \to 2^{\mathbb{E}}$. The policy π_P can then be defined forall $(S, t) \in \mathbb{S} \times \mathcal{T}$ as $\pi_P(S, t) = \{e \in P(t) | S \models pre(e)\}$.

Example 1 (continued). *The scenario described in the previous example corresponds to putting on the voltage source at time 0, then closing both switches at time 1. Since the electocution should happen as soon as its preconditions are met, it must be attempted at each time.*

The corresponding policy will thus be π_P obtained from plan P such that $P(0) = \{e_3, e_\psi\}$, $P(1) = \{e_1, e_2, e_\psi\}$ and $P(2) = \{e_\psi\}$. As $pre(e_\psi)$ is not satisfied in $S^\chi(0)$ nor $S^\chi(1)$ but is satified in $S^\chi(2)$, we get $\pi_P(S^\chi(0), 0) = \{e_3\} = E^\chi(0)$, $\pi_P(S^\chi(1), 1) = \{e_1, e_2\} = E^\chi(1)$ and $\pi_P(S^\chi(2), 2) = \{e_\psi\} = E^\chi(2)$.

Note that this allows us to represent triggered events *such as e_ψ as actions that should be attempted at each step.*

2.2 Actual Causality

Having formalised Example 1 as a LTS, we proceed to introduce the necessary causal relations to enable reasoning about overdetermination cases. Similar to the LTS, we will not define a specific definition of causality, but instead describe the type of relations that are needed. In that way, our typology can be used by any approach with a formalism describing an LTS, regardless of the specific causal definition that it adopts.

A causal relation can generally be seen as a binary relation that links a cause to a consequence. In a LTS, we can either want to know what are the causes of $\psi \in \mathcal{F}$ being true in a state—$S^\chi(t) \models \psi$ denoted (ψ, t)—or of an *occurrence of event*—$e \in E^\chi(t)$ denoted (e, t). Therefore, two causal relations are required. As a basic input, we abstractly define the basic causal relation, which we call $\mathcal{F}$-causes, as a binary relation between $\mathbb{E} \times \mathbb{T}$ and $\mathcal{F} \times \mathbb{T}$. Given a $\mathcal{F}$-causes, $\rightsquigarrow \subseteq (\mathbb{E} \times \mathbb{T}) \times (\mathcal{F} \times \mathbb{T})$, we denote by $(e, t) \rightsquigarrow (\psi, t)$ the fact that (e, t) is a cause of (ψ, t) wrt this relation.

The precise definition of a $\mathcal{F}$-cause relation $\rightsquigarrow$ is specific to each RAC approach of causality [4, 9, 22, 32]. By remaining abstract here, we aim to encompass as many different definitions of causality as possible. We thus do not further characterise this relation here, but there are a number of necessary or sufficient conditions that have been identified in the case of structural models [8]. Some of the properties that we can expect of this relation are precedence (the cause should not occur after the consequence), inclusion of counter-factual dependency (if the consequence is simply counterfactually dependent on the antecedent, it should be a cause), and contribution (the cause should belong to a set of events that is minimally sufficient to bring about the consequence).

Given an $\mathcal{F}$-cause relation $\rightsquigarrow$, we extend it to a relation between occurrences of events, as found in most approaches to actual causality, by considering that (e, t) is an *actual cause* (wrt $\rightsquigarrow$) of (e', t') if $(e, t) \rightsquigarrow (pre(e'), t')$. For the sake of simplicity, we will overload the notation and represent by $(e, t) \rightsquigarrow (e', t')$ the fact that (e, t) is an actual cause of (e',t') wrt $\mathcal{F}$-cause relation $\rightsquigarrow$.

Lastly, it is essential to define the concept of causal paths in our setting, as it appears to be fundamental in the discussion of overdetermination cases. This concept has been formalised for other objectives by various authors in different formalisms [5, 7].

Definition 4 (causal path). *Given a causal setting χ, an event $e_\psi \in E^\chi(t_\psi)$, with $\psi = pre(e_\psi)$, and a $\mathcal{F}$-cause relation $\rightsquigarrow$, the sequence of occurrence of events $\omega = (e_n, t_n), \ldots, (e_1, t_1)$ is a* causal path *linking (e_n, t_n) to (e_ψ, t_ψ) if:*

- $(e_n, t_n) \rightsquigarrow (pre(e_{n-1}), t_{n-1})$;
- $\ldots$;
- $(e_2, t_2) \rightsquigarrow (pre(e_1), t_1)$;
- $(e_1, t_1) \rightsquigarrow (\psi, t_\psi)$.

A complete causal path *linking (e_n, t_n) to (e_ψ, t_ψ) is one that cannot be derived by deleting elements from another sequence that is also a causal path linking (e_n, t_n) to (e_ψ, t_ψ).*

In other words, a causal path between a cause and a consequence is a sequence of event occurrences where each event contributes to trigger the next.

Proposition 1. *If $(e_n, t_n) \rightsquigarrow (e_\psi, t_\psi)$, there is at least one complete causal path linking (e_n, t_n) to (e_ψ, t_ψ).*

Proof. Existence of a causal path directly follows from the definition with $n = 1$. Then, either this causal path is complete or it is a sub-sequence of a complete causal path. □

If the precise definition of $\mathcal{F}$-causes used by an RAC approach is transitive, the reciprocal of Proposition 1 is true. Among other things, this means that if there is a causal path ω linking (e_n, t_n) to (e_ψ, t_ψ), (e_n, t_n) is an actual cause of (e_ψ, t_ψ). In the following sections, we will only be referring to complete causal paths when we are using causal paths.

3 OVERDETERMINATION IN CAUSALITY

Before presenting our formalisation proposal, we review the different categories in the existing typology. To the best of our knowledge, these categories have only been broadly defined using natural language, which may be subject to a variety of interpretations and thus confusion. We provide an example of each definition, along with a classical example and a figure showing a trace of Example 1 corresponding to each.

• Symmetric/Duplicative causation: situations where ‘a factor other than the specified act would have been sufficient to produce the injury in the absence of the specified act, but its effects [...] combined with or duplicated those of the specified act to jointly produce the injury’ [34].

• Trumping: situations ‘where the pre-empted factor runs its whole course–a course that normally produce some event e–yet that factor does not cause e on this occasion because some pre-empting cause

"trumps" it' [17].
• Early preemption: situations 'in which the pre-emption of the pre-empted causal sequence occurs before the completion of the completed causal sequence' [35].
• Late preemption: situations 'in which the pre-emption of the pre-empted causal sequence occurs at the same time as (or after) the completion of the completed causal sequence' [35].

[5] traces the history of preemption definition and shows different interpretations that exist:

> In recent years, there has been some variance in the literature as to what exactly the difference between early and late preemption amounts to. According to the understanding which hearkens back to [23] [...], the difference is that 'in cases of early preemption, the backup process is cut off before the effect occurs, whereas in cases of late preemption, the process is cut off by the effect itself' [21]. By contrast, e.g. [17] hold that the characteristic feature of early preemption is that a process is interrupted by another process, whereas in cases of late preemption no interruption takes place, rather, the preempted process just does not run to completion.

The above quote gives rise to two remarks. First, another type of late preemption can be considered. [17] variant specificity is that, from a RAC point of view, such variant can only be properly represented by durative processes. Indeed, if 'no interruption takes place' but the 'process just does not run to completion', the only explanation is that a durative event starts and never ends having the expected effect because this one is made true by another quicker process. We call such variant durative late preemption.

The second remark is that durative late preemption and trumping descriptions seem to describe the same cases. It appears from literature discussions that trumping is the most unclearly defined case. Note also that the terminology used to define these cases—'(completed) causal sequence', 'completion', 'process', 'backup process', 'cut off', 'interrupted'—confirms the existence of a common intuition about the presence of causal paths when describing overdetermination.

Example 2 (symmetric/duplicative). *From now on, we suppose that Example 1 was more complex than described. The agents do not control the switches directly, there is a multiplicity of mechanisms made of pulleys, ropes, and gears between their action and the event which modifies the switches state. The actions made by two of the agents are now denoted e^1_m and e^2_n. To each event the causal paths $\omega^1 = (e^1_m, t^1_m), \dots, (e^1_1, t^1_1)$ and $\omega^2 = (e^2_n, t^2_n), \dots, (e^2_1, t^2_1)$ can potentially be associated. Figure 2 shows traces of Example 1 corresponding to different cases of overdetermination. Note that the occurrence of e^1_m and e^2_n happening at the same time in all cases is simply a choice to simplify the illustration. In (a) we see that ω^1 and ω^2 jointly contribute to the occurrence of e_ψ through f_2. This corresponds to the example in the literature of two assassins pouring a lethal dose of poison into a victim's drink [21].*

Example 3 (trumping). *In (b) we see the same that in (a). However, in (b) occurrence of e^1_1 and e^2_1 are not simultaneous and thus e^2_1 has as effect a fluent that was already true. This corresponds to the example where a boat on a river is forced to stop because the river is blocked. A bridge A has collapsed in its path. It turns out that another bridge B has also collapsed a little further on and is also blocking the river [35].*

Example 4 (early preemption). *In (c) we can see that ω^1 interrupts ω^2 before the effect occurs. This interruption can be done by any element of ω^1 which makes false a necessary element for the triggering of an element of ω^2. This corresponds to the example where assassin A poisons a desert traveller canteen, but assassin B empties the canteen before the traveller can drink. The traveller is found dead from dehydration some days after [11].*

Example 5 (late preemption). *In (d) we can see that ω^2 is interrupted 'by the effect itself'. This would be the case if $\psi = (f_1 \wedge f_3 \wedge \neg d) \vee (f_2 \wedge f_3 \wedge \neg d)$. Indeed, as the occurrence (e_ψ, t_ψ) causes d to be true at $t_\psi+1$, ψ cannot be true at $t_\psi+1$ and thus ω^2 is interrupted. This corresponds to the example where two forest fires are set and each is sufficient to destroy a house. One of them arrives first and burns down the whole house just before the other one arrives [4]. Some details and discussions were abstracted on purpose, they will be discussed in Section 4.*

In the following, we first introduce a formal definition of overdetermination given the formalism described in Section 2. All categories in the existing typology can be seen as a special case of Definition 6. Second, we propose a formal typology of overdetermination cases.

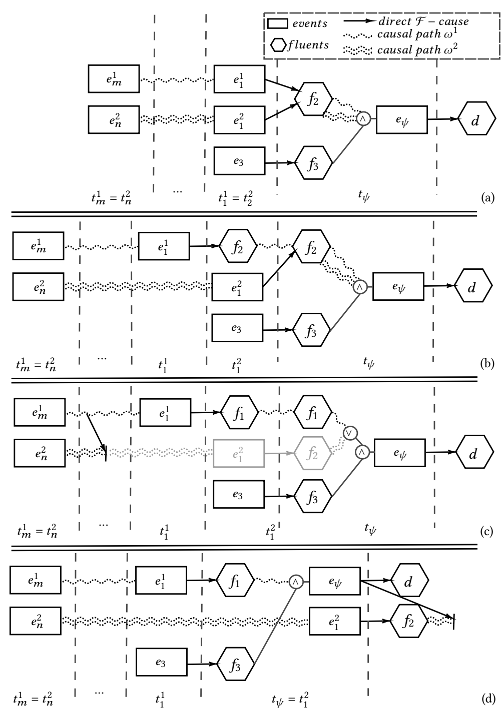

Figure 2: Illustration of (a) symmetric/duplicative, (b) trumping, (c) early preemption, and (d) late preemption overdetermination cases.

3.1 Overdetermination Formally Defined

Most formalisms used to reason about actual causality do not make a distinction between states of the world and events, but consider everything as events. In the famous Suzy and Billy example variables are *Suzy/Billy throws*, *Suzy/Billy hits*, and *bottle shatters*. Therefore, when actual causality is discussed most examples concern causal relations between two occurrences of events. Such choice appears intuitive. In many of the domains where actual causality could be used—as law or ethics—agents and their actions are at the centre of the thinking. However, distinction between states of the world and events seems deeply embedded in the way we reason about how the world evolves and thus appears useful to discuss the intricacies of causality. [25] emphasise 'that practical systems require epistemologically adequate systems in which those facts [commonsense concepts] which are actually ascertainable can be expressed'. The formalism described in Section 2 makes a strong distinction between states of the world and events. When speaking about causal overdetermination, we could use $\mathcal{F}$-causes as well as actual causes. In order to keep comparison as easy as possible with classical causality approaches, we chose to focus on actual causes. In other words, we are interested in the causes of the event *bottle shatters* occurrence rather than in the causes of the fluent *broken bottle* being true.

To define overdetermination we must engage in counterfactual reasoning. Therefore, to build counterfactual policies we define, for a given set of event occurrences $X \subseteq \mathbb{E}\times\mathbb{T}$, $\pi\setminus X \overset{\text{def}}{=} \forall S, \forall t,\ \pi(S,t) = \pi(S,t) \setminus \{e \in \mathbb{E} | (e,t) \in X\}$. This amounts to removing the occurrences of X from each possible state. For such a counterfactual policy to be correct, we must ensure that there will be a transition labeled by the set of remaining events. We characterize that with the following definition:

Definition 5 (Simple omissibility). *Given a LTS $\langle \mathbb{S}, \tau\rangle$, an event $e \in \mathbb{E}$ is* simply omissible *iff for all $(S, E, S') \in \tau$, there exists S'' such that $(S, E \setminus \{e\}, S'') \in \tau$.*

This means that it is always possible for such an event not to occur in a given state without changing any other event occurrence. By extension, if e is simply omissible, we would also say that any occurrence (e,t) is simply omissible.

Proposition 2. *If X is a set of simply omissible event occurrences and π is a correct policy, then $\pi \setminus X$ is a correct policy.*

Proof. For all S and t, π correct gives us S' such that $(S, \pi(S,t), S') \in \tau$. Then, denoting $e_1, \ldots, e_p$ the elements of $\{e \in \mathbb{E} | (e,t) \in X\}$, e_1 simply omissible gives S'_1 such that $(S, \pi(S,t) \setminus \{e_1\}, S'_1) \in \tau$, e_2 simply omissible thus gives S'_2 such that $(S, \pi(S,t) \setminus \{e_1, e_2\}, S'_2) \in \tau$ and so on until $(S, \pi(S,t) \setminus \{e_1, e_2, \ldots e_n\}, S'_n) \in \tau$ which proves $\pi \setminus X$ correct. □

Let us consider the case of an event e that is not simply omissible. There are two possible cases. First case happens when there is a state that contains e in all its outgoing transitions. This means that our model considers that this event is mandatory, it is not possible to consider a scenario in which it would not happen in this state. Then, it seems normal that we cannot consider a counterfactual case where it does not occur as such a scenario would not belong to any of the possible evolution of the world that we modeled. In such a case, counterfactual reasoning is prohibited by the model. The second case however, would be a case in which there can be transitions without e, but these transitions would then have other event that should occur. We would say that such an event is *indirectly omissible*. For example, it could be the case that if you cannot turn right, you must turn left and doing nothing is not an option. In such case, we need a more refined counterfactual policy that can choose, among the transition where e does not happen in this state, which one should be considered. Such cases are outside the scope of our current approach, but they could be handled by considering more expressive policies that order possible transitions rather than selecting only one. Then the counterfactual policy of removing e would select the best transition among those that do not contains it. In the following definition we stick to simple omissibility, but as explained above, extension to indirect omissibility could be considered with more expressive policies.

Definition 6 (Direct overdetermination case). *Let $\chi = (\pi, \kappa)$ be the causal setting and $(a^1, t^1), (a^2, t^2), (e_\psi, t_\psi)$ three occurrences of events, where a^1 and a^2 are simply omissible events. Given three counterfactual causal settings $\chi^1_I = (\pi \setminus \{(a^2, t^2)\}, \kappa)$, $\chi^2_I = (\pi \setminus \{(a^1, t^1)\}, \kappa)$—two Individual causal settings—and $\chi_- = (\pi \setminus \{(a^1, t^1), (a^2, t^2)\}, \kappa)$, we have an* overdetermination case *between (a^1, t^1) and (a^2, t^2) in χ if: $e_\psi \in E^{\chi}(t_\psi)$, $e_\psi \notin E^{\chi_-}(t_\psi)$, $e_\psi \in E^{\chi^1_I}(t_\psi)$, and $e_\psi \in E^{\chi^2_I}(t_\psi)$.*

For the sake of brevity and clarity we only consider cases of overdetermination involving two overdetermining events. However, in cases involving more overdetermining events, it is possible to construct a policy where the case can be dealt with pairwise. To handle it properly, it would be necessary to define that there is indirect overdetermination between (a^1, t^1) and (a^2, t^2) in (χ, π) if there exists a set of occurrences of omissible events $X \subset \mathbb{E} \times \mathbb{T}$ such that a^1 and a^2 are in overdetermination in $(\chi, \pi \setminus X)$. From a computational point of view, we get the following result on complexity.

Proposition 3.

- *Determining whether two given events occurrences are in direct overdetermination is linear in the number of time steps ($|\mathbb{T}|$).*
- *The complexity of determining all pairs of events occurrences that are in direct overdetermination is in $O(n_o^2.|\mathbb{T}|)$ where n_o is the number of event occurrences in the main trace (which is at most $|\mathbb{E}|.|\mathbb{T}|$).*

Proof. Determining the trace from a causal setting is linear in the number of time steps. Determining the trace from a counterfactual causal setting is equivalent as the computation of the counterfactual policy can be done on the fly. Determining whether two given event occurrences are in direct overdetermination can be done by computing 4 traces, so it stays linear in $|\mathbb{T}|$.

Then to determine all pairs of events that are in direct overdetermination, we need to compute the trace for the main scenario, for all scenario with one omission, and for the scenario with two omission of events that were not removed in previous step. In the worst case, this last step requires to check all pairs, and thus n_o^2 traces to compute. □

3.2 Formal Typology of Overdetermination

In order to formalise the typology, we place ourselves in the causal setting χ and we consider we are in an overdetermination case. Thus, we have $e_\psi \in E^\chi(t_\psi)$ with $pre(e_\psi) = \psi$, (e^1_m, t^1_m), (e^2_n, t^2_n) two occurrences of events, and $\chi^1_I = (\pi \setminus \{e^2_n\}, \kappa)$, $\chi^2_I = (\pi \setminus \{e^1_m\}, \kappa)$, $\chi_- = (\pi \setminus \{e^1_m, e^2_n\}, \kappa)$ three counterfactual causal settings. Given $i \in \{1, 2\}$, we denote Ω^i_I the set of all causal paths ω which link (e^i_k, t^i_k) to (e_ψ, t_ψ) in χ^i_I, and Ω^i the set of all causal paths ω which link (e^i_k, t^i_k) to (e_ψ, t_ψ) in χ.

Being in an overdetermination case, we deduce by Definition 6 that $\Omega^1_I \neq \emptyset$ and $\Omega^2_I \neq \emptyset$ as in each case, (e^i_k, t^i_k) is the only difference between χ^i_I—in which $e_\psi \in E^{\chi^i_I}(t_\psi)$—and χ_-—in which $e_\psi \notin E^{\chi_-}(t_\psi)$. With a similar reasoning, we deduce that $\Omega^1 \cup \Omega^2 \neq \emptyset$. Without loss of generality, we can assume to simplify that $\Omega^1 \neq \emptyset$—i.e. that there is always a path in χ linking (e^1_m, t^1_m) to (e_ψ, t_ψ). Ω^2 can then be either empty or not.

For the sake of brevity and clarity, we simplify the current analysis by making two assumptions: $|\Omega^1_I| = |\Omega^2_I| = 1$—i.e. there is a unique causal path in the individual causal path set— and $\Omega^i \setminus \Omega^i_I = \emptyset$—i.e. no causal paths are created. From such hypothesis we can deduce $\Omega^1_I = \Omega^1 = \{\omega^1\}$, $\Omega^2_I = \{\omega^2\}$, and either $\Omega^2 = \{\omega^2\}$, or $\Omega^2 = \emptyset$. If $\Omega^2 = \{\omega^2\}$ it means $\forall (e^2, t^2) \in \omega^2,\ e^2 \in E^\chi(t^2)$, while if $\Omega^2 = \emptyset$ it means that ω^2 is not a path in χ given that $\exists (e^2, t^2) \in \omega^2,\ e^2 \notin E^\chi(t^2)$. This distinction is crucial as it enables the differentiation of preemption cases from others.

The existence of multiple causal paths between two occurrences of events raises potentially interesting new cases but is beyond the scope of the paper. For example, some causal paths in Ω^i_I may no longer be in Ω^i because interrupted—$\Omega^i_I \setminus \Omega^i$—others may be conserved—$\Omega^i_I \cap \Omega^i$—and others may be created by modifying individual ones—$\Omega^i \setminus \Omega^i_I$. Additionally, if $|\Omega^1| = 3$ and $|\Omega^2| = 1$, do we consider that overdetermination cases of different nature coexist, or are there rules of subsumption? As far as we know, such interesting cases have not been discussed in the field before.

When studying how examples are described and analysed in the literature, it appears that temporal relations between the causal paths are ubiquitous. In addition, if there is a difference between durative late preemption and late preemption, beyond the durative formalisation, it is to be found in time. For those reasons, we need to be exhaustive when studying time. To achieve such an exhaustive analysis, we extract from each causal path ω^i its corresponding time interval which starts at t^i_n and ends at t^i_1, and we study the situation of each typology column for all thirteen possible relations given by Allen's interval algebra [2]. Note that the names of the Allen's relations should not be interpreted causally in the understanding of the situation. For example in line 'ω^1 ended by ω^2', we only consider both interval time relation and do not consider that in such case ω^2 had an influence on ω^1.

In Table 1 first two columns, we consider $\Omega^2 = \emptyset$. In such case we can deduce that the addition of (e^1_m, t^1_m), that changes χ^2_I into χ, interrupts ω_2—an occurrence invalidates the triggering of one element of ω_2. The relevant question according to the literature is whether this element is the consequence itself—$\exists (e^2_j, t^2_j) \in \omega^2, (e_\psi, t_\psi) \rightsquigarrow \neg pre(e^2_j), t^2_j)$—or an occurrence of the causal path ω_1—$\exists \left((e^1_i, t^1_i) \in \omega^1, (e^2_j, t^2_j) \in \omega^2\right), (e^1_i, t^1_i) \rightsquigarrow (\neg pre(e^2_j), t^2_j)$. Table 1 first two columns consider each one of these cases.

Definition 7 (early preemption). *Let us be in an overdetermination case. It is considered an early preemption case if* $\Omega^1 = \{\omega^1\}$, $\Omega^2 = \emptyset$, *and* $\exists ((e^1_i, t^1_i) \in \omega^1, (e^2_j, t^2_j) \in \omega^2), (e^1_i, t^1_i) \rightsquigarrow (\neg pre(e^2_j), t^2_j)$.

Definition 8 (late preemption). *Let us be in an overdetermination case. It is considered a late preemption case if* $\Omega^1 = \{\omega^1\}$, $\Omega^2 = \emptyset$, *and* $\exists (e^2_j, t^2_j) \in \omega^2, (e_\psi, t_\psi) \rightsquigarrow (\neg pre(e^2_j), t^2_j)$.

In Table 1 last two columns, we consider $\Omega^2 = \{\omega^2\}$. In such case the relevant question seems to be the way the occurrence of (e_ψ, t_ψ) is achieved—is the traveller dead due to poisoning or dehydration? In the LTS we are in, 'the way the occurrence of (e_ψ, t_ψ) is achieved' corresponds to the concept of support, intrinsically linked with $\psi = pre(e_\psi)$. Indeed, a *support* W of ψ is a prime implicant of ψ. In Example 1, two ways of causing (e_ψ, t_ψ) are possible, either through $W = \{f_1, f_3\}$, or through $W' = \{f_2, f_3\}$. It is thus a question of knowing whether the supports related to ω^1 and ω^2 are identical—$W^1 = W^2$ as in Figure 2.a—or not—$W^1 \neq W^2$ as in Figure 2.c.

By exhaustively studying the literature examples corresponding to Table 1 third and fourth column, it appears that not all time relations are relevant. The relation between latest time of each causal path—corresponding to t^1_1 and t^2_1—is the parameter that seems to be discriminatory. By looking to this parameter some nuances can be done in duplicative/symmetric overdetermination cases. We propose a subdivision of this category into three more refined categories.

Definition 9 (synchronous duplicative). *Let us be in an overdetermination case. It is considered a synchronous duplicative case if* $\Omega^1 = \{\omega^1\}$, $\Omega^2 = \{\omega^2\}$, $W^1 \neq W^2$, *and* $t^1_1 = t^2_1$.

Definition 10 (asynchronous duplicative). *Let us be in an overdetermination case. It is considered an asynchronous duplicative case if* $\Omega^1 = \{\omega^1\}$, $\Omega^2 = \{\omega^2\}$, $W^1 \neq W^2$, *and* $t^1_1 < t^2_1$.

Definition 11 (symmetric). *Let us be in an overdetermination case. It is considered a symmetric case if* $\Omega^1 = \{\omega^1\}$, $\Omega^2 = \{\omega^2\}$, $W^1 = W^2$, *and* $t^1_1 = t^2_1$.

The main difference between the first two lies in time. The last one takes its name from the fact that what characterises it, is that the way in which it is caused is the same for both paths and occurs at the same time. An example of each is given in Figures 3.e, 3.f, and 2.a respectively.

For [17] 'cases of trumping turn out on inspection to be nothing more than either cases of symmetric overdetermination in disguise or cases of late pre-emption in disguise'. In this paper we propose a trumping case definition which fits the classical examples for this case and which is distinct from 'symmetric overdetermination' and 'late pre-emption'. In the typology, the trumping definition we propose is closer to Hitchcock's [2007] position who follows '[26] and [19] in thinking that trumping is a species of overdetermination [symmetric] and not of preemption' as shown in Figure 2.b.

Allen's interval algebra		$\Omega^1 = \{\omega^1\}, \Omega^2 = \emptyset$		$\Omega^1 = \{\omega^1\}, \Omega^2 = \{\omega^2\}$	
		$(e_i^1, t_i^1) \rightarrow (\neg pre(e_j^2), t_j^2)$	$(e_\psi, t_\psi) \rightarrow (\neg pre(e_j^2), t_j^2)$	$W^1 \neq W^2$	$W^1 = W^2$
ω^1 equal to ω^2		Early Preemption	*	Synchronous Duplicative	Symmetric
ω^1 ends ω^2		Early Preemption	*	Synchronous Duplicative	Symmetric
ω^1 ended by ω^2		Early Preemption	*	Synchronous Duplicative	Symmetric
ω^1 overlaps ω^2		Early Preemption	Late Preemption/Durative	Asynchronous Duplicative	Trumping
ω^1 overlapped by ω^2		Early Preemption	*	**	**
ω^1 starts ω^2		Early Preemption	Late Preemption/Durative	Asynchronous Duplicative	Trumping
ω^1 started by ω^2		Early Preemption	*	**	**
ω^1 during ω^2		Early Preemption	Late Preemption/Durative	Asynchronous Duplicative	Trumping
ω^1 contains ω^2		Early Preemption	*	**	**
ω^1 meets ω^2		Early Preemption	Late Preemption	Asynchronous Duplicative	Trumping
ω^1 met by ω^2		*	*	**	**
ω^1 proceeds by ω^2		Early Preemption	Late Preemption	Asynchronous Duplicative	Trumping
ω^1 proceeded by ω^2		*	*	**	**

Table 1: Formal typology of overdetermination cases given all possible time relations between two causal paths. (*) Incoherence between the causal relation at the origin of ω^2 interruption and time relation between intervals. (**) Incoherence between the assumption that ω^1 is always the first to achieve to the occurrence of (e_ψ, t_ψ) and time relation between intervals.

Definition 12 (trumping). *Let us be in an overdetermination case. It is considered a trumping case if* $\Omega^1 = \{\omega^1\}$, $\Omega^2 = \{\omega^2\}$, $W^1 = W^2$, *and* $t_1^1 < t_1^2$.

Note that in the last two columns the nature of the concerned cases is not affected by which causal paths ends first. However, as such information can affect the causal relations deduced by different causal approaches, it is desirable for our later properties to eliminate such possibility. Thus, for the sake of brevity, we will suppose that in the last two columns ω^1 always ends first, which explains the grey cells. Note that while cells with '*' are grey because there is a real incoherence between the causal relation at the origin of ω^2 interruption and time relation between intervals, cells with '**' are grey by an incoherence between our assumption that ω^1 always achieves first and time relation between intervals.

Table 1 gives, for the thirteen possible time relations between two causal paths, the type of overdetermination case corresponding to each column situation. The typology formalisation allows to clearly define the mainly discussed categories of overdetermination in the literature. Indeed, the exhaustive analysis presented in Table 1 allows us to propose six well defined cases of overdetermination.

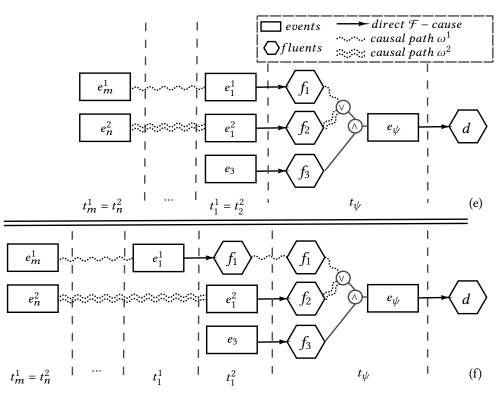

Figure 3: Illustration of (e) synchronous duplicative, and (f) asynchronous duplicative overdetermination cases.

In addition, we can deduce that cases of late preemption can be transformed to durative late preemption only if $t_1^1 > t_n^2$.

4 DISCUSSION

In this section we discuss examples showing how thin the boundary between different categories is, and thus how important is the problem formalisation. Then, we present an example of axiomatic properties of an RAC methods that deal with overdetermination.

4.1 On the Importance of the Formalisation

Let us start by discussing trumping cases which appear to be the most debated ones. Indeed, while some authors consider trumping as being part of the 'relevant cases of actual causation' [10], others consider that 'cases of trumping turn out on inspection to be nothing more than either cases of symmetric overdetermination in disguise or cases of late pre-emption in disguise' [17]. The classical trumping example is the one where a boat on a river is forced to stop because the river is blocked. A bridge A has collapsed in its path. It turns out that another bridge B has also collapsed a little further, also blocking the river.

In fact, this example can indeed belong to different categories according to how it is formalised. If what makes the boat stop is the approach of an obstacle—e_ψ corresponds to the boat stops and $\psi = obstacle\_in\_front$—then we are in a late preemption case. The causal path ω^1, corresponding to the bridge A's collapse, interrupts the causal path ω^2, corresponding to the same event for bridge B, by the effect of e_ψ which is the boat being motionless. This would correspond to Definition 8. However, if what makes the boat stop is the river being blocked, then we could be in one of the four cases of Table 1 third and fourth column. On one hand, if we consider that each bridge blocks the river in its own way—the triggering condition will be $\psi = river\_blocked\_A \lor river\_blocked\_B$—we are either in a synchronous duplicative case or in an asynchronous duplicative case. This would correspond to Definition 9 or 10. On the other hand, if we consider that each bridge blocks the river in the same way—$\psi = river\_blocked$—we will be either in a symmetric case corresponding to Definition 11—if both bridges collapse at the same time—or in a trumping case corresponding to Definition 12—if the collapsing is not simultaneous. Formalising such example implies that despite fluent *river_blocked* being true by one of the bridges collapse, the latter collapse will also have as effect fluent *river_blocked*. This simple example shows how thin boundaries are between cases in Table 1 third and fourth columns.

Boundaries between cases in first and second column can also be thin. Such fact is very well explained by [35]. If we consider the classical example where assassin A poisons the victim's drink, but assassin B kills the victim with a gun before the poison can take effect, granularity of the formalisation has an important role. From the lowest level of granularity, the case appears as being a durative late preemption or late preemption case. Either the ongoing poisoning process does not run to completion, or the death of the victim interrupts the causal path of the poison as in Figure 2.d. However, by increasing the granularity of the formalisation, 'the steps of the physical process that must occur inside the victim's body' [35] can be taken into account. In such a case, we are likely to find that assassin B's causal path interrupts the poison's causal path, for example by causing a haemorrhage that prevents the poison from spreading through the body. Such case will correspond to the early preemption case of Definition 7. Halpern's [2016] rewrite of Suzy and Billy example, which adds two variables to the problem, equates changing a durative late preemption case—not manageable in structural equations framework—into early preemption.

Note that frontiers are not just porous between the first two columns and between the last two columns, there is porosity between preemption cases and last two columns. First, in the case where the formalism used does not allow to cause a fluent if this fluent was already true when the event occurred, trumping cases cannot be represented and become early preemption cases. Second, slightly modifying the preconditions of e_ψ can make an asynchronous duplicative case be a late preemption case and vice versa. Indeed, the presence or not of $\neg d$ in $pre(e_\psi)$ makes all the difference. In Example 1, the intuition is not the same if the electrocution leads to death or not. Whereas in the first case, the first agent who closes the switch preempts any other path that might come after—$\neg d \in pre(e_\psi)$ which represents that the victim must be alive to be executed—in the second case, closing the second switch afterwards actually causes current to flow through that part of the circuit and reach the victim, thus allowing for cases other than preemption.

Most of the formalisation choices we have discussed only become explicit when one has to formalise the example. Otherwise, those assumptions remain implicit but affect the causal intuition we have of it—by changing asynchronous duplicative cases into late preemption cases for instance. Having shown that how one formalises the problem is critical, it seems necessary to use a representation which makes these nuances clear when dealing with overdetermination.

4.2 Axiomatic Properties in Overdetermination

Almost all papers in the Actual Causality field propose a definition of actual causality and then try to show how intuitive the causal relations obtained are by confronting them with examples. In [6], Beckers points out that 'It is unrealistic to expect that this [...] strategy in and on itself can deliver a satisfactory account of causation, because there are too many examples and even more intuitions'. The discussion of intuitions is beyond the scope of this paper. However, we would like provide researchers seeking to incorporate actual causality into their RAC methods with the means to effectively address overdetermination issues. Through the use of our typology, one can generalise beyond specific examples and propose properties that characterise a definition of actual causality. Rather than confronting definitions with multiple examples as classically done in causality, one can prove for all examples belonging to a category what events will be considered as causes by a causal definition.

Definition 13 (sensitivity of approaches). *Given an overdetermination case, an actual causality approach is sensitive:*

- *to preemption: if in such cases it considers (e_m^1, t_m^1) an actual cause of (e_ψ, t_ψ), but not (e_n^2, t_n^2).*
- *to duplicative/symmetric overdetermination: if in such cases it considers both (e_m^1, t_m^1) and (e_n^2, t_n^2) actual causes of (e_ψ, t_ψ).*

To illustrate how this characterization of overdetermination can be applied to specific definitions of a causality relation, we assess the sensitivity of the NESS-based definition from [32]. Though in [24], it is argued that NESS test cannot deal with preemption cases because it misses the nuances of structural and temporal order of events, this shows it is mainly because of the formalism.

Proposition 4. *The actual causality approach of [32] is sensitive to preemption and to duplicative/symmetric overdetermination.*

Detailed proofs are given as supplementary material. Definition 13 is an example of what an actual causality definition considers as causes of each category. Such specifications can formally be stated using the proposed formal typology.

As action language $\mathcal{AL}$ and situation calculus can also be represented by LTS, future works will assess the sensitivity of other RAC approaches to causality such as [22] and [4]. In addition, we are working on direct translations of situations described in a LTS into structural equations, allowing causal approaches formalised only in structural equations to benefit from the typology.

5 CONCLUSION

In this paper, we provide researchers seeking to incorporate actual causality into their RAC methods with the means to effectively address overdetermination issues by proposing a formalisation of the existing typology of overdetermination. This enables the description of properties describing how RAC methods integrating causality manage each overdetermination case. To achieve this, overdetermination and other relevant notions were defined in a LTS. The exhaustive study has revealed more detailed categories of overdetermination and shown that the classically discussed examples cover a narrow range of possibilities. In future work, we intend to explore the possibilities not yet considered and to use the typology to better address omission and negative causation.

Another interesting venue of future work is leveraging the principles of causation defined in a structural equation model setting by Beckers and Vennekens [8] in our LTS framework. Especially, translating the notion of contribution and production to compare them to our notion of sensitivity to overdetermination seems promising. Indeed that notion of production is based on enriching structural equation models with a notion of timing which is already present in our model.

REFERENCES

[1] Yuval Abrams. 2022. Omissive Overdetermination: Why the Act-Omission Distinction Makes a Difference for Causal Analysis. *University of Western Australia Law Review* 49, 57 (2022). https://papers.ssrn.com/abstract=4061990

[2] James F. Allen. 1990. Maintaining Knowledge about Temporal Intervals. In *Readings in Qualitative Reasoning About Physical Systems*. Morgan Kaufmann, 361–372.

[3] Holger Andreas and Mario Guenther. 2021. Regularity and Inferential Theories of Causation. In *The Stanford Encyclopedia of Philosophy*. Stanford University.

[4] Vitaliy Batusov and Mikhail Soutchanski. 2018. Situation Calculus Semantics for Actual Causality. In *Proc. of the Thirty-Second AAAI Conference on Artificial Intelligence, (AAAI-18)*. 1744–1752.

[5] Michael Baumgartner. 2013. A Regularity Theoretic Approach to Actual Causation. *Erkenntnis* 78, 1 (2013), 85–109. https://doi.org/10.1007/s10670-013-9438-3

[6] Sander Beckers. 2021. Causal Sufficiency and Actual Causation. *J. Philos. Log.* 50, 6 (2021), 1341–1374.

[7] Sander Beckers. 2021. The Counterfactual NESS Definition of Causation. *Proc. of the AAAI Conf. on Artificial Intelligence* 35, 7 (2021), 6210–6217. https://ojs.aaai.org/index.php/AAAI/article/view/16772

[8] Sander Beckers and Joost Vennekens. 2018. A Principled Approach to Defining Actual Causation. *Synthese* 195, 2 (2018), 835–862. https://doi.org/10.1007/s11229-016-1247-1

[9] Fiona Berreby, Gauvain Bourgne, and Jean-Gabriel Ganascia. 2018. Event-Based and Scenario-Based Causality for Computational Ethics. In *Proceedings of the Seventeenth International Conference on Autonomous Agents and MultiAgent Systems, AAMAS 2018*, Elisabeth André, Sven Koenig, Mehdi Dastani, and Gita Sukthankar (Eds.). International Foundation for Autonomous Agents and Multiagent Systems Richland, SC, USA / ACM, Stockholm, Sweden, 147–155. http://dl.acm.org/citation.cfm?id=3237412

[10] Alexander Bochman. 2018. Actual Causality in a Logical Setting. In *Proc. of the Twenty-Seventh International Joint Conference on Artificial Intelligence (IJCAI)*, Jérôme Lang (Ed.). 1730–1736.

[11] Alexander Bochman. 2018. On Laws and Counterfactuals in Causal Reasoning. In *Proc. of the Sixteenth International Conference on Principles of Knowledge Representation and Reasoning*. 494–503.

[12] Richard Fikes and Nils J. Nilsson. 1971. STRIPS: A New Approach to the Application of Theorem Proving to Problem Solving. *Artificial Intelligence* 2, 3/4 (1971), 189–208. https://doi.org/10.1016/0004-3702(71)90010-5

[13] Michael Gelfond and Vladimir Lifschitz. 1993. Representing Action and Change by Logic Programs. *Journal of Logic Programming* 17, 2/3&4 (1993), 301–321. https://doi.org/10.1016/0743-1066(93)90035-F

[14] Michael Gelfond and Vladimir Lifschitz. 1998. Action Languages. *Electron. Trans. Artificial Intelligence* 2 (1998), 193–210.

[15] Enrico Giunchiglia and Vladimir Lifschitz. 1998. An Action Language Based on Causal Explanation: Preliminary Report. In *Proc. of the 15th Nat. Conf. on Artificial Intelligence and 10th Innovative Applications of Artificial Intelligence Conf., AAAI 98, IAAI 98*. 623–630.

[16] Ned Hall. 2004. Two Concepts of Causation. In *Causation and Counterfactuals*. MIT Press, 225–276.

[17] Ned Hall and Laurie Ann Paul. 2003. Causation and Pre-emption. In *Philosophy of Science Today*. Oxford University Press.

[18] Joseph Y. Halpern. 2016. *Actual Causality*. The MIT Press.

[19] Joseph Y. Halpern and Judea Pearl. 2005. Causes and Explanations: A Structural-Model Approach. Part I: Causes. *The British Journal for the Philosophy of Science* 56, 4 (2005), 843–887.

[20] Patrik Haslum, Nir Lipovetzky, Daniele Magazzeni, and Christian Muise. 2019. *An Introduction to the Planning Domain Definition Language*. Morgan & Claypool Publishers. https://doi.org/10.2200/S00900ED2V01Y201902AIM042

[21] Christopher Hitchcock. 2007. Prevention, Preemption, and the Principle of Sufficient Reason. *Philosophical Review* 116, 4 (2007), 495–532.

[22] Emily C. LeBlanc, Marcello Balduccini, and Joost Vennekens. 2019. Explaining Actual Causation via Reasoning About Actions and Change. In *Proceedings of the Sixteenth European Conference on Logics in Artificial Intelligence, JELIA 2019 (Lecture Notes in Computer Science, Vol. 11468)*, Francesco Calimeri, Nicola Leone, and Marco Manna (Eds.). Springer, Rende, Italy, 231–246. https://doi.org/10.1007/978-3-030-19570-0_15

[23] David Lewis. 1986. Postscripts to 'Causation'. In *Philosophical Papers Vol. Ii*. Oxford University Press.

[24] Ruta Liepina, Giovanni Sartor, and Adam Wyner. 2020. Arguing about causes in law: a semi-formal framework for causal arguments. *Artif. Intell. Law* 28, 1 (2020), 69–89.

[25] John McCarthy and Patrick Hayes. 1969. Some Philosophical Problems from the Standpoint of Artificial Intelligence. In *Proceedings of the Fourth Machine Intelligence Workshop*, Bernard Metzler and Michie Donald (Eds.). Edinburgh University Press, Edinburgh, Scotland, UK, 463–502.

[26] Michael McDermott. 2002. Causation: Influence versus Sufficiency. *The Journal of Philosophy* 99, 2 (2002), 84–101.

[27] Peter Menzies and Helen Beebee. 2020. Counterfactual Theories of Causation. In *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University.

[28] Rob Miller and Murray Shanahan. 2002. Some Alternative Formulations of the Event Calculus. In *Computational Logic: Logic Programming and Beyond, Essays in Honour of Robert A. Kowalski, Part II (Lecture Notes in Computer Science, Vol. 2408)*, Antonis C. Kakas and Fariba Sadri (Eds.). Springer, 452–490. https://doi.org/10.1007/3-540-45632-5_17

[29] Michael Moore. 2019. Causation in the Law. In *The Stanford Encyclopedia of Philosophy* (Winter 2019 ed.). Metaphysics Research Lab, Stanford University.

[30] Erik T. Mueller. 2014. *Commonsense Reasoning: An Event Calculus-Based Approach* (second edition ed.). Morgan Kaufmann. https://doi.org/10.1016/C2014-0-00192-X

[31] Raymond Reiter. 2001. *Knowledge in Action: Logical Foundations for Specifying and Implementing Dynamical Systems*. The MIT Press. https://doi.org/10.7551/mitpress/4074.001.0001

[32] Camilo Sarmiento, Gauvain Bourgne, Katsumi Inoue, and Jean-Gabriel Ganascia. 2022. Action Languages Based Actual Causality in Decision Making Contexts. In *PRIMA 2022, Proc. (LNCS)*. Springer.

[33] Mark Timmons. 2012. *Moral Theory: An Introduction* (second edition ed.). Rowman & Littlefield Publishers.

[34] Richard W. Wright. 1985. Causation in Tort Law. *California Law Review* 73, 6 (1985), 1735–1828. https://doi.org/10.2307/3480373

[35] Richard W. Wright. 2011. The NESS Account of Natural Causation: A Response to Criticisms. In *Perspectives on Causation*. Social Science Research Network.