Preferential Semantics for Causal Systems

Pavlos Peppas and **Maurice Pagnucco**
Computational Reasoning Group
Department of Computing
Division of ICS
Macquarie University
NSW, 2109, Australia
{pavlos, morri}@ics.mq.edu.au

Mikhail Prokopenko
CSIRO Mathematical and Information Sciences
Locked Bag 17
North Ryde
NSW, 2113, Australia
mikhail.prokopenko@cmis.csiro.au

Norman Y. Foo
Knowledge Systems Group
Department of AI
School of CSE
The University of New South Wales
2052, NSW, Australia
norman@cse.unsw.edu.au

Abhaya Nayak
Computational Reasoning Group
Department of Computing
Division of ICS
Macquarie University
NSW, 2109, Australia
abhaya@ics.mq.edu.au

Abstract

In the present work we examine the causal theory of actions put forward by McCain and Turner [McCain and Turner, 1995] for determining *ramifications*. Our principal aim is to provide a characterisation of this causal theory of actions in terms of a Shoham-like *preferential semantics* [Shoham, 1988]. This would have a twofold advantage: it would place McCain and Turner's theory in perspective, allowing a comparison with other logics of action; and, it would allow us to glean further insights into the nature of causality underlying their work. We begin by showing that our aim is not attainable by a preferential mechanism alone. At this point we do not abandon preferential semantics altogether but augment it in order to arrive at the desired result. We draw the following moral which is at the heart of our paper: *two components — minimal change under a preferential structure and causality — are required to provide a concise solution to the frame and ramification problems.*

1 Introduction

One of the cornerstone developments in the field of reasoning about action (and nonmonotonic reasoning) was Shoham's *preferential semantics* [Shoham, 1988]. While work in this area has since progressed, preferential semantics remain an important and intuitively appealing concept. It furnishes a semantics for a class of nonmonotonic logics. Under this idea an ordering is placed over the class of interpretations. The models corresponding to a particular inference are then identified as the minimal models under this ordering that satisfy the premises. In an intuitive sense, the ordering represents a preference or plausibility ordering over interpretations with only the most preferred (most plausible) being countenanced as serious possibilities. At the heart of this approach lies the *principle of minimal change*: consider only the minimal (i.e., most preferred) models.

In more recent times the notion of causality has attracted much attention [McCain and Turner, 1995; Thielscher, 1997; Sandewall, 1996] in an attempt to provide a concise solution to the *frame problem* (and, more specifically, the *ramification problem*). This is mainly in response to the recognition that traditional *domain constraints* alone are not sufficient for providing compact solutions to these problems.

In this work we focus on a proposal by McCain and Turner [McCain and Turner, 1995] which includes a causal component. McCain and Turner introduce causal laws of the form $\phi \Rightarrow \psi$ where ϕ and ψ are fluent formulae (i.e., they do not contain further instances of $\Rightarrow$ but only classical truth functional connectives). ¿From an intuitive standpoint these formulae can be read as 'ϕ *causes* ψ'. In this way they express 'a relation of determination between states of affairs that make ϕ and ψ true' [McCain and Turner, 1995, p. 1979]. In fact, it is possible to do away with traditional domain constraints altogether in deference to these causal laws [McCain and Turner, 1995, Proposition 3]. An important point to note is that causal laws function as 'uni-directional' implications — the contrapositive ($\neg\psi \Rightarrow \neg\phi$) does not hold in general.

The principal aim of this paper is to determine whether it is possible to supply McCain and Turner's causal theory of actions with a preferential-style semantics. More specifically, the contributions of this paper are as follows. We show that it is not possible to characterise McCain and Turner's causal theory via a traditional preferential semantics applied to interpretations of the original (unaugmented) language. We rectify this, not by abandoning preferential semantics but by augmenting it with a further relational structure. A similar result was sought by Peppas et al. [Peppas *et al.*, 1997] but the counterexample they present assumes a transitive and total ordering where we only assume transitivity, and the semantics they develop only characterises a subset of the possible McCain and Turner causal systems whereas the semantics we present captures all.

In arriving at this result we introduce two state-selection mechanisms: *state elimination systems* and *state transition systems*. State transition systems enhance a preference structure, based on symmetric difference, with a binary relation on states. State elimination systems function as a way of tying together McCain and Turner causal systems and state transition systems. Both state elimination systems and state transition systems give a clearer insight into the nature of causality at play in McCain and Turner's causal theory of actions.

The results contained herein are significant on a number of fronts. They allow us to gain a deeper understanding of the nature of causality underpinning McCain and Turner's framework. They also permit us to place McCain and Turner's causal theory in the context of other nonmonotonic logics whose semantics make use of preferential structures. It has previously been suggested that minimal change and causality are interreducible (for instance [Gustafsson and Doherty, 1996]). However, such works make use of an expanded language. In doing so they move away from the ideal of a *concise* solution to the frame and ramification problems. We, on the contrary, claim that minimal change and causality can coexist in separate roles and, moreover, can complement each other.

This paper is structured as follows. In the next two sections we shall outline some basic terminology and notation followed by a brief overview of McCain and Turner's [McCain and Turner, 1995] causal theory of action. In section 4 we shall show that it is not possible to supply a straightforward preferential semantics to capture McCain and Turner's approach. The solution we suggest here is not to abandon preferential semantics entirely but rather to augment it. In sections 5—7 we investigate the different state selection mechanisms, giving the desired result. We end with a discussion of the significance of these results and our conclusions.

2 Technical Preliminaries

Throughout this paper we shall be working with a fixed finitary propositional language $\mathcal{L}$ whose propositional letters we shall call *fluents*. The set of all fluents is denoted by $\mathcal{F}$. A *literal* is a fluent or the negation of a fluent. A *state* (or *world*) is defined as a maximal consistent set of literals. The set of all literals will be denoted $\mathcal{N}$. The set of all states will be denoted $\mathcal{W}$. By $[\phi]$ we denote all states consistent with the sentence $\phi \in \mathcal{L}$ (i.e., $[\phi] = \{w \in \mathcal{W} : w \vdash \phi\}$). Occasionally we will refer to $[\phi]$ as the *ϕ-states* (or *ϕ-worlds*).

3 Causal Systems

In this section we briefly review McCain and Turner's [McCain and Turner, 1995] causal theory of actions. In so doing we shall introduce some further notation that will be useful for the remainder of the paper.

As outlined above, McCain and Turner introduce a new connective $\Rightarrow$ to stand for the existence of a causal relationship between sentences ϕ and ψ of the underlying language $\mathcal{L}$. This allows for expressions of the form $\phi \Rightarrow \psi$ (where $\phi,\ \psi \in \mathcal{L}$)[1] which are termed *causal laws* (or *casual rules*).[2] A set of causal laws $\mathcal{D}$ is referred to as a *causal system*. Given any set of sentences $\Gamma \subseteq \mathcal{L}$ and a causal system $\mathcal{D}$, the (causal) *closure* of Γ in $\mathcal{D}$ is denoted $C_{\mathcal{D}}(\Gamma)$ and defined to be the smallest superset of Γ closed under classical logical consequence and such that for any $\phi \Rightarrow \psi \in \mathcal{D}$, if $\phi \in C_{\mathcal{D}}(\Gamma)$, then $\psi \in C_{\mathcal{D}}(\Gamma)$. We also say that Γ *causally implies* ϕ with respect to $\mathcal{D}$ if and only if $\phi \in C_{\mathcal{D}}(\Gamma)$ and denote this $\Gamma \mid\!\sim_{\mathcal{D}} \phi$.

Another notion that will be of importance is that of a *legitimate state* with respect to a causal system $\mathcal{D}$. Any state r is legitimate with respect to $\mathcal{D}$ if and only if $r = C_{\mathcal{D}}(r) \cap \mathcal{N}$. That is, a state is legitimate if and only if it does not contravene any causal laws of $\mathcal{D}$. The set of legitimate states with respect to $\mathcal{D}$ is denoted by $\mathit{Legit}_{\mathcal{D}}$.

McCain and Turner's aim is to determine the set of possible next (or resultant) states $\mathtt{Res}_{\mathcal{D}}(E,\ w)$ given an initial state w and the direct effects (or post-conditions) of an action represented by the sentence E.[3] Formally speaking, we have for any causal system $\mathcal{D}$ a function $\mathtt{Res}_{\mathcal{D}}$ mapping a legitimate (initial) state w and sentence E (direct effects) to the set of states $\mathtt{Res}_{\mathcal{D}}(E,\ w)$ according to the definition [McCain and Turner, 1995]:

$$r \in \mathtt{Res}_{\mathcal{D}}(E,\ w) \text{ iff } r = \{\mathtt{f} \in \mathcal{N} :\ (w \cap r) \cup \{E\} \mid\!\sim_{\mathcal{D}} \mathtt{f}\}$$

We often refer to the elements of $\mathtt{Res}_{\mathcal{D}}(E,\ w)$ as *causal fixed-points*. Note that it follows from this definition that if $r \in \mathtt{Res}_{\mathcal{D}}(E,\ w)$, then $r \in [E]$ (i.e., r must satisfy the direct effects of the action). Intuitively speaking, the elements of $\mathtt{Res}_{\mathcal{D}}(E, w)$ are simply those E-states where all changes with respect to w can be justified by the underlying causal system.

We are now in a position to state our aims more clearly. The desire is to mimic McCain and Turner's fixed-point definition using a preference ordering over states and in such a way as not to introduce auxiliary sentences into our language. More specifically, we wish to investigate whether this is at all possible; whether we can provide a preferential-style semantics characterising $\mathtt{Res}_{\mathcal{D}}(E,\ w)$ for any legitimate state w and sentence E.

[1] Note that nesting of $\Rightarrow$ is not permitted.

[2] For the sake of simplicity we shall assume here that the antecedent of any causal law is consistent.

[3] We shall refer to actions only through their direct effects as they play no direct role in McCain and Turner's framework.

4 Impossibility Results

In this section we clearly specify what we mean by a preferential semantics. We then present an impossibility result showing that a traditional preferential semantics is not capable of characterising McCain and Turner's fixed-point definition.

We are given an initial state $w \in \mathcal{W}$ and a (strict) preference ordering $<_w \subseteq \mathcal{W} \times \mathcal{W}$ on states. The only restriction we place on $<_w$ is that it satisfy transitivity. Adhering to the essence of preferential semantics [Shoham, 1988] we seek to define those states resulting from the occurrence of an action with direct effects E at initial state w as the minimal E-states under $<_w$. The following condition expresses these desiderata:

(P) $\quad \mathtt{Res}_{\mathcal{D}}(E, w) = min([E], <_w)$

Now, according to McCain and Turner, there is no need to consider illegitimate states as possible resultant states since they contradict the causal laws. Hence, we begin by focusing on a variant of condition (P):

(P′) $\quad \mathtt{Res}_{\mathcal{D}}(E, w) = min([E], <_w) \cap Legit_{\mathcal{D}}$

We are now in a position to state a fundamental result of this section; that, in general, it is not possible to satisfy the condition (P′) (with transitive $<_w$). Note firstly that a *non-trivial language* is one with at least three fluents.

Theorem 4.1 *(First Impossibility Theorem)*
Given a non-trivial language $\mathcal{L}$, there exists a causal system $\mathcal{D}$ and (initial) state $w \in \mathcal{W}$ such that no ordering $<_w$ on states (generated by $\mathcal{L}$) satisfies (P′).

Proof: Assume that $\mathcal{L}$ has three propositional letters $\mathtt{a}, \mathtt{b}, \mathtt{c}$. Let the initial state be $w = \{\mathtt{a}, \mathtt{b}, \mathtt{c}\}$ and define s_1, s_2, s_3 and s_4 to be the following states: $s_1 = \{\neg\mathtt{a}, \mathtt{b}, \mathtt{c}\}$, $s_2 = \{\mathtt{a}, \neg\mathtt{b}, \mathtt{c}\}$, $s_3 = \{\neg\mathtt{a}, \neg\mathtt{b}, \mathtt{c}\}$ and $s_4 = \{\neg\mathtt{a}, \mathtt{b}, \neg\mathtt{c}\}$. Finally let $\mathcal{D}$ be the following causal system: $\mathcal{D} = \{\neg\mathtt{a} \wedge \mathtt{c} \Rightarrow \neg\mathtt{b}, \neg\mathtt{b} \wedge \mathtt{c} \Rightarrow \neg\mathtt{a}, \neg\mathtt{a} \wedge \mathtt{b} \Rightarrow \neg\mathtt{c}\}$.

Consider now the following direct effects (post-conditions) of actions. $\Delta_1 = \neg\mathtt{a} \wedge \mathtt{c}$, $\Delta_2 = \neg\mathtt{b} \wedge \mathtt{c}$, $\Delta_3 = (\mathtt{b} \leftrightarrow \neg\mathtt{c})$ and $\Delta_4 = (\bigwedge s_1) \vee (\bigwedge s_2) \vee (\bigwedge s_3) \vee (\bigwedge s_4)$. Clearly, states s_1 and s_3 satisfy Δ_1; s_2 and s_3 satisfy Δ_2; s_3 and s_4 satisfy Δ_3; and all four states s_1, s_2, s_3, s_4 satisfy Δ_4.

Suppose a (transitive) ordering on states $<_w$ satisfying condition (P′) exists. Now, the following is easily (albeit tediously) verified. $\mathtt{Res}_{\mathcal{D}}(\Delta_1, w) = \{s_3\}$ from which we conclude $s_1 \not<_w s_3$. $\mathtt{Res}_{\mathcal{D}}(\Delta_2, w) = \{s_3\}$, therefore $s_2 \not<_w s_3$. $\mathtt{Res}_{\mathcal{D}}(\Delta_3, w) = \{s_3, \{\mathtt{a}, \mathtt{b}, \neg\mathtt{c}\}\}$, therefore $s_4 \not<_w s_3$. Finally, $\mathtt{Res}_{\mathcal{D}}(\Delta_4, w) = \{s_4\}$ from which it follows that $(s_1 <_w s_3) \vee (s_2 <_w s_3) \vee (s_4 <_w s_3)$. This leads us to a contradiction. ∎

The following impossibility result now follows quite straightforwardly and is more appropriate for our purposes given that condition (P) is a more faithful rendering of the spirit of preferential semantics than condition (P′). It allows us to conclude that a traditional preferential semantics (captured by condition (P)) cannot, in general, be given to McCain and Turner's causal theory of actions.

Theorem 4.2 *(Second Impossibility Theorem)*
Given a non-trivial language $\mathcal{L}$, there exists a causal system $\mathcal{D}$ and (initial) state $w \in \mathcal{W}$ such that no ordering $<_w$ on states (generated by $\mathcal{L}$) satisfies (P).

We shall not give away preferential semantics entirely however. Our aim now becomes to retain as much of preferential semantics as possible and include a further mechanism to capture the influence of causality. To this end we investigate separate (though related) mechanisms for selecting possible resultant states.

5 State-Selection Mechanisms

Taking a step backwards for a moment, we can simply view McCain and Turner's approach as a *state-selection mechanism*. More specifically, McCain and Turner's causal theory of actions, given some domain knowledge in terms of a causal theory $\mathcal{D}$, specifies a way of selecting a subset $\mathtt{Res}_{\mathcal{D}}(E, w)$ of $[E]$ given an initial state w and direct effects E. $\mathtt{Res}_{\mathcal{D}}(E, w)$ returns exactly those states that are possible upon the occurrence of an action with direct effects E at state w.

Viewing this as a *selection function* however, we consider $\mathtt{Res}_{\mathcal{D}}(E, w)$ to be a function selecting the 'best' states from among $[E]$ (with respect to w). This is the view we shall adopt here in presenting two further state-selection mechanisms: *state elimination systems* and *state transition systems*. State transition systems will provide the augmented preferential semantics we seek in terms of our aims.

This desired result is achieved in two steps. We begin by showing how to intertranslate McCain and Turner causal systems and state elimination systems in a way that preserves the selection process. We then show how to intertranslate state elimination systems and state transition systems (again preserving the selection process). In truth, we could do away with state elimination systems and simply translate directly between causal systems and state transition systems. However, we choose not to do so because it simplifies the proofs and provides further insight into the nature of causality captured by McCain and Turner's approach.

6 Mechanism 1: State Elimination Systems

In this section we describe our first state-selection mechanism: state elimination systems. The underlying idea is to use *state elimination rules* to discard E-states from further consideration for we have noted above that in McCain and Turner's [McCain and Turner, 1995] causal theory $\mathtt{Res}_{\mathcal{D}}(E, w) \subseteq [E]$. A state rejected or eliminated by a state elimination rule is one which contravenes a causal relationship deemed to hold in the resultant state (in fact, in the causal system as a whole).

Definition 6.1 *(State elimination rule)*
A state elimination rule *(or simply,* elimination rule*) is an expression of the form* $\{r_1, r_2, \cdots, r_k, r_{k+1}, \cdots, r_n\} \rhd \{r_1, r_2, \cdots, r_k\}$ *where each r_i is a state.*

A state elimination system $\mathcal{S}$ is a set of state elimination rules. An elimination rule functions by rejecting certain states from among those currently considered possible. Suppose that according to an agent's current beliefs it considers the states that are possible to be among $\{r_1, \cdots, r_n\}$. An elimination rule like that in Definition 6.1 allows the agent to reject states $r_{k+1}, \cdots, r_n$.

Let us briefly consider the mechanics of a state elimination system. At any point we are working with the set of states currently being entertained (a subset of $[E]$). We repeatedly apply elimination rules to this set of states to reject the illegitimate ones (those not possible) focusing on the possible resultant states. All elimination rules need to be applied until no further states can be rejected to ensure that all illegitimate states have been purged and only definite possibilities remain. To put it another way, a state elimination system acts as a *filtering* mechanism; illegitimate states are successively filtered out through use of elimination rules.

Definition 6.2 *($\rightsquigarrow$ and $\overset{*}{\rightsquigarrow}$)*
In a state elimination system $\mathcal{S}$, we shall say that a set of states Q yields *a set of states R* in one step*, denoted by $Q \rightsquigarrow R$, iff there exists an elimination rule $X \rhd Y$ such that $Q \subseteq X$ and $R = Q \cap Y$. We define $\overset{*}{\rightsquigarrow}$ to be the reflexive transitive closure of $\rightsquigarrow$.*

After the application of certain elimination rules we find that any further application does not result in the rejection of additional states. At this point we reach a *compact set of states*; a point of equilibrium.

Definition 6.3 *(Compact state)*
A set of states Q is compact *(in S) iff for any R such that $Q \overset{*}{\rightsquigarrow} R$, it follows that $Q = R$. If Q is a singleton and compact, we will call the state in Q compact.*

One last notion that we require is that of an *E-predecessor* of a given state; those E-states preceding the given state with respect to an ordering based on symmetric difference. More formally:

Definition 6.4 *(E-predecessor)*
Given any two states w, r and any sentence E, the E-predecessors of r with respect to w is defined to be the set $\langle\!\langle r, E \rangle\!\rangle_w = \{r' : r' \in [E]$ and $\mathit{Diff}(w, r') \subseteq \mathit{Diff}(w, r)\}$ where $\mathit{Diff}(x, y)$ denotes the symmetric difference of states x and y (i.e., $(x \setminus y) \cup (y \setminus x)$) as in the PMA [Winslett, 1988].

It is clear that any $r \in [E]$ is an E-predecessor of itself with respect to w (i.e., $r \in \langle\!\langle r, E \rangle\!\rangle_w$). The E-predecessors of r with respect to w are just the E-states which agree with w on at least those fluents where w and r agree and possibly others. If one considers a PMA ordering [Winslett, 1988] of states $\leq_w$ (i.e., $r \leq_w s$ iff $\mathit{Diff}(w, r) \subseteq \mathit{Diff}(w, s)$), then the E-predecessors are those E-states at least as close to w as r.

We are now in a position to define a state-selection mechanism based on state elimination systems.

Definition 6.5 ($\mathtt{Next}_{\mathcal{S}}(E, w)$)
With any state elimination system $\mathcal{S}$ we associate a result function $\mathtt{Next}_{\mathcal{S}}$ (mapping a compact (in $\mathcal{S}$) state w and a sentence E to the set of states $\mathtt{Next}_{\mathcal{S}}(E, w)$) defined as follows: $\mathtt{Next}_{\mathcal{S}}(E, w) = \{r \in [E] :$ r is compact (in $\mathcal{S}$) and $\langle\!\langle r, E \rangle\!\rangle_w \overset{*}{\rightsquigarrow} \{r\}\}$.

In the following section we characterise $\mathtt{Res}_{\mathcal{D}}(E, w)$ in terms of $\mathtt{Next}_{\mathcal{S}}(E, w)$. First, however, let us briefly consider the definition of $\mathtt{Next}_{\mathcal{S}}(E, w)$. According to the definition above, a state r is a possible resultant state if and only if all its E-predecessors (with respect to w) are rejected by elimination rules in $\mathcal{S}$ but r and only r is retained. If r is retained along with some other state, then there is some closer state (one with 'less' change) consistent with the state elimination system $\mathcal{S}$ (and, therefore, causal system $\mathcal{D}$) under consideration. Moreover, it means that there is something in the state(s) for which causality cannot account. If r is rejected on the other hand, it must violate a causal relationship. For these reasons we only consider the E-predecessors of r to determine whether it belongs to $\mathtt{Next}_{\mathcal{S}}(E, w)$; we need to determine whether r is illegitimate or whether a 'closer' state satisfies the causal relationships. If either is the case, we can safely reject the state. Otherwise, we can retain the state.

6.1 Causal Systems and State Elimination Systems

We now establish the interrelationship between causal systems and state elimination systems. This will give us a way of moving back and forth between the two systems facilitating the final intertranslation between causal systems and state transition systems. The following definition will prove useful.

Definition 6.6 *(Selection-equivalent)*
A causal system $\mathcal{D}$ is selection-equivalent *to a state elimination system $\mathcal{S}$ iff $\mathtt{Res}_{\mathcal{D}}(E, w) = \mathtt{Next}_{\mathcal{S}}(E, w)$, for every sentence E and state w.*

The notion of selection-equivalence will be useful in relating causal systems and state elimination systems. Moreover, it will be useful in relating any two state-selection mechanisms (in an obvious way).

We now turn to the main result of this section. A state elimination system can exactly capture a causal system (and vice versa).

Theorem 6.7 *For every causal system there exists a selection-equivalent state elimination system. Conversely, for every state elimination system there exists a selection-equivalent causal system.*

Proof (Sketch)
($\Rightarrow$) Let $\mathcal{D}$ be an arbitrary causal system. For every causal law $\varphi \Rightarrow \psi$ in $\mathcal{D}$, produce the elimination rule $[\varphi] \rhd [\varphi \wedge \psi]$. Call $\mathcal{S}$ the set of elimination rules so produced. It is not difficult to verify that for any legitimate state w and sentence E, $\mathtt{Res}_{\mathcal{D}}(E, w) = \mathtt{Next}_{\mathcal{S}}(E, w)$ (simply notice that for any state r, $[(w \cap r) \cup \{E\}] = \langle\!\langle r, E \rangle\!\rangle_w$).
($\Leftarrow$) Let $\mathcal{S}$ be an arbitrary state elimination system. For every elimination rule $X \rhd Y$ produce the causal law $\varphi \Rightarrow \psi$, where φ, ψ are such that $[\varphi] = X$ and $[\psi] = Y$ (since our language is a finitary propositional one, such φ and ψ always exist). The set of causal laws so produced, call it $\mathcal{D}$, is selection-equivalent to $\mathcal{S}$. ■

Of particular note in this proof is the relationship between causal laws and elimination rules: $\phi \Rightarrow \psi$ if and only if $[\phi] \rhd [\phi \wedge \psi]$ (or, equivalently, $[\phi] \rhd [\phi] \cap [\psi]$).

We can also identify an important class of state elimination systems that will be useful later (in the proof of Theorem 7.3).

Definition 6.8 *($\mathcal{S}$ Unary)*
A state elimination system $\mathcal{S}$ is unary *iff every elimination rule eliminates precisely one state, i.e. for all $(X \rhd Y) \in \mathcal{S}$, $X \setminus Y$ is a singleton.*

The following result reveals an interesting and important aspect of unary state elimination systems.

Theorem 6.9 *Every state elimination system is selection-equivalent to a unary state elimination system.*

7 Mechanism 2: State Transition Systems

In this section we consider our second (and last) state-selection mechanism: state transition systems. Again, we shall obtain a direct characterisation of causal systems (and state elimination systems). In this case we have a preferential mechanism augmented by further structure to achieve the result we desire in this paper.

A state transition system consists of a binary relation on states intended to represent possible transitions between states due to the presence of causality. It is this relation that, together with a preferential ordering based on symmetric difference, will be used to determine possible resultant states.

We begin with some requisite definitions.

Definition 7.1 *A* state transition system $\mathcal{M}$ *is a binary relation on the set $\mathcal{W}$ of states (i.e., $\mathcal{M} \subseteq \mathcal{W} \times \mathcal{W}$). Whenever $\langle r, r' \rangle \in \mathcal{M}$ we will write $r \rightharpoonup r'$. We shall say that a state r is* final *(in $\mathcal{M}$) iff for any r' such that $r \rightharpoonup r'$, $r = r'$.*

The binary relation $\mathcal{M}$ can be considered to represent state transitions due to the influence of causality.

We are now in a position to define the mechanism for selecting possible resultant states (or *successor states*) $\texttt{Succ}_{\mathcal{M}}(E,\ w)$ for a state transition system $\mathcal{M}$ given initial state w and action with direct effects E.

Definition 7.2 ($\texttt{Succ}_{\mathcal{M}}(E, w)$)
To any state transition system $\mathcal{M}$ we associate a function $\texttt{Succ}_{\mathcal{M}}$ (mapping a final (in $\mathcal{M}$) state w and a sentence E to the set of states $\texttt{Succ}_{\mathcal{M}}(E, w)$) defined as follows: $\texttt{Succ}_{\mathcal{M}}(E, w) = \{r' \in [E] :\ r'$ is final (in $\mathcal{M}$) and there is a Hamiltonian path through states in $\langle\!\lbrack r', E \rbrack\!\rangle_w\}$.

A Hamiltonian path is one which traverses every vertex (here states) of a graph [Wilson, 1985]. In this case, the graph's vertices are the E-predecessors of r' and the edges are given by the binary relation $\mathcal{M}$ (i.e., there is an edge between states r' and s iff $\langle r',\ s \rangle \in \mathcal{M}$). The significance of a Hamiltonian path will be considered further in the next section.

It is important to notice that $\texttt{Succ}_{\mathcal{M}}(E,\ w)$ is determined by two components: a preference ordering on states, based on symmetric difference, used to derive the E-predecessors of r' with respect to w (i.e., $\langle\!\lbrack r', E \rbrack\!\rangle_w$) and the binary relation on states $\mathcal{M}$. We maintain that the preference ordering captures the principle of minimal change while the binary relation captures the effect of causality. Notice firstly that a state must be reachable via a Hamiltonian path through all E-predecessors ending in r'. If this is not possible, either a 'closer' E-state is consistent with the causal relationships that hold (absence of Hamiltonian path) or r' violates a causal relationship (path does not end at r'; i.e., r' is not final). Another important point is that, like state elimination systems, we only need consider E-predecessors of r' (with respect to w) to determine whether it is a possible resultant state.

7.1 State Elimination Systems and State Transition Systems

In this section we establish an intertranslation between state elimination systems and state transition systems. We can then use the results of Section 6.1 to establish a correspondence between causal systems and state transition systems. This gives us the result we seek: an augmented preferential semantics for McCain and Turner's causal theory of actions.

Theorem 7.3 *For every state elimination system $\mathcal{S}$ there is a selection-equivalent state transition system $\mathcal{M}$. Conversely, for every state transition system $\mathcal{M}$ there is a selection-equivalent state elimination system $\mathcal{S}$.*

Proof. (Sketch)
($\Rightarrow$) Let $\mathcal{S}$ be a state elimination system. Let $\mathcal{S}'$ be a unary state elimination system that is selection-equivalent to $\mathcal{S}$. From $\mathcal{S}'$ we construct a selection-equivalent state transition system $\mathcal{M}$. First we require some definitions.

Consider an arbitrary set of states Q with cardinality $n + 1$. We shall say that the string of elimination rules $\sigma_1; \sigma_2; \cdots; \sigma_n$ *dissolves* Q iff after applying these rules successively (in the order given), all but one of the states of Q are eliminated and, furthermore, the one remaining state is compact.

Assume that $\sigma_1; \sigma_2; \cdots; \sigma_n$ dissolves Q and for all $1 \leq i \leq n$, let r_i be the state of Q that is eliminated by the rule σ_i; let us also call w the one state of Q that is not eliminated. We shall call the sequence of states $r_1; r_2; \cdots; r_n; w$ a *trace* for Q (in $\mathcal{S}'$).

We now construct from $\mathcal{S}'$ a state transition system $\mathcal{M}$ in the following manner. For any two states $r,\ r'$, $\langle r, r' \rangle \in \mathcal{M}$ if and only if there is a dissolvable set of states Q containing r and r', such that for some trace of Q in $\mathcal{S}'$, r' appears immediately after r. It can be shown that $\mathcal{M}$ is selection-equivalent to $\mathcal{S}'$.

($\Leftarrow$) Essentially proved by reversing the construction presented above. ■

The central result of this paper, as expressed by the following corollary, is now obtained by combining theorems 6.7 and 7.3.

Corollary 7.4 *For every causal system $\mathcal{D}$ there exists a selection-equivalent state transition system $\mathcal{M}$. Conversely, for every state transition system $\mathcal{M}$ there exists a selection-equivalent causal system $\mathcal{D}$.*

This result states that it is possible to exactly characterise McCain and Turner's causal theory of actions via a preferential-style semantics (defined in terms of symmetric difference) augmented with a binary relation on states and through the notion of a Hamiltonian path.

8 Discussion

One of the morals that we have attempted to stress in this work is that it is possible to retain preferential semantics and augment it in capturing the causal theory of McCain and Turner. We maintain that the preferential component of state transition systems relates to the principle of minimal change while the binary relation on states relates to the causal system. It is our contention that both of these components — minimal change and causality — are required if one is to supply a *concise* solution to the frame problem; the two can co-exist and, in fact, complement each other.

It has been suggested that other preferential-style approaches to reasoning about action are capable of capturing McCain and Turner's causal theory of actions (for instance, [Gustafsson and Doherty, 1996]); an apparent contradiction of what is suggested by our impossibility theorems in Section 4. However, this and similar approaches are able to do so only by augmenting the original language $\mathcal{L}$ (with, for instance, predicates like *occludes* or so-called frame fluents). At the outset we made clear that we did not wish to adopt this tactic. To do so would be to call into question our adherence to the quest for conciseness in seeking a solution to the frame problem and, more specifically, the ramification problem. Moreover, the ontological status of added predicates is not always clear and places a huge burden on the designer who must determine whether and when to occlude predicates.

The notion of a Hamiltonian path through E-predecessors in a state transition system to determine possible resultant states is an interesting one. Essentially, a Hamiltonian path serves as a contextual mechanism much in the same way that augmenting the underlying language through the addition of extra predicates does. The additional information allows the effects of causality to contribute in certain situations and not in others.

9 Conclusion and Future Work

In this paper we set out to determine whether it is possible to furnish McCain and Turner's [McCain and Turner, 1995] causal theory of actions with a preferential semantics in the spirit of Shoham [Shoham, 1988]. We demonstrated, through use of an impossibility theorem (Theorem 4.2), that this is not possible in general when we do not embellish the original language $\mathcal{L}$ and assume the preferential ordering satisfies transitivity. Choosing not to abandon preferential semantics entirely, we then adopted an abstract view in terms of state-selection mechanisms introducing two such systems: state elimination systems and state transition systems. The latter of these provides the sought after semantics augmenting a preferential structure based on symmetric difference with a binary relation on states and making use of the notion of a Hamiltonian path. The former provides a stepping stone and, of equal importance, gives further insight into the nature of causality at play in McCain and Turner's approach. Significantly, we show that causal systems, state elimination systems and state transition systems, as defined here, are (selection) equivalent. That state transition systems augment a preferential structure with a binary relation on states demonstrates, we claim, that minimal change and causality — the former captured by preferential semantics and the latter by a binary relation — together are essential in furnishing a concise solution to the frame and ramification problems.

Providing McCain and Turner's causal theory of action with an augmented preferential semantics allows comparison with other logics of action and provides an interesting direction for future work. Another avenue for future work would be a contrast of Sandewall's *causal propagation semantics* [Sandewall, 1996] and our state transition systems. This would indirectly link Sandewall's semantics with McCain and Turner's causal theory of actions giving further insight into causal approaches to reasoning about action. It may also be possible to modify the approach here in terms of state transition systems and Hamiltonian paths to capture other causal approaches such as that of Thielscher [Thielscher, 1997]. Note, however, that Thielscher's system does not satisfy the property that the possible resultant states lie among the E-states. This suggests that such a result is not likely to be a straightforward extension of the current work.

Acknowledgements

The authors would like to thank members of the School of Computing Science at Simon Fraser University for their comments on this work.

References

[Gustafsson and Doherty, 1996] Joakim Gustafsson and Patrick Doherty. Embracing occlusion in specifying the indirect effects of actions. In L. Aiello, J. Doyle, and S. Shapiro, editors, *Proceedings of the Fifth International Conference on Knowledge Representation and Reasoning*. Morgan-Kaufmann, 1996.

[McCain and Turner, 1995] Norman McCain and Hudson Turner. A causal theory of ramifications and qualifications. In *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence*, pages 1978–1984. Montreal, 1995.

[Peppas *et al.*, 1997] Pavlos Peppas, Maurice Pagnucco, Mikhail Prokopenko, and Norman Foo. Preferential semantics for causal fixpoints. In Abdul Sattar, editor, *Proceedings of the Tenth Australian Joint Conference on Artificial Intelligence*, pages 197–206, Perth, Australia, 1997.

[Sandewall, 1996] Erik Sandewall. Assessments of ramification methods that use static domain constraints. In L. Aiello, J. Doyle, and S. Shapiro, editors, *Proceedings of the Fifth International Conference on Knowledge Representation and Reasoning*. Morgan-Kaufmann, 1996.

[Shoham, 1988] Yoav Shoham. *Reasoning About Change*. MIT Press, Cambridge, Massachusetts, 1988.

[Thielscher, 1997] Michael Thielscher. Ramification and causality. *Artificial Intelligence*, 89:317–364, 1997.

[Wilson, 1985] Robin J. Wilson. *Introduction to Graph Theory*. Longman Scientific and Technical, 3rd edition, 1985.

[Winslett, 1988] Marianne Winslett. Reasoning about actions using a possible models approach. In *Proceedings of the Seventh National Artificial Intelligence Conference*, San Mateo, CA., 1988. Morgan Kaufmann Publishers.