A Unifying Semantics for Causal Ramifications

Mikhail Prokopenko[1], Maurice Pagnucco[1], Pavlos Peppas[2], Abhaya Nayak[1]

[1]Computational Reasoning Group
Department of Computing, Macquarie University, NSW 2109, Australia
[2]Sisifou 27, Korinthos 20100, Greece

Abstract. A unifying semantic framework for different reasoning approaches provides an ideal tool to compare these competing alternatives. However, it has been shown recently that a pure preferential semantics alone is not capable of providing such a unifying framework. On the other hand, variants of preferential semantics augmented by additional structures on the state space have been successfully used to characterise some influential approaches to reasoning about action and causality. The primary aim of this paper is to provide an augmented preferential semantics that is general enough to unify two prominent frameworks for reasoning about action and causality — Sandewall's causal propagation semantics [4] and Thielscher's causal relationships approach [5]. There are indications that these and other different augmented preferential semantical approaches can by unified into a general framework, and provide the unified semantics that is lacking so far.

Preferential style semantics have always been seen as a critical step on the way towards a concise solution to the frame and ramification problems. However, it has been argued in recent literature that an explicit representation of causal information is required to solve these problems in a concise manner. It has been shown that some approaches demand a more complex semantics than a pure preferential semantics. For example, McCain and Turner's causal theory of action [1] was recently characterised by an augmented preferential semantics, using an appropriately constructed binary relation on states in addition to a preference relation [2]. This additional relation captured causal context of action systems by translating individual causal laws into state transitions. Another causal theory of action — that of Thielscher [5] — has been characterised by a variant of an augmented preferential semantics [3]. Here the minimality component was complemented by a binary relation on states of higher dimension. The standard state-space of possible worlds was extended to a hyper-space, and action effects (including indirect ones) were traced in the hyper-space. Again, the purpose of these hyper-states was to supply extra context to the process of causal propagation. The hyper-space semantics [3] can be clearly seen to employ a component of minimal change coupled with causality. On the other hand, another rather general semantical approach — the *causal propagation semantics* proposed by Sandewall [4] — deals with causal ramifications without explicitly relying on the principle of minimal change.

This work introduces a preferential style semantics augmented with a causal transition relation on states, that is general enough to unify two of the above-mentioned frameworks to reasoning about action and causality — Sandewall's causal propagation semantics [4] and Thielscher's causal relationships approach [5]. This is achieved by observing that the principle of minimal change is hidden behind action invocation and causal propagation in both proposals.

1 Causal Propagation Semantics

The causal propagation semantics introduced by Sandewall [4], uses the following basic concepts. The set of possible states of the world, formed as a Cartesian product of the finite sets of a finite number of state variables, is denoted as $\mathcal{R}$. E is the set of possible actions. The causal propagation semantics extends a basic state transition semantics with a *causal transition relation*. The causal transition relation C is a non-reflexive relation on states in $\mathcal{R}$. A state r is called *stable* if it does not have any successor s such that $C(r, s)$; we will denote the set of stable states $\{r \in \mathcal{R} : \neg\exists s \in \mathcal{R}, C(r, s)\}$ as $\mathcal{S}_c$. Another component, $\mathcal{R}_c \subseteq \mathcal{S}_c$, is a set of admitted states. Another important concept, introduced by Sandewall, is an *action invocation relation* $G(e, r, r')$, where $e \in E$ is an action, r is the state where the action e is invoked, and r' is "the new state where the instrumental part of the action has been executed" [4]. In other words, the state r' satisfies direct effects of the action e. It is required that every action is always invokable, that is, for every $e \in E$ and $r \in \mathcal{R}$ there must be at least one r' such that $G(e, r, r')$ holds. Of course, this requirement does not mean to guarantee that every action results in an admitted state—on the contrary, the intention is to trace the indirect effects of the action, possibly reaching an admitted (and, therefore, stable) state.

A finite (the infinite case is omitted) transition chain for a state $w \in \mathcal{R}_c$ and an action $e \in E$ is a finite sequence of states $r_1, r_2, ..., (r_k)$, where $G(e, w, r_1), C(r_i, r_{i+1})$ for every $i, 1 \leq i < k$, and where r_k is a stable state. The last element of a finite transition chain is called a result state of action e performed in state w.

These basic concepts define an *action system* as a tuple $\langle \mathcal{R}, E, C, \mathcal{R}_c, G \rangle$. The following definition strengthens action systems based on the causal propagation semantics.

Definition 1. If three states w, p, q are given, we say that the pair p, q respects w, denoted as $\lhd_w(p, q)$, if and only if $p(f) \neq q(f) \rightarrow p(f) = w(f)$ for every state variable f that is defined in $\mathcal{R}$, where $r(f)$ is a valuation of variable f in state r.

An action system $\langle \mathcal{R}, E, C, \mathcal{R}_c, G \rangle$ is called respectful if and only if, for every $w \in \mathcal{R}_c$, every $e \in E$, w is respected by every pair r_i, r_{i+1} in every transition chain for the state w, and the last element of the chain is a member of $\mathcal{R}_c$.

According to Sandewall [4], respectful action systems are intended to ensure that in each transition there cannot be changes in state variables which have changed previously upon invocation or in the causal propagation sequence. This requirement, of course, guarantees that a resultant state is always consistent with the direct effects of the action (which cannot be cancelled by indirect ones), and that there are no cycles in transition chains. As with many other state transition action systems, the intention is to characterise a result state in terms of an initial state w and action e, without "referring explicitly to the details of the intermediate states" [4]. In other words, it is desirable to define a selection function $Res(w, a)$. For a respectful action system $\langle \mathcal{R}, E, C, \mathcal{R}_c, G \rangle$, a selection function can be given as

$$Res_{C\mathcal{R}_cG}(w, e) = \{r_k \in \mathcal{R}_c : G(w, e, r_1), C(r_i, r_{i+1}), \lhd_w(r_i, r_{i+1}), 1 \leq i < k\}.$$

2 Causal Relationships Approach

Let $\mathcal{F}$ be a finite set of symbols from a fixed language $\mathcal{B}$, called fluent names. A fluent literal is either a fluent name $f \in \mathcal{F}$ or its negation, denoted by $\neg f$. Let $L_{\mathcal{F}}$ be the set of all fluent literals defined over the set of fluent names $\mathcal{F}$. We will adopt from Thielscher [5] the following notation. If $\epsilon \in L_{\mathcal{F}}$, then $|\epsilon|$ denotes its affirmative component, that is, $|f| = |\neg f| = f$, where $f \in \mathcal{F}$. This notation can be extended to sets of fluent literals as follows: $|S| = \{|f| : f \in S\}$. By the term *state* we intend a maximal consistent set of fluent literals. We will denote the set of all states as W, and call the number m of fluent names in $\mathcal{F}$ the dimension of W. By $[\phi]$ we denote all states consistent with the sentence $\phi \in \mathcal{B}$ (i.e., $[\phi] = \{w \in W : w \vdash \phi\}$). Domain constraints are sentences which have to be satisfied in all states.

Thielscher's [5] causal theory of action consists of two main components: *action laws* which describe direct effects of action performed in a given state, and *causal relationships* which determine indirect effects of action. Every action law contains a condition C, which is a set of fluent literals, all of which must be contained in an initial state where the action is intended to be applied; and a (direct) effect E, which is also a set of fluent literals, all of which must hold in the resulting state after having applied the action. An action may result in a number of state transitions.

Definition 2. Let $\mathcal{F}$ be the set of fluent names and let $\mathcal{A}$ be a finite set of symbols called action names, such that $\mathcal{F} \cap \mathcal{A} = \emptyset$. An action law is a triple $\langle C, a, E \rangle$ where C, called *condition*, and E, called *effect*, are individually consistent sets of fluent literals, composed of the very same set of fluent names (i.e., $|C| = |E|$) and $a \in \mathcal{A}$. If w is a state then an action law $\alpha = \langle C, a, E \rangle$ is applicable in w iff $C \subseteq w$. The application of α to w yields the state $(w \setminus C) \cup E$ (where $\setminus$ denotes set subtraction).

Causal relationships are specified as ϵ *causes* ρ *if* Φ, where ϵ and ρ are fluent literals and Φ is a fluent formula based on the set of fluent names $\mathcal{F}$.

Definition 3. Let (s, E) be a pair consisting of a state s and a set of fluent literals E. Then a causal relationship ϵ *causes* ρ *if* Φ is applicable to (s, E) iff $\Phi \wedge \neg\rho$ is true in s, and $\epsilon \in E$. Its application yields the pair (s', E'), denoted as $(s, E) \rightsquigarrow (s', E')$, where $s' = (s \setminus \{\neg\rho\}) \cup \{\rho\}$ and $E' = (E \setminus \{\neg\rho\}) \cup \{\rho\}$.

In other words, a causal relationship is applicable if Φ holds, the indirect effect ρ is false and the cause ϵ is among the current effects. A possible *successor state* is determined through repeated application of causal relationships. Specifically, given an initial state w and action a, the set of successor states $Res_{R\mathcal{D}\mathcal{L}}(w, a)$ is determined as follows.

Definition 4. Let $\mathcal{F}$ be the set of fluent names, A a set of action names, $\mathcal{L}$ a set of action laws, $\mathcal{D}$ a set of domain constraints, and R a set of causal relationships. Furthermore, let w be a state satisfying $\mathcal{D}$ and let $a \in A$ be an action name.

A state r is a *successor state* of w and a, denoted $r \in Res_{R\mathcal{D}\mathcal{L}}(w, a)$, iff there exists an applicable (with respect to w) action law $\alpha = \langle C, a, E \rangle \in \mathcal{L}$ such that $((w \setminus C) \cup E, E) \overset{*}{\rightsquigarrow} (r, E')$ for some E', r satisfies $\mathcal{D}$, and $E \subseteq r$, where $\overset{*}{\rightsquigarrow}$ denotes the transitive closure of $\rightsquigarrow$.

The last requirement ($E \subseteq r$) ensures that a (conservative) successor state does not violate action direct effects [3]. As mentioned before, an occurrence of a literal ϵ in a state s does not guarantee that a causal relationship ϵ *causes* ρ *if* Φ is applicable to a pair (s, E) — to ensure applicability, the literal ϵ has to belong to the current effects E. That is why, in order to trace causal propagation with causal relationships, one needs to keep an explicit (and changing) account of context-dependent action effects.

3 General Augmented Preferential Semantics

Let us now consider a general augmented preferential semantics, in terms of state transition systems. This system may be viewed as a tuple $\langle \mathcal{W},\ \Gamma,\ \mathcal{E},\ \mathcal{M},\ \mathcal{W}',\ \mathcal{O}\rangle$, where $\mathcal{W}$ is a set of states; Γ is a set whose elements are referred to as power-states (and Γ itself is referred to as power-space); $\mathcal{E}$ is a set of actions; $\mathcal{M}$ is a binary relation on Γ; $\mathcal{W}' \subseteq \mathcal{W}$ is the set of legitimate (admitted) states, and $\mathcal{O}$ is a set of orderings $<_\gamma$, each with respect to some $\gamma \in \Gamma$.

Some distinguishing features of this system are worth noting. The preference orderings in $\mathcal{O}$ capture minimality whereas the binary relation $\mathcal{M}$ captures causality, thus minimality and causality play distinct roles. In order to provide a concise solution for the frame and ramification problems, both $\mathcal{O}$ and $\mathcal{M}$ are defined over the power-space Γ instead of the normal state-space $\mathcal{W}$. In particular, use of the power space allows us to encode causal contextual information in the power-states themselves, thereby providing a means to avoid encoding of contextual information in (state, history) pairs as in Thielscher [5]. Intuitively, a successor state is an admitted state which is reachable (by means of some transition relation) from states nearest (to the initial one) among states satisfying post-conditions of the performed action. In the current framework, when an action e takes place, instead of using the minimal normal states that satisfy the post-conditions of e as the starting point, we start from the minimal elements (with respect to some $<_\gamma \in \mathcal{O}$) among the power-states whose projections in the normal space $\mathcal{W}$ satisfy the post-conditions of e. The relation $\mathcal{M}$ is then used for causal propagation in the power-space, and the "final" power-states are ultimately projected back to the normal state space in order to determine the successor states resulting from action e.

Let $[e]$ denote a set of states satisfying the post-conditions of an action e, and $\mathcal{P}$ a projection function from Γ to $\mathcal{W}$. A set of power-states $[e]^\Gamma$ is defined as $\{\gamma \in \Gamma : \mathcal{P}(\gamma) \in [e]\}$. In other words, $[e]^\Gamma$ denotes the set of power-states whose normal-space projections make up the set $[e]$. We also define a set $min(<_\gamma, [e]^\Gamma)$ as a subset of $[e]^\Gamma$ containing states nearest to the state γ in terms of the ordering $<_\gamma$. In other words, $min(<_\gamma, [e]^\Gamma) = \{\beta \in [e]^\Gamma, \neg\exists\alpha \in [e]^\Gamma, \alpha \neq \beta, \alpha <_\gamma \beta\}$. Sometimes, we will refer to an element of $min(<_\gamma, [e]^\Gamma)$ as a $<_\gamma$-minimal state in $[e]^\Gamma$.

Let $\mathcal{M}^*$ be a transitive closure of the relation $\mathcal{M}$. We shall say that a power-state β is $\mathcal{M}$-reachable from a power-state α, if $\mathcal{M}^*(\alpha, \beta)$. Finally, let us denote by $\mathcal{K}_\mathcal{M}$ the set $\{p \in \Gamma : \neg\exists q \in \Gamma, \mathcal{M}(p, q)\}$—an obvious counterpart of the set of stable states in the causal propagation semantics. In spirit of the latter we require that any admitted state is a projection of some stable power-state: if $r \in \mathcal{W}'$ then $r = \mathcal{P}(\beta)$ for some $\beta \in \mathcal{K}_\mathcal{M}$. Also, for simplicity, we denote by γ_w any power-state such that $\mathcal{P}(\gamma_w) = w$.

We say that an admitted state r satisfying direct action effects is a successor state, $r \in Res(w, e)$, if and only if r is a projection of some stable power-state β, which is $\mathcal{M}$-reachable from a power-state nearest to γ_w[1] among power-states in $[e]^\Gamma$. More precisely, a selection function of the action system $\langle \mathcal{W},\ \Gamma,\ \mathcal{E},\ \mathcal{M},\ \mathcal{W}',\ \mathcal{O}\rangle$ is given as

$$\begin{aligned} Res_{\mathcal{M}\mathcal{W}'\mathcal{O}}(w, e) = \{r \in \mathcal{W}' \cap [e] :\ & r = \mathcal{P}(\beta),\ \mathcal{M}^*(\alpha, \beta),\ \text{where} \\ & \alpha \in min(<_{\gamma_w},\ [e]^\Gamma)\ \text{and}\ \beta \in \mathcal{K}_\mathcal{M}\}. \end{aligned}$$

[1] It does not matter which of the power-states γ_w such that $\mathcal{P}(\gamma_w) = w$ is chosen.

The power-space concept is not always necessary. Sometimes, we may choose $\Gamma = \mathcal{W}$ and set the projection function as $\mathcal{P}(\lambda) = \lambda$. Then the set $[e]^{\Gamma}$ becomes $[e]$, $\gamma_w = w$, and $\mathcal{W}' \subseteq \mathcal{K}_{\mathcal{M}}$. This essentially specifies an action system $\langle \mathcal{W},\ \mathcal{E},\ \mathcal{M},\ \mathcal{W}',\ \mathcal{O} \rangle$, with the following selection function

$$Res_{\mathcal{M}\mathcal{W}'\mathcal{O}}(w, e) = \{r \in \mathcal{W}' \cap [e] : \mathcal{M}^*(\alpha, r), \alpha \in min(<_w, [e])\}.$$

Following [2], we say that an action system with a function Res_1 is *selection-equivalent* to an action system with a function Res_2 if and only if $Res_1(w, e) = Res_2(w, e)$, for every action e and state w. Now, our goal becomes clear: we intend to find under what conditions it is possible to achieve a selection-equivalence between a generalised action system and action systems based on the causal propagation semantics and the causal relationships approach. More precisely, we wish to identify conditions when $Res_{\mathcal{M}\mathcal{W}'\mathcal{O}}(w, e) = Res_{C\mathcal{R}_cG}(w, e)$ and $Res_{\mathcal{M}\mathcal{W}'\mathcal{O}}(w, e) = Res_{R\mathcal{D}\mathcal{L}}(w, e)$.

3.1 Invoking Minimal Change

In this section, we intend to analyse under what conditions it is possible to represent the Sandewall's causal propagation semantics as an instance $\langle \mathcal{W},\ \mathcal{E},\ \mathcal{M},\ \mathcal{W}',\ \mathcal{O} \rangle$ of the general augmented preferential semantics, where $\Gamma = \mathcal{W}$. This reduction will be carried out while staying within the same sets of states and actions ($\mathcal{W} = \mathcal{R}$, $\mathcal{W}' = \mathcal{R}_c$, and $\mathcal{E} = E$) and transforming the causal transition relation C into the binary relation $\mathcal{M}$. In other words, our primary focus will be discovering the nature of minimality hidden, as we believe, in the invocation relation G. Motivated by a preferential-style semantics, one may be tempted to suggest an ordering on states such that the invocation relation can be simply realised by selecting nearest states satisfying action post-condition. However, this does not appear to be possible without restricting relation G.

Lemma 5. *There is no ordering $<_w$ such that for every action e and state r, $G(e, w, r)$ if and only if $r \in min(<_w, [e])$.*

Consequently, our intention at this stage is to restrict the invocation relation G in such a way that, given an initial state and an action, the invoked states can be characterised precisely as states nearest to the initial one in terms of some appropriate minimality ordering. Before we identify required restrictions on the invocation relation G, we introduce some more abbreviations. If $S \subseteq [a]$, for a set of states S and an action $a \in E$, we call the action a an S-*covering* action. Furthermore, if there exists an S-covering action a such that $G(a, w, x)$ for states $w, x \in \mathcal{R}$, we say that the state x is S-*cover accessible* from state w. Also, we say that a state x is *not* S-*cover accessible* from state w, if there is no S-covering actions a such that $G(a, w, x)$.

Importantly, it follows that all states in a set S satisfy post-conditions of an S-covering action. It is worth pointing out that, given two states p and q satisfying post-conditions of some action a (that is, the action a is a $\{p, q\}$-covering action), the state p may be $\{p, q\}$-cover accessible from some state w, while state q is not $\{p, q\}$-cover accessible from w. The first restriction on invocation relation is given as

(G_1) if p is $\{p, q\}$-cover accessible from w but q is not, and
q is $\{q, x\}$-cover accessible from w but x is not
then p is $\{p, x\}$-cover accessible from w and x is not,
for arbitrary states w, p, q, x.

The premise of the implication is that, considering all actions whose post-conditions are satisfied by two states p and q, state p is chosen at least once by the invocation relation and state q is never chosen; and considering all actions whose post-conditions are satisfied by two states q and x, state q is chosen at least once, while state x is never chosen. This then necessitates that, considering all actions whose post-conditions are satisfied by states p and x, invocation of the state p must eventuate at least once, but state x cannot be invoked at all. Undoubtedly and not surprisingly, the condition (G_1) has a transitive flavour. Another condition is given as

(G_2) Given any two $\{p, q\}$-covering actions e' and e'',
if $G(e', w, p)$ and $G(e'', w, q)$ then $G(e', w, q)$.

This condition simply requires that if neither of two states p and q is chosen over the other in terms of the criterion implicitly used in the condition (G_1), then selection of either of them necessitates selection of the other. Finally, we reinforce the requirement that any action is invokable in principle.

(G_3) $\forall\, e \in E,\ w \in \mathcal{R},\ \exists\, p \in [e],\ G(e, w, p)$

As noted above, this condition does not guarantee that the invoked action will succeed—it may possibly be qualified by causal propagation ending in a non-admitted state. Now, we are ready to describe a set of orderings $\mathcal{O}$ corresponding to the invocation relation. Ideally, any ordering $<_w$ should satisfy only the transitivity property:

(M_1) if $p <_w q$ and $q <_w x$ then $p <_w x$.

However, it turns out that, given an action system, the related ordering has to satisfy, in addition, two other properties.

(M_2) if p is $<_w$-minimal in $[a]$ for some $\{p, q\}$-covering action a and
q is not $<_w$-minimal in $[e]$ for any $\{p, q\}$-covering action e then $p <_w q$.

(M_3) if $p <_w q$ then p is $<_w$-minimal in $[e]$ for some $\{p, q\}$-covering action e.

Basically, the second property (M_2) requires that any state p which is $<_w$-minimal in some set $[a]$ is preferred to any state q, where q belongs to the set $[a]$ as well, and which is not $<_w$-minimal in any set $[e]$ where post-conditions of e are satisfied by both p and q. The third property (M_3) posits that if a state p is preferred to a state q by a preference relation $<_w$, then there must exist an action e, whose post-conditions are satisfied by these two states, such that state p is $<_w$-minimal in $[e]$.

We intend to prove at this stage that there is a way to define the invocation relation in terms of a preference relation and vice versa, while preserving respective selections of states, satisfying direct action effects. The following two definitions will be shown to ensure such an equivalence.

Definition 6. A new invocation relation $G_<$ is defined as follows: $G_<(e, w, r)$ if and only if r is $<_w$-minimal in $[e]$, where $w, r \in \mathcal{R}, e \in E$.

Put simply, the new relation $G_<(e, w, r)$ specifies states r that are nearest among all states in $[e]$ to the initial state w, where the action e was invoked.

Definition 7. Given an invocation relation G, for each $w \in \mathcal{R}$ we define an ordering $<_{w,G}$ on states in $\mathcal{R}$ as follows: $p <_{w,G} q$ if and only if state p is $\{p, q\}$-cover accessible from w and state q is not $\{p, q\}$-cover accessible from w.

This definition specifies a preference relation on states driven by a given invocation relation—state p is nearer to an initial state w than state q if and only if for all actions whose direct effects are satisfied by both states p and q, the state q is never selected by the invocation relation G, while state p is selected at least once. The following lemma establishes the sought-after equivalence between invocation and preference relations.

Lemma 8. *If the relation G satisfies the conditions $(G_1) - (G_3)$, then for each $w \in \mathcal{R}$, the ordering $<_{w,G}$ satisfies conditions $(M_1) - (M_3)$.*
If each ordering $<_w$ for $w \in \mathcal{R}$ satisfies conditions $(M_1) - (M_3)$, then the relation $G_<$ satisfies the conditions $(G_1) - (G_3)$.

Having established the role of minimality in the process of action invocation, we shall analyse actual propagation in the state-space.First of all, it is interesting to observe that the respectfulness requirement in terms of states, is related to the notion of minimality as well. More precisely, the former can be achieved by a preference relation on states. We shall say that a state x is preferred to a state y in terms of the PMA ordering [6], denoted $x \prec_w y$, if and only if $Diff(x, w) \subset Diff(y, w)$, where $Diff(p, q)$ represents the symmetric difference of p and q, i.e., $(p \setminus q) \cup (q \setminus p)$. Formally, the following observation holds.

Lemma 9. *The pair p, q respects w, $\lhd_w(p, q)$, if and only if $p \prec_w q$ in the PMA ordering $\prec_w$ associated with w.*

This connection indicates a way to capture $respectful$ action systems, given a system based on general augmented preferential semantics. The following two conditions further constrain orderings in $\mathcal{O}$.

(M_4) if $p \prec_w q$ then $p <_w q$.

The additional condition (M_4) ensures that an ordering $<_w$ incorporates the PMA ordering, or, in other words, includes all pairs p, q such that $p \prec_w q$. Our next condition relates to a connectivity of the set $[e]$ in terms of an ordering $\prec_w$.

(M_5) For every action e and state w, the set $min(\prec_w, [e])$ is a singleton.

Now we are ready to specify conditions ensuring desired selection-equivalence.

Theorem 10. *For every respectful action system $\langle \mathcal{R}, E, C, \mathcal{R}_c, G\rangle$ there exists a selection-equivalent action system $\langle \mathcal{W}, \mathcal{E}, \mathcal{M}, \mathcal{W}', \mathcal{O}\rangle$, if the relation G satisfies conditions $(G_1)-(G_3)$. Conversely, for every action system $\langle \mathcal{W}, \mathcal{E}, \mathcal{M}, \mathcal{W}', \mathcal{O}\rangle$ there exists a selection-equivalent respectful action system $\langle \mathcal{R}, E, C, \mathcal{R}_c, G\rangle$, if the orderings in $\mathcal{O}$ satisfy conditions $(M_1) - (M_5)$.*

Basically, this observation establishes that, under considered conditions, we obtain

$$Res_{C\mathcal{R}_cG}(w, e) = Res_{\mathcal{M}\mathcal{W}'\mathcal{O}}(w, e)$$

for any action e and state w. The observed selection-equivalence has been achieved without extending the state-space, action domain, or causal transition relation—indicating that minimal change is present in the causal propagation semantics masked by the invocation relation.

3.2 Propagating Minimal Change with Causal Relationships

While the hyper-state space, proposed in [3], is a powerful concept that would allow us to completely characterise Thielscher's approach, we shall now introduce another novel concept that abstracts away certain elements of the hyper-state space semantics. It is this *power-state space* proposal that we believe can be unified with Sandewall's causal propagation semantics under the general augmented preferential semantics.

First of all, we need to re-iterate the main idea behind the hyper-state space semantics. This proposal augmented the underlying language by adding to the set of fluent names $\mathcal{F}$, and constructing the set of *justifier fluents* $\overset{\circ}{\mathcal{F}}$ which has the same cardinality as $\mathcal{F}$. The justifier fluents maintain contextual information that becomes important during causal propagation. The set of *justifier literals* $\overset{\circ}{L}_{\mathcal{F}} = \overset{\circ}{\mathcal{F}} \cup \{\neg q : q \in \overset{\circ}{\mathcal{F}}\}$ is mapped from $L_{\mathcal{F}}$ by the function $l : L_{\mathcal{F}} \to \overset{\circ}{L}_{\mathcal{F}}$ which intuitively provides an added space-dimension corresponding to each fluent $f \in \mathcal{F}$. We will use the abbreviation $\overset{\circ}{f}$ instead of $l(f)$ for simplicity. In addition, a justifier set $\overset{\circ}{J}$ for any set of fluent literals $J \subseteq L_{\mathcal{F}}$ is defined as $\overset{\circ}{J} = \cup_{f \in J}\{l(f)\} = \cup_{f \in J}\{\overset{\circ}{f}\}$. These constructions allow us to state more precisely what is meant by a hyper-state.

Definition 11. Given a set of fluents $\mathcal{F}$, a *hyper-state* is a maximal consistent set of literals from $L_{\mathcal{F}} \cup \overset{\circ}{L}_{\mathcal{F}}$.

That is, we produce "clones" or copies of all the fluent names in our language and use this expanded language in forming (hyper-)states. We will denote the set of all hyper-states as Ω, where the dimension of Ω is $2m$, m being the dimension of W. The following two functions map hyper-state space Ω to normal space W and vice versa.

Definition 12. A projection from Ω to W, $p : \Omega \to W$, is the function mapping a hyper-state $s = \{f_1, ..., f_m, \overset{\circ}{f}_1, ..., \overset{\circ}{f}_m\} \in \Omega$ to a state $r = \{f_1, ..., f_m\} \in W$.

Definition 13. A hyper-neighbourhood of a state $r \in W$, $N : W \to 2^{\Omega}$, is the function mapping a state r to a set of hyper-states: $N(r) = \{s \in \Omega : r = p(s)\}$.

Intuitively, justifier literals represent explicit causes for a state $r \in W$, and the set $N(r)$ contains states where all possible causes (i.e., justifier literals) vary, while the (proper) literals defined on $\mathcal{F}$ are fixed. We denote the hyper-part of a hyper-state $s \in \Omega$ as $h(s) = s \setminus p(s)$. Before we formally introduce the required notion of a binary causal relation on hyper-states, let us illustrate its purpose.

Suppose we have an action system with $\mathcal{F} = \{a, b, c\}$, $\mathcal{D} = \{\neg b \to \neg a\}$, $R = \{\neg b \textit{ causes } \neg a \textit{ if } \top\}$, and $\mathcal{L} = \{\langle\{b\}, x, \{\neg b\}\rangle\}$. Let us consider action x executed at initial state $w = \{a, b, c\}$. The action's direct effect is $\{\neg b\}$, yielding the intermediate state $\{a, \neg b, c\} = (w \setminus \{b\}) \cup \{\neg b\}$. This state contradicts the given domain constraint. However, the system's sole causal law applies: $(\{a, \neg b, c\}, \{\neg b\}) \leadsto (\{\neg a, \neg b, c\}, \{\neg a, \neg b\})$. The state component of the resultant pair obeys the domain constraint and satisfies direct effect. Therefore it is an element of $Res_{R\mathcal{D}\mathcal{L}}(w, a)$. It can be verified that $Res_{R\mathcal{D}\mathcal{L}}(w, a)$ is a singleton. We now indicate how this propagation can be traced in the hyper-state space. The hyper-neighbourhood $N(r)$ of the intermediate state $r = \{a, \neg b, c\}$ contains eight hyper-states (see Figure 1). Some of these represent the initial history component $\{\neg b\}$ — these hyper-states are exactly those in $N(r) \cap [\neg \mathring{b}]$. The hyper-neighbourhood of the successor state $r' = \{\neg a, \neg b, c\}$ contains some hyper-states accountable for the final history component $\{\neg a, \neg b\}$. These states are exactly those in $N(r') \cap [\neg \mathring{a} \wedge \neg \mathring{b}]$ The idea, then, is to construct just such a binary relation on hyper-states for an action system so that transitions in hyper-state space correspond to causal propagation.

Definition 14. A binary relation $\mathcal{C}$ is defined on $\Omega \times \Omega$. We say that $\mathcal{C}(s, s')$ if and only if there exists a causal relationship ϵ *causes* ρ *if* Φ such that

$$\begin{array}{lll} p(s) \vdash \epsilon \wedge \Phi \wedge \neg \rho & \text{and} & p(s') = (p(s) \setminus \{\neg \rho\}) \cup \{\rho\} \\ h(s) \vdash \mathring{\epsilon} & \text{and} & h(s') = (h(s) \setminus \{\neg \mathring{\rho}\}) \cup \{\mathring{\rho}\}. \end{array}$$

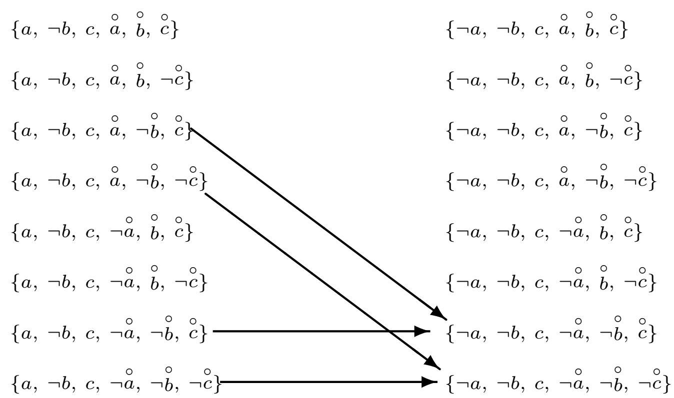

Fig. 1. The $\mathcal{C}$-links between hyper-neighbourhoods of the states $r = \{a, \neg b, c\}$ and $r' = \{\neg a, \neg b, c\}$, generated by a causal relationship $\neg b$ *causes* $\neg a$ *if* $\top$.

The fact that all the states in $N(r) \cap [\neg \mathring{b}]$ have $\mathcal{C}(s, s')$-links to the states in $N(r') \cap [\neg \mathring{a} \wedge \neg \mathring{b}]$ is not a coincidence, and is formally captured by a so-called *power-state space* semantics which allows us to concentrate more on the actual causal propagations occurring between one hyper-neighbourhood and another. First, however, the following definition will prove convenient.

Definition 15. For a state $q \in W$ and a set $z \subseteq N(q)$, a *partial* state $\gamma_q(z)$ is defined as $\cap_{s \in z} s$.

For example, $r = \{a, \neg b, c\}$ and $z = N(r) \cap [\neg \mathring{b}]$ yield a partial state $\gamma_r(z) = \{a, \neg b, c, \neg \mathring{b}\}$. Now let us consider a set Γ of cardinality equal to that of 2^{Ω}. We define a mapping $\gamma : 2^{\Omega} \to \Gamma$, such that $\gamma(z) = \gamma_q(z)$ if $z \subseteq N(q)$ for some $q \in W$, and $\gamma(z) = \emptyset$ otherwise. This set Γ will be referred to as *power-state space*, being isomorphic to the power set of the hyper-state space Ω. Having defined the function $\gamma(z)$ for every subset of the hyper-state space Ω, we construct a binary relation on elements of Γ.

Definition 16. A binary relation $\rightharpoonup$ is defined on $\Gamma \times \Gamma$. Given two elements $x_1, x_2 \in \Gamma$ such that $x_1 \neq \emptyset$ and $x_2 \neq \emptyset$, we say that $x_1 \rightharpoonup x_2$ if and only if $x_1 = \gamma(z_1)$ and $x_2 = \gamma(z_2)$ for some $z_1, z_2 \in 2^{\Omega}$ such that

$$\forall\, s \in z_1, \exists\, s' \in z_2, \text{such that}\, \mathcal{C}(s, s') \quad \text{and} \quad \forall\, s' \in z_2, \exists\, s \in z_1, \text{such that}\, \mathcal{C}(s, s').$$

We will abbreviate $x_1 \rightharpoonup x_2$, where $x_1 = \gamma(z_1)$ and $x_2 = \gamma(z_2)$, as $\gamma(z_1) \rightharpoonup \gamma(z_2)$. Intuitively, $\gamma(z_1) \rightharpoonup \gamma(z_2)$ means that there are no hyper-states in z_1 without an outcoming $\mathcal{C}-$link to some hyper-state in z_2, and there are no hyper-states in z_2 without an incoming $\mathcal{C}-$link from some hyper-state in z_1. Figure 1 exemplifies that $\gamma(N(r) \cap [\neg \mathring{b}]) \rightharpoonup \gamma(N(r') \cap [\neg \mathring{a} \wedge \neg \mathring{b}])$. It is easy to verify that the defined relation $\rightharpoonup$ is transitive. It is precisely the binary relation $\rightharpoonup$ that captures causal propagation in Thielscher's system. We begin by introducing a few notions that will be useful in analysing causal links $\mathcal{C}(s,\ s')$ and $x \rightharpoonup x'$.

Definition 17. A *trigger set* of states, denoted $\|E\|_w$, is defined for an initial state $w \in W$ and an action a, where $\langle C, a, E\rangle$ is the action law, as

$$\{s \in N(q) : q \in W, q \in min(\prec_w, [a]), h(s) \vdash \mathring{E}\}$$

where $\prec_w$ is the PMA ordering, and $[a]$ stands for a set of states consistent with $\wedge E$ (a conjunction of all literals in E).

That is, in terms of the PMA ordering, $\|E\|_w$ is the set contained in the hyper-neighbourhood $N(q)$ of state q nearest to the initial state w, and the states $s \in \|E\|_w$ represent the initial causal context, i.e., initial causally justified changes triggered by effects E. For instance, if an action law $\langle\{b\}, x, \{\neg b\}\rangle$, is applied to the initial state $\{a, b, c\}$, then the trigger set $\|\{\neg b\}\|_{\{a,b,c\}}$ contains exactly those states which happen to have out-coming $\mathcal{C}$-links in Figure 1. We can view changes triggered by the set $\|E\|_w$ as propagating in hyper-state space towards a hyper-neighbourhood of a possible successor state. The point where this propagation ends can now be defined explicitly.

Definition 18. A state $s \in \Omega$ is *final* iff $\{s' : \mathcal{C}(s, s')\} = \emptyset$.
A state $x \in \Gamma$ is *final* iff $\{x' : x \rightharpoonup x')\} = \emptyset$.

We are now in a position to define a selection function for power-space semantics specifying the set of possible successor states $Res_{\Gamma}(w, a)$. We shall see that this completely characterises Thielscher's resultant state set $Res_{\mathcal{RDL}}(w, a)$.

Definition 19. Let $\mathcal{F}$, A, $\mathcal{L}$, w, $\langle C, a, E\rangle$ be the same as in definition 4, Γ the set of power-states, and $\rightharpoonup$ binary relation defined by Definition 16. A state $r \in W$ is a *successor state* of w and a, that is, $r \in Res_{\Gamma}(w, a)$, if and only if $\gamma(\|E\|_w) \rightharpoonup \gamma(z)$, where $\gamma(z)$ is final and $p(z) = r$.

Intuitively, the causal propagation in power-state space starts in a power-state which corresponds to a trigger set of hyper-states and ends in a final power-state corresponding to a set transitively reachable from *all* initial justifier literals. In other words, this process propagates "minimal change" within a set of possible states of higher dimensions, instead of keeping an explicit (and changing) account of context-dependent action effects. The second central result of this study can now be obtained.

Theorem 20. $Res_{\mathcal{RDL}}(w, e) = Res_{\Gamma}(w, e)$.

It is quite clear that the power-space semantics is an instance of the general augmented preferential semantics. Specifically, we just need to construct an action system $\langle \mathcal{W},\ \Gamma,\ \mathcal{E},\ \mathcal{M},\ \mathcal{W}',\ \mathcal{O} \rangle$, where $\mathcal{E} = A$, $\mathcal{M} = \rightharpoonup$, and $\mathcal{W}'$ is a subset of $\mathcal{W}$ such that its elements satisfy constraints $\mathcal{D}$. The projection function $\mathcal{P}$ is defined via the hyper-space projection function p as $\mathcal{P}(\gamma(z)) = p(s)$, where $s \in z$. The set $\mathcal{O}$ is a set of orderings $\ll$ defined on power-states in such a way that respective projections satisfy the PMA ordering while preferring maximal subsets within each hyper-neighbourhood — more precisely, $\gamma(z_1) \ll_{\gamma(x)} \gamma(z_2)$ if and only if $p(z_1) \prec_{p(x)} p(z_2)$ and $z_1 \supseteq z_2$. This tiered preference relation ensures that, given an initial state w and an action law $\langle C, a, E \rangle$, the power-state $\gamma(\|E\|_w)$ is $\ll_{\gamma(N(w))}$-minimal state among all power-states in $[a]^{\Gamma}$.

4 Discussion and Conclusions

In this paper we considered a general augmented preferential semantics for reasoning about action and causality. This semantics appears to be intuitively simple and involves, as the only components, a state-space, its admitted subset of legitimates states, a power-space, an action domain, a (causal) binary relation on states, and a uniform set of orderings. Varying all these components allows us to specify different instances of the framework. For example, a pure preferential semantics can be obtained by requesting that the causal relation is an empty set. The nature of the distinction between Sandewall's and Thielscher's approaches to propagation-oriented ramification is uncovered and reduced to the variance in transition space dimensions and an employment of different preference metrics when identifying states nearest to the initial one. In light of the general semantics, both of these cases can be found to be very similar to McCain and Turner's causal theory of action using causal fix-points [1]. This was also characterised by an augmented preferential semantics relying on the PMA ordering and an appropriate binary relation on normal state-space [2]. The main difference is that the treatment in [2] requires a *Hamiltonian* path through certain states in a state transition system leading to a McCain and Turner causal fix-point. Essentially, such a Hamiltonian path serves as a context-oriented mechanism: the effects of causality are allowed to contribute in certain situations and not in others. Formal capture of such additional contextual requirements will be the subject of future research.

References

1. Norman McCain and Hudson Turner. A causal theory of ramifications and qualifications. In *Proceedings of the 14th Int. Joint Conf. on Artificial Intelligence*, pages 1978–1984. 1995.
2. Pavlos Peppas, Maurice Pagnucco, Mikhail Prokopenko, Abhaya Nayak and Norman Foo. Preferential Semantics for Causal Systems. In *Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence*, pages 118–123. Stockholm, 1999.
3. Mikhail Prokopenko, Maurice Pagnucco, Pavlos Peppas, Abhaya Nayak. Causal Propagation Semantics — A Study. In *Proceedings of the 12th Australian Joint Conference on Artificial Intelligence*, pages 378–392. Sydney, 1999.
4. 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.
5. Michael Thielscher. Ramification and Causality. Artificial Intelligence 89: 317–364, 1997.
6. Marianne Winslett. Reasoning about actions using a possible models approach. In *Proceedings of the Seventh National Artificial Intelligence Conference*, San Mateo, CA, 1988.