

# A Semantics for Means-End Relations

Jesse Hughes ([j.hughes@tm.tue.nl](mailto:j.hughes@tm.tue.nl))

*Technical University of Eindhoven*

Peter Kroes ([p.a.kroes@tbm.tudelft.nl](mailto:p.a.kroes@tbm.tudelft.nl))

*Technical University of Delft*

Sjoerd Zwart ([s.d.zwart@tm.tue.nl](mailto:s.d.zwart@tm.tue.nl))

*Technical University of Eindhoven and Technical University of Delft*

**Abstract.** There has been relatively much work on practical reasoning in artificial intelligence and also in philosophy. Typically, such reasoning includes premises regarding means-end relations. A clear semantics for such relations is needed in order to evaluate proposed syllogisms. In this paper, we provide a formal semantics for means-end relations, in particular for necessary and sufficient means-end relations. Our semantics includes a non-monotonic conditional operator, so that related practical reasoning is naturally defeasible. This work is primarily an exercise in conceptual analysis, aimed at clarifying and eventually evaluating existing theories of practical reasoning (pending a similar analysis regarding desires, intentions and other relevant concepts).

**Keywords:** Means-end relations, propositional dynamic logic, formal semantics, practical reasoning

“They were in conversation without speaking. They didn’t need to speak. They just changed reality so that they had spoken.”  
Terry Pratchett, *Reaper Man*

## 1. Introduction

The aim of this paper is to improve our understanding of means-end relations in practical reasoning. We take practical reasoning to be the process of deriving prescriptions for actions, typically from premises including means-end relations. Traditionally, practical reasoning has been closely related to ethical theory. Stephan Darwall, for instance writes in (1996) that, “Practical-reasoning theory is a kind of metaethical view ... that aims to understand ethics as rooted in practical reason.” This tradition, however, has changed lately and our conception of practical reasoning is closer to that of Elijah Millgram in (2004): “The current debate in practical reasoning focuses on the question: what inference patterns are legitimate methods of arriving at decisions or intentions to act, or other characteristically practical predecessors of actions such as evaluations, plans, policies, and judgments about what one ought to do?” Practical reasoning and the use of means-end relations are integral aspects of linguistic practices in general, and in particular, of engineering practices. We want to contribute to the understanding of these practices by providing a clear analysis of means-end relations.

The broad topic of practical reasoning has been the focus of much attention in artificial intelligence, largely starting with the seminal paper of (McCarthy and Hayes, 1969). This work aims at producing software agents capable of attaining goals by choosing appropriate actions. Some of this work has been cast explicitly in terms of means-end relations, including the engineering perspective presented in (Bratman



et al., 1988) and John Pollock’s work in (2002). Central issues include the epistemological problem of knowledge representation, and the heuristical problem regarding decision making and achieving goals, discussed in (McCarthy, 1999).

Our work is primarily inspired by an older tradition in philosophical circles, namely the investigation of practical syllogisms, dating back to Aristotle and enjoying renewed interest due to the work of Georg Henrik von Wright (1963). We are particularly indebted to von Wright’s work and take his analysis as a model for our semantics. Broadly, such syllogisms typically involve premises like the following:

- (1) an assertion that an agent  $A$  desires some *end*  $\varphi$ ,
- (2) an assertion that (possibly given some precondition  $\psi$ ) the action  $\alpha$  is related to the realization of  $\varphi$ ,
- (3) an assertion of some factual matter, such as that the precondition  $\psi$  is true.

Premises of type (2) express causal relations about the world (or, perhaps, *beliefs* about causal relations). Such premises are essential to practical reasoning, since they give the motivational force for the argument. The reason to *do* the action  $\alpha$  is that it is related in the right way to the desired condition  $\varphi$ . Because one wants  $\varphi$  to be realized, he will be motivated to do  $\alpha$ . We call such premises (*conditional*) *means-end relations*, since they assert that the action  $\alpha$  is a *means* to the *end*  $\varphi$ .

Our working example of such syllogisms comes from (von Wright, 1963), stated in the third-person form here.

$A$ wants to make the hut habitable.	
Unless $A$ heats the hut, it will not become habitable.	
Therefore $A$ must heat the hut.	

Such syllogisms conclude either in an action, an intention to act or a prescription to act, depending on the author. The premises are supposed to be sufficient to justify the conclusion, of course, but how should one evaluate this claim? For this, one must have an unambiguous understanding of each of the premises involved. A clear means-end semantics, however, seems to be largely lacking in the literature. We aim to contribute by providing a semantics for means-end relations which may help to evaluate practical syllogisms. We also hope that our analysis helps clarify AI work on means-end relations, but our focus is primarily the premises of practical syllogisms. We have chosen to present a formal semantics, because such formalisms help ensure clarity and allow one to indicate precisely which features are taken to be relevant to the analysis. Any such formalization requires some idealization as well as deviation from natural language expressions<sup>1</sup>, but we hope that our semantics reasonably approximates natural language in the relevant features.

Means-end talk in natural language exhibits many different features and our aim is to represent these features as faithfully as possible in our formal system. We present the following list of features of means-end talk that we wish to address, but we make no claims regarding completeness of this list.

- (1) the semantic distinction between means and ends
- (2) an end in one context may be a means in another

- (3) the distinction between necessary and sufficient means
- (4) the causal impact of means
- (5) means-end conditionals are defeasible, which relates to the so-called ramification problem
- (6) a means may be an end in itself, rather than a means for some distinct end
- (7) entities of different types, such as objects and actions, may constitute means
- (8) the distinction between effective and efficient means
- (9) the distinction between good and bad means

These features should be taken as *prima facie* features of means-end talk. Our formal reconstruction may show that some of these features have to be reconsidered due to the vagueness and sloppiness of natural language. For instance, feature (1) suggests that means and ends are distinct types, while features (2) and (6) claim that means can be ends and ends can be means. Presumably, any formal semantics will be unable to simultaneously satisfy these three features, but we accept this fact. While we aim to provide a semantics that approximates the natural language meaning of means-end relations, we are also prepared to differ from natural language usage at times.

In the end, our semantics satisfies or explains features (1)–(5). We believe we can give a good account on (6) and (7) and our work on efficacy in (Hughes et al., 2005) forms a partial analysis of (8), but we save these considerations for a later day.

Our most fundamental contribution comes in Sections 2 and 3, where we provide the basic semantics for *local* means-end relations — relations which express the sufficiency/necessity of a means to an end in *this* world. We use models of *Propositional Dynamic Logic* (PDL) (Harel, 1984) to provide the setting for these means-end relations. PDL is a natural setting for means-end relations, since it is motivated by reasoning about outcomes of actions. PDL has long been used for reasoning about program correctness but also has a healthy tradition in current AI research, surveyed in (Meyer, 2000), with additional examples in (Castilho et al., 2002; Castilho et al., 1999; Giacomo and Lenzerini, 1995; Giordano et al., 2000; Prendinger and Schurz, 1996; Zhang and Foo, 2002; Zhang and Foo, 2005). But where this work is concerned with feasibly deriving plans from goals or defeasibly deducing consequences of actions given partial information, we are interested in the conceptual analysis of certain natural language expressions via formal semantics.

An alternative tradition for practical reasoning involves temporal logic. Recently, Mark Brown (2005) suggested a means-end semantics involving such logics with *stit* (see-to-it-that) operators (see also (Horty and Belnap, 1995)). His logic includes sophisticated temporally defined ends, such as making  $\varphi$  true for a certain period, attainable for some time in the future and so on, and these are useful features lacking in our present account. However, he identifies means to an end as certain formulas expressing ability, which does not seem quite right. Indeed, since his logic has no place for actions as syntactic entities, it's hard to see how it can represent means at all. Rather, it seems closer to a logic of ability (very different from his prior account in (1988)). Thus, despite the attractive features of Brown's use of temporal logic, we prefer PDL

for means-end semantics, since we are committed to means as actions and, as Meyer says (2000), in PDL actions appear as “first-class citizens”.

Following our discussion of local means-end relations, we introduce conditional relations in Section 4. Conditional means-end relations more closely approximate natural language usage and introduce certain epistemological issues. A typical agent will not know which world is the actual world, but instead reasons about means to an end given certain features he believes to be true of the actual world. A conditional operator serves to represent this limitation. Moreover, due to well-known issues in practical reasoning (notably the ramification problem), we prefer a *non-monotonic* conditional operator. We sketch some of the features that such an operator should have and in Section 5, we apply this operator to yield sufficient and necessary preconditions for various conditional means-end relations.

## 2. Means-end relations in PDL

Means-end reasoning is about the adjustment of the actual world to realize a sought-after situation that may fail to be the case at present. Consequently, it concerns doing something that brings about a change in the present state of affairs such that some sentence  $\varphi$  describing this favorable end will be true. As our description suggests, we find possible world semantics to be an appropriate setting for reasoning about propositions with changing truth values. Then a means to  $\varphi$  is a way to change the current actual world to a world in which  $\varphi$  is realized. Clearly, a means involves a transition to some  $\varphi$ -world and, inasmuch as the agent may choose to make the change or not, it is natural to think of means as actions in a dynamic logic. Hence, we follow Krister Segerberg’s suggestion in (1992) and choose Propositional Dynamic Logic as our basic setting.

There are alternatives to PDL that may serve for a means-end semantics, including temporal logic (applied to means-end relations in (Brown, 2005)) and the modal  $\mu$ -calculus. The former does not seem well suited for our application, since it does not naturally include a syntax for actions. The latter is better suited, since it combines the explicit actions of PDL with many of the fixed point operators in temporal logic. Although some of these operators (while, until, etc.) may be useful for understanding complex ends, we felt that the simplicity of PDL sufficed for an introduction to means-end semantics.

Similarly, on grounds of simplicity, we rejected expansions of PDL found in the AI literature, including: EPDL used in (Zhang and Foo, 2001), *DIFR* in (Giacomo and Lenzerini, 1995) and AD in (Giordano et al., 2000). Because we are currently interested in practical syllogisms instead of automated agent reasoning, the motivations for these extensions are less pressing on our investigation. We prefer to focus presently on what means-end relations *mean* than to consider how to build agents capable of constructing plans to achieve goals.

### 2.1. PROPOSITIONAL DYNAMIC LOGIC

PDL is a logic of actions, typically used to reason about computer program behavior. It is a multi-modal propositional language where each atomic action corresponds to an

accessibility relation. The strong modal operator  $[\alpha]\varphi$  expresses that  $\varphi$  will necessarily be realized after performing  $\alpha$  (with no commitment that  $\alpha$  *can* be performed). The weak modal operator  $\langle\alpha\rangle\varphi$  is defined by conjugation as usual and means that it is possible to realize  $\varphi$  by doing  $\alpha$ . We refer the reader to (Harel, 1984), from which we take much of the following material. We simplify our presentation by omitting the iteration operation  $\alpha^*$ . For our introduction to means-end semantics, iteration is more distracting than necessary.

The syntax of PDL is based on two disjoint types: the set  $\Pi_0$  of atomic actions and the set  $\Phi_0$  of atomic propositions. From these two sets, we inductively define the sets  $\Pi$  of actions and  $\Phi$  of formulas as follows

- $\{\top\} \cup \Phi_0 \subseteq \Phi$ ;
- if  $\varphi, \psi \in \Phi$  then  $\neg\varphi$  and  $\varphi \wedge \psi$  are in  $\Phi$ ;
- if  $\alpha \in \Pi$  and  $\varphi \in \Phi$  then  $[\alpha]\varphi \in \Phi$ ;
- $\Pi_0 \subseteq \Pi$ ;
- if  $\alpha, \beta \in \Pi$  then  $\alpha;\beta$  and  $\alpha \cup \beta$  are in  $\Pi$ .
- if  $\varphi \in \Phi$  then  $\varphi? \in \Pi$ .

We introduce the propositional constant  $\perp$ , the connectives  $\neg$ ,  $\vee$  and  $\rightarrow$  and the weak operator  $\langle\alpha\rangle$  as usual. The action constructors are intended thus: the semicolon denotes sequential composition, the union  $\alpha \cup \beta$  of actions represents non-deterministic choice between  $\alpha$  and  $\beta$  and the test operator  $\varphi?$  allows one to form conditional actions dependent on the truth value of  $\varphi^2$ .

A *PDL frame*  $\mathcal{F}$  for  $\Pi_0$  consists of a set  $\mathcal{W}$  of worlds (or states) and a *dynamic interpretation*  $\llbracket - \rrbracket^{\mathcal{F}} : \Pi_0 \rightarrow (\mathcal{P}\mathcal{W})^{\mathcal{W}}$  of actions via non-deterministic transition systems. Here  $\mathcal{P}$  denotes the powerset functor and exponentiation  $A^B$  denotes the set of functions  $B \rightarrow A$ . Consequently, the interpretation  $\Pi_0 \rightarrow (\mathcal{P}\mathcal{W})^{\mathcal{W}}$  assigns to each  $m \in \Pi_0$  a function  $\llbracket m \rrbracket^{\mathcal{F}} : \mathcal{W} \rightarrow \mathcal{P}\mathcal{W}$ . For  $w \in \mathcal{W}$ , we interpret  $\llbracket m \rrbracket^{\mathcal{F}}(w)$  as the set of possible (or “normal” or “reasonably expected” or ...) outcomes of doing  $m$  in  $w$ . Clearly, a PDL frame is just the same as a labeled transition system with nodes  $w \in \mathcal{W}$  and labels  $m \in \Pi_0$ . We sometimes write  $w \xrightarrow{m} w'$  for  $w' \in \llbracket m \rrbracket^{\mathcal{F}}(w)$ .

A *PDL model* is a frame  $\mathcal{F}$  together with a *valuation*  $\llbracket - \rrbracket^{\mathcal{M}} : \Phi_0 \rightarrow \mathcal{P}\mathcal{W}$  of atomic propositions. We abuse notation by adopting Scott brackets for both the valuation of atomic propositions and the interpretation of atomic actions, but since our sets  $\Pi_0$  and  $\Phi_0$  are disjoint, no confusion should result. Hereafter, we omit the superscripts  $\mathcal{F}$  and  $\mathcal{M}$  and hope that context makes our meaning clear. A valuation assigns to each atomic proposition  $P \in \Phi_0$  a set  $\llbracket P \rrbracket \subseteq \mathcal{W}$  of worlds. We interpret  $\llbracket P \rrbracket$  as the set of worlds in which  $P$  is true. We extend the valuation of atomic propositions to a function  $\llbracket - \rrbracket : \Phi \rightarrow \mathcal{P}\mathcal{W}$  and the interpretation of atomic actions to a function  $\llbracket - \rrbracket : \Pi \rightarrow (\mathcal{P}\mathcal{W})^{\mathcal{W}}$  recursively as shown in Table I.

We say that  $w$  *satisfies*  $\varphi$  or that  $\varphi$  *is true* in  $w$  just in case  $w \in \llbracket \varphi \rrbracket$ . In this case, we write  $\mathcal{M}, w \models \varphi$  or just  $w \models \varphi$  when  $\mathcal{M}$  is understood by context. We write  $\mathcal{M} \models \varphi$  if for every  $w \in \mathcal{W}$  we have  $w \models \varphi$  and we write  $\models \varphi$  if  $\mathcal{M} \models \varphi$  for every model  $\mathcal{M}$ . In this case, we say that  $\varphi$  is *valid*.

Table I. Extension of valuation to  $\Phi$  and interpretation to  $\Pi$ .

<u>On formulas</u>
$\llbracket \top \rrbracket = \mathcal{W}$
$\llbracket \neg\varphi \rrbracket = \mathcal{W} \setminus \llbracket \varphi \rrbracket$
$\llbracket \varphi \wedge \psi \rrbracket = \llbracket \varphi \rrbracket \cap \llbracket \psi \rrbracket$
$\llbracket [\alpha]\varphi \rrbracket = \{w \in \mathcal{W} \mid \llbracket [\alpha] \rrbracket(w) \subseteq \llbracket \varphi \rrbracket\}$
<u>On actions</u>
$\llbracket \alpha; \beta \rrbracket(w) = \{w' \in \mathcal{W} \mid \exists w'' \in \mathcal{W} . w'' \in \llbracket [\alpha] \rrbracket(w) \text{ and } w' \in \llbracket [\beta] \rrbracket(w'')\}$
$\llbracket \alpha \cup \beta \rrbracket(w) = \llbracket [\alpha] \rrbracket(w) \cup \llbracket [\beta] \rrbracket(w)$
$\llbracket \varphi? \rrbracket(w) = \begin{cases} \{w\} & \text{if } w \in \llbracket \varphi \rrbracket; \\ \emptyset & \text{else.} \end{cases}$

We call an action  $\alpha$  *prohibited in  $w$*  if  $\llbracket [\alpha] \rrbracket(w) = \emptyset$ . Intuitively, such actions cannot be performed in  $w$ . If  $\alpha$  is prohibited in  $w$ , then  $w \models [\alpha]\varphi$  for any  $\varphi \in \Phi$  (including  $\perp$ ), but  $w \not\models \langle \alpha \rangle \varphi$  for any  $\varphi \in \Phi$  (not even  $\top$ ).

We call a formula  $\varphi$  *attainable in  $w$*  if there is some action  $\alpha$  such that  $w \models \langle \alpha \rangle \varphi$ . Otherwise,  $\varphi$  is *unattainable in  $w$* —there is no path from  $w$  to a world realizing  $\varphi$ .

*Example 2.1.* Consider the example of a footrace about to begin<sup>3</sup>. The starter has a (one-shot) pistol and the race will begin as soon as the pistol discharges a blank. We will construct a very simple model for this case consisting of only two atomic predicates:

**Started**    true if the race has started,  
**Loaded**    true if the pistol is loaded.

Our language will also include two atomic actions:

**load**    the starter loads the pistol,  
**fire**    the starter pulls the trigger.

Note that the action **fire** does not imply that the pistol discharges a blank, but only that the starter pulls the trigger. Our action name **fire** may be a bit misleading in this respect, but it is more suggestive than **pull** and less awkward than **pulltrigger**.

We consider a model of four worlds, so that each combination of atomic predicates is represented. See Figure 1, in which an arrow  $w \rightsquigarrow w'$  denotes that  $w' \in \llbracket \text{load} \rrbracket(w)$  and  $w \longrightarrow w'$  that  $w' \in \llbracket \text{fire} \rrbracket(w)$ . We assume that one cannot **load** an already loaded gun. Just to make the model more interesting, we assume that our starter pistol may misfire. When a loaded pistol misfires, nothing relevant in the world changes, so that **fire** has reflexive transitions in  $w_1$  and  $w_3$  in addition to the transitions representing

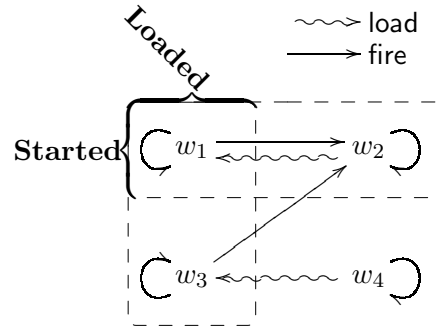


Figure 1. A sample PDL model.

successful discharge of a blank. The interpretation of a number of sample formulas is given in the figure. The reader may confirm the equations in Table II for himself.

Table II. Sample facts about the model in Figure 1.

Sample facts about the model in Figure 1	
$\llbracket \text{Started} \rrbracket$	$= \{w_1, w_2\}$
$\llbracket \text{Loaded} \rrbracket$	$= \{w_1, w_3\}$
$\llbracket \text{fire} \rrbracket \text{Started}$	$= \llbracket \text{Started} \rrbracket$
$\langle \text{fire} \rangle \text{Started}$	$= \llbracket \text{Started} \rrbracket \cup \llbracket \text{Loaded} \rrbracket$
$\llbracket \text{load} \rrbracket \text{Loaded}$	$= \mathcal{W}$
$\langle \text{load} \rangle \text{Loaded}$	$= \llbracket \neg \text{Loaded} \rrbracket$
$\llbracket \langle \text{Loaded} ? ; \text{fire} \rangle \text{Started} \rrbracket$	$= \llbracket \text{Started} \rrbracket \cup \llbracket \neg \text{Loaded} \rrbracket$
$\llbracket \langle \text{Loaded} ? ; \text{fire} \rangle \text{Started} \rrbracket$	$= \llbracket \text{Loaded} \rrbracket$

To complete our introduction to PDL, we present in Table III the standard axiom system for PDL, taken from (Harel, 1984). For rules of inference, we write  $\varphi/\psi$  to mean: From  $\varphi$  infer  $\psi$ . We omit the proof that this system is sound and complete for our semantics, i.e.  $\vdash \varphi$  iff  $\models \varphi$ .

Table III. The theory PDL.

Axioms	
Tautology	Every propositional tautology
Distributivity	$[\alpha](\varphi \wedge \psi) \leftrightarrow ([\alpha]\varphi \wedge [\alpha]\psi)$
Composition	$[\alpha; \beta]\varphi \leftrightarrow [\alpha][\beta]\varphi$
Choice	$[\alpha \cup \beta]\varphi \leftrightarrow ([\alpha]\varphi \wedge [\beta]\varphi)$
Test	$[\psi?]\varphi \leftrightarrow (\psi \rightarrow \varphi)$
K	$[\alpha](\varphi \rightarrow \psi) \rightarrow ([\alpha]\varphi \rightarrow [\alpha]\psi)$
Inference rules	
Modus Ponens	$\varphi, \varphi \rightarrow \psi \quad / \quad \psi$
Necessitation	$\varphi \quad / \quad [\alpha]\varphi$

## 2.2. SUFFICIENT MEANS FOR AN END

There are at least three distinct kinds of means-end relations that are relevant for practical reasoning. They are:

**weakly sufficient means:** doing  $\alpha$  *may* realize  $\varphi$ .

**(strongly) sufficient means:** doing  $\alpha$  *will* realize  $\varphi$ .

**necessary means:**  $\varphi$  will *not* be realized unless the agent does  $\alpha$ .

The different kinds of relations yield different motivational force for the agent that desires  $\varphi$ . In this section, we will provide semantics for the two sufficient means-end relations and sketch the kind of practical consequences they support.

When we say that an action  $\alpha$  is a (*strongly*) *sufficient means* for the end  $\varphi$  in  $w$ , we mean that, if one does  $\alpha$  in  $w$ , then  $\varphi$  will be realized. However, we must be careful to avoid trivial ascriptions, as when the action  $\alpha$  is prohibited in  $w$ . If one cannot do  $\alpha$ , then surely  $\alpha$  is not a means to any end at all.<sup>4</sup> Thus,  $\alpha$  is a sufficient means to  $\varphi$  in  $w$  if (a) doing  $\alpha$  in  $w$  ensures that  $\varphi$  and (b) one *can* do  $\alpha$ .

An action  $\alpha$  is *weakly sufficient* for  $\varphi$  in  $w$  just in case in  $w$  doing  $\alpha$  *might* realize  $\varphi$ . But this is exactly captured by the weak operator  $\langle \alpha \rangle$ . Thus, we suggest the following definition.

*Definition 2.2.* An action  $\alpha$  is a (*strongly*) *sufficient means* to  $\varphi$  in  $w$  iff

$$w \models [\alpha]\varphi \wedge \langle \alpha \rangle \top.$$

We say that  $\alpha$  is a *weakly sufficient means* to  $\varphi$  in  $w$  iff

$$w \models \langle \alpha \rangle \varphi.$$

Note that, because actions and formulas are disjoint, we see that means and ends are distinct, satisfying (1) in the introduction.

Semantically,  $\alpha$  is a sufficient means to  $\varphi$  in  $w$  iff  $\emptyset \subsetneq \llbracket \alpha \rrbracket(w) \subseteq \llbracket \varphi \rrbracket$  and is weakly sufficient iff  $\llbracket \alpha \rrbracket(w) \cap \llbracket \varphi \rrbracket \neq \emptyset$ . In case that one wants to realize  $\varphi$ , then one may be sure to do so by performing any sufficient means, but there may be reasons that he chooses not to perform any (strongly) sufficient means, of course. One *cannot* realize  $\varphi$ , however, *without* performing some weakly sufficient means.

*Remark 2.3.* As Sven Ove Hansson has noted, our definition of sufficient means neglects a common feature of natural language means-end talk: the relevancy of the means to its end. We should not call  $\alpha$  a means to  $\varphi$  if *every* action realizes  $\varphi$ . For instance, the action fire is surely not a means to realizing the condition that  $1 + 1 = 2$ , since that condition is inevitable. It does not depend on anything we can do. Von Wright agrees that relevance of the action is a central feature in means-end relations. Indeed, he describes such relations as *causal*—as mentioned in (4) from the introduction—and fire surely doesn't cause the mathematical fact.

This feature is central to the stit operators discussed in (Horty and Belnap, 1995). An agent can see to it that  $\varphi$  if he can perform an action that realizes  $\varphi$  and also can perform an action that *fails* to realize  $\varphi$ . Without the negative condition, the fact that



$\varphi$  is realized is irrelevant to the agent's actions. In that case, the agent can not see to it that  $\varphi$ ; instead,  $\varphi$  is simply inevitable.

We may amend Definition 2.2 by adding a negative condition as well. An obvious choice is to require that  $\alpha$  is a (weakly/strongly) sufficient means to  $\varphi$  in  $w$  only if there is some action  $\beta \neq \alpha$  such that  $w \models \neg[\beta]\varphi$  (equivalently  $w \models \langle\beta\rangle\neg\varphi$ ). This first approximation is suitable for atomic actions  $\alpha$ , if one restricts the quantifier to atomic  $\beta$ . However, it is easy to construct complex actions  $\alpha$  that satisfy this test and are intuitively irrelevant to  $\varphi$ . Further reflection is required to properly express the relevance condition for sufficient means-end relations.

The practical consequences of sufficient means are somewhat difficult to analyze. It is not the case that an agent, on pain of practical irrationality, say, should either give up his end or perform a given sufficient means. An agent may give up the certainty of realizing his end in order to avoid undesired consequences from strongly sufficient means. One might try to explain the motivation of (weakly/strongly) sufficient means in terms of defeasible reasons to do  $\alpha$ .

Alternatively, some might argue that our agent should give up his end or perform *some* weakly sufficient means. This latter claim is similar to reasoning involving necessary means, since (as we will see), if there are a finite number of weakly sufficient means  $\alpha_1, \dots, \alpha_k$  to  $\varphi$ , then their disjunction  $\alpha_1 \cup \dots \cup \alpha_k$  is a *necessary* means to  $\varphi$ . Generally, the motivational consequences of necessary means have seemed clearer than the consequences of sufficient means.

Thus, many treatments of practical reasoning, including von Wright's important contribution in (1963), spend considerable time on analyzing necessary means rather than sufficient means. Necessary means yield relatively clear practical conclusions. According to von Wright, for instance, if one wants  $\varphi$ , then one must be willing to do what is necessary to realize  $\varphi$ . Indeed, he writes (emphasis in original):

“Instead of saying ‘he will act’ I could also have said ‘he will necessarily act.’ This, moreover, is *logical* necessity. For, if action does not follow, we should have to describe the subject's case by saying either that he did not in fact *want* his professed object of desire or did not, after all, *think it necessary* to do the act in order to get the wanted thing.” (von Wright, 1963)

Regardless of whether one agrees with von Wright's strong claim, it supports the view that necessary means come with relatively clear practical consequences and that these consequences are simpler than the practical consequences of sufficient means.

### 3. Necessary Means and Complex Actions

It appears that the *semantics* of necessary means is considerably subtler than the semantics of sufficient means. Sufficiency is relatively straightforward:  $\alpha$  is sufficient just in case doing  $\alpha$  is sure to realize one's end. Necessary means, as they appear in the literature, are more complicated due, in part, to three features of such means.

- (A) A necessary means  $\alpha$  to  $\varphi$  need not be sufficient. Thus, necessity is not expressed in terms of  $[\alpha]$  or  $\langle\alpha\rangle$ .

- (B) A necessary means  $\alpha$  to  $\varphi$  need not be *immediately* necessary. One may do other things (relevant to  $\varphi$  or not) prior to performing  $\alpha$ .
- (C) A sequential necessary means  $\alpha_1; \dots; \alpha_n$  need not be performed “all at once” to realize its end. It may be the case that one can realize  $\varphi$  by performing  $\alpha_1; \beta_1; \dots; \beta_{n-1}; \alpha_n$  without refuting the necessity of  $\alpha_1; \dots; \alpha_n$ .

Features (A) and (B) are discussed explicitly in (von Wright, 1963). The third feature is not explicit there, but we believe that it is a reasonable feature of necessary means.

With these features in mind, let us present a rough working definition of necessary means.

*Definition 3.1 (Informal sketch).* An action  $\alpha$  is a necessary means to  $\varphi$  in  $w$  if the following hold.

1.  $\varphi$  is attainable in  $w$  and
2. every weakly sufficient means to  $\varphi$  in  $w$  involves  $\alpha$ .

Item (1) avoids trivial necessary means to unattainable ends. Without it, *every* action would be a necessary means to any unattainable end, but surely we do not conclude that an agent desiring an unattainable end ought to do *everything*.

The second item in Definition 3.1 depends on the undefined term “involves”<sup>5</sup>. Defining this term is our primary duty in this section. In doing so, we must ensure that features (A)–(C) are satisfied. This places certain restrictions on our definition of involvement and its interaction with composition. Because a necessary means  $\alpha$  need not be sufficient, our agent may satisfy his practical requirements by first doing  $\alpha$  and then some additional action,  $\beta$ . Thus,  $\alpha; \beta$  should involve  $\alpha$ . Since  $\alpha$  need not be immediately necessary, we need also to allow that  $\beta; \alpha$  involves  $\alpha$ . Similarly, it is plausible that if  $\alpha$  involves  $\gamma$ , then  $\alpha; \beta$  involves  $\gamma; \beta$  and  $\beta; \alpha$  involves  $\beta; \gamma$ . Finally, we need some structural rules, namely, distributivity and associativity.

Because of the test operator, our actual definition will be somewhat more complicated than sketched above. Indeed, we need involvement to be world-dependent. We introduce a family of relations  $\preceq_S$ , where  $S$  ranges over subsets of  $\mathcal{W}$ , and write  $\beta \preceq_w \alpha$  for  $\beta \preceq_{\{w\}} \alpha$ . Informally, we interpret this relation as: if one does  $\beta$  in  $w$ , then he also does  $\alpha$  “along the way”. Thus, for atomic actions  $m, n, o$ , we expect  $m; n; m; o; n \preceq_w n; o$ , since by performing  $m; n; m; o; n$ , one has also performed in order each of the actions in the sequence  $n; o$ .

The family  $\{\preceq_S \mid S \subseteq \mathcal{W}\}$  of pre-orders are defined by the deductive system<sup>6</sup> in Table IV. In the table, we use the abbreviation

$$\llbracket \beta \rrbracket(S) = \{w' \in \mathcal{W} \mid \exists w \in S . w' \in \llbracket \beta \rrbracket(w)\}$$

and write  $\alpha \approx_S \beta$  iff  $\alpha \preceq_S \beta$  and  $\beta \preceq_S \alpha$ . Interesting features include the properties involving sequential composition, described above. In addition, for every  $S \subseteq \mathcal{W}$ , the relation  $\preceq_S$  is a pre-order,  $\perp?$  is the least element,  $\top?$  the greatest and  $\cup$  the join.

For some systems, one might also include some additional axioms  $m \preceq_{\mathcal{W}} n$  representing relations among the atomic actions. We make no assumption about whether atomic actions should be related in this way. The appropriateness of such axioms may depend on the model and its interpretation.

Table IV. The deductive system for  $\preceq_S$ .

<u>Axioms</u>			
$\alpha \preceq_W \alpha$	$\varphi? \preceq_W \psi?$ if $S \models \varphi \rightarrow \psi$	$\perp?; \alpha \preceq_W \perp?$	$\top?; \alpha \preceq_W \alpha$
$\perp? \preceq_W \alpha$	$\alpha \preceq_W \top?$	$\alpha; \perp? \preceq_W \perp?$	$\alpha; \top? \preceq_W \alpha$
$\alpha \preceq_W \alpha \cup \beta$	$\beta \preceq_W \alpha \cup \beta$	$(\alpha \cup \beta); \gamma \approx_W (\alpha; \gamma) \cup (\beta; \gamma)$	$(\alpha; \beta); \gamma \approx_W \alpha; (\beta; \gamma)$
<u>Rules</u>			
$\frac{\alpha \preceq_S \beta \quad \beta \preceq_S \gamma}{\alpha \preceq_S \gamma}$		$\frac{\alpha \preceq_S \gamma}{\alpha; \beta \preceq_S \gamma; \beta}$	
$\frac{\alpha \preceq_S \gamma \quad \beta \preceq_S \gamma}{\alpha \cup \beta \preceq_S \gamma}$		$\frac{\alpha \preceq_{[[\beta]](S)} \gamma}{\beta; \alpha \preceq_S \beta; \gamma}$	
$\frac{\alpha \preceq_T \beta}{\alpha \preceq_S \beta} \quad \text{given } S \subseteq T$		$\frac{\alpha \preceq_S \beta \quad \alpha \preceq_T \beta}{\alpha \preceq_{S \cup T} \beta}$	

Note that  $\preceq_w$  is a relation in our meta-theory. We do not introduce this partial order in the syntax of PDL, because we see no convenient means of extending the language to do so. Thus, our definition of necessary means remains in the meta-theory, unlike the definitions of sufficient means in Section 2.2. But our goal here is to provide a semantics for necessary means to evaluate practical syllogisms and meta-theoretical definitions will suffice for this.

We are now prepared to give our definition of necessary means.

*Definition 3.2.* An action  $\alpha$  is a *necessary means* to  $\varphi$  in  $w$  if the following hold.

1. There is an action  $\beta$  such that  $w \models \langle \beta \rangle \varphi$  and  $\beta \preceq_w \alpha$ .
2. For every  $\cup$ -free action  $\beta$ , if  $w \models \langle \beta \rangle \varphi$  then  $\beta \preceq_w \alpha$ .

Each item above expresses explicitly the corresponding item from Definition 3.1, but (2) includes an additional technical constraint that it applies only to “ $\cup$ -free” actions  $\beta$ . An action  $\beta$  is  $\cup$ -free if it is constructed from atomic and test actions using sequential composition only. Without this constraint, actions  $m \cup n$  would trivially refute the necessity of  $m$  as a means to an end. But the constraint is conservative, since it is trivial to show (via a standard normal form theorem) that if there is a weakly sufficient means  $\beta$  to  $\varphi$ , then there is a  $\cup$ -free means  $\beta'$  which is constructed (in the obvious way) from the actions in  $\beta$ .

*Example 3.3.* We return to the footrace from Example 2.1 and investigate some sufficient and necessary means for starting the race.

For  $w_1$  and  $w_2$ , the action  $\top?$  is a necessary means for **Started** since  $\top?$  is involved in any action<sup>7</sup>. Furthermore, *any* non-prohibited action,  $\top?$  included, is a sufficient means for **Started** in these worlds since no actions lead to worlds in which **Started** is false. Thus, in  $w_2$ , the action **load** is a sufficient means for **Started**, since it leads to world  $w_1$  but in  $w_1$ , **load** is not sufficient since it is prohibited there.

For  $w_3$  and  $w_4$ , there is *no* (strongly) sufficient means for starting the race, since the possibility of misfire precludes any guarantee that **Started** will be realized. In  $w_4$ , the composite load; fire is weakly sufficient, since it first leads to  $w_3$  where fire is weakly sufficient. In  $w_3$  and  $w_4$ , the action fire is a necessary means to **Started**, and in  $w_4$ , the action load is also a necessary means, as is the sequence load; fire.

Finally, consider the action

$$\alpha \stackrel{\text{def}}{=} \mathbf{Started?} \cup (\neg \mathbf{Started?}; (\mathbf{Loaded?} \cup (\neg \mathbf{Loaded?}; \text{load})); \text{fire}).$$

This can be rewritten in terms of conditional actions as follows:

```

if Started then
  do nothing
else
  if Loaded then
    fire
  else
    load; fire

```

In every world,  $\alpha$  is a necessary and weakly sufficient means to **Started**. Moreover, it is maximally effective, in the sense that for all  $w$ , and for all actions  $\beta$ , if  $\beta$  is a (strongly) sufficient means to **Started** in  $w$ , then so is  $\alpha$ .

#### 4. Adding conditionals

Our semantics so far has involved *local* means-end relations: they have expressed causal relations in a given world (and hence each definition included a world  $w$  as a parameter). But this is a very narrow sense of means-end relations. Its primary advantage is simplicity (as we've seen), but it is in many ways insufficient as a semantics for natural language means-end relations. There are at least three shortcomings of local means-end relations.

- Local means-end relations are not useful in understanding the reasoning of epistemically limited agents. Such agents will not typically know which world is the actual world.
- Local means-end relations express a causal relation only about one particular world, but means-end relations in natural language are intended more broadly.
- Local means-end relations obscure the important role of intermediate ends.

This last limitation is of particular importance to us and is relevant to feature (2) from the introduction. An intermediate end is an end adopted by the agent primarily so that he can achieve another (primary or intermediate) end. For instance, an agent in world  $w_4$  from Example 2.1 knows that fire is a necessary means to **Started**, but it is not weakly sufficient unless **Loaded** holds. Thus, he adopts **Loaded** as an intermediate end. Such reasoning is the subject of much interest in the A.I. literature and is treated in some detail in (Pollock, 2002) and elsewhere.

In natural language, such intermediate ends are often confused with means. For example, an academic degree is sometimes called a means to better employment as well as an end. This is a common case of mistaking an intermediate end for a means. A bachelor's degree is not a means (since it is not an action that causes a desired condition), but it is a *precondition* for a related means-end relation and hence an *intermediate* end. Thus, we can explain the tension between features (2)—an end can be a means in different context—and (1)—ends and means are distinct types. Feature (2) is an artifact of sloppy language that confuses preconditions with means.

Such intermediate ends arise because our means-end relations are neither local nor global<sup>8</sup>, but conditional: *given* **Loaded**, the action fire is weakly sufficient for **Started**. Reasoning about intermediate ends is fraught with difficulties and these difficulties should be reflected in our semantics. In particular, we are interested in the *ramification* problem: the problem of indirect consequences of action. We motivate our interest via an example we call *The Shortsighted Suitor*.

If I had money then she might agree to my proposal for marriage.  
 Robbing her is a means to having money.  
 —————  
 If I robbed her then she might agree to my proposal for marriage.

This argument can be represented thus:

$$\frac{\mathbf{Money} \Rightarrow \langle \text{ask} \rangle \mathbf{Married} \quad \langle \text{rob} \rangle \mathbf{Money}}{\langle \text{rob}; \text{ask} \rangle \mathbf{Married}} \quad (4.1)$$

The argument fails, of course, because if I rob my sweetheart, she will hate me (let us assume such a clueless suitor will not mask his identity). We assume that

$$\mathbf{Hate} \Rightarrow \neg \langle \text{ask} \rangle \mathbf{Married}.$$

But this conditional is inconsistent with our first premise, unless we use a non-monotonic conditional. If our conditional is monotonic, then the first premise implies  $(\mathbf{Money} \wedge \mathbf{Hate}) \Rightarrow \langle \text{ask} \rangle \mathbf{Married}$  and our assumption yields  $(\mathbf{Money} \wedge \mathbf{Hate}) \Rightarrow \neg \langle \text{ask} \rangle \mathbf{Married}$ . Since we also suppose  $\langle \text{rob} \rangle (\mathbf{Money} \wedge \mathbf{Hate})$ , we would reach inconsistency.

Thus, in order to include intermediate ends, we add a conditional operator to our language and in order to represent some of the well-known problems of practical reasoning, we allow non-monotonicity. The literature on such conditional operators is broad, but we hope that a few simple definitions will satisfy our purposes. At present, we value flexibility over logical commitments. We propose the following (tentative) semantics for our conditional operator. We add to our PDL frames a “relevance” function  $r$  mapping pairs  $(w, S)$  to a set  $T \subseteq S$  of “normal”  $S$ -worlds (from the perspective of  $w$ ), explicitly  $r : \mathcal{W} \times \mathcal{PW} \rightarrow (\mathcal{PW})$  satisfying the constraint that for every world  $w$  and set  $S \subseteq \mathcal{W}$ ,<sup>9</sup>

$$r(w, S) \subseteq S.$$

We interpret  $r(w, S)$  to be the set of  $S$ -worlds that are reasonably close to (or *normal* from the perspective of)  $w$ . The idea is similar to the minimal-change or small-change conditionals discussed in (Nute, 1984), but one important difference is that we do not

require that  $w \in r(w, S)$  if  $w \in S$ . There's no requirement that the actual world is "normal".

Our conditionals are intended to capture a sense of normality:  $w \models \varphi \Rightarrow \psi$  iff, *normally*, given  $\varphi$ ,  $\psi$  is true, but the sense of "normally" may depend on the world  $w$ . We extend the semantics of Section 2.1 to include

$$\llbracket \varphi \Rightarrow \psi \rrbracket = \{w \in \mathcal{W} \mid r(w, \llbracket \varphi \rrbracket) \subseteq \llbracket \psi \rrbracket\}.$$

Thus,  $\varphi \Rightarrow \psi$  evaluates to true at  $w$  iff all the  $\text{normal}_w$   $\varphi$ -worlds also satisfy  $\psi$ .

Our models satisfy the following axioms and inference rules, taken from (Nute, 1984) and (Nute, 1994). This list is not minimal: axioms CC and CM, for instance, are derivable from the remainder.

Table V. Logical properties of  $\Rightarrow$ .

Axioms	
ID:	$\varphi \Rightarrow \varphi$
CC:	$((\varphi \Rightarrow \psi) \wedge (\varphi \Rightarrow \chi)) \rightarrow (\varphi \Rightarrow (\psi \wedge \chi))$
CM:	$(\varphi \Rightarrow (\psi \wedge \chi)) \rightarrow ((\varphi \Rightarrow \psi) \wedge (\varphi \Rightarrow \chi))$
Inference rules	
RCEC:	$\varphi \leftrightarrow \psi \ / \ (\chi \Rightarrow \varphi) \leftrightarrow (\chi \Rightarrow \psi)$
RCK:	$(\varphi_1 \wedge \dots \wedge \varphi_n) \rightarrow \psi \ / \ ((\chi \Rightarrow \varphi_1) \wedge \dots \wedge (\chi \Rightarrow \varphi_n)) \rightarrow (\chi \Rightarrow \psi) \ (n \geq 0)$
RCEA:	$\varphi \leftrightarrow \psi \ / \ (\varphi \Rightarrow \chi) \leftrightarrow (\psi \Rightarrow \chi)$
RCE:	$\varphi \rightarrow \psi \ / \ \varphi \Rightarrow \psi$

Clearly, one would like a more thorough investigation of our conditional semantics and its appropriateness for means-end reasoning. We consider the semantics presented here as fairly minimal in its commitments, so that later revisions may provide further commitments rather than retract existing commitments. This is in keeping with our present bias for flexibility.

## 5. Sufficient and Necessary Pre-Conditions

We have seen that in the presence of non-monotonic conditionals we should read a sentence  $\psi \Rightarrow \langle \alpha \rangle \varphi$  as, "*Normally*,  $\psi$  implies that doing  $\alpha$  will realize  $\varphi$ ." Combining the previous results, in this section we will use such conditional formulas to define conditional means-end relations.

*Definition 5.1.* In a world  $w$ , an action  $\alpha$  is a (*weakly, resp.*) *sufficient means to  $\varphi$  given  $\psi$*  if

$$w \models \psi \Rightarrow ([\alpha]\varphi \wedge \langle \alpha \rangle \top),$$

( $w \models \psi \Rightarrow \langle \alpha \rangle \varphi$ , resp.). Similarly,  $\alpha$  is a *necessary means to  $\varphi$  given  $\psi$*  if

1. there is some  $w' \in r(w, \llbracket \psi \rrbracket)$  and  $\beta \preceq_{w'} \alpha$  such that  $w' \models \langle \beta \rangle \varphi$ ;

2. For every  $\cup$ -free action  $\beta$  and world  $w' \in r(w, \llbracket \psi \rrbracket)$ , if  $w' \models \langle \beta \rangle \varphi$  then  $\beta \preceq_{w'} \alpha$ .

The formula  $\psi$  in Definition 5.1 is a *sufficient precondition* for the (sufficient or necessary) means-end relation involving  $\alpha$  and  $\varphi$ .

Such conditional relations can be used for creating tentative plans. If  $\alpha$  is a sufficient means to  $\varphi$  given  $\psi$ , then *normally*, one has the option of doing  $\alpha$  to realize  $\varphi$  when  $\psi$  holds. However, there may be certain circumstances in which  $\psi$  holds and doing  $\alpha$  fails to realize  $\varphi$ . Again, our focus is not on defeasible practical reasoning at present, but our means-end semantics should provide some indication as to *why* such reasoning is naturally defeasible.

Let us sketch where our shortsighted suitor went awry. We may suppose that in *every* world satisfying **Hate**, we have  $\neg \langle \text{ask} \rangle \text{Married}$ , i.e.,

$$\mathcal{M} \models \mathbf{Hate} \rightarrow \neg \langle \text{ask} \rangle \mathbf{Married}$$

(where  $\rightarrow$  is material implication, as usual). We further suppose that in the current world  $w$ , it is the case that every normal **Money**-world satisfies  $\langle \text{ask} \rangle \text{Married}$ , i.e. that

$$r(w, \llbracket \mathbf{Money} \rrbracket) \subseteq \llbracket \langle \text{ask} \rangle \mathbf{Married} \rrbracket.$$

Thus,  $w \models \mathbf{Money} \Rightarrow \langle \text{ask} \rangle \mathbf{Married}$ . These suppositions jointly satisfy our analysis from Section 4 and lead to no contradictions, provided that  $r(w, \llbracket \mathbf{Money} \rrbracket)$  is disjoint from  $\llbracket \mathbf{Hate} \rrbracket$ . Thus, we see how our non-monotonic conditional supports phenomena like the ramification problem.

We have discussed sufficient preconditions and so let us turn to *necessary preconditions* for means-end relations. It is easy to define this concept for sufficient means-end relations:  $\psi$  is a necessary precondition for  $\alpha$  to be a sufficient means to  $\varphi$ , just in case *normally*  $\alpha$  is a sufficient means for  $\varphi$  only if  $\psi$ .

*Definition 5.2.* We say that  $\psi$  is a *necessary precondition* for  $\alpha$  to be a (weakly, resp.) sufficient means for  $\varphi$  in a world  $w$  iff

$$w \models ([\alpha] \varphi \wedge \langle \alpha \rangle \top) \Rightarrow \psi,$$

( $w \models \langle \alpha \rangle \varphi \Rightarrow \psi$ , resp.).

We have not yet found a suitable corresponding definition of “necessary preconditions for necessary means-end relations,” but this notion does not arise as naturally in means-end talk as the other conditional means-end relations.

In Table VI we summarize our taxonomy of means-end relations.

To illustrate the flexibility and consequences of our conditional means-end relations, we return to the footrace and starter pistol example, and define three different relevance functions which may be added to the original example. In the third case the relevance function is also accompanied by new worlds in which the gun is malfunctioning.

*Example 5.3.* The relevance function may be used to interpret the conditional as material implication, to express epistemic limitations or to express the abnormality of complications like broken artifacts.

**Material implication:** First, we may define  $r_m(w, S) = S$  for every set  $S \subseteq \mathcal{W}$  and  $w \in \mathcal{W}$ . In that case, the conditional connective  $\Rightarrow$  coincides with material

Table VI. A summary of our means-end relations.

In $w$ , $\alpha$ is a .?. means to $\varphi$ iff. . .	Unconditional	Conditional (w.r.t $\psi$ )	
		Sufficient	Necessary
Sufficient	$w \models [\alpha]\varphi \wedge \langle \alpha \rangle \top$	$w \models \psi \Rightarrow ([\alpha]\varphi \wedge \langle \alpha \rangle \top)$	$w \models ([\alpha]\varphi \wedge \langle \alpha \rangle \top) \Rightarrow \psi$
Weakly sufficient	$w \models \langle \alpha \rangle \varphi$	$w \models \psi \Rightarrow \langle \alpha \rangle \varphi$	$w \models \langle \alpha \rangle \varphi \Rightarrow \psi$
Necessary	$\exists \beta \preceq_w \alpha,$ $w \models \langle \beta \rangle \varphi.$ $\forall \cup$ -free $\beta \not\preceq_w \alpha,$ $w \not\models \langle \beta \rangle \varphi.$	$\exists w' \in r(w, \llbracket \psi \rrbracket), \beta \preceq_w \alpha,$ $w' \models \langle \beta \rangle \varphi.$ $\forall w' \in r(w, \llbracket \psi \rrbracket),$ $\forall \cup$ -free $\beta \not\preceq_w \alpha,$ $w' \not\models \langle \beta \rangle \varphi.$	n/a

implication (the subscript  $m$  stands for “material implication”). Thus, in every world  $w$ , a formula  $\psi$  is a sufficient precondition for some means-end relation just in case every  $w' \in \llbracket \psi \rrbracket$  satisfies the means-end relation.

**Epistemic limitation:** Second, we may define  $r$  so that it reflects epistemic limitations of our agent. In our case, we suppose that the agent knows whether the race has started or not, but he does not know whether the gun is loaded. As a result, in  $w_4$  (say), he regards  $w_3$  as more relevant than  $w_1$  or  $w_2$  and as relevant as  $w_4$  itself. Our approach here is essentially a “nearest relevant worlds” system. In interpreting  $\varphi \Rightarrow \psi$  in  $w$ , we consider only the set of worlds *most relevant to  $w$*  that satisfy  $\varphi$ . Hence, we define

$$r_e(w_1, S) = r_e(w_2, S) = \begin{cases} S & \text{if } S \subseteq \{w_3, w_4\}; \\ S \cap \{w_1, w_2\} & \text{else} \end{cases}$$

$$r_e(w_3, S) = r_e(w_4, S) = \begin{cases} S & \text{if } S \subseteq \{w_1, w_2\}; \\ S \cap \{w_3, w_4\} & \text{else} \end{cases}$$

In this example, one can show that in  $w_4$ , **fire** is a necessary means to **Started** given the trivial precondition  $\top$ . This is not true for  $r_m$ , since  $r_m(w_4, \llbracket \top \rrbracket) = \mathcal{W}$  and **fire** is not a necessary means to **Started** in  $w_1$  or  $w_2$ .

**Broken gun:** For the third part of the example, we complicate our model by supposing that the gun may be broken. When the gun is broken, it always fails to fire. Thus, we add the worlds and transitions to our model as presented in Figure 2; we also include the new action **fix**.

We may suppose that the gun is not “normally” broken, regardless of which world is the actual world. Thus, we define for all  $w \in W$

$$r_b(w, S) = \begin{cases} S & \text{if } S \subseteq \llbracket \mathbf{Broken} \rrbracket; \\ S \setminus \llbracket \mathbf{Broken} \rrbracket & \text{else.} \end{cases}$$

With this definition, we assume that *even in worlds in which the gun is broken*, it is not “normally” broken. For instance,  $w_5 \notin r_b(w_5, \llbracket \top \rrbracket)$ .



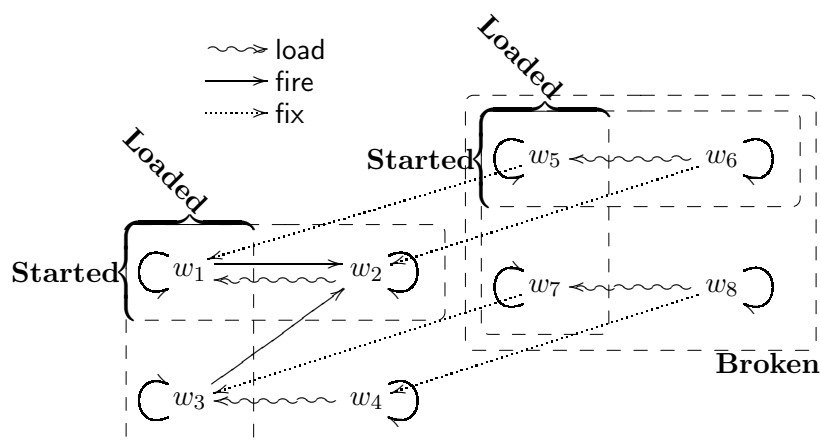


Figure 2. Additional worlds for the footrace model.

This model agrees with the material implication conditional defined by  $r_m$  whenever the antecedent  $\psi$  satisfies either  $\llbracket \psi \rrbracket \subseteq \llbracket \mathbf{Broken} \rrbracket$  or  $\llbracket \psi \rrbracket \subseteq \llbracket \neg \mathbf{Broken} \rrbracket$ . It differs from  $r_m$  just when  $\psi$  includes both **Broken** and  $\neg$ **Broken** worlds. The model also satisfies that, in every world  $w$ , fix is a necessary means to **Started**, given **Broken**. Moreover, in every world, fire is a weakly sufficient means to **Started** given **Loaded**, since then the “normal”-**Loaded** worlds regarding  $w$  are  $r_b(w, \llbracket \mathbf{Loaded} \rrbracket) = \{w_1, w_3\}$ , and in those worlds fire might lead to starting the race.

We have shown the considerable flexibility of our conditional semantics in these three examples. Of course, it may be that our restrictions are currently too loose—and the conditional semantics *too* flexible—to capture important features of conditional means-end relations. Nonetheless, we prefer to err on the side of flexibility for now.

This concludes our initial presentation of a semantics for means-end relations. In this paper, we have paid particular attention to the first five items on our list of features of means-end relations. In particular, we have argued that a proper analysis of means and ends involves taking the former as actions in a dynamic logic and the latter as formulas. Thus, since the two types are distinct (as required by (1)), it is not literally the case that an end may also be a means, as alleged in (2). This “feature” is an artifact of natural sloppiness in informal language, in our opinion. We have worked to distinguish sufficient means from necessary in Sections 2 and 3, in order to fulfill (3). We have touched on the causal impact of means (feature (4), especially regarding the relevance condition in Remark 2.3), although clearly more could be said. Finally, we have selected a non-monotonic conditional operator so that conditional means-end relations are defeasible, as anticipated by (5).

## Notes

<sup>1</sup> See (Hansson, 2000) for an insightful discussion of the role of formalisms in philosophy.

<sup>2</sup> The name “test operator” often creates more confusion than necessary. An action  $\varphi?$  does *not* consist in checking the truth condition of  $\varphi$ , updating one’s epistemic state or anything similar. Instead,  $\varphi?$  is defined precisely by the semantics given in Table I and the reader should avoid inferences about  $\varphi?$  based on observation, testing, and so on.

<sup>3</sup> This example is superficially similar to the Yale shooting problem (Hanks and McDermott, 1987) but we have a different purpose in mind. Where Hanks and McDermott are primarily interested in solutions to the frame problem, we postpone such considerations and instead concentrate on means-end semantics.

<sup>4</sup> Note that this is different than saying doing nothing cannot be a means at all. Letting things run their course may end in a desired situation; but the relevant action is “letting things run their course” (or something similar) and *not* some prohibited action.

<sup>5</sup> We find a notion of involvement in (Dignum et al., 1994) which the authors indicate can be easily extended to include sequential composition, but the obvious extension does not satisfy our requirements for composition. Also, they do not include the test operator, which leads to some of the complexity in our notation, specifically the need to subscript  $\preceq$  with sets of worlds.

<sup>6</sup> Note: we define the relation  $\preceq_w$  via a deductive system, but that doesn’t mean that the relation itself is inherently syntactic. It depends strongly on our model, as can be seen in the axiom for  $\varphi? \preceq_S \psi?$  and the rule of inference concluding with  $\beta; \alpha \preceq_S \beta; \gamma$ . The deductive system is simply a convenient means of defining our relation.

<sup>7</sup> This holds generally: if  $\varphi$  is attainable in  $w$ , then  $\top?$  is trivially a necessary means to  $\varphi$  in  $w$ . Admittedly, there is no natural language analogue to make this trivial ascription plausible, but we accept it as an artifact of our formalization.

<sup>8</sup> A *global* means-end relation is a local means-end relation that is true in every world.

<sup>9</sup> One probably wants some non-trivial relations to hold between the conditional operator and the dynamic operators, such as the axiom  $[\alpha](\varphi \Rightarrow \psi) \rightarrow ([\alpha]\varphi \Rightarrow [\alpha]\psi)$ . Such features can be introduced by adding appropriate restrictions to  $r$ , but we will not investigate them here.

## References

- Bratman, M. E., D. J. Israel, and M. E. Pollack: 1988, ‘Plans and Resource-Bounded Practical Reasoning’. *Computational Intelligence* **4**(4), 349–355.
- Brown, M. A.: 1988, ‘On the Logic of Ability’. *Journal of Philosophical Logic* **17**, 1–26.
- Brown, M. A.: 2005, ‘Means and Ends in Branching Time’. Presented at the Norms, Reasoning and Knowledge in Technology workshop.
- Castilho, M., A. Herzig, and I. Varzinczak: 2002, ‘It depends on the context! A decidable logic of actions and plans based on a ternary dependence relation’. In: *9th Intl. Workshop on Non-Monotonic Reasoning NMR’2002*.
- Castilho, M. A., O. Gasquet, and A. Herzig: 1999, ‘Formalizing action and change in modal logic I: the frame problem’. *Journal of Logic and Computation* **9**(5).
- Darwall, S.: 1996, ‘Practical Reason.’. In: D. M. Borchert (ed.): *The Encyclopedia of Philosophy, supplement*. MacMillan Reference USA, pp. 453–456.

- Dignum, F., J.-J.Ch.Meyer, and R. Wieringa: 1994, 'Contextual Permission: A Solution to the Free Choice Paradox'. In: A. J. Jones and M. Sergot (eds.): *DEON'94, Second International Workshop on Deontic Logic in Computer Science*. pp. 107–135.
- Giacomo, G. D. and M. Lenzerini: 1995, 'PDL-based framework for reasoning about actions'. In: *LNAI 1992*. pp. 103–114.
- Giordano, L., A. Martelli, and C. Schwind: 2000, 'Ramification and causality in a modal action logic'. *Journal of Logic and Computation* **10**(5), 615–662.
- Hanks, S. and D. McDermott: 1987, 'Default Reasoning, Nonmonotonic Logics, and the Frame Problem'. In: M. L. Ginsberg (ed.): *Readings in Nonmonotonic Reasoning*. Los Altos, CA: Kaufmann, pp. 390–395.
- Hansson, S. O.: 2000, 'Formalization in Philosophy'. *The Bulletin of Symbolic Logic* **6**(2), 162–175.
- Harel, D.: 1984, 'Dynamic Logic'. In: D. Gabbay and F. Guenther (eds.): *Handbook of Philosophical Logic*, Vol. II. D. Reidel Publishing Company, pp. 497–604.
- Horty, J. F. and N. Belnap: 1995, 'The Deliberative Stit: A Study of Action, Omission, Ability and Obligation'. *Journal of Philosophical Logic* **24**, 583–644.
- Hughes, J., A. Esterline, and B. Kimiaghalam: 2005, 'Means-end Semantics and a Measure of Efficacy'. *Journal of Logic, Language and Information*. Forthcoming.
- McCarthy, J.: 1999, 'Concepts of Logical AI'. <http://www-formal.stanford.edu/jmc/concepts-ai/concepts-ai.html>.
- McCarthy, J. and P. J. Hayes: 1969, 'Some Philosophical Problems from the Standpoint of Artificial Intelligence'. In: B. Meltzer and D. Michie (eds.): *Machine Intelligence 4*. Edinburgh University Press, pp. 463–502.
- Meyer, J.-J. C.: 2000, 'Dynamic logic for reasoning about actions and agents'. In: *Logic-based artificial intelligence*. Norwell, MA, USA: Kluwer Academic Publishers, pp. 281–311.
- Millgram, E.: 2004. <http://philosophy.uwaterloo.ca/MindDict/practicalreasoning.html>. Practical Reasoning entry in the online Dictionary of Philosophy of Mind, edited by Chris Eliasmith.
- Nute, D.: 1984, 'Conditional Logic'. In: D. Gabbay and F. Guenther (eds.): *Handbook of Philosophical Logic*, Vol. II. D. Reidel Publishing Company, pp. 387–439.
- Nute, D.: 1994, 'Defeasible Logic'. In: D. Gabbay, C. J. Hogger, and J. A. Robinson (eds.): *Handbook of Philosophical Logic*, Vol. III. D. Reidel Publishing Company, pp. 353–395.
- Pollock, J. L.: 2002, 'The Logical Foundations of Means-End Reasoning'. In: R. Elio (ed.): *Common Sense, Reasoning and Rationality*. Oxford University Press.
- Prendinger, H. and G. Schurz: 1996, 'Reasoning about Action and Change. A Dynamic Logic Approach.'. *Journal of Logic, Language and Information* **5**(2), 209–245.
- Seegerberg, K.: 1992, 'Getting Started: Beginnings in the Logic of Action'. *Studia Logica* **51**, 347–378.
- von Wright, G. H.: 1963, 'Practical Inference'. *The Philosophical Review* **72**(2), 159–179.
- Zhang, D. and N. Foo: 2002, 'Dealing with the ramification problem in the extended propositional dynamic logic'. In: F. Wolter, H. Wansing, M. de Rijke, and M. Zakharyashev (eds.): *Advances in Modal Logic*, Vol. 3. World Scientific, pp. 173–191.
- Zhang, D. and N. Y. Foo: 2001, 'EPDL: A Logic for Causal Reasoning.'. In: *Proceedings of the Seventeenth International Joint Conference on Artificial Intelligence, IJCAI 2001, Seattle, Washington, USA, August 4-10*. pp. 131–138.
- Zhang, D. and N. Y. Foo: 2005, 'Frame Problem in Dynamic Logic'. *Journal of Applied Non-Classical Logics* **15**(2), 215–239.

