"MEANS" MEANS *WHAT*? A SEMANTICS FOR MEANS-END RELATIONS

Abstract. Practical reasoning is the process of deriving actions or intentions from premises including means-end relations. In order to evaluate the appropriateness of practical reasoning theories, one wants a clear semantics of means-end relations. We offer an initial step in means-end semantics here.

We use *propositional dynamic logic* as the basic setting in which to analyze three kinds of so-called *local* means-end relations: weakly sufficient, sufficient and necessary means to a given end. We sketch the motivational consequences of each kind of relation for an agent desiring the given end. We also address a practical analogue to Ross's paradox and show a technical work-around to avoid this problem in our means-end semantics.

We also discuss *conditional* means-end relations, which more closely reflect the limited knowledge of an agent trying to achieve his end. We sketch some practical consequences of these means-end relations as well.

Throughout, we illustrate our semantics with a running example of a model with non-trivial means-end relations. We close with an indication of further developments which our semantics suggests.

This semantics forms a foundational step in an analysis of means-end reasoning.

1. Introduction

Practical reasoning is concerned with deriving actions (or intentions to act) from certain propositions. This distinct form of reasoning has been studied at least since Aristotle's time and enjoyed renewed interest recently beginning with von Wright's landmark article [9]. The topic has gained a wider audience in recent years, due to its application in artificial intelligence (in, e.g. , [8]). If one wants to create autonomous agents capable of rationally interacting with their environments, then one needs an algorithm for producing actions likely to achieve the agents' goals. It is natural to look to formal systems of practical reasoning for such an algorithm.

A typical argument in practical reasoning involves (some of) the following kinds of premises:

(1) an assertion that an agent A desires some *end* φ,

(2) an assertion that (possibly given some precondition ψ) the action α is related to the realization of φ,

(3) an assertion of some factual matter, such as that the precondition ψ 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 α is that it is related in the right way to the desired condition φ. Because one wants φ to be realized, he will be motivated to do α. We call such premises *(conditional) means-end relations*, since they assert that the action α is a *means* to the end φ. We will focus here on *local means-end relations*, which assert that *in this world*, the action α is related to the realization of φ, independent of any precondition.

We have been purposely vague about what sort of relation α should bear to φ. There are three distinct kinds of means-end relations that are relevant for practical reasoning. They are:

weakly sufficient means: doing α *may* realize φ.

(strongly) sufficient means: doing α *will* realize φ.

necessary means: φ will *not* be realized unless the agent does α.

The different kinds of relations yield different motivational force for the agent that desires φ. In the sequel, we will provide semantics for these relations and sketch the kind of practical consequences they support.

To justify an argument in practical reasoning, one must show exactly why the agent should be motivated to act on the basis of the premises. For this, it is essential that the meanings of the premises are clear and precise. In other words, one needs a semantics for the various premises found in practical reasoning. We find that the literature is lacking a clear analysis of the means-end relations that are so central to the endeavor[1]. We present an initial step in this direction.

We have chosen a formal semantics for means-end relations. The work done here should be considered an exercise in conceptual analysis via formal tools. Formal semantics permits a clear analysis with less ambiguity than natural language analyses. If we succeed in laying a semantic foundation for means-end relations, then our work will support evaluations of existing theories of practical reasoning and may also lead to new work in the area.

In choosing our semantics, there are several issues which we felt were essential to means-end relations as they appear in practical reasoning. Among these issues are the following.

(1) the distinction between means and their ends

(2) the distinction between the three types of means-end relations

(3) the role of intermediate ends in forming complex plans of action

(4) the difficulties of planning due to the frame problem

(5) efficacy of means with respect to their ends

(6) undesirable side effects of means

(7) the natural language role of objects-as-means

(8) that an action may be a means to mutually exclusive ends

We believe that the semantics we present may be extended to accommodate each of these issues, but in this preliminary presentation, we explicitly focus on only the first three issues. We also show how an analogue of Ross's paradox applies to our means-end relations and give a technical apparatus to avoid the undesirable consequences of this paradox.

[1] The stit logics of Belnap and Horty[3] come close to this, but they stress ends to the neglect of means from our perspective.

We do not discuss premises of the form "A wants φ to be realized" or the imperative conclusions "A must do α". These propositions also require a conceptual analysis if we are to understand practical reasoning, but these issues are distinct from the question we ask today: What does it mean that α is a means to φ?

2. Propositional Dynamic Logic

An end is a condition which some agent may desire. We take this in the broadest sense, so that any condition may be an end. Thus, it is reasonable to consider an end to be a formula in a formal language.

A means is a way to realize an end. Therefore, a means must be something one can do in order to change the world so that an end φ (which may currently be false) will become true. This suggests that means correspond to transitions between possible worlds. *Propositional Dynamic Logic* (PDL) is an appropriate language for modeling transitions between worlds via an agent's actions[2]. See [2] for an introduction to PDL. We will only sketch the semantics here.

The language of PDL is built from two non-empty disjoint atomic types: the set Φ_0 of atomic propositions and the set Π_0 of atomic actions. We use P, Q, ... to range over Φ_0 and m, n, ... to range over Π_0. The sets Φ of formulas and Π of actions are built via the following definitions, where φ, ψ, ... range over Φ and α, β, ... range over Π.

$$\Phi = P \mid \top \mid \varphi \wedge \psi \mid \neg\varphi \mid [\alpha]\varphi$$

$$\Pi = m \mid \alpha;\beta \mid \alpha \cup \beta$$

We have omitted the iteration α^* and test $\varphi?$ actions from our logic, since these are not essential to our present purposes. The sentence $[\alpha]\varphi$ expresses that, if one does α,

[2]It is common to refer to objects as means as well, which is opposed to our means-as-actions semantics. We hope to discuss how objects can be means in a later paper.

then φ will be realized. The construction $\alpha;\beta$ denotes sequential composition (first do α and then do β) and $\alpha \cup \beta$ denotes non-deterministic choice between α and β.

We introduce the connectives $\neg$, $\vee$ and $\rightarrow$ and the weak operator $\langle\alpha\rangle$ as usual.

A *PDL model* $\mathbf{F}$ for Π_0 consists of

- a set $\mathcal{W}$ of worlds (or states),
- an interpretation $v : \mathcal{W} \times \Phi_0 \rightarrow \{tt, ff\}$ assigning truth values to pairs of worlds and atomic propositions and
- a *dynamic interpretation* of actions. This dynamic interpretation consists of transitions between worlds, labeled by atomic actions. When an arrow $w \xrightarrow{m} w'$ exists, then w' is a possible outcome of doing m in world w.

The satisfaction relation $\models \;\subseteq\; \mathcal{W} \times \Phi$ is defined as usual for the boolean connectives. We write

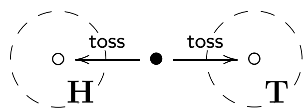
$$w \models [m]\varphi \quad \text{iff} \quad \text{for every } w', \text{ if } w \xrightarrow{m} w', \text{ then } w' \models \varphi.$$

Consequently,

$$w \models \langle m\rangle\varphi \quad \text{iff} \quad \text{there is a } w' \text{ such that } w \xrightarrow{m} w' \text{ and } w' \models \varphi.$$

Thus, $w \models [m]\varphi$ just in case doing m ensures that φ will be true in whatever world results and $w \models \langle m\rangle\varphi$ just in case it is possible that φ will be true in the world that results from doing m.

For example, consider a world in which one may toss a coin. If we neglect all of the features but the coin toss, there are two possible outcomes: the coin may come up heads or it may come up tails. This is modeled by three worlds and two atomic propositions with the following dynamic structure, where the actual world is denoted by the filled circle[3].



The actual world satisfies $\langle \mathsf{toss} \rangle \mathbf{H}$, but not $[\mathsf{toss}]\mathbf{H}$.

The action constructions $\alpha; \beta$ and $\alpha \cup \beta$ may be defined by the following axioms.

$$[\alpha; \beta]\varphi \leftrightarrow [\alpha][\beta]\varphi$$

$$[\alpha \cup \beta]\varphi \leftrightarrow [\alpha]\varphi \wedge [\beta]\varphi$$

The second axiom looks more natural in terms of the weak operator:

$$\langle \alpha \cup \beta \rangle \varphi \leftrightarrow \langle \alpha \rangle \varphi \vee \langle \beta \rangle \varphi.$$

Since both operations are clearly associative, we will drop parentheses indicating association hereafter.

We call an action α *prohibited in* w if there is no w' such that $w \xrightarrow{\alpha} w'$. Intuitively, such actions cannot be performed in w. If α is prohibited in w, then $w \models [\alpha]\varphi$ for any $\varphi \in \Phi$ (including $\bot$), but $w \not\models \langle \alpha \rangle \varphi$ for any $\varphi \in \Phi$ (not even $\top$).

We close the introduction with an example loosely based on von Wright's hut[4].

Example 1. Suppose that one has a hut that is not yet habitable and that will not be habitable unless two conditions are met: it must have a reliable heat source and it must have broadband internet access, conditions denoted by **Heated** and **NetAccess**, respectively. Further suppose that the hut is locked by a combination lock and that in order to provide a heat source, the agent must unlock it (the condition that the door is unlocked is denoted **Open**). However, our agent is unsure of the combination of the lock, so that an attempt to unlock it may fail.

[3]This example would be better handled by a semantics involving probabilities instead of non-determinism.

[4]It is not always clear whether von Wright's example is about actions (like heating the hut) or about conditions one brings about (that the hut is heated). We assume that it is about the former, and so is an example of means-end reasoning in our sense. If it is about the latter instead, then perhaps stit logic [3] is better suited to analyze his argument.

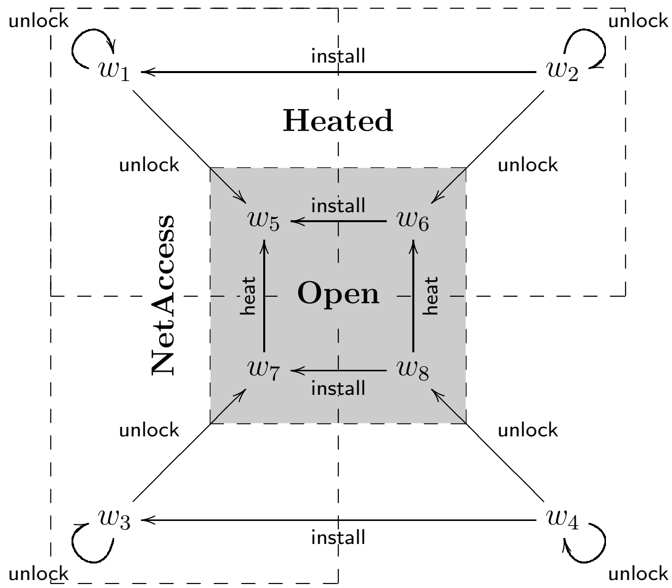


FIGURE 1. How to make a hut habitable: A PDL model. The current world is w_4, satisfying ¬**Heated**, ¬**NetAccess** and ¬**Open**. The agent desires to realize **Heated** ∧ **NetAccess**, satisfied by w_1 and w_5.

Our agents atomic actions, then, are unlock (attempt to unlock the door), heat (provide a heat source) and install (contact the local broadband provider to install internet access). Most models would also have a do-nothing action, but we omit that here.

The situation is given in Figure 1 and we calculate the sets of worlds satisfying various formulas in Table 1.

Formula	Worlds that satisfy φ
Heated	w_1, w_2, w_5, w_6
NetAccess	w_1, w_3, w_5, w_7
Open	w_5, w_6, w_7, w_8
Heated ∧ **NetAccess**	w_1, w_5
⟨install⟩**NetAccess**	w_2, w_4, w_6, w_8
[install]**NetAccess**	$w_1, w_2, w_3, w_4, w_5, w_6, w_7, w_8$
⟨unlock⟩**Open**	w_1, w_2, w_3, w_4
[unlock]**Open**	w_5, w_6, w_7, w_8
⟨heat⟩**Heated**	w_7, w_8
[heat]**Heated**	$w_1, w_2, w_3, w_4, w_5, w_6, w_7, w_8$

TABLE 1. Some satisfaction relations for Figure 1.

3. Sufficient means

An action α is a weakly sufficient means for φ in world w just in case doing α in w may bring about φ. This is easily captured in PDL by the weak dynamic operator. Thus, we define:

Definition 2. α is a weakly sufficient means for φ in w iff

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

We say that φ is *attainable* in w if there is some weakly sufficient means to φ in w and otherwise it is *unattainable* in w. (We do not speculate whether it is irrational to desire some unattainable end.)

That α is a weakly sufficient means to φ provides only weak motivation for the agent desiring φ to do α. Doing α may not guarantee that φ will be realized. Moreover, α may not be the only weakly sufficient means to φ. Thus, our agent has a weak, defeasible reason to do α in this case.

While our agent has only a weak reason to do a particular weakly sufficient means, he is strongly motivated to do *some* weakly sufficient means. If he does not do any weakly sufficient means, then φ will *not* be realized. Perhaps, if we follow von Wright's analysis, one may say that our agent must either do some weakly sufficient means or change his desires, on pain of practical irrationality.

We must be careful here to allow for the fact that sometimes, the world changes through no active effort on the part of our agent. If our agent desires light to read by and it is nearly dawn, then doing nothing is a means to his end. The action of doing nothing does not necessarily leave the world as it is. The world has a habit of changing on its own (and also due to the actions of other agents).

Strongly sufficient means are slightly subtler, but only slightly. We suggested in the introduction that α is a sufficient means to φ in w just in case doing α ensures that φ, i.e. that $w \models [\alpha]\varphi$. But this yields unfortunate consequences with respect to prohibited actions. If our agent cannot do α, then trivially $w \models [\alpha]\varphi$. But surely, α is not a sufficient means to φ in this case—at least not in any sense relevant for practical reasoning.

Thus, we amend the definition. We say that α is strongly sufficient means to φ in w just in case $w \models [\alpha]\varphi$ and α is not prohibited. Thus:

Definition 3. α is a *(strongly) sufficient means* for φ in w iff

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

Intuitively, one should have greater motivation to perform strongly sufficient means for his end than to perform weakly sufficient means, but it is unclear how to express this intuition. Certainly, there are cases in which one prefers a weak means to a strong, because the strong means has undesirable side effects. The practical difference between strong and weak sufficiency is not clear to us at present.

Example 4. We return to Example 1 to give some consequences of our definitions. The following table gives some sufficient means-end relations for our model.

Action	End	Weakly sufficient	Strongly sufficient
install	**NetAccess**	w_2, w_4, w_6, w_8	w_2, w_4, w_6, w_8
heat	**Heated**	w_7, w_8	w_7, w_8
unlock	**Open**	w_1, w_2, w_3, w_4	—
unlock; heat	**Heated**	w_3, w_4	—
install; unlock; heat	**Heated**	w_4	—
unlock; heat; install	**Heated**	w_4	—

TABLE 2. Worlds in which certain weakly/strongly sufficient means-end relations hold.

The hut is habitable in a world w iff $w \models$ **Heated** $\wedge$ **NetAccess**, and we assume that the current world is w_4 (in which none of the propositions **Heated**, **NetAccess** or **Open** are true). Then it is easy to see that any sequence of actions involving all of install, unlock and heat is weakly sufficient to make the hut habitable, provided that each occurrence of unlock occurs prior to any occurrence of heat. Thus, unlock; unlock; install; heat is weakly sufficient, but not unlock; heat; unlock; install.

No action is strongly sufficient, since unlock is not strongly sufficient for **Open**.

4. Necessary means

As far as practical reasoning is concerned, necessary means seem to be the easy case. The practical consequences of a necessary means-end relation seem strong and clear: if α is a necessary means to φ, then the agent desiring to realize φ must do α (or fail to achieve his end and, according to von Wright, suffer the embarrassment of practical irrationality). However, neither the consequences nor the meaning of necessity are as clear as they first appear.

Let us take von Wright's favorite example of a first-person inference involving a necessary means.

I want to make the hut habitable.
Unless I heat the hut, it will not become habitable.
———
Therefore I must heat the hut.

Von Wright claims that the "must" in the conclusion expresses a logical necessity, that in fact (echoing Aristotle) the conclusion of the syllogism is properly the act itself[5]. However, he is explicit that, in some cases, the act need not be immediately undertaken, but rather done "sooner or later".

The open-ended nature of necessary means complicates the semantics, as we will see. That a means may be necessary but not immediately required must be reflected in our definition. This leads us to a subtle notion of "involvement" as part of our analysis of necessary means. In order to motivate necessity and involvement, we consider what actions count as a counterexample to the claim that α is a necessary means to φ in w.

[5] We find the idea that the act is a conclusion to an inference very difficult to understand, but let us press on in trying to understand von Wright without becoming too distracted by this point.

If one wants to refute this claim, he must show that φ can be realized without doing α. Thus, he must show that $w \models \langle\beta\rangle\varphi$ for some β distinct from α. But distinctiveness is not enough: there are some β different than α that should not serve as counterexamples.

Consider an agent that desires a neater lawn (**Neat**) and suppose that he is told mowing the lawn (mow) is a necessary means to **Neat**. He may respond that mow is *not* necessary, since he could first read a magazine (read) and *then* mow the lawn to achieve the same effect, i.e. that the composite $\mathsf{read};\mathsf{mow}$ is sufficient for **Neat**, so mow is not necessary.

This refutation is spurious. That mow is necessary means that the condition **Neat** will not be realized unless one does mow. When one does $\mathsf{read};\mathsf{mow}$, one does mow as part of the sequence, so this is no counterexample at all[6].

Turning to choice, if our agent claimed that $\mathsf{read} \cup \mathsf{mow}$ is weakly sufficient and therefore a counterexample, we would not take his argument seriously, since the weak sufficiency comes from the fact that one may choose to do mow itself. The choice $\mathsf{read} \cup \mathsf{mow}$ refutes the necessity of mow iff the action read refutes it.

In order to eliminate these spurious counterexamples, we introduce a notion of *involvement*, where an act β *involves* α if by doing β one might[7] do α as a "sub-action". In this case, we write $\beta \preccurlyeq \alpha$. A counterexample to the claim that α is a necessary means to φ in w would consist of an action β such that

(1) $w \models \langle\beta\rangle\varphi$ and

(2) $\beta \not\preccurlyeq \alpha$.

If α is necessary then there must be no such β.

The pre-order $\preccurlyeq$ is axiomatized in Table 3.

[6]There may be meanings of necessity (immediate necessity) in which $\mathsf{read};\mathsf{mow}$ should be a counterexample, but these meanings do not support von Wright's "sooner or later" conclusions.

[7]"Might", not "must", due to non-deterministic choice.

Axioms

$$\alpha \preccurlyeq \alpha \qquad \alpha \preccurlyeq \alpha \cup \beta \qquad \alpha;\beta \preccurlyeq \alpha$$
$$\beta \preccurlyeq \alpha \cup \beta \qquad \alpha;\beta \preccurlyeq \beta$$

Rules

$$\frac{\alpha \preccurlyeq \beta \quad \beta \preccurlyeq \gamma}{\alpha \preccurlyeq \gamma} \qquad \frac{\alpha \preccurlyeq \gamma}{\alpha;\beta \preccurlyeq \gamma;\beta}$$

$$\frac{\alpha \preccurlyeq \gamma \quad \beta \preccurlyeq \gamma}{\alpha \cup \beta \preccurlyeq \gamma} \qquad \frac{\alpha \preccurlyeq \gamma}{\beta;\alpha \preccurlyeq \beta;\gamma}$$

TABLE 3. The deductive system for $\preccurlyeq$.

In addition to the requirement that there is no counterexample, we require that φ is attainable. Otherwise, *every* action would be a necessary means to any unattainable end (such as $\bot$). But we have no motivation to perform any action for an unattainable end, and so necessary means would not play the right motivational role in practical reasoning.

Thus, we offer the following definition.

Definition 5. α is a necessary means for φ iff

(1) there is β such that $w \models \langle\beta\rangle\varphi$;
(2) for every β, if $w \models \langle\beta\rangle\varphi$ then $\beta \preccurlyeq \alpha$.

It follows that an atomic action m is necessary for φ in w iff φ is attainable in w and every path from w to some φ world includes an edge labeled m.

Example 6. We sketch a few examples of necessary means-end relations involving our model from Figure 1. We give in Table 4 some examples of necessary means-end relations.

In world w_4 (the "current" world), each of the atomic actions $\mathsf{install}$, heat and unlock are necessary to make the hut habitable. So is the sequence $\mathsf{unlock};\mathsf{heat}$ and

Action	End	Necessary in worlds
install	**NetAccess**	w_2, w_4, w_6, w_8
heat	**Heated**	w_3, w_4, w_7, w_8
unlock	**Open**	w_1, w_2, w_3, w_4
unlock; heat	**Heated**	w_3, w_4
(install; heat) ∪ (heat; install)	**Heated** ∧ **NetAccess**	w_4, w_8

TABLE 4. Worlds in which certain necessary means-end relations hold.

the complex action (install; unlock; heat) ∪ (unlock; install; heat) ∪ (unlock; heat; install). This complex action is also weakly sufficient.

What is the practical consequence of necessary means-end relations? If our agent wants to realize φ, then he must perform some action β which involves every necessary action. In this sense, he must "do" every necessary action, but this does not mean that he immediately performs any of the necessary actions. Rather, it is acceptable that the necessary actions are performed as part of a long sequence of actions.

This practical consequence is not satisfied by doing just any action β involving every necessary action. The agent is still required to do some weakly sufficient means and, in fact, any weakly sufficient means involves every necessary means. In this respect, necessary means add little to the practical commitments of our agent, despite their central role in von Wright. They nonetheless play a role in deciding whether one is willing to pursue his end: they allow one to state more clearly the piecemeal acts which one must perform to achieve his end and to allow the agent to judge his willingness to do what is necessary.

5. Ross's paradox

Ross's paradox is one of a list of undesirable properties of standard deontic logics (see [4]). In such logics, if one ought to mail a letter, then he ought to mail it or burn it. This consequence conflicts with our intuitions about one's duties.

An analogue of Ross’s paradox applies to our definitions of weakly sufficient and necessary means: If α is a weakly sufficient (necessary, resp.) means to φ, then so is $\alpha \cup \beta$. If mailing a letter is a weakly sufficient (or necessary) means of sending a message someone, then so is mailing the letter or burning it, and so our agent has a reason to do the action “mail it or burn it.”

As with the deontic version, one can ignore the problem as a minor mismatch between natural and formal languages. We sketch here an alternative response, somewhat complicated but yielding a narrower set of means-end relations in which Ross’s paradox does not arise. For this, we define a canonical normal form for actions.

We call a sequence $m_1; \ldots; m_k$ of atomic actions an *atom-sequence*. An action is *normal* if it is a disjunction $\beta_1 \cup \ldots \cup \beta_j$ of atom-sequences. Let $\alpha \equiv \beta$ whenever $w \xrightarrow{\alpha} w'$ iff $w \xrightarrow{\beta} w'$. It is easy to confirm the two distributive identities:

$$\alpha; (\beta \cup \gamma) \equiv (\alpha; \beta) \cup (\alpha; \gamma)$$

$$(\alpha \cup \beta); \gamma \equiv (\alpha; \gamma) \cup (\beta; \gamma)$$

Using these identities, one can construct for each α a normal action β such that $\alpha \equiv \beta$, but we omit the details of the recursive construction. We call the constructed normal action the *canonical normal form* of α, denoted $\text{cnf}(\alpha)$.

We say that an action $\gamma_1 \cup \ldots \cup \gamma_k$ is a *proper sub-disjunction* of $\beta_1 \cup \ldots \cup \beta_l$ if each γ_i is identical to some β_j, but some β_j is not equal to any γ_i.

Canonical normal forms give a technical means of avoiding Ross’s paradox, giving new recursive definitions of weakly sufficient and necessary means. These definitions are given in Table 5.

These revised definitions avoid Ross’s paradox. For instance, in our running example, $\mathsf{unlock}; \mathsf{heat}$ is a weakly sufficient means to **Heated** in w_4, but $\mathsf{install}$ is not. Thus, $(\mathsf{unlock}; \mathsf{heat}) \cup \mathsf{install}$ is not weakly sufficient.

Action α is ...	Weakly sufficient for φ in w	Necessary for φ in w
atom-sequence	Definition 2	Definition 5
normal disjunction	Each disjunct is weakly sufficient as above.	α satisfies Definition 5 and no proper sub-disjunction satisfies Definition 5.
non-normal	$\text{cnf}(\alpha)$ is weakly sufficient as above.	$\text{cnf}(\alpha)$ is necessary as above.

Table 5. Revised definitions of weak sufficient and necessary means. These versions avoid Ross's paradox.

Similarly, heat is a necessary means to **Heated** in w_4, but $\mathsf{install}$ is not. Hence $\mathsf{heat} \cup \mathsf{install}$ is not a necessary means, because a proper sub-disjunction (heat) is necessary.

The examples in Tables 2 and 4 are also correct for the revised definitions of this section.

We regard the revised definitions as an interesting alternative to Definitions 2 and 5 and they strengthen the relationship between weak sufficiency/necessity and reasons for acting, but they are a bit cumbersome in practice. We will return to the subject of means-end analogues of deontic paradoxes in subsequent work.

6. Means-end relations and practical reasoning

We have sketched the semantics of the various kinds of local means-end relations, but our ultimate end is to understand practical reasoning. For this, we do not think that local means-end relations play a central role: they require too much knowledge about the causal structure of the world. We do not expect that the agent immediately grasps the sufficiency of a long sequence of acts in attaining his goal, but instead he must proceed step-by-step. While we are not prepared to give a theory of practical reasoning at this point, we would like to sketch some of the features that such a theory would involve as well as some sample reasoning that such a theory may support.

We expect that an agent engaged in practical reasoning will not possess all of the information contained in our possible worlds model. He does not know the particular world-to-world transitions, but he will know certain conditional means-end relations and certain facts about the actual world and must form his plans given this data.

By a *conditional means-end relation*, we mean a statement of the form

(1) Given ψ, the action α is (weakly sufficient/sufficient/necessary) for φ.

We may interpret this in terms of any handy conditional operator $\Rightarrow$, but for simplicity's sake, let us use material implication for now. Thus, we interpret (1) as true iff

For every w, if $w \models \psi$, then α is a (weakly sufficient/sufficient/necessary) means for φ in w.

(In more sophisticated treatments, we may choose a non-monotonic conditional operator so that our semantics includes the frame problem as discussed in [8].)

Example 7. Our agent desires to realize **Heated** $\wedge$ **NetAccess** and so he wants to realize **NetAccess**. If he knows that, given $\neg$**NetAccess**, the action install is necessary and sufficient for **NetAccess**, then he has a strong reason to do install. Similarly, if he knows that, given $\neg$**Heated**, the action heat is necessary for **Heated**, he has reason to do heat.

In some cases, actions cannot be performed without realizing certain preconditions. If our agent is to reason about means to ends, he should be aware of some of these relations.

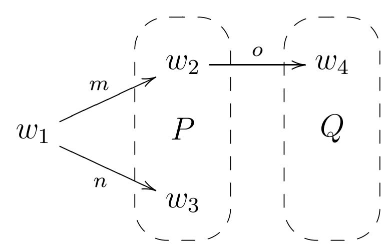
Note that we again require that necessity implies weak sufficiency. We do so because without weak sufficiency, there is no motivational force for necessary preconditions to a given action.

ψ is .?. for α	Definition
Weakly sufficient	$\exists w \,.\, w \models \psi \wedge \langle\alpha\rangle\top$
Strongly sufficient	$\forall w \,.\, w \models \psi \Rightarrow \langle\alpha\rangle\top$
Necessary	$\forall w \,.\, w \models \neg\psi \Rightarrow [\alpha]\bot$ and $\exists w \,.\, w \models \psi \wedge \langle\alpha\rangle\top$

TABLE 6. Preconditions for actions

Example 8. Suppose our agent knows that, given $\neg$**Heated**, the action heat is necessary for his end and also that the condition **Open** is both necessary and sufficient for doing heat. Then he has reason to realize **Open** and so he adopts **Open** as an intermediate end. If he also knows that, given $\neg$**Open**, the action unlock is a necessary means to **Open**, then he should deduce that he must first perform unlock and then heat.

Whether an agent can reliably form plans of actions that realize his goals depends on which causal relationships he knows. Consider the following transition system, in which our agent desires Q.



Suppose he knows that P is necessary to do o and that o is a necessary means to Q. He may conclude that the act $m \cup n$ is therefore necessary for his end, but if he chooses to do n then he will have chosen badly. If he cannot distinguish the worlds w_2 and w_3 in this respect, then he cannot reliably realize Q.

Obviously, reasoning with such conditional means-end relations is hard. Nonetheless, we think that this is a reasonable approximation of the kind of premises one encounters in practical reasoning. Our aim is not to simplify the process so that the reasoning is easy. It is to be clear on what the premises mean so that the reasonableness of proposed arguments can be evaluated.

7. Conclusion

This work forms a foundation for further development of means-end semantics and practical reasoning. We have focused on the kernel of such semantics here, but there are many extensions to this work one may pursue.

- We may interpret the conditional as a non-monotonic operator (as discussed in [5]), so that the frame problem (discussed in [8] and [1]) is a feature of our semantics. Consequently, any practical reasoning involving our semantics will be *defeasible* [6].
- We may introduce a measure of efficacy by adding probabilistic features to our semantics.
- We may include objects-as-means by adding appropriate actions "use o" for each object o.
- We may include a means α to mutually exclusive ends (a thermostat is a means to both heating and cooling a room) by using monotone neighborhood semantics (like the game logic of [7]) in place of Kripke semantics.

We have not the space to develop each of these topics in a first presentation of means-end semantics, but we hope that this list gives some idea of the flexibility of our proposal.

We believe that a working semantics for means-end relations is a necessary first step in evaluating existing theories of practical reasoning and may suggest new theories of same.

References

[1] Daniel C. Dennett. Cognitive wheels: The frame problem of AI. In Margaret A. Boden, editor, *The Philosophy of Artificial Intelligence*, Oxford Readings in Philosophy, pages 147 – 170. Oxford University Press, 1990.

[2] David Harel. Dynamic logic. In D. Gabbay and F. Guenthner, editors, *Handbook of Philosophical Logic*, volume II, pages 497–604. D. Reidel Publishing Company, 1984.

[3] John F. Horty and Nuel Belnap. The deliberative stit: A study of action, omission, ability and obligation. *Journal of Philosophical Logic*, 24:583–644, 1995.

[4] J.-J. Ch. Meyer, F.P.M. Dignum, and R.J. Wieringa. The paradoxes of deontic logic revisited: A computer science perspective. Technical Report UU-CS-1994-38, Utrecht University, 1994.
[5] Donald Nute. Conditional logic. In D. Gabbay and F. Guenthner, editors, *Handbook of Philosophical Logic*, volume II, pages 387–439. D. Reidel Publishing Company, 1984.
[6] Donald Nute. Defeasible logic. In D. Gabbay, C. J. Hogger, and J. A. Robinson, editors, *Handbook of Philosophical Logic*, volume III, pages 353–395. D. Reidel Publishing Company, 1994.
[7] H. Parikh. The logic of games and its applications. In M. Karpinski and J. van Leeuwen, editors, *Topics in the Theory of Computation*, volume 24 of *Annals of Discrete Mathematics*. Elsevier, 1985.
[8] John L. Pollock. The logical foundations of means-end reasoning. In Renée Elio, editor, *Common Sense, Reasoning and Rationality*. Oxford University Press, 2002.
[9] Georg Henrik von Wright. Practical inference. *The Philosophical Review*, 72(2):159–179, Apr. 1963.