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

JESSE HUGHES

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 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 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 [8]. The topic has gained a wider audience in recent years, due to its application in artificial intelligence (in, e.g. , [7]). 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. 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

[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.

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.

Our emphasis is on semantics for means-end relations, but we also give some tentative practical consequences for each of the kinds of means-end relations. These consequences clearly require a more careful analysis than they have received in this proposal.

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

[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.

over Π.

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

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

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, 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

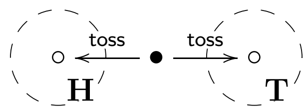
$$w \models [m]\varphi \quad \text{iff} \quad \text{for every } w'\text{, if } w \xrightarrow{m} w'\text{, then } w' \models \varphi.$$

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.

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

$$\alpha \text{ is a weakly sufficient means for } \varphi \text{ in } w \quad \text{iff} \quad w \models \langle \alpha \rangle \varphi.$$

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,

[3]This example would be better handled by a semantics involving probabilities instead of non-determinism. That is the subject of forthcoming research on fuzzy set semantics for PDL.

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.

However, our agent 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).

An action α is a (strongly) sufficient means to φ if (1) doing α ensures that φ and (2) one can do α (avoiding trivial means-end relations). Thus,

$$\alpha \text{ is a } \textit{(strongly) sufficient means} \text{ for } \varphi \text{ in } w \quad \text{iff} \quad w \models [\alpha]\varphi \wedge \langle\alpha\rangle\top.$$

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.

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[4] 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.

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

[4]It is not clear what von Wright means by "I heat the hut." It might be taken as bringing about a condition rather than doing an action, and so von Wright would be doing something closer to stit logic than our means-end semantics. We assume he meant it as an action, possibly involving choices, and continue our analysis.

[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.

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.

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 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 1.

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 1. 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.

$$\alpha \text{ is a necessary means for } \varphi \text{ iff } \begin{array}{l} (1) \text{ there is } \beta \text{ such that } w \models \langle\beta\rangle\varphi; \\ (2) \text{ for every } \beta, \text{ if } w \models \langle\beta\rangle\varphi \text{ then } \beta \preccurlyeq \alpha. \end{array}$$

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.

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.

Example. Consider the following situation: Jill and Fred are in a room with two locked boxes. Jill has a key for box 2 and Fred has a key for one of the boxes, but Jill doesn't know for which box. Jill can open box 2 or she can ask Fred to open a box.

Assume further that Fred is a petty man and if he doesn't get to open a box first, he will refuse to open any box and he will only open one box (his choice which).

We denote by **B1** (**B2**, resp.) the condition that box 1 (box 2, resp.) has been opened. The action jill denotes that Jill opens box 2 and fred that she has asked Fred to open a box.

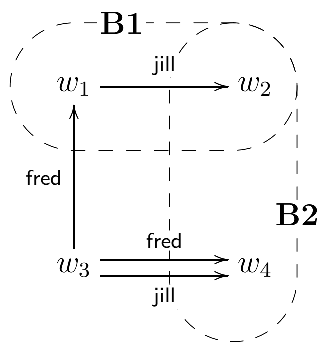


FIGURE 1. Two locked boxes: a PDL model.

The transition structure for this situation is given in the figure above. One can confirm the following relations hold in the "actual" world, w_3.

- fred is necessary and weakly sufficient for **B1**. It is not strongly sufficient.
- jill is strongly sufficient for **B2**, but not necessary.
- Both fred and jill are necessary for $\mathbf{B1} \wedge \mathbf{B2}$, but neither is weakly sufficient.
- The action fred; jill is necessary and weakly sufficient for $\mathbf{B1} \wedge \mathbf{B2}$.

5. 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 extend the semantics to include a conditional operator for conditional relations and to explain intermediate ends and the formation of complex plans of actions as means.

- We may interpret the conditional as a non-monotonic operator (as discussed in [4]), so that the frame problem (discussed in [7] and [1]) is a feature of our semantics. Consequently, any practical reasoning involving our semantics will be *defeasible* [5].
- 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 [6]) 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.

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] 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.
[5] 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.
[6] 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.
[7] John L. Pollock. The logical foundations of means-end reasoning. In Renée Elio, editor, *Common Sense, Reasoning and Rationality*. Oxford University Press, 2002.
[8] Georg Henrik von Wright. Practical inference. *The Philosophical Review*, 72(2):159–179, Apr. 1963.

Technical University of Eindhoven, The Netherlands, J.Hughes@tm.tue.nl