Gerard Allwein, Yingrui Yang
William L. Harrison

QUALITATIVE DECISION THEORY
VIA CHANNEL THEORY

Abstract. We recast parts of decision theory in terms of channel theory concentrating on qualitative issues. Channel theory allows one to move between model theoretic and language theoretic notions as is necessary for an adequate covering. Doing so clarifies decision theory and presents the opportunity to investigate alternative formulations. As an example, we take some of Savage's notions of decision theory and recast them within channel theory. In place of probabilities, we use a particular logic of preference. We introduce a logic for describing actions separate from the logic of preference over actions. The structures introduced by channel theory that represent the decision problems can be seen to be an abstract framework. This framework is very accommodating to changing the nature of the decision problems to handle different aspects or theories about decision making.

Keywords: Decision theory, channel theory, logics of preference, world decision problems.

1. Introduction

Savage [15] showed that certain axioms were able to restrict a decision problem so that a solution involving maximizing the expectation of utility functions was forced. In doing so, he used a certain ontology on different kinds of interrelated entities. Barwise and Seligman [2] demonstrated how logical problems can be organized into components using a modular architecture for logics called (colloquially) channel theory. We formulate the ontological classes of Savage's decision theories in terms of channel theory. In doing so, we divide the work up using three different

logics: classical logic, a logic of actions and a preferential entailment logic. This simplifies the decision theory problem by building a component architecture and isolating and identifying the components.

This paper presents a view of decision theory [9, 8, 15] with a logic of preference taking the place of probabilities. Using Joyce's [9] rendition of Savage's work, we show a consequence relation connecting states and outcomes of actions satisfying Joyce's criteria for such a consequence relation. We do not adopt Savage's restrictions on preferences but instead adopt the preference relations for preferential entailment [3, 11, 10, 5]. We show that this preferential entailment satisfies many of the properties Joyce asserts for a proper conditional. Joyce seems to make a distinction between an entailment as a general conditional and an entailment connecting states and actions. We make this distinction as well but separate the two concerns as component logics within channel theory.

There are two stages to Savage's notion of decision making: refining a decision problem and using quantitative measures to solve a decision problem. This paper addresses the first problem. Joyce and Savage envision there being a sequence of refinements of a small world decision problem leading to a grand world decision problem. We show how to encapsulate each refinement stage as, in the parlance of channel theory, an "information channel". A channel modeling a connection between a small and grand world allows information to flow in either direction.

Section 2 begins with Savage's notion of decision problems [15] using Joyce's [9] approach. Joyce, following Jeffrey [8], formulates a decision theory in terms of a logical system. Joyce's goal is to localize the reasoning so that one can reason about individual actions and states within the same logic. Channel theory achieves the same localization goal with the added advantage that the logic of actions can be different than the logic of states. That is, the two logics need not even be in the same language or have the same axioms and rules. Channel theory does require, however, that one be precisely clear about their relationship. We present the preferential entailment relation as the logic for worlds and show how it fulfills much of Joyce's prescription for such a logic. Section 3 shows how to move information from a small world to grand world, or more aptly move information from a coarse world to a more refined world. The main construction is the notion of simulation taken from modal logic and system theory. The refined world simulates the coarse world just when the refined world's theorems validate the coarse world's theorems. These theorems are entailments of the preferential logic.

2. Decision Theory and Channel Theory

One has a decision problem when one cannot make a decision among alternatives. In explicating this notion, it becomes apparent that "alternatives" might be a euphemism for a (partial) state of the world, an outcome, and an action one could take given the state to reach the outcome. Savage's formalization of the problem is an expected utility computed with a utility function on (outcome,action) pairs weighted subjectively by the decider according to the state.

Decision theory has a long history touching many fields, cognitive science, psychology, economics, etc. One could argue that decision theory should include bits of all these fields. However, the result would quickly get out of hand due to attempting to keep track of too many influences. The result is that abstractions must be used that can be modified depending upon the area of interest. Channel theory is to be thought of with respect to decision problems as a different way to manage the abstractions and ultimately, provide the groundwork for new ones.

Grand World Decision Problem (GWDP) A decision problem rests upon four components: $\mathbf{D} = (\mathbf{\Omega}, \mathbf{A}, \mathbf{S}, \mathbf{O})$. $\mathbf{S}$ is a set of states; each state can be seen as a possible description of the world. $\mathbf{A}$ is a set of action functions, such that for any $f \in \mathbf{A}$, $f : \mathbf{S} \longrightarrow \mathbf{O}$, where $\mathbf{O}$ is a set of outcomes. $\mathbf{\Omega}$ is a partition function. $\mathbf{\Omega}(\mathbf{S})$ is a partition of $\mathbf{S}$; each element in $\mathbf{\Omega}(\mathbf{S})$ is a non-empty subset of $\mathbf{S}$, called an event.[1] The events in $\mathbf{\Omega}(\mathbf{S})$ are mutually exclusive, and collectively exhaustive to $\mathbf{S}$. In symbolic form, for any $E_i, E_j \in \mathbf{\Omega}(\mathbf{S})$ with $i \neq j$, $E_i \cap E_j = \emptyset$, and $\bigcup \mathbf{\Omega}(\mathbf{S}) = \mathbf{S}$. An agent faces a grand world decision problem when there is no proposition X, in $\mathbf{\Omega}$, or otherwise such that the agent strictly prefers $O \wedge X$ to $O \wedge \neg X$ for some outcome $O \in \mathbf{O}$.[2]

Mathematical logic, as most of us learn it, is not modular. In standard presentations of logic (e.g. Mendelson [14]), there is no structural notion of logic "components" and "composition" akin to modules or classes in programming languages. The standard approach presents a logic as a monolith containing a single logical language, a single set

---

[1] An event as a set of states is from probability theory and thought of as the event of those states obtaining.

[2] "Some" should be read as "any", O should be a free variable in this expression and hence universally quantified.

of inference rules and a single class of structures or interpretations for the language. Channel theory is a framework for formulating logical specifications in a modular (i.e., "distributed") manner. The subtitle of Barwise and Seligman's seminal text on channel theory [2] is, in fact, the "logic of distributed systems," where "distributed system" should be understood in the broadest sense. Such a system is distributed in that it is composed of interconnected parts. In particular, it does not refer to the computer science research area known also as distributed systems.

Let us now take this basic GWDP framework and express it in channel theory. The structures of channel theory are deceptively simple. Channel theory is a theory of distributed systems. Here, the distribution is directed by the ontology of states, actions, and outcomes. Each of these ontological classes will be describe by components called *classifications* which model information contexts. The classifications are connected by *infomorphisms*. A classification which is connected to two or more classifications in a particular way is a *channel*.

2.1. Basic Definitions

A classification contains two collections of objects, tokens and types. In this paper, the tokens-types pair will refer to different kinds of entities depending upon what the classification is modeling. At this point, one should not ascribe any preconceived notions to the words "tokens" and "types"; when modeling a particular context, they acquire an external meaning.

Shortly, an entailment relation will be ascribed to types. The relation does not prescribe what kinds of entities are related, merely that they be types. The relation must observe a few minimal rules but it is not the same kind of an entailment relation one finds in a particular logic unless the classification in which it lives is modeling that particular logic. Hence types are usually thought of as propositions, predicates, etc. However, they can be also be sets if one wishes to conceive of the types as constituting a Boolean σ-algebra supporting probability measurements as one would find in work of Savage and Joyce. We do not do so in this paper, but the option exists.

DEFINITION 2.1.1. A *classification*, X, is a pair of sets, $Tok(X)$, and $Typ(X)$, and a relation, $\models_X \subseteq Tok(X) \times Typ(X)$ written in infix, e.g., $x \models_X A$. $x \models_X A$ is the qualitative unit of information that flows in channel theory.

Classifications occur wherever models of systems are found, e.g., the classification of states, Boolean propositions, and $x \models A$ says the token x satisfies the type A; in this example, the tokens are states, the types are Boolean propositions, and $\models$ is the usual satisfaction relation. The term "satisfies" is taken from logic, but $\models$ is merely a relation between tokens and types. To call it satisfaction is to have a convenient name which is slightly less loaded than the term "models".

Another classification is states (as tokens) in a computer and propositions (or first-order logic statements) (as types) with the usual satisfaction relation between them. One can also think of the tokens as being states of the world, actions, messages in a network, etc. They are the sorts of entities that carry information. Generally speaking, types usually have some sort of linguistic, algebraic, and/or set structure and the tokens are modeling structures.

Formally, the objects are the same as Chu spaces [1]. Scott [17] uses similar structures that are specialized for computation. There is also an extensive literature on institutions; Goguen et al. [7] integrates institutions with the work of Barwise and Seligman. An institution is a functor from an abstract category to the category of classifications.

Remark 2.1.2. We will interpret Joyce's **A**, **S**, and **O** with classifications *A*, *S*, and *O*. The former are bold face Roman, the latter are slanted Roman.

Assuming a fixed classification *C*, a *Gentzen sequent*, $\Gamma \Vdash_C \Delta$ is two sets of types connected by a relation $\Vdash$. A *valid* sequent has the force of a meta-level implication of form: for all (normal, see below) tokens x, if $x \models_C A$ for all of the types A in Γ, then $x \models_C B$ for at least one type in Δ. Simple sequents with single left and right sides are written in the form $A \Vdash_C B$. A classification's valid sequents are the classification's *theory*, also called the classification's *constraints*.

DEFINITION 2.1.3. A *local logic* $\mathcal{L} = \langle C, \Vdash_{\mathcal{L}}, N_{\mathcal{L}} \rangle$ consists of a classification *C*, a set $\Vdash_{\mathcal{L}}$ of sequents involving the types of *C*, and a subset $N_{\mathcal{L}} \subseteq Tok(C)$ called the *normal tokens* of $\mathcal{L}$, which satisfy all the constraints $\Vdash_{\mathcal{L}}$. A local logic $\mathcal{L}$ is sound if every token is normal; it is complete if every sequent that holds of all normal tokens is in the consequence relation $\Vdash_{\mathcal{L}}$.

Each classification supports a local logic, including cores of channels (see below). A non-normal token represents a counter-example to the

theory. In this paper, only normal tokens are used. Non-normal tokens can be used to introduce probabilities for types and conditional probabilities associated with the sequents. Typically, the sequents are required to follow certain *structural rules* which we will elide here. Entailment systems of Gentzen [6] had finite left and right hand sides. Scott [16] also uses finite sides. Shoesmith and Smiley [18] introduce infinite sides but we are unsure if this is original with them.

2.2. States, Outcomes, Propositions

Joyce, and before him Jeffrey [8], treat states as propositions. An elegant way to achieve the same end is to construct a classification, S, of states where tokens are states as cross sections of the world or maybe small sections of the world, and types are propositions over states. It is to types in S to which Joyce and Jeffrey would like to refer in talking about states. That is, they code states as infinitary sentences. They do this because they wish to use a notion of entailment for states. Using propositions to talk about states as we do in this paper allows for a continuity of language between traditional philosophy and decision theory.

Another challenge for Joyce and Jeffrey is that states can be quite large. A state can be realized as a predicate in a language, but a state is an abstract thing whereas they attempt to use it just like a predicate. They also wish to use logic with states. There is no guarantee that their states are finitely describable, so they are at least committed to an infinitary language. It makes more sense to simply give the state the abstract nature it deserves (call it a proposition if you like) and allow it to satisfy as many predicates as is necessary. The predicates are the things you can say about a state. One could represent a state as being the collection of all predicates true of it. In channel theory, this can be accommodated by allowing sets of types in entailments or infinitely long formulas.

Outcomes for Joyce and Jeffrey are conjoined collections of coarse outcomes of Savage. They are treated linguistically as conjoined pairs of states and actions. To be ambiguous between the two, we allow outcomes to be anything you like. You can construct them within channel theory to satisfy your own requirements. We treat in this them as abstract, basic entities and allow propositions over collections of outcomes.

2.3. Formulating Decisions as Infomorphisms

Channel theory has its own notion of morphism, called an infomorphism. It is similar to a pair of adjoint functors from category theory [12] in that it is a pair of opposing arrows with a condition similar to the adjoint's bijection.

DEFINITION 2.3.1. An *infomorphism* $\iota: \boldsymbol{X} \longrightarrow \boldsymbol{Y}$ of classifications is a pair of contravariant maps, ι and π such that $\iota: Typ(\boldsymbol{X}) \longrightarrow Typ(\boldsymbol{Y})$ and $\pi: Tok(\boldsymbol{Y}) \longrightarrow Tok(\boldsymbol{X})$, and for all x and A, the following condition is satisfied, $\pi(x) \models_{\boldsymbol{X}} A$ iff $x \models_{\boldsymbol{Y}} \iota(A)$. For ease of presentation, we sometimes use a single letter h for an infomorphism; $h(x)$ is displayed as x^h and $h(A)$ as A^h with the argument determining whether the tokens or types arrow is meant.

Now we give a simple example illustrating how a decision can be encapsulated as an infomorphism. Let $\boldsymbol{S}$ be the classification of propositional logic and its models (states). Let f be a decision which evaluates a state and the agent executing the decision decides to either spend 3 dollars on a salad or 1 on a doughnut. Let $\boldsymbol{O}$ be the classification of outcomes. s^f represents a particular outcome of choosing either a salad or a doughnut. A proposition over outcomes categorizes them. Let one proposition be *EatsHealthy*, call this Q. Q^f is the proposition categorizing all the states in which a healthy food will be chosen. We wish

$$Typ(\boldsymbol{S}) \xleftarrow{\quad f \quad} Typ(\boldsymbol{O})$$

$$\models_S \Big| \qquad\qquad \Big| \models_O$$

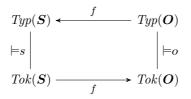$$Tok(\boldsymbol{S}) \xrightarrow[\quad f \quad]{} Tok(\boldsymbol{O})$$

Figure 1. State-Outcome Infomorphism

the infomorphism in Figure 1 to be valid, where both arrows are labeled with f since no confusion can arise if their source and destination are known. The infomorphism condition is then

$$s \models_S Q^f \text{ iff } s^f \models_O Q.$$

Intuitively, s is a state which will induce a healthy choice iff the choice made was a salad (and not a doughnut).

2.4. Connecting Components with Channels

Later on we will use an action classification A to connect a state clas-
sification S with an outcome classification O; this will remove the re-
striction that an action be a function. Connecting classifications in a
particular way constructs a channel in channel theory. To illustrate a
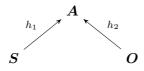channel, consider the following diagrams of a two-sided channel:

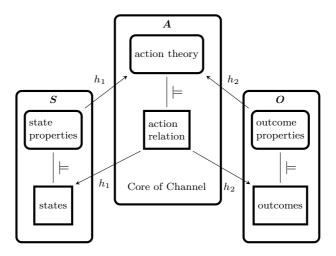Figure 2. Condensed channel diagram

Figure 3. Exploded channel diagram

There are two infomorphisms, h_1 and h_2 (again using the same symbol
for both the types and the tokens maps); the diagram on the left is the
short version. In the diagram on the right, the lower h_i are projections,
and the upper h_i are injections into a disjoint sum. The rule for the
morphisms becomes: $h_i\langle s, o\rangle \models A$ iff $\langle s, o\rangle \models h_i(A)$ with s being a state
and o being an outcome. The state and outcome languages are the type
sets and are allowed to be different. The action (channel) theory contains
the rules for translation in the form of sequents.

A channel's sequents may be used to underwrite information flow through a channel where the pieces of information are tokens and the information they carry are properties. Using the channel in the diagram, let s be a token of $\boldsymbol{S}$, y a token of $\boldsymbol{O}$ and $\langle s, o \rangle$ a token of the channel $\boldsymbol{A}$. Further, let $\Gamma \subseteq Typ(\boldsymbol{S})$ and $\Delta \subseteq Typ(\boldsymbol{O})$ and Γ^{h_1} and Δ^{h_2} refer to the forward images of these sets under h_1 and h_2 respectively. If the sequent $\Gamma^{h_1} \Vdash_{\boldsymbol{A}} \Delta^{h_2}$ as a constraint of the channel, it will relate tokens from $\boldsymbol{S}$ to tokens from $\boldsymbol{O}$ using the following form of reasoning:

$$
\begin{aligned}
s \models_{\boldsymbol{S}} \Gamma \text{ iff } & h_1 \langle s, o \rangle \models_{\boldsymbol{S}} \Gamma & & \text{assuming } h_1 \langle s, o \rangle \text{ for some } o \\
\text{iff } & \langle s, o \rangle \models_{\boldsymbol{A}} \Gamma^{h_1} & & \text{infomorphism condition} \\
\text{implies } & \langle s, o \rangle \models_{\boldsymbol{A}} \Delta^{h_2} & & \text{channel constraint} \\
\text{iff } & h_2 \langle s, o \rangle \models_{\boldsymbol{O}} \Delta & & \text{infomorphism condition} \\
\text{iff } & o \models_{\boldsymbol{O}} \Delta & & \text{assumption}
\end{aligned}
$$

The two-sided channel is technically a cocone in the category of classifications and infomorphisms. All that means is that the diagram on the right has the structure on the left. The definition of a cocone (see Appendix) also allows for arrows between the classifications in the *base*, here the $\boldsymbol{S}$ and $\boldsymbol{O}$. The point or vertex of the cocone is obviously $\boldsymbol{A}$. A cocone in, say, a lattice has the same structure and appears as a join or a least upper bound.

DEFINITION 2.4.1. An *information channel* is a cocone in the category of classifications and infomorphisms. An *information cochannel* is a cone (just reverse the arrows) in the category of classifications and infomorphisms. $\boldsymbol{A}$ in Figure 3 is called the core of the channel. A *channel's theory* refers to the theory in the core. In general, there may be many classifications connected to a core. See Appendix for definitions of cocone and cone.

The smallest channel over a base is a colimit. Frequently, the smallest channel is not the most useful because a channel is used as a model. The smallest channel would simply connect the base with no additional modeling apparatus. A colimit in the category of classifications is a colimit on types and a limit on tokens.

There are mechanisms for moving sequents along an infomorphism. Let $f \colon \boldsymbol{X} \longrightarrow \boldsymbol{Y}$ be an infomorphism, the following rules can be used

with some caveats:

$$\frac{\Gamma \Vdash_X \Delta}{\Gamma^f \Vdash_Y \Delta^f}(f - Intro) \quad \frac{\Gamma^{-f} \Vdash_X \Delta^{-f}}{\Gamma \Vdash_Y \Delta}(f - Intro)$$

$$\frac{\Gamma \Vdash_Y \Delta}{\Gamma^{-f} \Vdash_X \Delta^{-f}}(f - Elim) \quad \frac{\Gamma^f \Vdash_Y \Delta^f}{\Gamma \Vdash_X \Delta}(f - Elim)$$

where Γ^f (Γ^{-f}) is the forward (inverse) image of Γ along f (f^{-1}). There are two (equivalent) versions for each rule. The f-*Intro* rules preserve validity; the f-*Elim* rules preserve non-validity.

Channels and cochannels are used to hold channel logics. A logic in the core of a channel is used to underwrite or authorize information transfer among the side classifications. Cochannels, as they are used in this paper, are used to distribute a common logic in the core to the side classifications.

To model actions, worlds, and connections between worlds, we do not use classical propositional logic (theory) but instead use theories tuned to their roles as models. For the classification of states and the classification for outcomes, we will use classical propositional logic. When we meet up with preferential entailment for worlds, we will use it as the theory of a particular classification. There are two key properties for theories: (1) they can be moved from classification to classification, and (2) the theory in a channel underwrites information flow through a channel.

2.5. States, Outcomes, and Actions

We represent the state set $\mathbf{S}$ with the classification $\boldsymbol{S}$, the tokens are the states in $\mathbf{S}$ and the types are propositions of states. The outcomes $\mathbf{O}$ are represented by the classification $\boldsymbol{O}$, the tokens are the elements of $\mathbf{O}$ and the types are propositions of outcomes. We will model $\mathbf{A}$ as a collection of actions, each action modeled by a channel $\boldsymbol{A}_i$. We will let the logics for $\boldsymbol{S}$ and $\boldsymbol{O}$ be classical propositional logic with $\Vdash$ being the usual entailment relation. The connectives of the logic fill in for the sets of typical sequents and so each sequent will have the form $P \Vdash Q$ with single sides. Notice, there are two logics, one for $\boldsymbol{S}$ and one for $\boldsymbol{O}$ since they have different type sets.

Let the channel $\boldsymbol{A}_i$ connect propositions about states with propositions about outcomes. To model Savage's notion of action as having a single outcome for each state, we let f be an action function. The

following "diagram" is in the category of **Set**; at least the arrows are
arrows of **Set**, the relations, $\models_{\boldsymbol{S}}, \models_{\boldsymbol{A}_i}, \models_{\boldsymbol{O}}$ are clearly not arrows of **Set**:
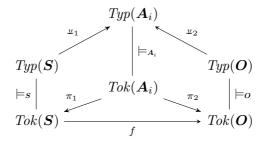

Figure 4. First Approximation of an Action

The tokens are elements of the form $\langle s, f(s) \rangle$. The projections π_i
project either the first or the second elements of these pairs, the lower
triangle commutes. The injections $\bot_i$ (pronounced "ip of i") inject propo-
sitions about states and outcomes into the channel. The channel's theory
has sequents of the form $\bot_1(P) \Vdash_{\boldsymbol{A}_i} \bot_2(Q)$. The flow of information is
left to right when one is looking for the effect of some action on states.
The flow of information, using sequents of the form $\bot_2(Q) \Vdash_{\boldsymbol{A}_i} \bot_1(P)$,
is right to left when one is looking for information about the state given
some outcome under the action. This latter option is available to us but
we do not use it in this paper.

Savage used actual functions to represent actions. A function is not
an infomorphism, so the question remains how to represent this notion
within channel theory. There are two ways to go: (1) a channel as a
cocone of an infomorphism whose arrow on tokens is the action function
and whose types function is a generalization of the inverse of the action
function, and (2) merely a channel.

DEFINITION 2.5.1. Let an *action infomorphism* α_i be appropriate for
an arrow function f just when α_i agrees with α_i on tokens and that α_i
respects the infomorphism condition.

When each α_i is a function on tokens, we let $\boldsymbol{A}_i$ be a colimit where
for an action as an infomorphism $\alpha_i \colon \boldsymbol{O} \longrightarrow \boldsymbol{S}$ (recall the token arrow
of α_i is from $Tok(\boldsymbol{S})$ to $Tok(\boldsymbol{O})$), the limit is over α_i as the base. When
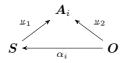the action is a relation on tokens (i.e., not a Savage action), we simply
use a channel.

$$\begin{array}{ccc} & \boldsymbol{A_i} & \\ {}_{\mu_1}\nearrow & & \nwarrow{}_{\mu_2} \\ \boldsymbol{S} & \xleftarrow{\quad \alpha_i \quad} & \boldsymbol{O} \end{array}$$

Figure 5. Single Action as a Colimit

$$\begin{array}{ccc} & \boldsymbol{A_i} & \\ {}_{\mu_1}\nearrow & & \nwarrow{}_{\mu_2} \\ \boldsymbol{S} & & \boldsymbol{O} \end{array}$$

Figure 6. Single Action as a Mere Channel

Actions form a cochannel:

$$\begin{array}{ccccccc} & & \boldsymbol{A} & & & & \\ & {}^{a_i}\swarrow & & \searrow{}^{a_j} & & & \\ \cdots & \boldsymbol{A_i} & \cdots & & \boldsymbol{A_j} & \cdots \end{array}$$
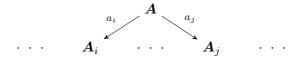
Figure 7. Actions as a Cochannel

where the cardinality of the collection of $\boldsymbol{A_i}$'s is not specified (the i and j are merely indices in some collection, not necessarily natural numbers). The cochannel is a limit which collects together all the actions. The tokens of each $\boldsymbol{A_i}$ are injected into the cochannel with an action label inserted so that $Tok(\boldsymbol{A_i})$ is a disjoint sum, i.e., $\langle s, o \rangle \xmapsto{a_i} \langle s, \alpha_i, o \rangle$. Note that α_i is not a_i, the former is an action as an infomorphism in the deterministic case and a label in the relational case, the latter is the arrow a_i in the cochannel diagram.

There is a choice to made for $Typ(\boldsymbol{A})$. One could choose $Typ(\boldsymbol{A}) = Typ(\boldsymbol{A_i})$ for all i thereby identifying all the type sets $Typ(\boldsymbol{A_i})$. One could also choose $Typ(\boldsymbol{A})$ a subset of the powerset of $Tok(\boldsymbol{A})$. There are other choices but we will choose to the former.

Allowing actions to be relations allows for actions to be partial maps. There is no reason to think an action, say, if modeling personal actions, need have an outcome in every state. Relations also allow for nonde-

terminate outcomes to model imprecise knowledge over the result of an action that could be taken.

Savage uses a mix rule. Using our notation, if α_i and α_j are actions, then the action α_{ij} is defined as $\alpha_i \downarrow_E \alpha_j$ with

$$\langle s, o \rangle \in \alpha_i \downarrow_E \alpha_j \text{ iff } \begin{cases} \langle s, o \rangle \in \mathit{Tok}(\boldsymbol{A}_i) & s \in E \text{ or} \\ \langle s, o \rangle \in \mathit{Tok}(\boldsymbol{A}_j) & s \in \neg E. \end{cases}$$

Consequently, an action channel $\boldsymbol{A}_{ij}$ can be defined using the above prescription. Please see the Appendix for how to do this more formally in the category of channel theory.

2.6. Joyce's Conditions

Joyce states three conditions that an action conditional, $\Rightarrow$, should observe. However, he does this on the basis of some definitions which will need to be interpreted in channel theory.

Definition 2.6.1. P and P^* are contrary when there is no entity x such that $x \models_{\boldsymbol{X}} P$ and $x \models_{\boldsymbol{X}} P^*$ where $P, P^* \in \mathit{Typ}(\boldsymbol{X})$. $P \Vdash_{\boldsymbol{X}} Q$ and $P \Vdash_{\boldsymbol{X}} Q^*$ are contrary if for any $x \models_{\boldsymbol{X}} P$, x satisfies $P \Vdash_{\boldsymbol{X}} Q$ iff x fails to satisfy $P \Vdash_{\boldsymbol{X}} Q^*$ and that there is at least one x for which this is true, i.e., the condition cannot hold vacuously simply because there are no x satisfying P.

Definition 2.6.2. $(P \Rightarrow Q) \wedge (P^* \Rightarrow Q^*)$ means $P \Vdash_{\boldsymbol{X}} Q$ and $P^* \Vdash_{\boldsymbol{X}} Q^*$ are both part of the theory of some ambient classification $\boldsymbol{X}$.

In Joyce [9] pp. 65, a "Savage Conditional" is defined as one that has a state for the antecedent and an outcome for the consequent. Also, "...we cannot adequately formulate Savage's theory without assuming the existence of a conditional $\Rightarrow$ that obeys the three principles in question in cases where the antecedent is a state in **S** and its consequent is an outcome in **O**."

This is easily handed in channel theory.

Definition 2.6.3. A *Savage Conditional* is interpreted in channel theory as a channel $\boldsymbol{A}_i$.

The following properties are required by Joyce for $\Rightarrow$ acting as a linkage between states and outcomes. Our interpretation is use a channel entailment to answer to $\Rightarrow$. The channel entailment provides the connection and no more.

Conditional Contradiction If Q and $Q^\perp$ are contraries and P is not a contradiction, then $P \Rightarrow Q$ and $P \Rightarrow Q^\perp$ are contraries.

Harmony If P and $P^\perp$ are contraries, and neither $P \Rightarrow Q$ nor $P^\perp \Rightarrow Q^\perp$ is a logical falsehood, then $(P \Rightarrow Q) \wedge (P^\perp \Rightarrow Q^\perp)$ is not a logical falsehood.

Conditional Excluded Middle $(P \Rightarrow Q) \vee (P \Rightarrow \neg Q)$ is a logical truth.

There is a oddity in the way Joyce has presented Harmony, there is actually no connection between Q and $Q^\perp$; the symbol $^\perp$ is not an operation. This should instead read

Harmony If P and $P^\perp$ are contraries, and neither $P \Rightarrow Q$ nor $P^\perp \Rightarrow R$ is a logical falsehood, then $(P \Rightarrow Q) \wedge (P^\perp \Rightarrow R)$ is not a logical falsehood.

We will interpret the formula $P \Rightarrow Q$ as $\varkappa_1(P) \Vdash_{A_i} \varkappa_2(Q)$ where $\varkappa_1$ is injecting a state proposition from $Typ(S)$ into the channel A_i and $\varkappa_2$ is injecting an outcome proposition from $Typ(O)$ into the channel.

Conditional Contradiction and Conditional Extended Middle are satisfied merely because actions are represented as channels. Harmony requires the following condition on actions.

DEFINITION 2.6.4. The channel A_i is harmonious iff

$\langle x, o \rangle$ failing to satisfy $\varkappa_1(P) \Vdash_{A_i} \varkappa_2(R)$ implies
$$\exists y, o'(y \equiv_P x \text{ and } \langle y, o' \rangle \text{ satisfies } \varkappa_1(P) \Vdash_{A_i} \varkappa_2(R)).$$

where $y \equiv_P x$ means $y \models_S P$ iff $x \models_S P$.

THEOREM 2.6.5. *A harmonious channel entailment in A_i observes Conditional Contradiction, Harmony, and Conditional Extended Middle.*

PROOF. We only evaluate over the channel, hence there could be states and outcomes that are not connected in the channel to anything. This will be the case if $Tok(A_i)$ is a relation and not the graph of an action function between states and outcomes. Let $\varkappa_2(Q)$ and $\varkappa_2(Q^\perp)$ be contraries in the channel, i.e., $\langle s, o \rangle \Vdash_{A_i} \varkappa_2(Q)$ iff $\langle s, o \rangle \nVdash_{A_i} \varkappa_2(Q^\perp)$. Let $\varkappa_1(P) \Vdash_{A_i} \varkappa_2(Q)$ and $\langle s, o \rangle \Vdash_{A_i} \varkappa_1(P)$ for some $\langle s, o \rangle \in Tok(A_i)$. Hence $s \models P$ from the infomorphism condition and $\langle s, o \rangle \models_{A_i} \varkappa_2(Q)$ from $\varkappa_1(P) \Vdash_{A_i} \varkappa_2(Q)$. If $\varkappa_1(P) \Vdash_{A_i} \varkappa_2(Q^\perp)$ were to hold, then $\langle s, o \rangle \models_{A_i} \varkappa_1(Q)$ and $\langle s, o \rangle \models_{A_i} \varkappa_2(Q^\perp)$. From the infomorphism condition, $o \models_O Q$ and $o \models_O Q^\perp$ which is a contradiction. Since the contrary

definition is symmetric with respect to Q and $Q^\perp$, the symmetric argument with $Q^\perp$ for Q and Q for $Q^\perp$ holds as well and Conditional Contradiction holds.

Let P and $P^\perp$ be contraries and neither $\varkappa_1(P) \Vdash_{\boldsymbol{A}_i} \varkappa_2(Q)$ nor $\varkappa_1(P^\perp) \Vdash_{\boldsymbol{A}_i} \varkappa_2(R)$ are contradicted by all tokens. Hence there is some $\langle x, o \rangle$ that satisfies $\varkappa_1(P) \Vdash_{\boldsymbol{A}_i} \varkappa_2(Q)$. Suppose $\langle x, o \rangle$ fails to satisfy $\varkappa_1(P^\perp) \Vdash_{\boldsymbol{A}_i} \varkappa_2(R)$, then $\langle x, o \rangle \models_{\boldsymbol{A}_i} \varkappa_1(P^\perp)$ and $\langle x, o \rangle \not\models_{\boldsymbol{A}_i} \varkappa_2(R)$. Since $\boldsymbol{A}_i$ is harmonious, there is some y such that y agrees with x on P and y satisfies $\varkappa_1(P^\perp) \Vdash_{\boldsymbol{A}_i} \varkappa_2(R)$. Since $\langle x, o \rangle$ was a counterexample to this sequent, $x \models_{\boldsymbol{S}} P^\perp$ and $x \not\models_{\boldsymbol{S}} P$. So $y \not\models_{\boldsymbol{S}} P$, and $\langle y, o' \rangle$ satisfies $\varkappa_1(P) \Vdash_{\boldsymbol{A}_i} \varkappa_2(Q)$. Hence Harmony holds.

Assume $\langle s, o \rangle$ fails to satisfy $\varkappa_1(P) \models_{\boldsymbol{A}_i} \varkappa_2(Q)$, then $\langle s, o \rangle \models_{\boldsymbol{A}_i} \varkappa_1(P)$ and $\langle s, o \rangle \not\models_{\boldsymbol{A}_i} \varkappa_2(Q)$. From the infomorphism condition, $o \not\models_{\boldsymbol{O}} Q$. Since Q and $Q^\perp$ are contraries, $o \models_{\boldsymbol{O}} Q^\perp$ and hence $\langle s, o \rangle \models_{\boldsymbol{A}_i} \varkappa_2(Q^\perp)$. Since the argument is symmetric, either $\langle s, o \rangle$ satisfies either $\varkappa_1(P) \models_{\boldsymbol{A}_i} \varkappa_2(Q)$ or $\varkappa_1(P^\perp) \models_{\boldsymbol{A}_i} \varkappa_2(Q^\perp)$. Hence Conditional Extended Middle holds $\dashv$

In the last condition, one does not normally join sequents together in a theory, hence the $\vee$ cannot be represented directly. The join must be looked upon as a metalevel theorem about the system. The intent of the $\Rightarrow$ was to formalize the notion of action as a mediator between states and outcomes, there is no need to work with nested $\Rightarrow$. So $\Vdash_{\boldsymbol{A}_i}$ is an adequate gloss on this notion.

Joyce goes on to say that regardless of the other properties one might like to assume for $\Rightarrow$, it must observe these three. In this respect, we posit that using $\Vdash_{\boldsymbol{A}_i}$ in the indicated channel for $\Rightarrow$ adequately addresses Joyce's concerns. Not only that, one is free to ascribe any logic one feels is appropriate for an entailment connective separately in $\boldsymbol{S}$ or $\boldsymbol{O}$. Each of $\boldsymbol{S}$ and $\boldsymbol{O}$ can have their own quite different logic, not the classical propositional logic we have chosen.

2.7. Worlds

Joyce states the grand world decision problem thusly [9]:

$\boldsymbol{D} = (\boldsymbol{\Omega}, \boldsymbol{A}, \boldsymbol{S}, \boldsymbol{O})$ is the *grand world decision problem* that an agent faces if and only if there is no proposition X, whether in $\boldsymbol{\Omega}$ or not, such that she strictly prefers $(O \wedge X)$ to $(O \wedge \neg X)$ for some outcome $O \in \boldsymbol{O}$.

At this point, the minimal requirement on $\mathbf{\Omega}$ is that it be the union, or better, injection of $Typ(\mathbf{S})$, $Typ(\mathbf{A})$, and $Typ(\mathbf{O})$ since those are the only kinds of entities mentioned. We will model $\mathbf{\Omega}$ with a classification $\mathbf{W}$.

We construct $\mathbf{W}$ as the channel on Figure 8.

Classi-fication	Types	Tokens
$\mathbf{S}$	Propositions of states	states
$\mathbf{O}$	Propositions of outcomes	outcomes
$\mathbf{A}_i$	Injections of state and outcome propositions	$\langle s, o \rangle$ for $s \in Tok(\mathbf{S})$, $o \in Tok(\mathbf{O})$
$\mathbf{A}$	Injections of state and outcome propositions	$\langle s, \alpha_i, o \rangle$ for $s \in Tok(\mathbf{S})$, $o \in Tok(\mathbf{O})$, and α_i is the action for $\mathbf{A}_i$
$\mathbf{W}$	Injections of state, actions, and outcome propositions	$\langle s, \alpha_i, o \rangle$ for $s \in Tok(\mathbf{S})$, $o \in Tok(\mathbf{O})$

Figure 8. A World

The channels $\mathbf{A}_i$ and cochannel $\mathbf{A}$ are as described in Section 2.5. There are many channels $\mathbf{A}_i$, we only show one here to keep the diagram manageable. The arrows from $\mathbf{S}$, $\mathbf{A}$, and $\mathbf{O}$ to $\mathbf{W}$ are injections on types and projections on tokens where the projection from $Tok(\mathbf{W})$ to $Tok(\mathbf{A})$ is the identity function.

The set $Typ(\mathbf{W})$ only need hold conjunctions of the form $\pi_1(P) \wedge \pi_2(Q)$ for $P \in Typ(\mathbf{S})$ and $Q \in Typ(\mathbf{O})$ in order to state a grand world decision problem. One can also let $Typ(\mathbf{W})$ consist of all propositions formed over the injected propositions if one is interested in Jeffrey's set up where the logics for these propositions are injected into a common logic.

We have the following condition

$$\langle s, \alpha_i, o \rangle \models_{\boldsymbol{W}} \varphi \text{ iff } \begin{cases} \varphi \equiv \langle \amalg_1(P), \amalg_2(Q) \rangle \text{ and } \langle s, \alpha_i, o \rangle \models_{\boldsymbol{A}} \amalg_1(P) \\ \quad \text{and } \langle s, \alpha_i, o \rangle \models_{\boldsymbol{XA}} \amalg_2(Q) \quad \text{or} \\ \varphi \equiv \amalg_1(P) \text{ and } \langle s, \alpha_i, o \rangle \models_{\boldsymbol{XA}} \amalg_1(P) \quad \text{or} \\ \varphi \equiv \amalg_2(Q) \text{ and } \langle s, \alpha_i, o \rangle \models_{\boldsymbol{XA}} \amalg_2(Q) \end{cases}$$

THEOREM 2.7.1. *The inner triangles of Figure 8 commute.*

2.8. Preferential Entailments

The relation "prefers" in Joyce is taken to mean a partial order: a reflexive, antisymmetric, and transitive binary relation on propositions in $\boldsymbol{\Omega}$. Logics with some sort of preference relation on propositions, rules, and/or interpretations are very familiar to the non-monotonic reasoning community. There are way too many to canvass here. We refer the reader to [4], [13], [10]. In our setup, the preference relation $\prec$ is on $Tok(\boldsymbol{W})$ and we will use it to spawn a $\prec$ relation on $Typ(\boldsymbol{W})$.

We choose to use the preferential entailment relation defined in [3, 11]. They use what is called a rational entailment relation., $\varphi \mathrel{|\!\sim} \psi$. This relation is induced by the semantics for preferential entailment. One cautionary note, [3] have used the converse of the preference relations in [11]. It is merely a matter of taste and we follow [3]; $x \prec y$ if one prefers y to x.

The preference relation is assumed to be *smooth* according to the following prescription:

> For any proposition φ, the set elements of $Tok(\boldsymbol{W})$ satisfying φ is such that there is no infinite ascending chain in the precedence order.

The preference relation will be assumed to be *modular*, then the relation is termed *rational*. A modular relation is a partial order satisfying any of the following equivalent conditions.

LEMMA 2.8.1 ([11]). *If $\leq$ is a partial order on $Tok(\boldsymbol{W})$, the following conditions are equivalent*

1. *for any $x, y, z \in Tok(\boldsymbol{W})$, if $x \nprec y$ and $y \nprec x$ and $x \leq z$, then $y \prec z$.*
2. *for any $x, y, z \in Tok(\boldsymbol{W})$, if $x \prec y$, then either $z \prec y$ or $x \prec z$.*
3. *for any $x, y, z \in Tok(\boldsymbol{W})$, if $x \nprec y$ and $y \nprec z$, then $x \nprec z$.*

4. *there is a totally ordered set Υ (the strict order on Υ will be denoted by $<$) and a function $r\colon Tok(\boldsymbol{W}) \longrightarrow \Upsilon$ (the ranking function) such that $s \prec t$ iff $r(s) < r(t)$.*

The last condition is a bit odd until one realizes what it is really talking about is the *not comparable* relation. Suppose $x \not\prec y$ and $y \not\prec x$, then it cannot be the case that $r(x) < r(y)$ or $r(y) < r(x)$. Since Υ is a total order, it must be that $r(x) = r(y)$. So the picture of the original order is that it looks like a string of pearls (not a loop) where each pearl is a collection of pairwise incomparable elements. If any one element x is less that an element y of another pearl, then all the elements incomparable to x are less than y. The last condition shows that rational preference relations come very close to Savage's original conception of the preference relation being a total order.

Definition 2.8.2. Given a classical propositional language, a preferential model $\langle M, l, \prec \rangle$ is a collection of models M, a function l from models to valuations, and a smooth, modular partial order.

We will follow [3] in using $x \models_{\boldsymbol{W}} \varphi$ for the valuation $l(x)$ satisfying φ. The semantics of the preferential consequence relation relies on maximal tokens in the $Tok(\varphi)$, where $Tok(\varphi)$ is the set $\{x \mid x \in Tok(\boldsymbol{W})$ and $x \models_{\boldsymbol{W}} \varphi\}$, for proposition φ. The smoothness condition will force any non-empty collection of tokens satisfying a proposition to have non-empty collection of maximal elements:

$$\max(\varphi) \stackrel{\mathrm{df}}{=} \{x \mid x \models \varphi \text{ and for all } y, x \prec y \text{ implies } x = y \text{ or } y \not\models \varphi\}.$$

It turns out there is another preferential consequence relation which is the contra-positive of the $\mathbin{|\!\sim}$ and it will require the following function

$$\overline{\max}(\varphi) = M - \max(\neg\varphi).$$

The two consequence relations are defined as follows

$$\varphi \mathbin{|\!\sim} \psi \text{ iff } \max(\varphi) \subseteq Tok(\psi), \quad \varphi \mathbin{|\!\sim}^* \psi \text{ iff } Tok(\psi) \subseteq \overline{\max}(\varphi).$$

It then turns out that

$$\neg\varphi \mathbin{|\!\sim} \neg\psi \text{ iff } \max(\neg\varphi) \subseteq M - Tok(\psi)$$
$$\text{iff } Tok(\psi) \subseteq M - \max(\neg\varphi)$$
$$\text{iff } Tok(\psi) \subseteq \overline{\max}(\varphi)$$
$$\text{iff } \psi \mathbin{|\!\sim}^* \varphi.$$

The preferential consequence relations determined by a preferential model satisfy the following:

(1) Reflexivity $\varphi \mathrel{\vert\!\sim} \varphi$ $\varphi \mathrel{\vert\!\sim}^* \varphi$

(2) And $\dfrac{\varphi \mathrel{\vert\!\sim} \psi \quad \varphi \mathrel{\vert\!\sim} \theta}{\varphi \mathrel{\vert\!\sim} \psi \wedge \theta}$ $\dfrac{\varphi \mathrel{\vert\!\sim}^* \psi \quad \varphi \mathrel{\vert\!\sim}^* \theta}{\varphi \mathrel{\vert\!\sim}^* \psi \wedge \theta}$

(3) Or $\dfrac{\varphi \mathrel{\vert\!\sim} \theta \quad \psi \mathrel{\vert\!\sim} \theta}{\varphi \vee \psi \mathrel{\vert\!\sim} \theta}$ $\dfrac{\varphi \mathrel{\vert\!\sim}^* \theta \quad \psi \mathrel{\vert\!\sim}^* \theta}{\varphi \vee \psi \mathrel{\vert\!\sim}^* \theta}$

(4) Logical Equivalence

$$\dfrac{\models \varphi \leftrightarrow \psi \quad \varphi \mathrel{\vert\!\sim} \theta}{\psi \mathrel{\vert\!\sim} \theta}$$

$$\dfrac{\models \varphi \leftrightarrow \psi \quad \theta \mathrel{\vert\!\sim}^* \varphi}{\theta \mathrel{\vert\!\sim}^* \psi}$$

(5) Right Weakening
Monotonicity

$$\dfrac{\varphi \mathrel{\vert\!\sim} \psi \quad \models \psi \to \theta}{\varphi \mathrel{\vert\!\sim} \theta} \text{ Right W.}$$

$$\dfrac{\psi \mathrel{\vert\!\sim}^* \theta \quad \models \varphi \to \psi}{\varphi \mathrel{\vert\!\sim}^* \theta} \text{ Mono.}$$

(6) Cautious Monotonicity
Cautious Right Weakening

$$\dfrac{\varphi \mathrel{\vert\!\sim} \psi \quad \varphi \mathrel{\vert\!\sim} \theta}{\varphi \wedge \psi \mathrel{\vert\!\sim} \theta} \text{ Caut. Mono.}$$

$$\dfrac{\varphi \mathrel{\vert\!\sim}^* \psi \quad \theta \mathrel{\vert\!\sim}^* \psi}{\varphi \mathrel{\vert\!\sim}^* \psi \vee \theta} \text{ Caut. R. W.}$$

(7) Rational Monotonicity
Rational Right Weakening

$$\dfrac{\varphi \mathrel{\vert\!\sim} \psi \quad \varphi \mathrel{\not\vert\!\sim} \neg\theta}{\varphi \wedge \theta \mathrel{\vert\!\sim} \psi} \text{ Rat. Mono.}$$

$$\dfrac{\varphi \mathrel{\vert\!\sim}^* \psi \quad \neg\theta \mathrel{\not\vert\!\sim}^* \psi}{\varphi \mathrel{\vert\!\sim}^* \psi \vee \theta} \text{ Rat. R. W.}$$

where $\models \varphi \leftrightarrow \psi$ means $\varphi \leftrightarrow \psi$ is true everywhere.

The rational preferential entailment almost satisfies a list of eight general requirements in Joyce [9] on a conditional, in addition to Conditional Contradiction, Harmony, and Conditional Extended Middle. However, the latter three conditions are explained just to cater to actions whereas the eight are to be general requirements. We have separated out the general requirements as being mostly satisfied by preferential entailment and relegated the latter three just for actions.

Joyce's eight requirements are stated as first-degree entailments and hence we will interpret his conditional $\Rightarrow$ as rational preferential entailment. The first condition he states as modus ponens whereas our first condition is slightly weaker. The rest are almost verbatim were we interpret $\Rightarrow$ as $\mathrel|\!\sim$ and & as meta-level conjunction:

Cond$_1$ (*Guarded Entailment*) If $\varphi \wedge \psi \mathrel|\!\sim \theta$, then $\varphi \mathrel|\!\sim \psi \to \theta$

Cond$_2$ (*Centering*) If $\models \varphi$ and $\varphi \mathrel|\!\sim \psi$ is logically equivalent to $\varphi \wedge \psi$.

Cond$_3$ (*Weakening the Consequent*) $\varphi \mathrel|\!\sim \psi$ implies $\varphi \mathrel|\!\sim \psi \vee \theta$.

Cond$_4$ (*Conditional Conjunction*) $\varphi \mathrel|\!\sim \psi$ and $\varphi \mathrel|\!\sim \theta$ implies $\varphi \mathrel|\!\sim \psi \wedge \theta$.

Cond$_5$ (*Dilemma*) $\varphi \mathrel|\!\sim \theta$ and $\psi \mathrel|\!\sim \theta$ implies $\varphi \vee \psi \mathrel|\!\sim \theta$.

Cond$_6$ (*Weak Strengthening of the Antecedent*) $\varphi \mathrel|\!\sim \psi$ and $\varphi \mathrel|\!\!\not\sim \neg\theta$ implies $\varphi \wedge \theta \mathrel|\!\sim \psi$.

Cond$_7$ (*Reductio*) $\varphi \mathrel|\!\sim \psi \wedge \neg\psi$ implies $\theta \mathrel|\!\sim \varphi$.

Cond$_8$ (*Conditional Equivalence*) $\varphi \mathrel|\!\sim \psi$ and $\psi \mathrel|\!\sim \varphi$ implies $\varphi \mathrel|\!\sim \theta$ iff $\psi \mathrel|\!\sim \theta$.

Joyce's first condition reads φ entails ψ implies $\varphi \Rightarrow \psi$. The guard condition says that the left side can never be left empty. This is a result of an empty left side not being equivalent to *true*. Instead, an empty left side acts like $\bigcup\{\max(Tok(\varphi))\}$ where φ ranges over all propositions of the classification holding the token set.

THEOREM 2.8.3. *Rational preferential entailment satisfies* Cond$_1$ *and* Cond$_3$ $-$ Cond$_8$.

The relation $\prec$ can be extended to cover actual propositions via the following auxiliary relation

$\varphi \prec \psi$ iff for all $x \models_X \varphi$ there exists some y such that $y \models_X \psi$ and $x \prec y$.

The auxiliary $\prec$ relation has the following relationship with $\mathrel|\!\sim$:

THEOREM 2.8.4.
$$\frac{\varphi \mathrel|\!\sim \psi \quad \psi \prec \theta}{\varphi \prec \theta}$$

The auxiliary $\prec$ relation can be moved along infomorphisms:

THEOREM 2.8.5. *For* $k\colon \boldsymbol{X} \to \boldsymbol{Y}$, $\varphi \prec \psi$ *implies* $\varphi^k \prec \psi^k$ *if infomorphism* k *satisfies*

$$x^k \prec y \text{ implies } \exists z \ (z^k = y \text{ and } x \prec z).$$

We will use the preferential entailment relation for the classically defined entailment $\Vdash$ when we construct grand and small worlds in the sequel. One feature of $\Vdash$ is that it is preserved under infomorphism. To preserve $\mathrel{|\!\!\sim}$, a further condition on infomorphisms is required:

THEOREM 2.8.6. *An infomorphism* $k\colon \boldsymbol{X} \to \boldsymbol{Y}$ *preserves the preferential entailment relation in the sense of* $\varphi \mathrel{|\!\!\sim} \psi$ *implies* $\varphi^k \mathrel{|\!\!\sim} \psi^k$ *if*

$$x \in \max(\varphi^k) \text{ implies } x^k \in \max(\varphi).$$

THEOREM 2.8.7. *An infomorphism* $k\colon \boldsymbol{X} \to \boldsymbol{Y}$ *preserves the dual preferential entailment relation in the sense of* $\varphi \mathrel{|\!\!\sim^*} \psi$ *implies* $\varphi^k \mathrel{|\!\!\sim^*} \psi^k$ *if in addition to preserving* $\mathrel{|\!\!\sim}$, k *commutes with* $\neg$, *i.e.,* $Tok(\neg(\varphi^k)) = Tok((\neg\varphi)^k)$.

These results show that the usual entailment $\Vdash$ of channel theory can be replaced with preferential entailment $\mathrel{|\!\!\sim}$.

3. Moving Between Worlds

A Grand World Decision Problem (GWDP) can be thought of as the problem one has making a decision given all information. A Small World Decision Problem (SWDP) is the problem one has when some information is unknown. Savage envisioned there to be a series of decision problems in between at different levels of information. The question arises as to how well the SWDP represents the relevant information in the sense that a decision made assuming the small amount of information would change the extra information in the GWDP were available.

The distinction between grand and small worlds in Savage is couched in some rather draconian conditions for relating the two. We think that it makes sense to open up the relationship a bit and we do this by forming a channel. One can then talk about reasoning in one world using the devices of the other. We capture this by using the notion of a simulation from modal logic but altered to work for channel theory and decision theory. Our treatment is one sided in that the grand world simulates the small world. One can run the simulation the other way or use two simulations in the form a bisimulation. These latter two can easily be obtained by seeing how it works in the one-sided case. We use the one-sided case due to space considerations.

The appellation "small" is somewhat of a misnomer, it might better be termed "smaller". In Savage, decision making is a two-stage process. Stage 1 is to refine a decision problem, which is syntactic stage. Stage 2 is to solve the problem using utility functions. The analysis in this section should be seen as opening up Stage 1 to more logical devices. Joyce notes that some researchers begin with grand worlds and refine to get small worlds, and others begin with small worlds and generalize to get grand worlds. A channel between a world and smaller world should be seen a device for bridging the gap between worlds.

3.1. Moving Actions from Grand Worlds to Small Worlds

Let GA_i and SA_j be grand world and small world action channels respectively. The grand world action simulates the small world action just when there is the following channel configuration where all arrows are projection, injection pairs:

Figure 9. Grand World Action Simulates Small World Action

where SS and GS are the small and grand world state classifications, SO and GO are the small and grand world outcome classifications, CS and CO are channels, and $\rho_k, \tau_k, \nu_k, \sigma_k$ are injection projection pairs and with the following conditions (note: we will use the same function name for injection as projection and let its domain determine whether the injection on types or projection on tokens is meant):

C1: The CS theory has sequents of the form $\rho_1(P) \Vdash_{CS} \rho_2(P')$ where P is injected from $Typ(SS)$ and P' is injected from $Typ(GS)$.

C2: The $\boldsymbol{CO}$ theory has sequents of the form $\sigma_1(Q') \Vdash_{\boldsymbol{CO}} \sigma_2(Q')$ where Q is injected from $Typ(\boldsymbol{SO})$ and Q' is injected from $Typ(\boldsymbol{GO})$.

C3: The projection ρ_1 must cover $Tok(\boldsymbol{SS})$, i.e., $x \models_{\boldsymbol{SS}} P$ implies there exists some y and $\langle x, y \rangle \in Tok(\boldsymbol{CS})$.

C4: The two channels $\boldsymbol{CS}$ and $\boldsymbol{CO}$ satisfy

Figure 10. Simulation Condition

which expresses the condition that for $x \in Tok(\boldsymbol{SS})$, $y \in Tok(\boldsymbol{GS})$, $p \in Tok(\boldsymbol{SO})$

$\langle x, y \rangle \in Tok(\boldsymbol{CS})$ and $x\alpha_j p$ implies
$$\exists q \ \in Tok(\boldsymbol{GO}) \ (y\alpha_i q \text{ and } \langle p, q \rangle \in Tok(\boldsymbol{CO})).$$

Note that α_j is referring to $Tok(\boldsymbol{SA}_j)$ and α_i to $Tok(\boldsymbol{GA}_i)$ as binary relations in infix form, this being less clumsy than $\langle x, p \rangle \in Tok(\boldsymbol{A}_j)$ and $\langle y, q \rangle \in Tok(\boldsymbol{A}_i)$.

THEOREM 3.1.1. *If grand world action α_i simulates the small world action α_j, $\rho_1(P) \Vdash_{\boldsymbol{CS}} \rho_2(P')$ and $\sigma_2(Q') \Vdash_{\boldsymbol{CO}} \sigma_1(Q)$, then the sequent $\nu_1(P') \Vdash_{\boldsymbol{GA}_i} \nu_2(Q')$ validates the sequent $\tau_1(P) \Vdash_{\boldsymbol{SA}_j} \tau_2(Q)$.*

PROOF. Assume $x \models_{\boldsymbol{SS}} P$ and $x\alpha_j p$, i.e., $\langle x, p \rangle \in Tok(\boldsymbol{SA}_j)$ and $\tau_1 \langle x, p \rangle = x$. Since $Tok(P)$ is covered by ρ_1, there is some y such that $\langle x, y \rangle \in \boldsymbol{CS}$. From the infomorphism condition, $\langle x, y \rangle \models_{\boldsymbol{CS}} \rho_1(P)$. From the theory in the channel $\boldsymbol{CS}$, $\rho_1(P) \Vdash_{\boldsymbol{CS}} \rho_2(P')$ and hence $\langle x, y \rangle \models_{\boldsymbol{CS}} \rho_2(P')$. From the infomorphism condition, $\rho_2 \langle x, y \rangle = y$ and $y \models_{\boldsymbol{GS}_i} P'$. The premises of C4 are met and so there is some q such that $y\alpha_i q$ and $\langle p, q \rangle \in Tok(\boldsymbol{CO})$. Using the channel $\boldsymbol{GA}_i$, $\langle y, q \rangle \in \boldsymbol{GA}_i$ and so $\nu_1 \langle y, q \rangle = y$ and $\langle y, q \rangle \models_{\boldsymbol{GA}_i} \nu_1(P')$. Using the condition $\nu_1(P') \Vdash_{\boldsymbol{GA}_i} \nu_2(Q')$, $\langle y, q \rangle \models_{\boldsymbol{GA}_i} \nu_2(Q')$. From the infomorphism condition, $q \models_{\boldsymbol{GO}} Q'$. Now use similar reasoning back through the $\boldsymbol{CO}$ channel using $\sigma_2(Q') \Vdash_{\boldsymbol{CO}} \sigma_1(Q)$ to conclude that $p \models_{\boldsymbol{SO}} Q$. $\dashv$

3.2. Moving Preferential Entailment Relations

We will use the following channel diagram to connect grand and small worlds. The channel C does the actual connection. The classification SW is the small world and GW is the grand world:
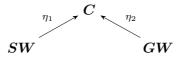

Figure 11. Grand World Simulates Small World

The types of C are injections of $Typ(SW)$ and $Typ(GW)$ and the tokens are a subset of $Tok(SW) \times Tok(GW)$, i.e., a relation. The tokens are the unions of the simulation relations between the pairs SA_j and GA_i that have been related by simulations. The $\models_C$ relation is determined by the infomorphisms from SW and GW. The token maps for η_i are the left and right projections.

To transfer the preferential entailment relation from the grand world to the small world, we will need a few conditions.

D1: The C theory has sequents of the form $\eta_1(\varphi) \Vdash_C \eta_2(\varphi')$ and $\eta_2(\psi') \Vdash_C \eta_1(\psi)$

D2: The projection η_1 must cover $Tok(SW)$, i.e., $x \models_{SW} \varphi$ implies there exists some y and $\langle x, y \rangle \in Tok(C)$.

D3: The channel C is such that $\langle x, y \rangle \in Tok(C)$ and $y \prec y'$ implies $\langle x, y' \rangle \in Tok(C)$.

THEOREM 3.2.1. *If the grand world simulates the small world, $\eta_1(\varphi) \Vdash_C \eta_2(\varphi')$, and $\eta_2(\psi') \Vdash_C \eta_1(\psi)$, then $\varphi' \mathrel{\vdash\mkern-7mu\sim}_{GW} \psi'$ validates $\varphi \mathrel{\vdash\mkern-7mu\sim}_{SW} \psi$.*

4. Conclusions and Future Work

We have shown that channel theory is very adaptable to the ontology of decision theory and flexible enough to add some generality to the structures. Our use of a preferential entailment logic is related to [5]. They fitted a cut down version of preferential entailment into the rather rigid structures of Savage. Our use was to simply adopt the full preferential entailment and ask what must change qualitatively in order for it to work as a logic of decisions.

Decision theory via channel theory can be understood as introducing a meta-theoretic view on decision theory. Channel theory is a way to formulate the overall structure but allow for changing the pieces. We do not go so far as to define quantitative functions forced by the pieces, but much of Savage, Jeffrey, and Joyce's quantitative work can also be expressed in the new architecture. We have represented actions as relations rather than as functions for the sake of additional generality. To capture Savage's framework, one may always restrict the action relations to be the graphs of functions, restrict the small world outcomes to be collections of grand world actions, small world states to be equivalence classes of grand world states, and small world events to be disjunctions of grand world events. In place of the preference logic we use, he would use a different logic, but there would still be a logic. One may similarly restrict the channel theoretic formulation for Jeffrey's theory by identifying the interpretations of states, actions, and outcomes as collections of the same entities.

The use of channel theory is mainly to relate small and grand worlds. In doing so, we found that the stock entailment logic provided by Barwise-Seligman's version of channel theory [2] had to be changed. This represents an addition to channel theory, i.e., allowing for the theory of a classification to be variable. Our use of cochannels, while in Barwise-Seligman, stems from realizing that cochannels work well in collecting common entailments which are then distributed to the legs.

The expansion of the notion of action to a relation allows for partiality and nondeterminism. These notions seems to be somewhat overlooked within decision theory since decision theory is very tied to quantitative modeling. It is plausible that an action need not have an output at every state for the simple reason that some actions can easily be prohibited in some states, say, in turning left at an intersection where there is no road leading to the left. Nondeterminism can arise when full knowledge of alternatives is lacking, say, in a lottery drawing when any one of several possibilities might occur. One can always add more and more information to the set up until full determinism is reached, but the resulting theory is then so large and disjoint from common sense as to not be very helpful.

The quantitative aspect of decision theory has not been treated in this paper as we were exploring what kinds of qualitative structures must be accommodated in order to have a theory that is not too divorced from everyday decisions. One of our goals is to now explore the interaction

between decision theories by translating more of them into channel theory where their assumptions can be compared in terms of their effect on a common structure. This is very close to Joyce's, and before him, Jeffrey's, outlook. They translated into a common logic. The problem we see with that approach is that every proposition no matter whether it is a proposition of states or actions must be interpreted by the same collection of models. This does not allow for building decision theories with components. There is no reason to think a logic of actions should be the same a logic of propositions.

Channel theory allows for one to choose logics for each component based on the job that component is there to model. Channel theory then forces one to be explicit about the interactions of the components. We think this is an exciting new way to view decision theory. The components or ontology of decision theory is well-known. We chose to take the components seriously as components that combine to form a system. In doing so within channel theory, we now have a tool box upon which quantitative measures can be applied and their commitments evaluated.

5. Appendix

5.1. Cocones and Cones

A commuting *finite cocone* consists of a graph homomorphism G from a finite graph to the category of classifications, a vertex classification $\boldsymbol{C}$, and a collection of arrows $g_i \colon G(i) \longrightarrow \boldsymbol{C}$.
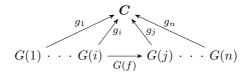

Figure 12. Cocone of Infomorphisms

It is required that for all $f \colon i \longrightarrow j$, $g_i = g_j \circ G(f)$. The base of the cocone is the objects and arrows identified by G. There will also be a need for cones: a *finite cone* consists of a graph homomorphism G from a finite graph to the category of classifications, a vertex classification $\boldsymbol{C}$, and a collection of arrows $g_i \colon G(i) \longrightarrow \boldsymbol{C}$. It is required that for all $f \colon j \longrightarrow i$, $g_i = G(f) \circ g_j$. Just reverse all the arrows in the diagram.

5.2. Mix

The exposition in Joyce is a bit hard to follow; using the notation of this paper, if α_i and α_j are actions, then $\alpha_1 \downarrow_E \alpha_j$ defined

$$\alpha_i \downarrow_E \alpha_j = \begin{cases} \alpha_i(s) & s \models_{\boldsymbol{S}} E \\ \alpha_j(s) & s \models_{\boldsymbol{S}} \neg E \end{cases}$$

Notice that this assumes $Typ(\boldsymbol{S})$ has at least a Boolean negation. The collection of arrows is to closed under this construction for any α_i and α_j.

5.2.1. Mix Using Action Infomorphisms

Each $\boldsymbol{A}_i$ is a colimit where for an action as an infomorphism $\alpha_i \colon \boldsymbol{O} \longrightarrow \boldsymbol{S}$, the colimit is over α_i as the base:
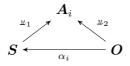

Figure 13. Single Action as a Channel

and has a rather pretty exposition in channel theory: given the two restricted actions in Figure 14, there is a pushout in the category of classifications in Figure 15:
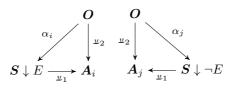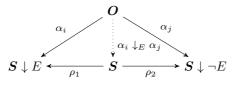

Figure 14. Two Restricted Actions

Figure 15. The Constructed Action

where $\alpha \downarrow_E \beta$ is the unique arrow which is a product in the category of tokens and a coproduct in the category of types. The combined action is the colimit channel
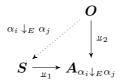
$$
\begin{array}{ccc}
 & & \boldsymbol{O} \\
{\scriptstyle \alpha_i \downarrow_E \alpha_j} \nearrow & & \downarrow {\scriptstyle \amalg_2} \\
\boldsymbol{S} & \xrightarrow[\amalg_1]{} & \boldsymbol{A}_{\alpha_i \downarrow_E \alpha_j}
\end{array}
$$

Figure 16. Action as a Colimit

5.2.2. Mix Using Just Action Channels

The conditions are on the left for the basic diagram in Figure 17:

$$
\begin{aligned}
Tok(\boldsymbol{A}_i \downarrow E) = \{\langle s, o \rangle \mid \langle s, o \rangle \in Tok(\boldsymbol{A}_i) \text{ and} \\
s \in Tok(E)\} \\
Tok(\boldsymbol{A}_j \downarrow \neg E) = \{\langle s, o \rangle \mid \langle s, o \rangle \in Tok(\boldsymbol{A}_j) \text{ and} \\
s \in Tok(\neg E)\},
\end{aligned}
$$

and

$$
\langle s, o \rangle \in Tok(\boldsymbol{A}_{ij}) \text{ iff } \begin{cases} \langle s, o \rangle \in Tok(\boldsymbol{A}_i) & s \in E \text{ or} \\ \langle s, o \rangle \in Tok(\boldsymbol{A}_j) & s \in \neg E. \end{cases}
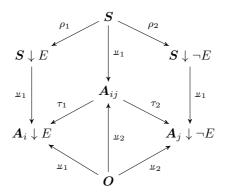$$

Figure 17. Single Action as a Channel

where $Tok(\boldsymbol{S} \downarrow E) = Tok(\boldsymbol{S}) \cap Tok(E)$ and similarly for $\boldsymbol{S} \downarrow \neg E$,

Hence $\tau_1 \colon \boldsymbol{A}_{ij} \longrightarrow \boldsymbol{A}_i$ injects $Tok(\boldsymbol{A}_i)$ into $Tok(\boldsymbol{A}_{ij})$. τ_2 is a similar injection on tokens. Both are the identity function on types. The ρ_i are injections on tokens and identity functions on types. This has the effect of constructing $Tok(\boldsymbol{S})$ as the union $Tok(E) \cup Tok(\neg E)$.

Any arrow labeled with $\underline{\nu}_i$ works just as it did before, i.e., injecting types into a disjoint sum and projecting tokens into their constituents.

References

[1] Barr, M., "∗-autonomous categories and linear logic", *Mathematical Structures in Computer Science* 1 (1991): 159–178.

[2] Barwise, J., and J. Seligman, *Information Flow: The Logic of Distributed Systems*, Cambridge University Press, 1997, Cambridge Tracts in Theoretical Computer Science 44.

[3] Britz, K., J. Heidema, and W. Labuschagne, "Semantics for dual preferential entailment", *Journal of Philosophical Logic* 38 (2009): 433–446.

[4] Delgrande, J., T. Schaub, H. Tompits, and K. Wang, "A classiŢcation and survey of preference handling approaches in non-monotonic reasoning", *Computational Intelligence* 20 (2004): 308–334.

[5] Dubois, D., H. Fargier, H. Prade, and P. Penny, "Qualitative decision theory: From savage's axioms to nonmonotonic reasoning", *Journal of Association of Computing Machinery* 49 (2002): 455–495.

[6] Gentzen, G., "Untersuchungen über das logische Schliessen", *Mathematische Zeitschrift* 39 (1934): 176–210, 405–431.

[7] Goguen, J., "Information integration in institutions" pages 1–48 in: L. Moss, editor, *Thinking Logically: a Memorial Volume for Jon Barwise*, Indiana University Press, 201x.

[8] Jeffrey, R. C., *The Logic of Decision*, University Of Chicago Press, first edition, 1990.

[9] Joyce, J., *The Foundations of Causal Decision Theory*, Cambridge University Press, 1999.

[10] Kraus, S., D. Lehmann, and M. Magidor, "Nonmonotonic reasoning, preferential models and cumulative logics", *Artificial Intelligence* 44 (1990): 167–207.

[11] Lehmann, D., and M. Magidor, "What does a conditional knowledge base entail?", *Artificial Intelligence* 55, 1992.

[12] Mac Lane, S., *Categories for the Working Mathematician*, volume 5, Springer-Verlag, second edition, 1998.

[13] Makinson, D., "How to go nonmonotonic", pages 175–278 in: *Handbook of Philosophical Logic*, vol. 12, Springer, second edition, 2005.

[14] Mendelson, E., *Introduction to Mathematical Logic*, Shapman and Hall, fourth edition, 1997.

[15] Savage, L., *The Foundations of Statistics*, Dover, second edition, 1972.

[16] Scott, D. S., "Background to formalization", pages 411–435 in: *Truth, Syntax, and Modality*, North-Holland, 1973.

[17] Scott, D. S., "Domains for denotational semantics", pages 1–47 in: *An extended version of the paper prepared for ICALP'82*, Springer-Verlag, 1982.

[18] Shoesmith, D. J., and T. J. Smiley, *Multiple-Conclusion Logic*, Cambridge University Press, 1978.

GERARD ALLWEIN
Center for High Assurance Computer Systems, Code 5540
Naval Research Laboratory
Washington, DC 20375, USA
allwein@itd.nrl.navy.mil

YINGRUI YANG
Department of Cognitive Science
Rensselaer Polytechnic Institute
Troy, New York, USA
yangyri @ rpi.edu

WILLIAM L. HARRISON
Department of Computer Science
University of Missouri
Columbia, Missouri, USA
harrisonwl@missouri.edu